Lagrange–d'Alembert Integrators

A Thesis Submitted to the

College of Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the degree of Doctor of Philosophy

in the Department of Mathematics and Statistics

University of Saskatchewan

Saskatoon

By

Charles L. Cuell

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics and Statistics
University of Saskatchewan
106 Wiggins Road
Saskatoon, SK S7N 1W7
Canada

Abstract

A Lagrange–d'Alembert integrator is a geometric numerical method for finding numerical solutions to the Lagrange–d'Alembert equations for mechanical systems with nonholonomic constraints that are linear in the velocities. The integrator is developed from geometry and principles that are analogues of the continuous theory.

Using discrete analogues of the symplectic form and momentum map, the resulting methods are symplectic and momentum preserving whenever the continuous system is symplectic and momentum preserving. In addition, it is possible to, in principle, generate Lagrange–d'Alembert integrators of any method order.

Acknowledgements

I would like to thank Prof. George Patrick for supervising the work done in this thesis. My wife, Sally, deserves praise for her infinite patience and support, as do my children Samuel, Margaret and Kathleen.

For financial support:

- The department of Mathematics and Statistics at the University of Saskatchewan for teaching assistantships and sessional lectureships.
- Prof. George Patrick for financial assistance.
- Dawn and John Anderson for financial assistance, apple pies and expert babysitting.
- Bill and Dorothy Cuell for computers and take–out meals.

For me, because nobody else would want it.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The purpose of this thesis is to describe a new class of numerical simulation methods for constrained mechanical systems. The methods, called *Lagrange–d'Alembert integrators*, include variational integrators in the case that there are no constraints. Marsden and West [18] give a broad overview of variational integrators and a large number of examples.

Standard integration methods rely on a discretization of the equations of motion, being essentially discrete analogues of Newton's second law of motion. Lagrange–d'Alembert integrators, on the other hand, are derived from a discrete analogue of the Lagrange–d'Alembert principle (Marsden, Patrick, Shkoller [17]). As such, a Lagrange–d'Alembert integrator may be discretized *without the equations of motion.* This paradigm treats Lagrange–d'Alembert integrators as resulting from a fundamental principle so that the development of the theory proceeds without reference to continuous Lagrangian systems.

The advantage of this approach is that geometric structures present in continuos systems are also present in discrete Lagrange–d'Alembert systems. For example, discrete Lagrange–d'Alembert systems have an associated symplectic form that evolves in the same manner as the symplectic form associated to continuos systems. In the case where the discrete symplectic form is preserved by the discrete evolution, there is also preservation of phase volume. If the discrete system exhibits a symmetry, then there is a discrete momentum equation.

In practice, Lagrange–d'Alembert integrators exhibit long time energy and, if applicable, momentum "near" conservation. That is, the energy and momentum of the continuous system are not exactly conserved, but oscillate about the true value and remain bounded for very long time integrations.

This thesis contributes new theory and results in the following ways:

1. The discretization is explicitly studied and strongly tied to the definition of a *discrete tangent bundle*.

2. The theory and integrators are set in phase space, whereas the usual setting is a direct product of configuration space. The phase space setting clarifies the link between the continuous and discrete systems.

3. Numerical methods of arbitrary order can be developed, with higher order methods no more difficult to set up than lower order methods.

4. A discrete analogue of the semi-Hamilton equations are presented. See Bates and Sniatycki [4] for the distributional version and Patrick [20] for the variational development and a large set of references.

5. Discrete holonomic subsystems are introduced, in analogy to continuous holonomic subsystems as developed by Patrick [20].

6. A discrete momentum equation is derived that reduces to the one given by Cortés and Martínez [7] for integrators of Moser–Veselov type.

The fundamental discretization is achieved by approximating the action function for the Lagrangian, L, over curves γ in configuration space Q

$$S(\gamma) = \int_a^b L(\gamma'(t))\, dt,$$

to

$$S_d(q_d) = \sum_{k=0}^{N} L_h(v_k)\, h,$$

where q_d is a sequence of points in Q and the v_k form a sequence in TQ by taking a *discrete derivative* of q_d (Chapter 3, Definition 4). The form of L_h is the key to the accuracy of the integrator.

The continuous Lagrange–d'Alembert principle for systems constrained by a distribution $\mathcal{D}$ states that the curve γ is a physical trajectory if and only if the following equations are satisfied:

$$
\begin{aligned}
&dS(\gamma)\,\delta\gamma = 0,\\
&\gamma'(t) \in \mathcal{D}_{\gamma(t)}, \quad t \in [a,b],\\
&\delta\gamma(a) = 0, \qquad \delta\gamma(b) = 0,\\
&\delta\gamma(t) \in \mathcal{D}_{\gamma(t)}, \quad t \in [a,b].
\end{aligned}
$$

The variational derivative of S is computed first, and then applied as a one form to the infinitesimal variations $\delta\gamma \in \mathcal{D}$. This type of problem is called a *skew critical problem*

The discrete Lagrange–d'Alembert principle is analogous for a uniform discretization of $[a,b]$. The $N+1$ element sequence q_d in Q^{N+1} is a discrete trajectory if and only if the following equations are satisfied:

$$
\begin{aligned}
&dS_d(q_d)\,\delta q_d = 0,\\
&q_d' \in \mathcal{D}^N,\\
&\delta q_0 = 0, \qquad \delta q_{N+1} = 0,\\
&\delta q_d \in \mathcal{D}^{N+1}.
\end{aligned}
$$

where q_d' is the discrete derivative of q_d in TQ and $\delta q_d = \{\delta q_k\}_{k=0}^{N+1}$ is a sequence of tangent vectors. This is a skew critical problem for the derivative of S_d. It will be described in detail in Chapter 3.

In the special case that $\mathcal{D}$ is integrable, the discrete flow is symplectic, with respect to a discrete symplectic two form. This form can be identified along with a pair of Lagrange one forms by examining dS_d along the solution sequence, much in the same way the continuous Lagrange form is identified in Marsden, Patrick and Shkoller [17]. The foliation of Q by $\mathcal{D}$ is exactly preserved by the discrete evolution.

The papers by Cortés and Martínez [7] and McLachlan and Perlmutter [19] set the integrators in $Q \times Q$ with the discrete evolution being a map $(q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$. This leads to the impression that the evolution is on Q and hence the subsequent integrator is a two–step method requiring an additional integration in order to start the integrator from the initial condition q_0. The approach in this thesis is to set the discretization in TQ so that the evolution map is $v_k \mapsto v_{k+1}$ with an initial condition v_0, giving a one–step method.

In de León, de Diego and Santamaria–Merino [11], nonholonomic integrators are developed from generating functions, an approach not investigated in this thesis.

Chapter 2 contains the necessary background material for continuous Lagrange–d'Alembert systems, including their formalization as skew variational problems.

Chapter 3 contains the development of discrete Lagrange–d'Alembert systems. The notion of *discrete tangent bundle*, P, of Q is introduced in order to develop the theory without referring to TQ. This clarifies the structures by forcing the theory to evolve without reference to the derivative of curves. The discrete Lagrangian is defined on P as is the discrete action function. It is also shown that Moser–Veselov integrators are Lagrange–d'Alembert integrators. This chapter is laid out in a manner very similar to Chapter 2 in order to stress the strong analogy between the discrete and continuous theories.

Chapter 4 is a special adaptation of Chapter 3 in that the discrete tangent bundle is now taken to be TQ and the discrete Lagrangian, L_h, an approximation to the continuous Lagrangian. These discretizations are made using curve segments that exactly satisfy the constraints and solve the exact Euler–Lagrange equations up to order r. It is then shown that the resulting discrete Euler–Lagrange equations are a numerical method of order r as well.

Chapter 5 goes one step further and describes a method for constructing the necessary curve segments and a class of implicit Lagrange–d'Alembert methods.

Chapter 6 applies the theory of Chapters 4 and 5 to various mechanical systems in order to provide a proof of concept for the theory.

Finally, there are appendices that deal with necessary theorems that are not a part of the flow of the thesis.

The notation used throughout the thesis is described in Table 1.1.

Q	Smooth n dimensional configuration manifold
TQ	Tangent bundle of Q
τ_Q	Bundle projection $TQ \to Q$
$\mathcal{D}$	Smooth d dimensional distribution on Q
$\mathcal{D}^0$	Annihilator of $\mathcal{D}$ in T^*Q
$X, Y, Z, \ldots$	Vector fields
$\mathcal{D}_Y$	Affine constraint submanifold (C.2)
X_i	Basis vector field of T_qQ adapted to $\mathcal{D}$
c^b	Constraint functions (C.3)
c	Constraint function (C.4)
(U, μ)	Coordinate chart for Q
$(TU, T\mu)$	Natural coordinates for TQ
$\hat{U}$	$\mu(U)$
$T\hat{U}$	$T\mu(TU)$
$\hat{\mathcal{D}}$	$T\mu(\mathcal{D} \cap TU)$
$\hat{\mathcal{D}}_Y$	$T\mu(\mathcal{D}_Y \cap TU)$
$\hat{c}^b$	Local representation of c^b
$\hat{c}$	Local representation of c
$\hat{X}_i$	Local representative of X_i
$\hat{\phi}^i$	Local representative of ϕ^i

Table 1.1: Notation

Chapter 2

Classical Lagrangian Mechanics

2.1 Introduction

The formal setup for the Lagrange–d'Alembert principle will take the form of a skew–variational problem in analogy to the skew–critical point problem found in Section B.2. Abraham and Marsden [1] contain some of the background theory and references for the infinite dimensional technicalities regarding manifolds of curves.

2.2 The Classical Lagrange–d'Alembert Principle

Skew–critical point problems require a one form on a manifold, a submanifold and a distribution on the manifold. In order to develop the Lagrange–d'Alembert principle as a skew problem, these components all need to be identified. It should be noted, however, that the setting for the Lagrange–d'Alembert principle is infinite dimensional with the derivatives involved being *variational derivatives* (Gelfand and Fomin [12]) and thus the theory of Section B.2 can be referred to only in analogy. The comparison is quite strong, however, and very instructive.

Let Q be an n dimensional configuration manifold and TQ its tangent bundle. Define the manifold of curves in Q as

$$\Omega = \{\gamma\colon [a,b] \to Q \mid \gamma \text{ is } C^2\}.$$

The tangent space at $\gamma \in \Omega$ is

$$T_\gamma \Omega = \{\eta \colon [a,b] \to TQ \,|\, \tau_Q \circ \eta = \gamma\}.$$

Write $\gamma' \equiv T\gamma$ and $\dot\gamma(t) \equiv \frac{d}{dt}\gamma(t) \equiv T_t\gamma\, 1$. For a constrained system, the set of *admissible curves* is

$$\mathcal{N} = \{\gamma \in \Omega \,|\, \gamma' \in \mathcal{D}\},$$

and the set of *admissible variations*, at $\gamma \in \mathcal{N}$, is

$$\mathcal{D}_\gamma = \{\delta\gamma \colon [a,b] \to \mathcal{D} \,|\, \tau_Q \circ \delta\gamma = \gamma, \delta\gamma(a) = \delta\gamma(b) = 0\}.$$

Let $L : TQ \to \mathbb{R}$ be the Lagrangian for a classical mechanical system. The action functional on curves γ in Ω over the interval $[a,b]$ is defined as

$$S \colon \Omega \to \mathbb{R}, \qquad S(\gamma) = \int_a^b L(\gamma'(t))dt. \tag{2.1}$$

The Lagrange–d'Alembert principle is then the skew problem of finding the curves, γ, that satisfy the equations

$$dS(\gamma)\,\mathcal{D}_\gamma = 0, \tag{2.2}$$

$$\gamma \in \mathcal{N}. \tag{2.3}$$

The Lagrange–d'Alembert principle gives a necessary condition for a curve to be a physical trajectory of the Lagrangian system over the time interval $[a,b]$.

Equation (2.2) is evaluated by first taking the derivative of S on Ω and then evaluating on vectors in $\mathcal{D}_\gamma$. In particular, $\delta\gamma$ is not restricted in any way and therefore is in $T_\gamma\Omega$.

Let $\gamma \in \Omega$ and $\delta\gamma \in T_\gamma\Omega$. Let γ_ϵ be a curve in Ω such that $\gamma = \gamma_0$ and $\delta\gamma = \frac{d}{d\epsilon}\big|_{\epsilon=0}\gamma_\epsilon$. To be clear, γ_ϵ is a curve in the set of curves in Q, so that γ_ϵ is in fact a one parameter set of curves. In local coordinates $(q^i, \dot q^i)$ of TQ, compute

$$dS(\gamma)\,\delta\gamma = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} S(\gamma_\epsilon)$$

$$
\begin{aligned}
&= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b L(\gamma_\epsilon'(t))\,dt \qquad \text{Equation(2.1) for } S \\
&= \int_a^b dL(\gamma'(t)) \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma_\epsilon'(t)\,dt \qquad \text{chain rule} \\
&= \int_a^b \left(\frac{\partial L}{\partial q^i}(\gamma'(t))\,\delta\gamma^i(t) + \frac{\partial L}{\partial \dot{q}^i}(\gamma'(t)) \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \frac{d}{dt}\gamma_\epsilon^i(t) \right) dt \qquad \text{expanding } dL \\
&= \int_a^b \left(\frac{\partial L}{\partial q^i}(\gamma'(t))\,\delta\gamma^i(t) + \frac{\partial L}{\partial \dot{q}^i}(q'(t)) \frac{d}{dt}\delta\gamma^i(t) \right) dt \qquad \text{commuting derivatives} \\
&= \int_a^b \left(\frac{\partial L}{\partial q^i}(\gamma'(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\gamma'(t)) \right) \right) \delta\gamma^i(t)\,dt + \\
&\qquad + \left.\frac{\partial L}{\partial \dot{q}^i}(\gamma'(t))\,\delta\gamma^i(t)\right|_a^b \qquad \text{integration by parts}
\end{aligned} \tag{2.4}
$$

Restricting $\delta\gamma \in \mathcal{D}_\gamma$, the boundary terms in Equation (2.4) vanish and Equation (2.2) becomes,

$$
\int_a^b \left(\frac{\partial L}{\partial q^i}(\gamma'(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\gamma'(t)) \right) \right) \delta\gamma^i(t)\,dt = 0. \tag{2.5}
$$

Proposition 1 is based on Lemma 1, Chapter 1, from Gelfand and Fomin [12].

Proposition 1. *Let F be a continuous one form on Q, $\mathcal{D}$ a continuous distribution on Q and γ a curve in Q such that*

$$
\int_a^b F(\gamma(t))\,Y(\gamma(t))\,dt = 0,
$$

for all continuous vector fields Y taking their values in $\mathcal{D}$. Then $F(\gamma(t))\,Y(\gamma(t)) = 0$.

Proof. Let (U, μ) be a local chart on Q containing $\gamma([a, b])$, shrinking the interval $[a, b]$ if necessary. Let $\{X_a\}_{a=1}^n$ be a local basis for TU such that $\{X_a\}_{a=1}^d$ is a local basis for $\mathcal{D}$. Let $\{\phi^a\}_{a=1}^n$ be the dual basis for T^*U. Then $\{\phi^a\}_{a=d+1}^n$ is a local basis for $\mathcal{D}^0$, the annihilator of $\mathcal{D}$. For $c = 1 \ldots d$,

$$
F(\gamma(t))\,Y(\gamma(t)) = F_a(\gamma(t))\,\phi^a(\gamma(t))\,Y^c(\gamma(t))\,X_c(\gamma(t)) = F_c(\gamma(t))\,Y^c(\gamma(t)),
$$

where Y^c are arbitrary functions. For each c, assume that there is a $\tau \in [a, b]$ such that $F_c(\gamma(\tau)) \neq 0$. Then there is an open set, $V \subset U$, containing $\gamma(\tau)$ such that $F_c(q) > 0$ or $F_c(q) < 0$ for $q \in V$. Without loss of generality, assume $F_c(q) > 0$ on V. Set the principal part of Y to be

$$Y^c(q) = H(q),$$

where $H(q)$ is a bump function supported in V. Set all other components of Y to 0. Then,

$$\int_a^b F(\gamma(t))\, Y(\gamma(t))\, dt = \int_a^b F_c(\gamma(t))\, H(\gamma(t))\, dt > 0,$$

contradicting the initial hypothesis. Therefore, $F_c(\gamma(t)) = 0$ for each $c = 1 \ldots d$ and $F(\gamma(t)) = F_a(\gamma(t))\, \phi^a(\gamma(t))$ for $a = d+1 \ldots n$. Then, $F \in \mathcal{D}^0$ and $F(\gamma(t))\, Y(\gamma(t)) = 0$ for all $Y \in \mathcal{D}$. □

Applying Proposition 1 to Equation (2.5) gives the resulting coordinate version of the equations implied by the Lagrange–d'Alembert principle as

$$\left(\frac{\partial L}{\partial q^i}(\gamma'(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\gamma'(t)) \right) \right) \delta\gamma^i(t) = 0, \tag{2.6}$$

$$\delta\gamma \in \mathcal{D}_\gamma, \tag{2.7}$$

$$\gamma \in \mathcal{N}. \tag{2.8}$$

Equation (2.6) is a differential equation for curves $\eta = \gamma'$ in TQ. The fact that the curves, γ', are derivatives of curves in Q is expressed as the *second order condition* $(\tau_Q \circ \eta)' = \eta$. In coordinates $(q, \dot{q})$ for TQ, let $\gamma'(t) = (q(t), \dot{q}(t))$. The second order condition implies that $\dot{q}(t) = \frac{d}{dt} q(t)$.

Equation (2.7) restricts the $\delta\gamma(t)$ to a $\dim \mathcal{D} = d$ subspace for each $t \in [a, b]$ so that Equation (2.6) consists of d independent equations. The second order condition on the curves γ' gives another n equations for a total of $n + d$ equations for the curves γ'. Equation (2.8) restricts γ' to be in the $n + d$ dimensional submanifold $\mathcal{D}$ of TQ.

Issues of existence and uniqueness of solutions of the Lagrange–d'Alembert equations will be addressed in Section 2.3. Assume for the moment, however, that Equations (2.6), (2.7) and (2.8) determine a set of second order differential equations on $\mathcal{D}$. Should fixed endpoints $q_a = \gamma(a)$ and $q_b = \gamma(b)$ be imposed, there is only $n + d$ degrees of freedom from which to choose, and in such a way as to guarantee a solution of the resulting boundary value problem. Alternatively, an initial condition in $\mathcal{D}$ may be chosen. Either way, it is not necessary to specify a boundary condition beyond $\delta\gamma(a) = \delta\gamma(b) = 0$ until a particular solution is sought.

As in the usual skew–critical point situation (Appendix B), the skew–variational problem will reduce to a variational problem when $\mathcal{D}$ is integrable. In this case (to be proven shortly), $T_\gamma\mathcal{N} = \mathcal{D}_\gamma$, where, for γ_ϵ a curve in $\mathcal{N}$,

$$T_\gamma\mathcal{N} = \left\{ \eta \colon [a,b] \to TQ \mid \eta = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma_\epsilon, \tau_Q \circ \eta = \gamma \right\}.$$

The distinction between $\mathcal{D}_\gamma$ and $T_\gamma\mathcal{N}$ is not a vacuous one, as Example 1 shows.

Example 1. Let $Q = \mathbb{R}^3$ with coordinates (x, y, z) and $TQ = T\mathbb{R}^3$ with coordinates $(x, y, z, \dot{x}, \dot{y}, \dot{z})$. Let $\mathcal{D} = \{\, \partial_x + y\partial_z,\ \partial_y \,\}$. The distribution $\mathcal{D}$ shows up in various references, for example, in Rosenberg [22] and Cuell [8] as the nonholonomic constraint forcing a disk to roll on a rotating disk without slipping. For fixed $(x, y, z, \dot{x}, \dot{y}, \dot{z})$, let

$$\gamma_\epsilon(t) = \left(x + t\dot{x}, y + t\dot{y}, z + t\dot{x}y + \frac{t^2}{2}\dot{x}\dot{y} \right) + \epsilon \left(0, \sin(2\pi t), \dot{x} \int_0^t \sin(2\pi\tau)\, d\tau \right).$$

The curve γ_ϵ is in $\mathcal{N}$ where

$$\begin{aligned} \gamma_\epsilon(0) &= (x, y, z), \\ \gamma_\epsilon(1) &= \left(x + \dot{x}, y + \dot{y}, z + \dot{x}y + \frac{1}{2}\dot{x}\dot{y} \right) \end{aligned}$$

are fixed points that do not depend on ϵ. However,

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma_\epsilon(t) = \left(0, \sin(2\pi t), \dot{x} \int_0^t \sin(2\pi\tau)\, d\tau \right),$$

which is not in $\mathcal{D}$ for all $t \in [0, 1]$. This shows that $T_\gamma\mathcal{N} \neq \mathcal{D}_\gamma$.

When $\mathcal{D}$ is integrable, the mechanical system is holonomic and the Lagrange–d'Alembert principle reduces to a variational problem in which

$$dS(\gamma)\, \mathcal{D}_\gamma = d\, S|_{\mathcal{N}}(\gamma) = 0.$$

That is, the Lagrange–d'Alembert principle becomes a constrained variational principle. The necessary requirement is that $\mathcal{D}_\gamma = T_\gamma \mathcal{N}$, which occurs when $\mathcal{D}$ is integrable.

Proposition 2. *Let $\mathcal{D}$ be a C^1 distribution on a manifold Q. $\mathcal{D}$ is integrable if and only if $\mathcal{D}_\gamma = T_\gamma \mathcal{N}$ for all $\gamma \in \Omega$.*

Proof. First, fix a and b and let $\gamma \in \Omega$.

If $\mathcal{D}$ is integrable, $\mathcal{N}$ is a set of curves in the leaves of $\mathcal{D}$ so that, for γ_ϵ a curve in $\mathcal{N}$, $\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma_\epsilon(t) = \delta\gamma(t) \in \mathcal{D}$. This implies that $T_\gamma \mathcal{N} = \mathcal{D}_\gamma$.

Now, suppose $T_\gamma \mathcal{N} = \mathcal{D}_\gamma$. Let X and Y be vector fields on Q taking their values in $\mathcal{D}$ with flows F_t^X and F_t^Y respectively.

Let $q \in Q$, C a closed set containing q and V an open set containing C such that $F_a^X(q)$, $F_b^X(q)$, $F_a^Y(q)$ and $F_b^Y(q)$ are not in V. Using an appropriate bump function, vector fields $\tilde{X}$ and $\tilde{Y}$ can be arranged so that on C, $\tilde{X} = X$, $\tilde{Y} = Y$ and $\tilde{X} = \tilde{Y} = 0$ outside V.

Without loss of generality, then, it can be assumed that the flow of X satisfies $X(F_a^X(q)) = X(F_b^X(q)) = 0$ and the flow of Y satisfies $Y(F_a^Y(q)) = Y(F_b^Y(q)) = 0$.

The curve

$$\gamma(t) = (F_t^X \circ F_\epsilon^Y)(q),$$

is in $\mathcal{N}$ since, for ϵ small enough,

$$\frac{d}{dt}\gamma(t) = \frac{d}{dt}\left(F_t^X \circ F_\epsilon^Y\right)(q) = X(F_t^X(F_\epsilon^Y(q))).$$

By hypothesis, $T_q F_t^X Y \in \mathcal{D}$ since

$$T_q F_t^X Y = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \left(F_t^X \circ F_\epsilon^Y\right)(q). \tag{2.9}$$

Equation (2.9) holds for arbitrary t and q so that

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \left(F_{-t}^X \circ F_\epsilon^Y\right)\left(F_t^X(q)\right) = T_{F_t^X(q)} F_{-t}^X \, Y(F_t^X(q)) = (F_t^X)^* Y(q) \in \mathcal{D},$$

and, therefore,

$$\left.\frac{d}{dt}\right|_{t=0} \left(F_t^X\right)^* Y(q) = [X, Y] \in \mathcal{D}.$$

$\square$

2.3 Equations of Motion

Define the following submanifold of TTQ:

$$\ddot{Q} = \left\{ \, T_a^2 \gamma \;\middle|\; \gamma \in \Omega \, \right\}.$$

In coordinates, $\ddot{Q}$ contains elements of the form $(q, \dot{q}, \dot{q}, \ddot{q}) \in TTQ$.

Let γ_ϵ be a curve in Ω, $\delta\gamma = \frac{d}{d\epsilon}\big|_{\epsilon=0} \gamma_\epsilon$, $\dot{q} = \frac{d\gamma}{dt}$ and $\ddot{q} = \frac{d^2\gamma}{dt^2}$. The derivative of the action, Equation (2.1), as calculated in Section 2.2, is

$$\begin{aligned} dS(\gamma)\,\delta\gamma &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} S(\gamma_\epsilon) \\ &= \int_a^b \left(\frac{\partial L}{\partial q^i}(\gamma'(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\gamma'(t)) \right) \right) \delta\gamma^i(t)\, dt + \frac{\partial L}{\partial \dot{q}^i}(\gamma'(t))\, \delta\gamma^i(t) \bigg|_a^b . \end{aligned}$$

Expanding the time derivative in the integral gives

$$\begin{aligned} dS(\gamma)\delta\gamma = \int_a^b \left(\frac{\partial L}{\partial q^i}(\gamma'(t)) - \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i}(\gamma'(t))\, \dot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}(\gamma'(t))\, \ddot{q}^j \right) \delta\gamma^i(t)\, dt + \\ + \frac{\partial L}{\partial \dot{q}^i}(\gamma'(t))\, \delta\gamma^i(t) \bigg|_a^b . \end{aligned} \tag{2.10}$$

The expression under the integral in Equation (2.10) is a linear form defined on $\ddot{Q}$ acting on a vector in TQ. That is, define, in coordinates,

$$D_{EL}L(q, \dot{q}, \dot{q}, \ddot{q})\,\delta q = \left(\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i}\, \dot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}\, \ddot{q}^j \right) \delta q^i. \tag{2.11}$$

The boundary term in Equation (2.10) is a one form on TQ and is used to motivate the definition of the *Lagrange one form* on TQ,

$$\Theta_L(q, \dot{q})\,(\delta q, \dot{\delta q}) = \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q})\, \delta q^i, \tag{2.12}$$

for $\left(\delta q, \dot{\delta q}\right) \in TTQ$. The definitions in equations (2.11) and (2.12) are formalized in Theorem 1, see Marsden, Patrick, Shkoller [17].

Theorem 1. *Let L be a C^k Lagrangian, $k > 2$. There exists a unique C^{k-2} mapping $D_{EL}L\colon \ddot{Q} \to T^*Q$ and a unique C^{k-1} one form Θ_L on TQ such that, for all curves γ_ϵ in Ω,*

$$dS(\gamma)\,\delta\gamma = \int_a^b D_{EL}L\left(T_t^2\gamma\right)\,\delta\gamma(t)\,dt + \Theta_L\left(T_t\gamma\right)\,\widehat{\delta\gamma}(t)\Big|_a^b,$$

where

$$\delta\gamma = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\gamma_\epsilon \quad and \quad \widehat{\delta\gamma}(t) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\frac{d}{dt}\gamma_\epsilon(t).$$

The *fiber derivative* of L is a fiber preserving mapping from TQ to T^*Q defined by

$$\mathbb{F}L(v)\,w = \frac{d}{dt}\Big|_{t=0} L(v+tw),$$

where v and w are in T_qQ. See Marsden and Ratiu [21], Chapter 7.2. The second fiber derivative is

$$\mathbb{F}^2L(v)(w_1,w_2) = \frac{d}{ds}\Big|_{s=0}\frac{d}{dt}\Big|_{t=0} L(v+tw_1+sw_2),$$

for v, w_1 and w_2 in T_qQ. In coordinates $(q,\dot{q})$ for TQ, these are

$$\mathbb{F}L(q,\dot{q}) = \frac{\partial L}{\partial \dot{q}^i}dq^i,$$
$$\mathbb{F}^2L(q,\dot{q}) = \frac{\partial^2 L}{\partial \dot{q}^j\partial \dot{q}^i}dq^j \otimes dq^i.$$

Definition 1. A Lagrangian, L, is *$\mathcal{D}$–regular* at v_q if $F^2L(v_q)$ is nonsingular when restricted to $\mathcal{D}$.

When L is $\mathcal{D}$–regular, $\mathbb{F}L$ is locally invertible. If $\mathbb{F}L$ is globally invertible, then L is said to be *$\mathcal{D}$–hyperregular.*

Proposition 3 gives a condition for Equations (2.6), (2.7) and (2.8) to form a set of second order differential equations. Recall that a vector field X on TQ is said to be *second order* if $T_{v_q}\tau_Q X = v_q$ for all v_q in the domain of X.

Proposition 3. *Let $\mathcal{D}$ be C^{k-2} distribution on Q and $\mathcal{D}^0$ the annihilator of $\mathcal{D}$. Let $L\colon TQ \to \mathbb{R}$ be a C^k $\mathcal{D}$–regular Lagrangian. Then there is a unique C^{k-2} second order vector field X_E on $\mathcal{D}$ such that $D_{EL}L \circ X_E \in \mathcal{D}^0$.*

Proof. Let $\{\, \phi^b \,\}_{b=d+1}^n$ be a local basis for $\mathcal{D}^0$ and $\{X_a\}_{a=1}^d$ a local basis for $\mathcal{D}$. The Lagrange–d'Alembert principle applied to the basis of $\mathcal{D}$ gives

$$D_{EL}L(q, \dot{q}, \dot{q}, \ddot{q})\, X_a = 0, \qquad a = 1, \ldots, d, \tag{2.13}$$

$$\phi^b(q)\, \dot{q} = 0, \qquad b = d+1, \ldots, n. \tag{2.14}$$

Using the definition of $D_{EL}L$ in Equation (2.11) and differentiating equation (2.14) with respect to t, the above equations become

$$\left(\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i} \dot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i} \ddot{q}^j \right) X_a^i = 0, \tag{2.15}$$

$$\frac{\partial \phi_i^b}{\partial q^j}(q)\, \dot{q}^j \dot{q}^i + \phi_i^b(q)\, \ddot{q}^i = 0. \tag{2.16}$$

Equations (2.15) and (2.16) are n linear equations in $\ddot{q}$ with coefficient matrix

$$A = \begin{bmatrix} \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i} X_a^i \\ \\ \phi_j^b(q) \end{bmatrix}.$$

In coordinates, the principle part of X_E is $X_E(q, \dot{q}) = (\dot{q}, \ddot{q})$ so that X_E will be unique if Equations (2.15) and (2.16) have a unique solution for $\ddot{q}$. This will be the case if and only if $\ker A = \{0\}$, which will now be shown.

Suppose $w \in T_qQ$ such that $Aw = 0$. The lower block of A gives $\phi^b w = 0$ for $b = d+1, \ldots, n$ so that $w \in \mathcal{D}$. Therefore, $w = w^c X_c$, $c = 1, \ldots d$ and the upper block of A gives

$$\frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}(q, \dot{q}) X_a^i\, X_c^j\, w^c = 0. \tag{2.17}$$

Since L is assumed to be $\mathcal{D}$–regular and the X_a and X_c are in $\mathcal{D}$, Equation 2.17 is satisfied only if $w^c = 0$, $c = d+1, \ldots n$. □

The second order differential equations associated to X_E, along with the constraints, give the following equations of motion for the constrained system as

$$\dot{q} = \frac{d\gamma}{dt}, \qquad \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i} \right) \right) \delta\gamma^i = 0, \tag{2.18}$$

$$\phi_i(q)\, \dot{q}^i = 0,, \qquad \delta\gamma \in \mathcal{D}, \qquad \phi \in \mathcal{D}^0.$$

These equations are called the *constrained Euler–Lagrange* equations and the vector field, X_E, for the second order set of differential equations is called the *Euler–Lagrange vector field*. A complete set of equations is generated by choosing d linearly independent $\delta\gamma$s from $\mathcal{D}$ and $n-d$ linearly independent ϕs from $\mathcal{D}^0$.

The space of solutions of the constrained Euler–Lagrange equations is parameterized by the set of initial conditions by $v_q \mapsto F_t^{X_E}(v_q)$, where $v_q \in \mathcal{D}$ and $F_t^{X_E}$ is the flow of the Euler–Lagrange vector field X_E. Since X_E is a second order vector field, $F_t^{X_E}(v_q) = \frac{d}{dt}\left(\tau_Q \circ F_t^{X_E}(v_q)\right)$ and

$$\begin{aligned} T_{v_q}F_t^{X_E}\, w_{v_q} &= T_{v_q}\frac{d}{dt}\left(\tau_Q \circ F_t^{X_E}(v_q)\right) w_{v_q} \\ &= \frac{d}{dt}T_{v_q}\left(\tau_Q \circ F_t^{X_E}(v_q)\right) w_{v_q}. \end{aligned} \tag{2.19}$$

Define $S_t\colon TQ \to \mathbb{R}$ by

$$S_t(v_q) = \int_0^t L\left(F_s^{X_E}(v_q)\right)\, ds. \tag{2.20}$$

Equation (2.19) shows that the curve $T_{v_q}F_t^{X_E}\, w_{v_q}$ in TTQ has the form (locally) $\left(\delta\gamma(t), \frac{d}{dt}\delta\gamma(t)\right)$ for $\delta\gamma(t) \in TQ$. Then,

$$\begin{aligned} dS_t(v_q)\, w_{v_q} &= \int_0^t dL(F_s^{X_E}(v_q))\, T_{v_q}F_s^{X_E}\, w_{v_q}\, ds \\ &= \int_0^t D_{EL}L(X_E)\, T_{v_q}(\tau_Q \circ F_s^{X_E})\, w_{v_q}\, ds + \\ &\qquad + \Theta_L(F_s^{X_E}(v_q))\, T_{v_q}F_s^{X_E}\, w_{v_q}\Big|_{s=0}^{t} \qquad (2.21) \\ &= \alpha_t(v_q)\, w_{v_q} + (F_t^{X_E})^*\,\Theta_L(v_q)\, w_{v_q} - \Theta_L(v_q)\, w_{v_q}, \end{aligned}$$

where α_t is defined by, and replaces, the integral term in equation (2.21) so that

$$dS_t = \alpha_t + (F_t^{X_E})^*\,\Theta_L - \Theta_L. \tag{2.22}$$

The *symplectic form associated to* L is defined by $\omega_L = -d\Theta_L$. Another derivative of Equation (2.22) gives

$$0 = d^2S_t = d\alpha_t + (F_t^{X_E})^* d\Theta_L - d\Theta_L,$$

which implies

$$(F_t^{X_E})^* \omega_L = \omega_L + d\alpha_t, \tag{2.23}$$

giving an evolution equation for the symplectic form associated to L.

Proposition 3 implies

$$\left.\frac{d}{dt}\right|_{t=0} \alpha_t(v_q) = D_{EL}L(X_E)\, T\tau_Q \in \mathcal{D}^\circ,$$

so that, locally,

$$\left.\frac{d}{dt}\right|_{t=0} \alpha_t = \lambda_a \tau_Q^* \phi^a, \tag{2.24}$$

for some functions $\lambda_a \colon TQ \to \mathbb{R}$.

2.4 Special Cases

2.4.1 Holonomic Constraints

If the constraints are holonomic, then $\mathcal{D}$ is integrable and Proposition 2 implies that the system is equivalent to unconstrained motion on the leaves of $\mathcal{D}$.

Let Q_0 be a leaf of $\mathcal{D}$ and TQ_0 its tangent bundle with coordinates $(r^1, \ldots, r^d, \dot{r}^1, \ldots, \dot{r}^d)$. The definition of a leaf gives $TQ_0 = \mathcal{D}|_{Q_0}$. Write $L_0 = L|_{TQ_0}$. The action is

$$S_0(\gamma) = \int_a^b L_0(\gamma'(t))\, dt, \qquad \gamma \text{ a curve in } Q_0.$$

The Lagrange–d'Alembert principle reduces to

$$dS_0(\gamma) = 0, \qquad \gamma \text{ a curve in } Q_0,$$

which is equivalent to

$$D_{EL}L_0(X_E)\,\delta\gamma = 0, \qquad \delta\gamma \in TQ_0. \tag{2.25}$$

In the coordinates $(r^\alpha, \dot{r}^\alpha)$, this is

$$\frac{\partial L}{\partial r^\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}^\alpha}\right) = 0, \qquad \alpha = 1, \ldots, d.$$

Since the Euler–Lagrange vector field preserves TQ_0, $T_{v_q}(\tau_Q \circ F_s^{X_E})\, w_{v_q} \in TQ_0$ for $w_{v_q} \in T_{v_q}TQ_0$ and

$$\alpha_t(v_q)\, w_{v_q} = \int_0^t D_{EL}L_0(X_E)\, T_{v_q}(\tau_Q \circ F_s^{X_E})\, w_{v_q}\, ds = 0,$$

since the integrand satisfies Equation (2.25). This reduces Equation (2.23) to

$$(F_t^{X_E})^* \omega_L = \omega_L,$$

so that the flow of holonomic systems is symplectic.

2.4.2 Symplectic Subsystems

Proposition 4 is found in Patrick [20].

Proposition 4. *Let $\mathcal{D}$ be a C^{k-2} distribution on a manifold Q, $k > 2$. Let $\mathcal{K} = (T\tau_Q)^{-1}\mathcal{D} \cap T\mathcal{D}$, $L\colon TQ \to \mathbb{R}$ a C^k $\mathcal{D}$–regular Lagrangian, X_E the Euler–Lagrange vector field guaranteed by Proposition 3 and $F_t^{X_E}$ the flow of X_E. Let $\mathcal{D}_0$ be a submanifold of $\mathcal{D}$ and let $\mathcal{K}_0$ be a subbundle of $\mathcal{K}$ on $\mathcal{D}_0$. Suppose $TF_t^{X_E}\mathcal{K}_0 \subseteq \mathcal{K}_0$ on $\mathcal{D}_0$ and $[Y_1, Y_2] \in \mathcal{K}_0$ for vector fields Y_1 and Y_2 in $\mathcal{K}_0$. Then $F_t^{X_E}$ is symplectic on $\mathcal{K}_0$.*

Proof. Recall equation (2.22),

$$dS_t = \alpha_t + (F_t^{X_E})^*\, \Theta_L - \Theta_L,$$

and, differentiating, (recall $\omega_L = -d\Theta_L$),

$$0 = d^2 S_t = d\alpha_t - (F_t^{X_E})^*\, \omega_L + \omega_L.$$

Evidently, $F_t^{X_E}$ will preserve ω_L when $d\alpha_t$ is zero. The exterior derivative of α_t can be computed as (see Abraham, Marsden and Ratiu [2] Section 6.4)

$$d\alpha_t(v_q)\,(Y_1, Y_2) = Y_1\,(\alpha_t(Y_2))(v_q) - Y_2\,(\alpha_t(Y_1))(v_q) - \alpha_t([Y_1, Y_2])(v_q), \tag{2.26}$$

so that $d\alpha_t$ will be zero when each term is zero. Recall,

$$\alpha_t(v_q)\, w_{v_q} = \int_0^t D_{EL}L(X_E)\, T_{v_q}(\tau_Q \circ F_s^{X_E})\, w_{v_q}\, ds.$$

If $w_{v_q} \in \mathcal{K}_0$ and $T_{v_q} F_t^{X_E} \, w_{v_q} \in \mathcal{K}_0$, the integrand will vanish identically. By hypothesis, then, each term on the right hand side of equation (2.26) is zero. Therefore, $(F_t^{X_E})^* \, \omega_L = \omega_L$ on $\mathcal{K}_0$. Further, since $d\omega_L(v_q) = 0$ on $\mathcal{K}_0$ and $v_q \in \mathcal{D}_0$, ω_L is symplectic on $\mathcal{K}_0$. □

2.5 Semi - Hamilton's Equations

Using the definition of S_t in Equation (2.20), and $w_{v_q} \in T_{v_q}TQ$, Equation (2.22) is

$$\int_0^t (F_s^{X_E})^* dL(v_q) \, w_{v_q} \, ds = \alpha_t(v_q) \, w_{v_q} + (F_t^{X_E})^* \Theta_L(v_q) \, w_{v_q} - \Theta_L(v_q) \, w_{v_q}. \tag{2.27}$$

Taking one time derivative of Equation (2.27), and using the definition of α_t,

$$(F_t^{X_E})^* dL(v_q) \, w_{v_q} = D_{EL}L(X_E) \, T_{v_q}(\tau_Q \circ F_t^{X_E}) \, w_{v_q} + \frac{d}{dt}\Big((F_t^{X_E})^* \Theta_L(v_q)\Big) \, w_{v_q}. \tag{2.28}$$

Let $v_q \in \mathcal{D}$ and let $w(s)$ be a curve in $T_{v_q}\mathcal{D}$ such that

$$T_{v_q}(\tau_Q \circ F_s^{X_E}) \, w(s) \in \mathcal{D}_{(\tau_Q \circ F_s^{X_E})(v_q)}. \tag{2.29}$$

Then $D_{EL}(X_E) \, T_{v_q}(\tau_Q \circ F_t^{X_E}) \, w(t) = 0$ by the Lagrange–d'Alembert principle and Equation (2.28) becomes, for $s = t$,

$$(F_t^{X_E})^* dL(v_q) \, w(t) = \frac{d}{dt}\Big((F_t^{X_E})^* \Theta_L(v_q)\Big) \, w(t). \tag{2.30}$$

Using the definition of the Lie derivative, $\mathcal{L}_{X_E}$, this gives

$$(F_t^{X_E})^* dL(v_q) \, w(t) = (F_t^{X_E})^* \mathcal{L}_{X_E} \Theta_L(v_q) \, w(t).$$

The definition of the pullback implies

$$dL(F_t^{X_E}(v_q)) \, T_{v_q} F_t^{X_E} \, w(t) = \mathcal{L}_{X_E} \Theta_L(F_t^{X_E}(v_q)) \, T_{v_q} F_t^{X_E} \, w(t).$$

The Cartan identity then gives

$$dL(F_t^{X_E}(v_q)) \, T_{v_q} F_t^{X_E} \, w(t) = (X_E \,\lrcorner\, d\Theta_L + d(X_E \,\lrcorner\, \Theta_L))(F_t^{X_E}(v_q)) \, T_{v_q} F_t^{X_E} \, w(t).$$

Using $d\theta_L = -\omega_L$,

$$dL(F_t^{X_E}(v_q))\, T_{v_q} F_t^{X_E}\, w(t) = (-X_E \lrcorner\, d\omega_L + d(X_E \lrcorner\, \Theta_L))(F_t^{X_E}(v_q))\, T_{v_q} F_t^{X_E}\, w(t).$$

Rearrange and let $E = X_E \lrcorner\, \Theta_L - L$ to get

$$(X_E \lrcorner\, \omega_L)(F_t^{X_E}(v_q))\, T_{v_q} F_t^{X_E}\, w(t) = dE(F_t^{X_E}(v_q))\, T_{v_q} F_t^{X_E}\, w(t). \tag{2.31}$$

Since $w(t)$ is defined so that $T_{v_q}(\tau_Q \circ F_t^{X_E})\, w(t)$ is in the constraint distribution, $w(0) = w_{v_q}$ satisfies $T\tau_Q\, w_{v_q} \in \mathcal{D}_q$ so that $w_{v_q} \in \mathcal{K}$ (Section 2.4.2). Then, at $t = 0$ on $\mathcal{D}$, Equation (2.31) is

$$(X_E \lrcorner\, \omega_L)(v_q)\, w_{v_q} = dE(v_q)\, w_{v_q}, \qquad w_{v_q} \in \mathcal{K}. \tag{2.32}$$

The flow preserves $\mathcal{D}$ and $T\tau_Q X_E \in \mathcal{D}$, so $X_E \in \mathcal{K}$. Equation (2.32) can therefore be considered as a linear equation in X_E on $\mathcal{K}$. Equation (2.32) is uniquely solvable for X_E as long as ω_L is nondegenerate on $\mathcal{K}$.

Proposition 5. *Let $\mathcal{D}$ be a C^{k-2} distribution on a manifold Q, $k > 2$. Let $L\colon TQ \to \mathbb{R}$ be C^k Lagrangian. Then $\omega_L = -d\Theta_L$ is nondegenerate on $\mathcal{K}$ (Section 2.4.2) if and only if L is $\mathcal{D}$–regular.*

Proof. First, construct a suitable basis for $\mathcal{K}$. Let $\{X_a\}_{a=0}^d$ be a basis for $\mathcal{D}$ and $\{\phi^b\}_{b=d+1}^n$ be a basis for $\mathcal{D}^0$ at q. Let $(q, \dot{q})$ be local coordinates for TQ and let $(q(t), \dot{q}(t))$ be a curve in $\mathcal{D}$ such that $(q(0), \dot{q}(0)) = (q_0, \dot{q}_0)$ and $\frac{d}{dt}\big|_{t=0}(q(t), \dot{q}(t)) = (v, \dot{v}) \in \mathcal{K}$. Then $\phi_i^b(q(t))\, \dot{q}^i(t) = 0$ so that, differentiating with respect to t at $t = 0$,

$$\frac{\partial \phi_i^b}{\partial q^j}(q_0)\, \dot{q}_0^i\, v^j + \phi_i^b(q_0) \dot{v}^i = 0. \tag{2.33}$$

Equation (2.33) is the condition that $(v, \dot{v}) \in T\mathcal{D}$. The requirement that $(v, \dot{v}) \in T\tau_Q^{-1}\mathcal{D}$ is

$$\phi_i^b(q_0)\, v^i = 0. \tag{2.34}$$

Equations (2.33) and (2.34) form a linear system of $2(n-d)$ equations in the $2n$ unknowns $(v, \dot{v})$. Since the ϕ^b are pointwise linearly independent, the equations have full rank and hence the solution set has dimension $2d$. Therefore, $\dim \mathcal{K} = 2d$.

A dimension d subspace of $\mathcal{K}$ is spanned by the vectors

$$Z_a = X_a^i \partial_{q^i} + \dot{v}^i \partial_{\dot{q}^i}, \qquad a = 1, \ldots, d, \tag{2.35}$$

where $\dot{v}$ satisfies Equation (2.33) for $v = X_a$, for each $a = 1, \ldots, d$ in turn.

Another dimension d subspace is found by setting $v = 0$, so that Equation (2.34) is trivially satisfied and Equation (2.33) becomes $\phi_i^b(q_0)\dot{v}^i = 0$, which has solution set $\operatorname{span}\{X_a^i \partial_{\dot{q}^i}\}$. A basis for $\mathcal{K}$ is therefore given by

$$\operatorname{span}\{Z_a, X_a^i \partial_{\dot{q}^i}\}, \qquad a = 1, \ldots, d, \tag{2.36}$$

and the Z_a are given by Equation (2.35).

In coordinates,

$$\begin{aligned}
\omega_L &= -d\Theta_L \\
&= -d\left(\frac{\partial L}{\partial \dot{q}^i} dq^i\right) \\
&= \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j} dq^i \wedge dq^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} dq^i \wedge d\dot{q}^j.
\end{aligned}$$

The block form of the matrix for ω_L with respect to the basis of $\mathcal{K}$ in Equation (2.36) is

$$[\omega_L] = \begin{bmatrix} A & B \\ -B & 0 \end{bmatrix},$$

where B is $\dfrac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$ with respect to the basis in Equation (2.36). That is,

$$\begin{aligned}
B_{ab} &= \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \, dq^i \wedge d\dot{q}^j (X_a^k \partial_{q^k} + \dot{v}^k \partial_{\dot{q}^k}, X_b^\ell \partial_{\dot{q}^\ell}) \\
&= \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \, X_a^i X_b^j,
\end{aligned}$$

so that ω_L is nondegenerate on $\mathcal{K}$ if and only if L is $\mathcal{D}$–regular. □

Equation (2.32) and Proposition 5 establish the nonholonomic equations of motion, X_E, as a *distributional Hamiltonian system*. See Bates and Sniatycki [5] and Patrick [20] for more information.

2.6 Conservation of Energy

Let L be a $\mathcal{D}$–regular Lagrangian and define the *Energy* $E\colon TQ \to \mathbb{R}$ by

$$E(v_q) = \Theta_L(v_q)\, X_E(v_q) - L(v_q). \tag{2.37}$$

As a remark, if L is a *natural Lagrangian*,

$$L(v_q) = \frac{1}{2}T(q)\,(v_q, v_q) - V(q)$$

with T a metric on Q and $V\colon Q \to \mathbb{R}$ a potential, then

$$E(v_q) = \frac{1}{2}T(q)\,(v_q, v_q) + V(q).$$

Proposition 6 shows that E is conserved by solutions of the Euler–Lagrange equations.

Proposition 6. *Let $\mathcal{D}$ be a C^{k-2} distribution, $k > 2$, $L\colon TQ \to \mathbb{R}$ a C^k $\mathcal{D}$–regular Lagrangian and X_E the Euler–Lagrange vector field of Equation (2.32) with flow $F_t^{X_E}$. Then E is preserved by $F_t^{X_E}$.*

Proof. Compute

$$\begin{aligned}
\frac{d}{dt}E(F_t^{X_E}) &= \left(F_t^{X_E}\right)^* \mathcal{L}_{X_E}\, E(v_q) \\
&= \left(F_t^{X_E}\right)^* (X_E \,\lrcorner\, dE)\,(v_q) \qquad \text{Lie derivative of a function} \\
&= \left(F_t^{X_E}\right)^* (X_E \,\lrcorner\, \omega_L)(v_q)\, X_E(v_q), \qquad \text{from Equation (2.32)} \\
&= 0. \qquad \text{since } \omega_L \text{ is skew symmetric}
\end{aligned}$$

Therefore, E is constant along $F_t^{X_E}$. □

2.7 Symmetry and Momentum Equations

The notation and definitions follow Abraham and Marsden [1] and Block, Krishnaprasad, Marsden and Murray [6].

Let $\mathcal{G}$ be a Lie group, with Lie algebra $\mathfrak{g}$, acting on Q by ϕ_g such that L and $\mathcal{D}$ are invariant. The infinitesimal generator on Q associated to $\xi \in \mathfrak{g}$ is

$$\xi_Q(q) = \left.\frac{d}{ds}\right|_{s=0} \phi_{\exp(s\xi)} q.$$

The tangent lifted action, $\Phi_g v_q = T_q \phi_g v_q$ has infinitesimal generator on TQ associated to $\xi \in \mathfrak{g}$

$$\xi_{TQ}(v_q) = \left.\frac{d}{ds}\right|_{s=0} \Phi_{\exp(s\xi)} v_q.$$

The infinitesimal generators on Q and TQ are τ_Q related so that

$$T_{v_q}\tau_Q \, \xi_{TQ} = \xi_Q(\tau_Q(v_q)).$$

The momentum function, $J_\xi \colon TQ \to \mathbb{R}$, associated to $\xi \in \mathfrak{q}$ is defined as

$$J_\xi(v_q) = \Theta_L(v_q) \, \xi_{TQ}(v_q),$$

and the momentum map $J \colon TQ \to \mathfrak{g}^*$ is

$$J(v_q)\xi = J_\xi(v_q).$$

Let $\xi(t)$ be a curve in $\mathfrak{g}$ and let ξ_Q^t be the infinitesimal generator on Q associated to $\xi(t)$. Further, let $\xi(t)$ be such that ξ_Q^t satisfies $\xi_Q^t\left(\tau_Q \circ F_t^{X_E}(v_q)\right) \in \mathcal{D}$ for all t. Let $w(t)$ be a curve in $T_{v_q}TQ$ such that $T_{v_q}F_t^{X_E}\, w(t) = \xi_{TQ}^t(F_t^{X_E}(v_q))$. In fact, since $F_t^{X_E} \circ \Phi_g = \Phi_g \circ F_t^{X_E}$, $w(t) = (F_t^{X_E})^* \xi_{TQ}^t(v_q) = \xi_{TQ}^t(v_q)$.

Equation (2.30), repeated here, is

$$(F_t^{X_E})^* dL(v_q)\, w(t) = \frac{d}{dt}\left((F_t^{X_E})^*\Theta_L(v_q)\right) w(t). \tag{2.38}$$

Since L is $\mathcal{G}$ invariant, Equation (2.38) evaluated at $t = t_0$ gives

$$\begin{aligned} 0 &= \left.\frac{d}{dt}\right|_{t=t_0} \left((F_t^{X_E})^*\Theta_L(v_q)\right) w(t_0) \\ &= \left.\frac{d}{dt}\right|_{t=t_0} (F_t^{X_E})^*\Theta_L(v_q)(T_{v_q}F_{t_0}^{X_E})^{-1}\xi_{TQ}^{t_0}(v_q) \qquad \text{by definition of } w(t) \end{aligned}$$

$$
\begin{aligned}
&= \frac{d}{dt}\bigg|_{t=t_0} \Theta_L(F_t^{X_E}(v_q))\, T_{v_q}F_t(T_{F_{t_0}(v_q)}F_{-t_0}^{X_E})\, \xi_{TQ}^{t_0}(v_q) \qquad \text{by definition of pullback} \\
&= \frac{d}{dt}\bigg|_{t=t_0} \Theta_L(F_t^{X_E}(v_q))\, T_{F_{t_0}(v_q)}F_{t-t_0}^{X_E}\, \xi_{TQ}^{t_0}(v_q) \qquad \text{by the chain rule} \\
&= \frac{d}{dt}\bigg|_{t=t_0} \Theta_L(F_{t-t_0}^{X_E}(F_{t_0}(v_q)))\, T_{F_{t_0}(v_q)}F_{t-t_0}^{X_E}\, \xi_{TQ}^{t_0}(v_q) \qquad \text{by properties of the flow} \\
&= \frac{d}{dt}\bigg|_{t=t_0} (F_{t-t_0}^{X_E})^*\Theta_L(F_{t_0}(v_q))\, \xi_{TQ}^{t_0}(v_q) \qquad \text{by definition of the pullback} \\
&= \frac{d}{dt}\bigg|_{t=t_0} (F_{t-t_0}^{X_E})^* J_{\xi(t_0)}(F_{t_0}(v_q)) \qquad \text{by definition of } J_\xi \\
&= \frac{d}{dt}\bigg|_{t=t_0} J_{\xi(t_0)}(F_t(v_q)) \qquad \text{by definition of the pullback} \\
&= \frac{d}{dt}\bigg|_{t=t_0} J_{\xi(t)}(F_t(v_q)) - \frac{d}{dt}\bigg|_{t=t_0} J_{\xi(t)}(F_{t_0}(v_q)) \\
&= \frac{d}{dt}\bigg|_{t=t_0} J_{\xi(t)}(F_t(v_q)) - \frac{d}{dt}\bigg|_{t=t_0} J(F_{t_0}(v_q))\, \xi(t) \qquad \text{by definition of } J \\
&= \frac{d}{dt}\bigg|_{t=t_0} J_{\xi(t)}(F_t(v_q)) - J(F_{t_0}(v_q))\, \frac{d}{dt}\bigg|_{t=t_0} \xi(t).
\end{aligned}
$$

Replacing t_0 with t gives the momentum equations as found in Bloch, Krishnaprasad, Marsden and Murray [6],

$$
\frac{d}{dt} J_{\xi(t)}(v(t)) = J(v(t)) \frac{d}{dt}\xi(t). \tag{2.39}
$$

If $\xi(t) = \xi$ is constant, then Equation (2.39) implies that $J_\xi(v(t))$, and hence, $J(v(t))\,\xi$ is constant.

Chapter 3

Discrete Lagrangian Mechanics

3.1 Introduction

Discretizing a mechanical system means replacing the time interval $[a, b]$ with the sequence $\{a = t_0, t_1, \ldots, t_{N-1}, t_N = b\}$ or just the sequence $\{0, 1, \ldots, N-1, N\}$. The configuration space of a discrete mechanical system will be a manifold Q, however, the discrete tangent bundle need not be TQ since the connection between configurations and velocities via the derivative of curves has been destroyed by the discretization of the time interval.

This section presents a new structure that replaces the derivative and tangent bundle in such a way as to recover a discrete analogue of the continuous Lagrange–d'Alembert principle in Equation (2.2). A definition of a discrete constraint is proposed that separates the constraint into a distribution and manifold pair. A discrete symplectic form is found, as well as a pair of Lagrange one forms. It is shown that the symplectic form evolves in a manner similar to the continuous symplectic form in Chapter 2. Finally, a discrete momentum equation is obtained that generalizes the Moser–Veselov discrete momentum equation of Cortés and Martínez [7].

3.2 Discrete Tangent Bundle and Lagrangian Systems

Definition 2. A *discrete tangent bundle of* Q is the triple $(P, \partial^+, \partial^-)$, where P is a manifold, $\dim P = 2 \dim Q$ and $\partial^+ : P \to Q$, $\partial^- : P \to Q$ are maps with the following properties:

1. ∂^+, ∂^- are submersions and $TP = \ker T\partial^+ \oplus \ker T\partial^-$.

2. For each $q \in Q$, $T_q^+Q = (\partial^+)^{-1}(q)$ and $T_q^-Q = (\partial^-)^{-1}(q)$ intersect in exactly one point.

Let $\partial^\pm : P \to Q \times Q$ be defined by $\partial^\pm(p) = (\,\partial^+(p)\,,\,\partial^-(p)\,)$. Item 1 of Definition 2 then implies that $T\partial^\pm$ is a bijection and therefore a local diffeomorphism.

Definition 3. The *discrete zero section* is $0_P = (\partial^\pm)^{-1}\Delta(Q \times Q)$, where $\Delta(Q \times Q)$ is the diagonal of $Q \times Q$.

Item 2 of Definition 2 implies that $\partial^\pm$ is one to one from 0_P to $\Delta(Q \times Q)$ so that $\partial^\pm$ is a diffeomorphism of 0_p to $\Delta(Q \times Q)$. Since $\Delta(Q \times Q)$ is a closed submanifold of $Q \times Q$, 0_P is a closed submanifold of P and Lemma 2 in Appendix A provides open neighbourhoods U of 0_P and V of $\Delta(Q \times Q)$ for which $\partial^\pm$ is a diffeomorphism. Without loss of generality, it will be assumed that that $U = P$ and $V = Q \times Q$ in order to make statements and notation more compact.

Definition 4. Let $q_d = \{q_k\}_{k=0}^N$ be an $N+1$ term sequence in Q. The *discrete derivative of* q_d is the N term sequence q_d' in P defined by $(q_d')_k = (\partial^\pm)^{-1}(q_{k+1}, q_k)$.

Definition 5. A sequence $p_d = \{p_k\}_{k=0}^{N-1}$ in P is called *first order* if $\partial^\pm(p_k) = (q_{k+1}, q_k)$ for some sequence $q_d = \{q_k\}_{k=0}^N$ in Q.

A first order sequence, therefore, has the property that $\partial^+(p_k) = q_{k+1} = \partial^-(p_{k+1})$.

Definition 6. The *base* of a first order sequence p_d in P is the sequence q_d in Q given by

$$q_0 = \partial^-(p_0),\ q_1 = \partial^-(p_1),\ \ldots,\ q_k = \partial^-(p_k), \ldots, \tag{3.1}$$

$$q_{N-1} = \partial^-(p_{N-1}),\ q_N = \partial^+(p_{N-1}). \tag{3.2}$$

It follows immediately from the definitions that the discrete derivative of q_d is a first order sequence. Also, it follows from the fact that $\partial^\pm$ is a diffeomorphism that the discrete derivative of q_d is unique.

Define $E_p^+ = \ker T_p\partial^-$ and $E_p^- = \ker T_p\partial^+$. Further, define $E^+ = \cup_{p\in P}E_p^+$ and $E^- = \cup_{p\in P}E_p^-$. Item 1 of Definition 2 gives $T_pP = E_p^+ \oplus E_p^-$. Therefore, every $w_p \in T_pP$ decomposes uniquely into $w_p^+ \in E_p^+$ and $w_p^- \in E_p^-$ such that $w_p = w_p^+ + w_p^-$.

Define the following projections:

$$\pi_+ \colon T_pP \to E_p^+, \qquad \pi_+\, w_p = w_p^+,$$

$$\pi_- \colon T_pP \to E_p^-, \qquad \pi_-\, w_p = w_p^-.$$

These projections can be calculated by

$$\pi_+ w_p = (T_p\partial^\pm)^{-1}(T_p\partial^+ w_p, 0), \qquad \pi_- w_p = (T_p\partial^\pm)^{-1}(0, T_p\partial^- w_p). \tag{3.3}$$

Proposition 7. *Let $\delta q \in T_qQ$ and $p_0, p_1 \in P$ such that $\partial^+(p_0) = q = \partial^-(p_1)$. Then there are unique vectors $\delta p_0^+ \in E_{p_0}^+$ and $\delta p_1^- \in E_{p_1}^-$ such that $T_{p_0}\partial^+\delta p_0^+ = \delta q = T_{p_1}\partial^-\delta p_1^-$.*

Proof. Existence and uniqueness of the vectors follows directly from the fact that $T_p\partial^\pm$ is an isomorphism of T_pP with $T_{\partial^\pm(p)}(Q\times Q)$. □

Proposition 7 provides a linear isomorphism from $E_{p_0}^+$ to $E_{p_1}^-$ in the following way: let $\delta p_0^+ \in E_{p_0}^+$ and $\delta q = T_{p_0}\partial^+\delta p_0^+$. Then δp_1^- is the unique vector in $E_{p_1}^-$ such that $T_{p_1}\partial^-\delta p_1^- = \delta q$. Denote this map as $\tau_{p_1,p_0} : E_{p_0}^+ \to E_{p_1}^-$, so that $\tau_{p_1,p_0}\, \delta p_0^+ = \delta p_1^-$.

Let q_d be a sequence in Q and $\delta q_d = \{\, \delta q_k \,\}_{k=0}^N$ a sequence in TQ such that $\delta q_k \in T_{q_k}Q$. Let p_d be the discrete derivative of q_d and $\delta p_d = \{\, \delta p_k \,\}_{k=0}^{N-1}$ the sequence in TP such that $\delta p_k \in T_{p_k}P$ and $\delta p_k = \delta p_k^+ + \delta p_k^-$ where δp_{k-1}^+ and δp_k^- are the vectors guaranteed by Proposition 7. That is, $T_{p_{k-1}}\partial^+\delta p_{k-1}^+ = \delta q_k = T_{p_k}\partial^-\delta p_k^-$.

Definition 7. The sequence δp_d described above is called the *discrete lift of δq_d with respect to the discrete derivative.*

Definition 8. (Sussmann [23]). Let $\mathcal{D}$ be a distribution on Q. Two points q_a and q_b are in an *orbit* of $\mathcal{D}$ if there is a continuous, piecewise differentiable curve $\gamma\colon [a,b] \to Q$, such that $\gamma(a) = q_a$, $\gamma(b) = q_b$ and $\dfrac{d\gamma(t)}{dt} \in \mathcal{D}_{\gamma(t)}$ wherever γ is differentiable.

Definition 9. A *discrete constraint* is the tuple $(\mathcal{D}, \mathcal{D}_d)$ where $\mathcal{D}$ is a d dimensional distribution (not necessarily integrable) on the n dimensional manifold Q and $\mathcal{D}_d$ is an $n+d$ dimensional submanifold of P such that

1. ∂^+ and ∂^- restricted to $\mathcal{D}_d$ are submersions,

2. $p \in \mathcal{D}_d$ implies that $\partial^+(p)$ and $\partial^-(p)$ are in the same orbit of $\mathcal{D}$ (Definition 8),

3. The discrete zero section, $0_P \subset \mathcal{D}_d$ (Definition 3).

The discrete constraint $(\mathcal{D}, \mathcal{D}_d)$ is said to be C^k if $\mathcal{D}$ is a C^k distribution.

If $\mathcal{D}$ is integrable, then the dimension of $\mathcal{D}_d$ is automatically $n+d$. The Global Frobenius Theorem, Abraham, Marsden and Ratiu [2] gives a foliation of Q by the leaves of $\mathcal{D}$. Let this foliation be denoted $\{Q_\alpha\}_{\alpha \in \Lambda}$. There is then a chart, (U, μ) containing q such that $\mu(U \cap Q_\alpha) = V \times \{c_\alpha\}$ for some open set $V \subset \mathbb{R}^d$ and constant $c_\alpha \in \mathbb{R}^{n-d}$. The point $(q^+, q^-) \in U \times U$ gives a point $(\partial^{\pm})^{-1}(q^+, q^-) \in \mathcal{D}_d$ if and only if q^+ and q^- are in $U \cap Q_\beta$ for some β. Equivalently, $(\mu \times \mu)(q^+, q^-) \in V \times \{c_\beta\} \times V \times \{c_\beta\}$. This constrains the q^+ and q^- to be in a submanifold parameterized by the set $V \times V \times \Delta(\mathbb{R}^{n-d} \times \mathbb{R}^{n-d})$ which has dimension $d + d + n - d = n + d$.

Definition 10. Let $(\mathcal{D}, \mathcal{D}_d)$ be a discrete constraint. If $\mathcal{D}$ is integrable, then the discrete constraint is *holonomic*. If $\mathcal{D}$ is not integrable, then the discrete constraint is *nonholonomic*.

3.3 Formal Setup for the Discrete Lagrange–d'Alembert Principle

The discrete Lagrange–d'Alembert principle will be developed on a discrete tangent bundle in a manner analogous to the continuous Lagrange–d'Alembert principle. The discrete version is a skew critical point problem as described in Appendix B.

Let Q be a configuration manifold and $(P, \partial^+, \partial^-)$ a discrete tangent bundle of Q. The set of $N+1$ element sequences in Q is

$$\Omega_N = \left\{ q_d \in Q^{N+1} \right\},$$

The tangent space to Ω_N at q_d is

$$T_{q_d}\Omega_N = \left\{ v_d \in TQ^{N+1} \,|\, \tau_Q \circ v_d = q_d \right\}.$$

For a system constrained by $(\mathcal{D}, \mathcal{D}_d)$, the set of *admissible sequences* is

$$\mathcal{N}_N = \left\{ q_d \in \Omega_N \,|\, q_d' \in \mathcal{D}_d^N \right\},$$

and the set of *admissible variations* at q_d is

$$\left(\mathcal{D}_N\right)_{q_d} = \left\{ \delta q_d \in TQ^{N+1} \,|\, \delta q_d \in \mathcal{D}^{N+1}, \tau_Q \circ \delta q_d = q_d, \delta q_0 = 0, \delta q_N = 0 \right\}.$$

A *discrete Lagrangian* is a C^2 function $L_d \colon P \to \mathbb{R}$ that characterizes the discrete mechanical system. Chapter 4 contains the construction of a discrete Lagrangian associated to a continuous Lagrangian..

The tuple $(L_d, P, \partial^+, \partial^-, Q, \mathcal{D}, \mathcal{D}_d)$ is a *discrete constrained Lagrangian system.* Without the constraint, the tuple $(L_d, P, \partial^+, \partial^-, Q)$ is simply a *discrete Lagrangian system.*

The *discrete action* is defined as

$$S_d \colon \Omega_N \to \mathbb{R}, \qquad S_d(q_d) = \sum_{k=0}^{N-1} L_d((q_d')_k), \tag{3.4}$$

and the *discrete Lagrange–d'Alembert principle* is the skew problem

$$dS_d(q_d)\ \left(\mathcal{D}_N\right)_{q_d} = 0, \tag{3.5}$$

$$q_d \in \mathcal{N}_N. \tag{3.6}$$

The discrete Lagrange–d'Alembert principle is a necessary and sufficient condition for a sequence to be a discrete trajectory of the discrete Lagrangian system.

Equation (3.5) is evaluated by first taking the derivative of S_d on Ω_N and then evaluating on vectors in $(\mathcal{D}_N)_{q_d}$.

Let $q_d \in \Omega_N$ and q_d' be its discrete derivative. Let $q_d(\epsilon)$ be a curve in Ω_N and $q_d'(\epsilon)$ be the discrete derivative of $q_d(\epsilon)$ such that $q_d(0) = q_d$ and $q_d'(0) = q_d'$. Let $\delta q_d = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q_d(\epsilon)$. Then,

$$\begin{aligned}
dS_d(q_d)\,\delta q_d &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S_d(q_d(\epsilon)) \\
&= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \sum_{k=0}^{N-1} L_d(q_d'(\epsilon)_k) \qquad \text{definition of } S_d \\
&= \sum_{k=0}^{N-1} dL_d((q_d')_k)\, \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (q_d'(\epsilon))_k \qquad \text{chain rule} \\
&= \sum_{k=0}^{N-1} dL_d((q_d')_k)\,\delta p_k \qquad \text{for some } \delta p_k \in T_{(q_d')_k} P \\
&= \sum_{k=0}^{N-1} \big(dL_d((q_d')_k)\,\delta p_k^+ + dL_d((q_d')_k)\,\delta p_k^-\big) \qquad \text{decompose } \delta p_k \text{ with} \\
&\qquad\qquad \text{respect to } E_p^+ \text{ and } E_p^- \\
&= \sum_{k=0}^{N-2} dL_d((q_d')_k)\,\delta p_k^+ + \sum_{k=1}^{N-1} dL_d((q_d')_k)\,\delta p_k^- + \\
&\qquad + dL_d((q_d')_0)\,\delta p_0^- + dL_d((q_d')_{N-1})\,\delta p_{N-1}^+ \qquad \text{separate out first} \\
&\qquad\qquad \text{and last terms} \\
&= \sum_{k=0}^{N-2} dL_d((q_d')_k)\,\delta p_k^+ + \sum_{k=1}^{N-1} dL_d((q_d')_k)\,\tau_{p_k,p_{k-1}}\delta p_{k-1}^+ + \\
&\qquad + dL_d((q_d')_0)\,\delta p_0^- + dL_d((q_d')_{N-1})\,\delta p_{N-1}^+ \\
&\qquad\qquad \text{since } \tau_{p_k,p_{k-1}}\delta p_{k-1}^+ = \delta p_k^- \\
&= \sum_{k=0}^{N-2} dL_d((q_d')_k)\,\delta p_k^+ + \sum_{k=0}^{N-2} dL_d((q_d')_{k+1})\,\tau_{p_{k+1},p_k}\delta p_k^+ + \\
&\qquad + dL_d((q_d')_0)\,\delta p_0^- + dL_d((q_d')_{N-1})\,\delta p_{N-1}^+ \\
&\qquad\qquad \text{reindexing the second sum} \\
&= \sum_{k=0}^{N-2} \big(\, dL_d((q_d')_k) + dL_d((q_d')_{k+1})\,\tau_{p_{k+1},p_k}\,\big)\,\delta p_k^+ +
\end{aligned}$$

$$+ \, dL_d((q_d')_0)\, \delta p_0^- + dL_d((q_d')_{N-1})\, \delta p_{N-1}^+ \tag{3.7}$$

Restricting $\delta q_d \in (\mathcal{D}_N)_{q_d}$, the boundary terms in Equation (3.7) satisfy $T_{p_0}\partial^- \delta p_0^- = \delta q_0 = 0$ and $T_{p_{N-1}}\partial^+ \delta p_{N-1}^+ = \delta q_N = 0.$, Equation (3.5) becomes

$$\sum_{k=0}^{N-2} (\, dL_d((q_d')_k) + dL_d((q_d')_{k+1})\, \tau_{p_{k+1},p_k}\,)\, \delta p_k^+ = 0, \tag{3.8}$$

where $T_{p_{k-1}}\partial^+ \delta p_{k-1}^+ = \delta q_k \in \mathcal{D}$. Choosing $\delta p_i^+ = 0$ for all $i \neq k$ in turn, Equation (3.8) is

$$(\, dL_d((q_d')_k) + dL_d((q_d')_{k+1})\, \tau_{p_{k+1},p_k}\,)\, \delta p_k^+ = 0, \qquad k = 0 \ldots N-2.$$

Since $\tau_{p_{k+1},p_k} \delta p_k^+ = \delta p_{k+1}^-$, and $(q_d')_k = p_k$, a more explicit form of the discrete Lagrange–d'Alembert principle is, for $k = 0, \ldots N-2$,

$$dL_d(p_k)\, \delta p_k^+ + dL_d(p_{k+1})\, \delta p_{k+1}^- = 0, \tag{3.9}$$

$$T_{p_k}\partial^+ \delta p_k^+ = T_{p_{k+1}}\partial^- \delta p_{k+1}^-, \tag{3.10}$$

$$T_{p_k}\partial^+ \delta p_k^+ \in \mathcal{D}, \tag{3.11}$$

$$p_k, p_{k+1} \in \mathcal{D}_d. \tag{3.12}$$

Equations (3.10) and (3.11) restrict the vectors δp_k^+ and δp_{k+1}^- to d degrees of freedom, giving d independent equations in Equation (3.9). The condition $\partial^+(p_k) = \partial^-(p_{k+1})$, implied by Equation (3.10), gives n more equations for a total of $n+d$ equations. Equation (3.12) implies that p_k and p_{k+1} are each variables in an $n+d$ dimensional manifold, leaving $n+d$ degrees of freedom in the set of equations, assuming the system has full rank. As in the continuous case, fixed endpoints or an initial condition for the entire sequence may be chosen, so long as Equation (3.12) is satisfied, leaving $n+d$ variables to be solved for, for each k.

Equations (3.5) and (3.6) give the discrete Lagrange–d'Alembert principle as a skew problem on sequences in Q with infinitesimal variations a sequence in $\mathcal{D}$. Equations (3.9), (3.10), (3.11) and (3.12) recast the discrete Lagrange–d'Alembert principle on the direct product of two discrete tangent bundles, $P \times P$. Define the d dimensional distribution

$$\mathcal{D}_P = \{(\delta p, \delta p') \in T(P \times P) \,|\, T\partial^+ \,\delta p = T\partial^- \,\delta p' \in \mathcal{D}, T\partial^- \,\delta p = 0, T\partial^+ \,\delta p' = 0\}, \tag{3.13}$$

and the $n+2d$ dimensional submanifold,

$$\mathcal{A} = \{(p, p') \in \mathcal{D}_d \times \mathcal{D}_d \,|\, \partial^+(p) = \partial^-(p')\}.$$

Let M be an $n+d$ dimensional manifold and $g\colon \mathcal{A} \to M$ a C^1 mapping with $m \in M$ a regular value. Then the discrete Lagrange–d'Alembert principle is the following skew critical point problem on $\mathcal{A}$ with respect to $\mathcal{D}_P$ and g:

$$(dL_d(p) + dL_d(p'))\, \mathcal{D}_P = 0, \tag{3.14}$$

$$(p, p') \in \mathcal{A}, \tag{3.15}$$

$$g(p, p') = m. \tag{3.16}$$

The purpose of the mapping g is to constrain the equations to a submanifold of $\mathcal{A}$. A useful choice is $M = \mathcal{D}_d$ and $g(p, p') = p$. Then the above equations are skew critical equations for p'.

Equations (3.14), (3.15) and (3.16) lift the discrete Euler Lagrange Equations from $Q \times Q$ to P. In Chapter 4, where $P = TQ$, this corresponds to writing the discrete equations of motion on TQ. This is in contrast to the approach taken by Cortés and Martínez [7], de León et. al. [11] and McLachlan and Perlmutter [19] where the discrete Lagrange–d'Alembert equations are developed on $Q \times Q$. See Section 3.5.3 for the connection between Moser–Veselov integrators and the Lagrange–d'Alembert integrators developed here.

This skew problem becomes a critical point problem when $\mathcal{D}$ is integrable, in which case $T_{q_d}\mathcal{N}_N = (\mathcal{D}_N)_{q_d}$, where, for $q_d(\epsilon)$ a curve in $\mathcal{N}_N$,

$$T_{q_d}\mathcal{N}_N = \left\{ v_d \in TQ^{N+1} \,|\, \tau_Q \circ v_d = q_d, v_d = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q_d(\epsilon) \right\}.$$

Proposition 8. *Let $(\mathcal{D}, \mathcal{D}_d)$ be a C^1 discrete constraint. $\mathcal{D}$ is integrable if and only if $(\mathcal{D}_N)_{q_d} = T_{q_d}\mathcal{N}_N$.*

Proof. If $\mathcal{D}$ is integrable then $\mathcal{N}_N$ is the set of sequences in the leaves of $\mathcal{D}$ so that, for $q_d(\epsilon) \in \mathcal{N}_N$, $\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q_d(\epsilon) = \delta q_d \in \mathcal{D}^{N+1}$. As described following Definition 9, $\mathcal{D}_d$ must be an

$n + d$ dimensional submanifold of P.

Suppose $T_{q_d}\mathcal{N}_N = (\mathcal{D}_N)_{q_d}$. Let X and Y be vector fields on Q taking their values in $\mathcal{D}$. Let the flows of X and Y be F_t^X and F_t^Y respectively.

Let $q \in Q$, C a closed set containing q and V an open set containing C such that $F_a^X(q)$, $F_b^X(q)$, $F_a^Y(q)$ and $F_b^Y(q)$ are not in V. Using an appropriate bump function, vector fields $\tilde{X}$ and $\tilde{Y}$ can be arranged so that on C, $\tilde{X} = X$, $\tilde{Y} = Y$ and $\tilde{X} = \tilde{Y} = 0$ outside V.

Without loss of generality, then, it can be assumed that the flow of X satisfies $X(F_a^X(q)) = X(F_b^X(q)) = 0$ and the flow of Y satisfies $Y(F_a^Y(q)) = Y(F_b^Y(q)) = 0$.

Let $[a, b]$ be an interval, $b > a$ and $\{t_k\}_{k=0}^N$ an increasing sequence in $[a, b]$. Then

$$k \mapsto \left\{F_{t_k}^X \circ F_\epsilon^Y(q)\right\}$$

is in $\mathcal{N}_N$ since it is in an orbit of $\mathcal{D}$ by virtue of it being in an integral curve of X. By hypothesis, $T_q F_{t_k}^X Y$ is in $\mathcal{D}$ since

$$T_q F_{t_k}^X Y = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \left(F_{t_k}^X \circ F_\epsilon^Y\right)(q).$$

This holds for arbitrary t_k and q so that,

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (F_{-t}^X \circ F_\epsilon^Y)(F_t(q)) = T_{F_t^X(q)} F_{-t}^X\, Y(F_t^X(q)) = (F_t^X)^* Y(q) \in \mathcal{D},$$

and, therefore

$$\left.\frac{d}{dt}\right|_{t=0} \left(F_t^X\right)^* Y = [X, Y] \in \mathcal{D}.$$

□

3.4 Discrete Equations of Motion

Define the following submanifold of $P \times P$:

$$\ddot{Q}_d = \{\, (\, (q_d')_0, (q_d')_1 \,) \;|\; q_d \text{ a sequence in } Q\,\},$$

where $(q_d')_i$ is the ith element of the discrete derivative of q_d, $i = 1, 2$. $\ddot{Q}_d$ contains elements (p_0, p_1) such that $\partial^+(p_0) = \partial^-(p_1)$. The set $\ddot{Q}$ appears in Marsden and West [18] for Moser–Veselov integrators.

Let $q_d(\epsilon)$ be a curve in Ω_N with $q_d(0) = q_d$, $\delta q_d = \frac{d}{d\epsilon}\big|_{\epsilon=0} q_d(\epsilon)$ and $\delta p_d = \frac{d}{d\epsilon}\big|_{\epsilon=0} q_d'(\epsilon)$. The derivative of the discrete action, Equation (3.4) is then, as calculated in Section 3.3

$$\begin{aligned} dS_d(q_d)\,\delta q_d &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} S_d(q_d(\epsilon)) \\ &= \sum_{k=0}^{N-2} \big(\, dL_d((q_d')_k)\,\delta p_k^+ + dL_d((q_d')_{k+1})\,\delta p_{k+1}^- \big) + \\ &\qquad + dL_d((q_d')_0)\,\delta p_0^- + dL_d((q_d')_{N-1})\,\delta p_{N-1}^+ \end{aligned} \tag{3.17}$$

The expression under the summation in Equation (3.17) is a linear form defined on $\ddot{Q}_d$ acting on a vector $\delta q \in T_q Q$ in the following way: Let $(p_0, p_1) \in \ddot{Q}_d$ such that $\partial^+(p_0) = q = \partial^-(p_1)$. Also let $\delta p_0^+ \in E_{p_0}^+$ and $\delta p_1^- \in E_{p_1}^-$ such that $T_{p_0}\partial^+ \delta p_0^+ = \delta q = T_{p_1}\partial^- \delta p^-$. Define

$$D_{EL}L_d \colon \ddot{Q} \to T^*Q, \qquad D_{EL}L_d(p_0, p_1)\,\delta q = dL_d(p_0)\,\delta p_0^+ + dL_d(p_1)\,\delta p_1^-.$$

The boundary terms in equation (3.17) motivate the definitions of the two *discrete Lagrange* one forms:

$$\Theta_{L_d}^+(p)\,\delta p = dL_d(p)\,\pi_+\,\delta p, \tag{3.18}$$

$$\Theta_{L_d}^-(p)\,\delta p = -dL_d(p)\,\pi_-\,\delta p, \tag{3.19}$$

where the minus sign in the definition of Θ_{L_d} is to promote the analogy with the continuous theory.

Theorem 2. *Let $(L_d, P, \partial^+, \partial^-, Q)$ be a discrete Lagrangian system with L_d a C^k discrete Lagrangian, $k > 1$. Then there exists a unique C^{k-1} mapping $D_{EL}L_d \colon \ddot{Q}_d \to T^*Q$ and two C^{k-1} one forms $\Theta_{L_d}^+$ and $\Theta_{L_d}^-$ on P such that, for all curves $q_d(\epsilon)$ in Ω_N,*

$$dS_d(q_d)\,\delta q_d = \sum_{k=0}^{N-2} D_{EL}L_d((q_d')_k, (q_d')_{k+1})\,\delta q_k +$$

$$- \Theta^-_{L_d}((q'_d)_0)\, \delta p_0 + \Theta^+_{L_d}((q'_d)_{N-1})\, \delta p_{N-1},$$

where

$$\delta q_d = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q_d(\epsilon) \;\; and \;\delta p_d = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} q'_d(\epsilon).$$

The discrete Lagrange–d'Alembert principle (3.9), (3.10), (3.11) and (3.12) can be rewritten as

$$D_{EL}L_d(p_k, p_{k+1}) \in \mathcal{D}^0, \tag{3.20}$$

$$p_k, p_{k+1} \in \mathcal{D}_d, \tag{3.21}$$

where $k = 0 \ldots N-2$.

Define the distributions

$$\mathcal{K}_d = (T\partial^-)^{-1}\mathcal{D} \cap (T\partial^+)^{-1}\mathcal{D}$$

and

$$\mathcal{D}^- = \mathcal{K}_d \cap \ker T\partial^+, \tag{3.22}$$

$$\mathcal{D}^+ = \mathcal{K}_d \cap \ker T\partial^-. \tag{3.23}$$

In fact, the distributions are set up so that $\mathcal{K}_d = \mathcal{D}^+ \oplus \mathcal{D}^-$, $T\partial^+\mathcal{D}^+ = \mathcal{D}$ and $T\partial^-\mathcal{D}^- = \mathcal{D}$. The fibre dimension of $\mathcal{K}_d$ is $2d$, since $X \in \mathcal{K}_d$ if $T\partial^+X \in \mathcal{D}$ and $T\partial^-X \in \mathcal{D}$.

$\mathcal{K}_d$ serves as an analogue of the continuous $\mathcal{K}_\mathcal{D}$ of Chapter 2 by identifying the distribution of admissible variations on the discrete tangent space.

Let $f\colon Q \times Q \to \mathbb{R}$ and let X and Y be vector fields on $Q \times Q$ such that, for all $(\tilde{q}^+, \tilde{q}^-)$ in a neighbourhood U of (q^+, q^-),

1. X is in $TQ \times \{0\}$ and $X(q^+, \tilde{q}^-) = X(q^+, q^-)$,

2. Y is in $\{0\} \times TQ$ and $Y(\tilde{q}^+, q^-) = Y(q^+, q^-)$.

Then $(X, Y) \mapsto X(Yf)$ is a bilinear form, since, in local coordinates (r^i, s^i) on U

$$\begin{aligned}
\hat{X}(\hat{Y}\hat{f}) &= \hat{X}\hat{Y}^i \frac{\partial \hat{f}}{\partial s^i} \\
&= \hat{X}^j \frac{\partial \hat{Y}^i}{\partial r^j} \frac{\partial \hat{f}}{\partial s^i} + \hat{X}^j \hat{Y}^i \frac{\partial^2 \hat{f}}{\partial r^j \partial s^i} \\
&= \hat{X}^j \hat{Y}^i \frac{\partial^2 \hat{f}}{\partial r^j \partial s^i}.
\end{aligned}$$

This bilinear form induces another bilinear form on P as follows. Let $f\colon P \to \mathbb{R}$ and $f_{Q\times Q} = f \circ (\partial^{\pm})^{-1}$. Let X^+ and X^- be vector fields on P such that, for all $\tilde{p}$ in a neighbourhood U of p,

1. X^+ takes its values in $\mathcal{D}^+$ and $X^+(\tilde{p}) = X^+(p)$ when $\partial^+(\tilde{p}) = \partial^+(p)$,

2. X^- takes its values in $\mathcal{D}^-$ and $X^-(\tilde{p}) = X^-(p)$ when $\partial^-(\tilde{p}) = \partial^-(p)$.

Let $Z = T\partial^{\pm} X^+$ and $Y = T\partial^{\pm} X^-$ and define

$$X^+(X^- f)(p) = Z(Y f_{Q\times Q})(q^+, q^-),$$

where $\partial^{\pm}(p) = (q^+, q^-)$.

Definition 11. Let $(L_d, P, \partial^+, \partial^-, Q, \mathcal{D}, \mathcal{D}_d)$ be a discrete constrained Lagrangian system. Let V and W be open in $\mathcal{D}_d$. L_d is called $(V, W) - \mathcal{D}$*–regular* if,

1. for all $p_0 \in V$ there is a unique $p_1 \in W$ such that (p_0, p_1) satisfies Equations (3.14) and (3.15),

2. $(\mathcal{K}_d)_p \subset T_p\mathcal{D}_d$ for all $p_0 \in V$,

3. For $X^+ \in \mathcal{D}^+$ and $X^- \in \mathcal{D}^-$ (see Equations (3.22) and (3.23)), the bilinear form $X^+(X^- L_d)(p_1)$ is nondegenerate for all $p_1 \in W$ such that there is a $p_0 \in V$ so that (p_0, p_1) satisfy Item 1.

If L_d is $(V, W) - \mathcal{D}$–regular and $\mathcal{D} = TQ$, then L_d will be called (V, W)–regular.

Item 1 of Definition 11 provides the existence of a map,

$$F_{L_d}\colon V \to W, \qquad F_{L_d}(p_0) = p_1. \tag{3.24}$$

Proposition 9. *Let $(\mathcal{D}, \mathcal{D}_d)$ be a C^k discrete constraint, $k > 1$, and let $L_d\colon P \to \mathbb{R}$ be a C^k, $(V, W) - \mathcal{D}$–regular discrete Lagrangian. Then F_{L_d}, Equation (3.24), is C^{k-1}.*

Proof. Let $(p_0, p_1) \in V \times W$ such that $F_{L_d}(p_0) = p_1$.

The implicit function theorem will be used to prove the theorem, which will require differentiating Equation (3.14) with respect to p_1 in directions in $\mathcal{D}^+$ on $(\partial^-)^{-1}(\partial^+(p_0))$. To this end, let Z be a vector field in $\mathcal{D}_P$ (Equation (3.13)) such that $Z = (0, Y)$ for some Y. Then, $Y \in \mathcal{D}^+$ on $(\partial^-)^{-1}(\partial^+(p_0))$ and

$$Z\,(dL_d(p_0)\,\delta p_0 + dL_d(p_1)\,\delta p_1) = Y(dL_d(p_1)\delta p_1) = (Y(\delta p_1 L_d))(p_1).$$

Since $\delta p_1 \in \mathcal{D}^+$, the above derivative is nonsingular by hypothesis.

The implicit function theorem therefore guarantees the existence of a unique C^{k-1} mapping F and appropriate neighbourhoods of p_0 and p_1 such that $F(p_0) = p_1$. By uniqueness, this must be the F_{L_d} already defined, and hence F_{L_d} is C^{k-1}. □

The discrete Lagrange–d'Alembert principle can be more explicitly rewritten as

$$dL_d(p_k)\,\delta p_k + dL_d(p_{k+1})\,\delta p_{k+1} = 0,$$

$$T_{p_k}\partial^-\,\delta p_k = 0,$$

$$T_{p_{k+1}}\partial^+\,\delta p_{k+1} = 0,$$

$$T_{p_k}\partial^+\,\delta p_k = T_{p_{k+1}}\partial^-\,\delta p_{k+1} \in \mathcal{D}_{\partial^+(p_k)},$$

$$\partial^+(p_k) = \partial^-(p_{k+1}),$$

$$p_k, p_{k+1} \in \mathcal{D}_d.$$

These equations will be called the *discrete constrained Euler–Lagrange equations.*

If $(L_d, P, \partial^+, \partial^-, Q, \mathcal{D}, \mathcal{D}_d)$ is $(V, W) - \mathcal{D}$–regular then there is a unique sequence of solutions to the discrete constrained Euler–Lagrange equations for each $p_0 \in \mathcal{D}_d$. Define

$$F_{L_d} \colon \mathcal{D}_d \to \mathcal{D}_d, \qquad F_{L_d}(p_k) = p_{k+1},$$

where (p_k, p_{k+1}) satisfy the discrete Euler–Lagrange equations. F_{L_d} is called the *discrete flow map*.

The space of solutions of the discrete constrained Euler–Lagrange equations is parameterized by the set of initial conditions as

$$p \mapsto \{\, F_{L_d}^k(p) \,\}_{k=0}^{N-1},$$

for $p \in \mathcal{D}_d$. Define $S_N \colon P \to \mathbb{R}$ by

$$S_N(p) = \sum_{k=0}^{N-1} L_d(F_{L_d}^k(p)), \tag{3.25}$$

and then, for $p \in D_d$ and $w_p \in T_pP$, using Theorem 2,

$$dS_N(p)\, w_p = \sum_{k=0}^{N-2} D_{EL} L_d(F_{L_d}^k(p), F_{L_d}^{k+1}(p))\, v_k + \\ - \Theta_{L_d}^-(p)\, w_p + \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_p, \tag{3.26}$$

where $v_k = T_p(\partial^+ \circ F_{L_d}^k)\, w_p = T_p(\partial^- \circ F_{L_d}^{k+1})\, w_p$. Write

$$dS_N = \alpha_N + (F_{L_d}^{N-1})^* \Theta_{L_d}^+ - \Theta_{L_d}^-, \tag{3.27}$$

where α_N is defined by and replaces the summation term in equation (3.26).

The definitions of Θ^+ and Θ^- in Equations (3.18) and (3.19) give

$$dL_d = \Theta_{L_d}^+ - \Theta_{L_d}^-.$$

Differentiating gives

$$0 = d^2 L_d = d\Theta_{L_d}^+ - d\Theta_{L_d}^-,$$

so that $d\Theta_{L_d}^+ = d\Theta_{L_d}^-$. Define $\omega_{L_d} = -d\Theta_{L_d}^+ = -d\Theta_{L_d}^-$. Differentiating Equation (3.27) gives

$$0 = d^2 S_N = d\alpha_N + (F_{L_d}^{N-1})^* d\Theta_{L_d}^+ - d\Theta_{L_d}^-,$$

which is equivalent to

$$(F_{L_d}^{N-1})^* \omega_{L_d} = \omega_{L_d} + d\alpha_N, \tag{3.28}$$

giving an evolution equation for ω_{L_d}.

Neglecting the constraint for a moment, the two–form ω_{L_d} only needs to be shown to be nondegenerate in order to be symplectic, which is the subject of Proposition 10. This does not imply, however, that the discrete flow is symplectic as the term $d\alpha_N$ in Equation (3.28) obstructs this.

Proposition 10. *Let $L_d\colon P \to \mathbb{R}$ be a C^k, $k > 1$ (V, W)–regular discrete Lagrangian. Then the two form ω_{L_d} defined above is nondegenerate.*

Proof. Let X_0, X_1 be vector fields in TP. According to Section 3.2 they can be decomposed into $X_0 = X_0^+ + X_0^-$ where $X_0^+ \in E^+$ and $X_0^- \in E^-$ and similarly for X_1. Using $\omega_{L_d} = -d\Theta_{L_d}^+$, calculate

$$\begin{aligned} d\Theta^+_{L_d}(\,X_0\,,\,X_1\,) &= X_0\,(d\Theta^+_{L_d}\,X_1) - X_1\,(d\Theta^+_{L_d}\,X_0) - d\Theta^+_{L_d}\,[\,X_0\,,\,X_1\,] \\ &= X_0\,(dL_d\,X_1^+) - X_1\,(dL_d\,X_0^+) + dL_d\,[\,X_0\,,\,X_1\,]^+ \qquad \text{definition of } d\Theta^+_{L_d} \\ &= X_0^+\,(dL_d\,X_1^+) + X_0^-\,(dL_d\,X_1^+) - X_1^+\,(dL_d\,X_0^+) + \\ &\qquad - X_1^-\,(dL_d\,X_0^+) + dL_d\,[\,X_0^+\,,\,X_1^+\,] \qquad \text{decompose } X_0 \text{ and } X_1 \\ &= d^2L_d\,(\,X_0^+\,,\,X_1^+\,) + X_0^-\,(dL_d\,X_1^+) - X_1^-\,(dL_d\,X_0^+) \\ &= X_0^-\,(dL_d\,X_1^+) - X_1^-\,(dL_d\,X_0^+). \end{aligned}$$

Which proves the theorem. □

Note that $\alpha_2(p_0)\,w_{p_0} = D_{EL}L_d(p_0,p_1)\,v_q$, where

$$\partial^+(p_0) = q = \partial^-(p_1),$$

$$T_{p_0}\partial^+\,w_{p_0} = v_q = T_{p_0}(\partial^- \circ F_{L_d})\,w_{p_0},$$

so that $D_{EL}L_d(p_0, p_1)\, v_q = 0$ if and only if $\alpha_2(p_0)\, w_{p_0} = 0$. This implies

$$\alpha_2(p_0) = \lambda_a(p_0)(\partial^+)^* \phi^a(p_0) = \lambda_a(p_0)(\partial^- \circ F_{L_d})^* \phi^a(p_0)$$

for some functions $\lambda_a \colon P \to \mathbb{R}$ and some ϕ^a in $\mathcal{D}^0$.

3.5 Special Cases

3.5.1 Holonomic Constraints

If the constraints are holonomic, then $\mathcal{D}$ is integrable and Proposition 8 implies that the system is equivalent to unconstrained evolution on the leaves of $\mathcal{D}$. Let Q_0 be a leaf of $\mathcal{D}$ and $\mathcal{D}_{d0} = (\partial^{\pm})^{-1}(Q_0 \times Q_0)$. Write $\partial_0^+ = \partial^+|_{\mathcal{D}_{d0}}$ and $\partial_0^- = \partial^-|_{\mathcal{D}_{d0}}$. Then, $(\mathcal{D}_{d0}, \partial_0^+, \partial_0^-)$ is a discrete tangent bundle of Q_0.

Let $L_{d0} = L|_{D_{d0}}$. Then $\Theta^+_{L_{d0}} = dL_{d0}\, \pi_+$, $\Theta^-_{L_{d0}} = -dL_{d0}\, \pi_-$ and $\omega_{L_d0} = -d\Theta^+_{L_{d0}} = -d\Theta^-_{L_{d0}}$ and

$$S_{d0}(q_d) = \sum_{k=0}^{N-1} L_{d0}(p_k), \qquad q_d \in Q_0^{N+1}, \qquad \partial^{\pm}(p_k) = (q_{k+1}, q_k).$$

The discrete Lagrange–d'Alembert principle is then

$$dS_{d0}(q_d) = 0, \qquad q_d \text{ in } Q_0^{N+1}.$$

For each pair of points, $(p_0, p_1) \in P \times P$,

$$D_{EL}L_{d0}(p_0, p_1)\, \delta q = 0, \qquad \delta q \in TQ_0. \tag{3.29}$$

Which is equivalent to

$$dL_{d0}(p_0)\delta p_0^+ + dL_{d0}(p_1)\delta p_1^- = 0,$$

$$\partial^+(p_0) = \partial^-(p_1),$$

$$p_0, p_1 \in D_{d0},$$

$$T_{p_0}\partial^+\delta p_0^+ = T_{p_1}\partial^-\delta p_0^- = \delta q \in TQ_0.$$

Since $F^k_{L_d}$ preserves D_{d0}, $T_p(\partial^+ \circ F^k_{L_d})\, w_p \in TQ_0$ and $T_p(\partial^- \circ F^k_{L_d})\, w_p \in TQ_0$, implying

$$\alpha_N(p)\, w_p = \sum_{k=0}^{N-2} D_{EL}L_{d0}(F^k_{L_d0}(p), F^{k+1}_{L_d0})\, v_k = 0,$$

since the summand satisfies Equation (3.29) for each k. This reduces Equation (3.28) to

$$(F^{N-1}_{L_{d0}})^*\omega_{L_{d0}} = \omega_{L_{d0}},$$

showing that the discrete flow is symplectic.

3.5.2 Special Symplectic Solutions

Proposition 11 gives conditions for the discrete evolution to be symplectic, even if the discrete constraint is nonholonomic.

Proposition 11. *Let $(\mathcal{D}, \mathcal{D}_d)$ be a discrete constraint, $\mathcal{D}_{d0}$ a submanifold of $\mathcal{D}_d$ and $\mathcal{K}_{d0}$ a subbundle of $\mathcal{K}_d$ on $\mathcal{D}_{d0}$. Suppose that $TF^k_{L_d}\,\mathcal{K}_{d0} \subseteq \mathcal{K}_{d0}$ on $\mathcal{D}_{d0}$ and $[Y_1, Y_2] \subset \mathcal{K}_{d0}$ for vector fields Y_1 and Y_2 in $\mathcal{K}_{d0}$. Then F_{L_d} is symplectic on $\mathcal{K}_{d0}$.*

Proof. Recall Equation (3.28),

$$(F^{N-1}_{L_d})^*\omega_{L_d} = \omega_{L_d} + d\alpha_N,$$

Evidently, F_{L_d} will preserve ω_{L_d} when $d\alpha_N$ is zero. Compute

$$d\alpha_N(p)\,(Y_1, Y_2) = Y_1(\alpha_N(Y_2))(p) - Y_2(\alpha_N(Y_1))(p) - \alpha_N([Y_1, Y_2])(p), \tag{3.30}$$

so that $d\alpha_N$ will be zero when each term is zero. Recall,

$$\alpha_N(p)\, w_p = \sum_{k=1}^{N-2} D_{EL}L_d(F^k_{L_d}(p), F^{k+1}_{L_d}(p))\, v_k,$$

where $v_k = T_p(\partial^+ \circ F^k_{L_d})\, w_p = T_p(\partial^- \circ F^{k+1}_{L_d})\, w_p$. If $w_p \in \mathcal{K}_{d0}$ and $T_pF^k_{L_d}\, w_p \in \mathcal{K}_{d0}$, the summand will vanish identically. By hypothesis, then, each term on the right hand side of Equation (3.30) is zero. Therefore $(F^{N-1}_{L_d})^*\,\omega_{L_d} = \omega_{L_d}$ on $\mathcal{K}_{d0}$. □

3.5.3 Moser–Veselov Systems

A Moser–Veselov system is a discrete Lagrange–d'Alembert system where $P = Q \times Q$ and $\partial^+(q^+, q^-) = q^+$ and $\partial^-(q^+, q^-) = q^-$. Then $E^+_{(q^+,q^-)} = T_{q^+}Q \times \{0_{q^-}\}$ and $E^-_{(q^+,q^-)} = \{0_{q^+}\} \times T_{q^-}Q$.

Let q_d be a sequence in Q. The discrete derivative of q_d is the sequence $\{(q_k^+, q_k^-)\}$ such that $\partial^+(q_k^+, q_k^-) = \partial^-(q_{k+1}^+, q_{k+1}^-)$. That is, $q_k^+ = q_{k+1}^-$.

The discrete lift $\{(\delta q_k^+, \delta q_k^-)\}$ of a sequence of tangent vectors δq_d satisfies $T_{(q_k^+,q_k^-)}\partial^+ (\delta q_k^+, \delta q_k^-) = T_{(q_{k+1}^+,q_{k+1}^-)}\partial^- (\delta q_{k+1}^+, \delta q_{k+1}^-)$. That is $\delta q_k^+ = \delta q_{k+1}^-$.

The discrete Lagrange–d'Alembert principle, from Equations (3.9), (3.10), (3.11), and (3.12), becomes

$$dL_d(q_k^+, q_{k-1}^+)\,(\delta q_k^+, 0) + dL_d(q_{k+1}^+, q_k^+)\,(0, \delta q_k^+) = 0,$$

$$\delta q_k^+ \in \mathcal{D}_{q_k^+},$$

$$q_d \in \mathcal{N}_N.$$

Writing q_k for q_k^+ and δq_k for δq_k^+ , these equations are

$$(D_1L_d(q_k, q_{k-1}) + D_2L_d(q_{k+1}, q_k))\,\delta q_k = 0,$$

$$\delta q_k \in \mathcal{D}_{q_k},$$

$$q_d \in \mathcal{N}_N.$$

These equations appear in Cortés and Martínez [7] for Moser–Veselov integrators.

3.6 Discrete Semi–Hamilton's Equations

Using the definition of S_N in Equation (3.25), Equation (3.26) is, for $w_p \in T_pP$,

$$\sum_{k=0}^{N-1} (F^k_{L_d})^* dL_d(p) w_p = \sum_{k=0}^{N-2} D_{EL}L_d(F^k_{L_d}(p), F^{k+1}_{L_d}(p)) v_k +$$

$$- (F_{L_d}^{N-1})^* \Theta_{L_d}^+(p) w_p - \Theta_{L_d}(p) w_p, \quad (3.31)$$

where $v_k = T_p(\partial^+ \circ F_{L_d}^k)\, w_p = T_p(\partial^- \circ F_{L_d}^{k+1})\, w_p$. The same solution sequence extended to $N+1$ terms (counting the q_ks) produces

$$\sum_{k=0}^{N} (F_{L_d}^k)^* dL_d(p) w_p = \sum_{k=0}^{N-1} D_{EL} L_d(F_{L_d}^k(p), F_{L_d}^{k+1}) v_k +$$

$$- (F_{L_d}^{N})^* \Theta_{L_d}^+(p) w_p - \Theta_{L_d}(p) w_p. \quad (3.32)$$

Subtracting Equation (3.31) from Equation (3.32) isolates the Nth term as

$$(F_{L_d}^N)^* dL_d(p)\, w_p = D_{EL} L_d(F_{L_d}^{N-1}(p), F_{L_d}^N(p))\, v_{N-1} +$$

$$+ (F_{L_d}^N)^* \Theta_{L_d}^+(p)\, w_p - (F_{L_d}^{N-1})^* \Theta_{L_d}^+(p)\, w_p. \quad (3.33)$$

Let $\{w_i\}$ be a sequence in T_pP such that $T_p(\partial^+ \circ F_{L_d}^i)\, w_p \in \mathcal{D}_{\partial^+ \circ F_{L_d}^i(p)}$. Then $D_{EL} L_d(F_{L_d}^{N-1}(p), F_{L_d}^N(p)) v_{N-1} = 0$ by the discrete Lagrange–d'Alembert principle. Equation (3.33) is, for $i = N$,

$$(F_{L_d}^N)^* dL_d(p)\, w_{N-1} = (F_{L_d}^N)^* \Theta_{L_d}^+(p)\, w_{N-1} - (F_{L_d}^{N-1})^* \Theta_{L_d}^+(p)\, w_{N-1}.$$

The definition of the pullback, for $F_{L_d}^{N-1}$, then gives

$$F_{L_d}^* dL_d(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1} = F_{L_d}^* \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1} +$$

$$- \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1}. \quad (3.34)$$

Then, expanding dL_d and using $\Theta_{L_d}^+ = dL_d\, \pi_+$ and $\Theta_{L_d}^- = -dL_d\, \pi_-$ from Section 3.2,

$$F_{L_d}^* (dL_d\, \pi_+ + dL_d\, \pi_-)(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1} =$$

$$F_{L_d}^* dL_d(F_{L_d}^{N-1}(p))\, \pi_+\, T_p F_{L_d}\, w_{N-1} - \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1},$$

which simplifies to

$$F_{L_d}^* dL_d\, \pi_-(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1} = -\Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_p F_{L_d}^{N-1}\, w_{N-1},$$

and again using $\Theta^-_{L_d} = -dL_d\pi_-$,

$$F^*_{L_d}\Theta^-_{L_d}(F^{N-1}_{L_d}(p))\,T_pF^{N-1}_{L_d}\,w_{N-1} = \Theta^+_{L_d}(F^{N-1}_{L_d}(p))\,T_pF^{N-1}_{L_d}\,w_{N-1}. \tag{3.35}$$

Taking $N = 1$, $w_p = w_0 \in T_pP$ satisfies $T_p(\partial^+ \circ F_{L_d})\,w_p \in \mathcal{D}_{\partial^+\circ F_{L_d}(p)}$. Then, on $\mathcal{D}_d$, Equation (3.35) is

$$F^*_{L_d}\Theta^-_{L_d}(p)\,w_p = \Theta^+_{L_d}(p)\,w_p. \tag{3.36}$$

Equation (3.36) is the discrete analogue of the semi–Hamilton equations in Chapter 2 and Proposition 12 is the discrete analogue of Proposition 5 in Chapter 2.

Proposition 12. *Let $(\mathcal{D}, \mathcal{D}_d)$ be a C^k discrete constraint, $k > 1$, and $L_d\colon P \to \mathbb{R}$ a C^k, $(V, W) - \mathcal{D}$–regular discrete Lagrangian. Equation (3.36) restricted to $\mathcal{K}_d$ (Definition 2) completely determines F_{L_d}.*

Proof. Using the definitions of Θ^+ and Θ^-, Equation (3.36) is

$$-F^*_{L_d}(dL_d(p)\,\pi_-)w_p = dL_d\pi_+\,w_p.$$

Using the definition of pullback and rearranging, this is

$$dL_d(F_{L_d}(p))\,\pi_-T_pF_{L_d}\,w_p + dL_d(p)\,\pi_+\,w_p = 0. \tag{3.37}$$

Write $p_0 = p, p_1 = F_{L_d}(p), w^+ = \pi_+\,w_p$ and $w^- = \pi_-\,T_pF_{L_d}\,w_p$. Then, Equation (3.37) is

$$dL_d(p_1)\,w^- + dL_d(p_0)\,w^+ = 0,$$

which are the discrete Euler–Lagrange equations. The existence of the solution and the determination of the map F_{L_d} are given by the requirement that L_d be $(V, W) - \mathcal{D}$–regular. □

3.7 Symmetry and Momentum Equations

Let $\mathcal{G}$ be a Lie group with Lie algebra $\mathfrak{g}$. Let $\mathcal{G}$ act on Q by ϕ_g and on $Q \times Q$ by the diagonal action, $\phi_g(q^+, q^-) = (\phi_g(q^+), \phi_g(q^-))$. $\mathcal{G}$ acts on P by *discrete lift*, $\Phi_g(p) = (\partial^\pm)^{-1} \circ$

$\phi_g(\partial^+(p), \partial^-(p))$. The infinitesimal generator associated to $\xi \in \mathfrak{g}$ of the group action on Q is $\xi_Q(q) = \frac{d}{d\epsilon}\big|_{\epsilon=0} \phi_{\exp(\epsilon\,\xi)}(q)$. Then the infinitesimal generator of the diagonal action is

$$\xi_{Q\times Q}(q^+, q^-) = (\xi_Q(q^+), \xi_Q(q^-))$$

and for the discrete lifted action is $\xi_P(p) = T_{(q^+,q^-)}(\partial^{\pm})^{-1}\, \xi_{Q\times Q}(q^+, q^-)$ where $(q^+, q^-) = \partial^{\pm}(p)$ is written for convenience.

Suppose L_d, $\mathcal{D}$ and $\mathcal{D}_d$ are invariant under the discrete lifted action. Define the discrete momentum function

$$J_\xi^{L_d} \colon P \to \mathbb{R}, \qquad J_\xi^{L_d}(p) = \left\langle \Theta^+_{L_d}, \xi_P \right\rangle(p).$$

There is another momentum function defined by using $\Theta^-_{L_d}$ instead of $\Theta^+_{L_d}$, but these evaluate the same since

$$\left\langle \Theta^+_{L_d} - \Theta^-_{L_d}, \xi_P \right\rangle(p) = \left\langle dL_d, \xi_P \right\rangle(p) = 0,$$

since L_d is $\mathcal{G}$ invariant. The discrete momentum mapping is defined

$$J^{L_d} \colon P \to \mathfrak{g}^*, \qquad J^{L_d}(p)\,\xi = J_\xi^{L_d}(p).$$

Proposition 13. *Let ϕ_g be a Lie group action of $\mathcal{G}$ on Q, and abusing notation, ϕ_g the diagonal action on $Q \times Q$. Let Φ_g be the discrete lift of ϕ_g. Then $\pi_+ T_p\Phi_g = T_p\Phi_g\pi_+$ and $\pi_- T_p\Phi_g = T_p\Phi_g\pi_-$.*

Proof. The definitions of π_+ and π_- are in Equations (3.3).

The proof will be shown for only π_+, as the computations are similar for π_-.

$$\begin{aligned}
\pi_+ T_p\Phi_g\, X &= T_{\partial^{\pm}(\Phi_g(p))}(\partial^{\pm})^{-1}\, (T_{\Phi_g(p)}\partial^+\, T_p\Phi_g\, X, 0) && \text{definition of } \pi_+ \\
&= T_{\partial^{\pm}(\Phi_g(p))}(\partial^{\pm})^{-1}\, (T_p(\partial^+ \circ \Phi_g)\, X, 0) && \text{chain rule} \\
&= T_{\partial^{\pm}(\Phi_g(p))}(\partial^{\pm})^{-1}\, (T_p(\phi_g \circ \partial^+)\, X, 0) && \text{by definition of } \Phi_g \\
&= T_{\partial^{\pm}(\Phi_g(p))}(\partial^{\pm})^{-1}\, (T_{\partial^+(p)}\phi_g\, T_p\partial^+\, X, 0) && \text{chain rule} \\
&= T_{\partial^{\pm}(\Phi_g(p))}(\partial^{\pm})^{-1}\, T_{\partial^{\pm}(p)}\phi_g\, (T_p\partial^+\, X, 0) && \text{diagonal action}
\end{aligned}$$

$$= T_{\partial^\pm(\Phi_g(p))}(\partial^\pm)^{-1}\,T_{\partial^\pm(p)}\phi_g\,T_p\partial^\pm\,(T_p\partial^\pm)^{-1}\,(T_p\partial^+\,X,0)$$

$$= T_p\Phi_g\,\pi_+\,X. \qquad \text{definition of } \Phi_g \text{ and chain rule}$$

□

Proposition 14. *Let L_d be $\mathcal{G}$ invariant. Then $\Theta^+_{L_d}$ and $\Theta^-_{L_d}$ are $\mathcal{G}$ invariant.*

Proof. The proof will be shown only for $\Theta^+_{L_d}$.

$$\begin{aligned}(\Phi_g^*\Theta^+_{L_d})(p)\,X &= \Theta^+_{L_d}(\Phi_g(p))\,T_p\Phi_g\,X \qquad \text{definition of pullback}\\ &= dL_d(\Phi_g(p))\,\pi_+\,T_p\Phi_g\,X \qquad \text{definition of } \Theta^+_{L_d}\\ &= dL_d(\Phi_g(p))\,T_p\Phi_g\,\pi_+\,X \qquad \text{Prop. 13}\\ &= (\Phi_g^* dL_d)(p)\pi_+\,X \qquad \text{definition of pullback}\\ &= dL_d(p)\pi_+\,X \qquad L_d \text{ is } G \text{ invariant}\\ &= \Theta^+_{L_d}\,X. \qquad \text{definition of } \Theta^+_{L_d}\end{aligned}$$

□

Proposition 15. *Let $\mathcal{G}$ be a Lie group acting on P by discrete lift such that L_d is invariant. Then J^{L_d} is CoAd equivariant (see Abraham and Marsden [1], 4.1).*

Proof. It needs to be shown that $\mathrm{CoAd}_g \circ J^{L_d} = J^{L_d} \circ \Phi_{g^{-1}}$,

$$\begin{aligned}\mathrm{CoAd}_g \circ J^{L_d}(p)\,\xi &= J^{L_d}(p)\,\mathrm{Ad}_g\,\xi \qquad \text{definition of } \mathrm{CoAd}_g\\ &= J^{L_d}_{\mathrm{Ad}_g\,\xi}(p) \qquad \text{definition of } J^{L_d}\\ &= \Theta^+_{L_d}(p)\,(\mathrm{Ad}_q\,\xi)_P(p) \qquad \text{definition of } J^{L_d}_\xi\\ &= \Theta^+_{L_d}(p)\,\Phi^*_{g^{-1}}\xi_P(p) \qquad \text{Abraham and Marsden [1] Prop. 4.1.26}\\ &= \Theta^+_{L_d}(p)\,T_{\Phi_{g^{-1}}(p)}\Phi_g\,\xi_P(\Phi_{g^{-1}}(p)) \qquad \text{definition of pullback}\\ &= \Phi_g^*\Theta^+_{L_d}(\Phi_{g^{-1}}(p))\,\xi_P(\Phi_{g^{-1}}(p)) \qquad \text{definition of pullback}\\ &= \Theta^+_{L_d}(\Phi_{g^{-1}}(p))\,\xi_P(\Phi_{g^{-1}}(p)) \qquad \text{Prop. 14}\end{aligned}$$

$$= J_\xi^{L_d}(\Phi_{g^{-1}}(p)) \qquad \text{definition of } J_\xi^{L_d}$$

$$= J^{L_d}(\Phi_{g^{-1}})\,\xi, \qquad \text{definition of } J^{L_d}$$

which proves the proposition. □

Let $\{\xi_i\}$ be a sequence in $\mathfrak{g}$ with infinitesimal generators denoted by ξ_Q^i and ξ_P^i so that $\xi_Q^i(\partial^+ \circ F_{L_d}^i(p)) \in \mathcal{D}_{(\partial^+ \circ F_{L_d}^i)(p)}$. Let $\{w_i\}$ be a sequence in T_pP such that $T_pF_{L_d}^i\, w_i = \xi_P^i(F_{L_d}^i(p))$. Then,

$$T_p(\partial^+ \circ F_{L_d}^i)\, w_i = T_{F_{L_d}^i(p)}\partial^+ \xi_P^i(F_{L_d}^i(p)) = \xi_Q^i(\partial^+ \circ F_{L_d}^i(p)).$$

If L_d is $\mathcal{G}$ invariant, Equation (3.34) becomes

$$F_{L_d}^* dL_d(F_{L_d}^{N-1}(p))\, T_pF_{L_d}^{N-1}\, w_{N-1} =$$
$$F_{L_d}^*\Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_pF_{L_d}^{N-1}\, w_{N-1} - \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, T_pF_{L_d}^{N-1}\, w_{N-1}.$$

Using the definition of w_{N-1}, this becomes

$$F_{L_d}^* dL_d(F_{L_d}^{N-1}(p))\, \xi_P^{N-1}(F_{L_d}^{N-1}(p)) =$$
$$F_{L_d}^*\Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, \xi_P^{N-1}(F_{L_d}^{N-1}(p)) - \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, \xi_P^{N-1}(F_{L_d}^{N-1}(p)).$$

Using $dL_d = \Theta_{L_d}^+ - \Theta_{L_d}^-$ and simplifying gives

$$F_{L_d}^*\Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, \xi_P^{N-1}(F_{L_d}^{N-1}(p)) = \Theta_{L_d}^+(F_{L_d}^{N-1}(p))\, \xi_P^{N-1}(F_{L_d}^{N-1}(p)).$$

The definition of $J_\xi^{L_d}$ gives

$$F_{L_d}^* J_{\xi^{N-1}}^{L_d}(F_{L_d}^{N-1}(p)) = J_{\xi^{N-1}}^{L_d}(F_{L_d}^{N-1}(p)).$$

Adding and subtracting $J_{\xi^N}^{L_d}(F_{L_d}^N(p))$ to the right hand side gives

$$F_{L_d}^* J^{L_d}(F_{L_d}^{N-1}(p))\xi^{N-1} = J_{\xi^{N-1}}^{L_d}(F_{L_d}^{N-1}(p)) - J_{\xi^N}^{L_d}(F_{L_d}^N(p)) + J_{\xi^N}^{L_d}(F_{L_d}^N(p)).$$

Rearrange and use the definition of pullback to get

$$J_{\xi^N}^{L_d}(F_{L_d}^N(p)) - J^{L_d}(F_{L_d}^N(p))\xi^{N-1} = J_{\xi^N}^{L_d}(F_{L_d}^N(p)) - J_{\xi^{N-1}}^{L_d}(F_{L_d}^{N-1}(p)).$$

Finally, the definition of J^{L_d} gives

$$J^{L_d}(F^N_{L_d}(p))\,(\xi^N - \xi^{N-1}) = J^{L_d}_{\xi^N}(F^N_{L_d}(p)) - J^{L_d}_{\xi^{N-1}}(F^{N-1}_{L_d}(p)). \tag{3.38}$$

Equation (3.38) with $N = 1$ is the *discrete momentum* equation

$$J^{L_d}(F_{L_d}(p))\,(\xi^1 - \xi^0) = J^{L_d}_{\xi^1}(F_{L_d}(p)) - J^{L_d}_{\xi^0}(p). \tag{3.39}$$

These equations appear in Cortés and Martínez [7] for Moser–Veselov integrators.

If $\{\xi^i\}$ can be chosen to be a constant sequence, $\xi^i = \xi$, then the discrete momentum equation reduces to

$$J^{L_d}_{\xi}(F_{L_d}(p)) = J^{L_d}_{\xi}(p),$$

which is equivalent to

$$J^{L_d}(F_{L_d}(p))\,\xi = J^{L_d}(p)\,\xi,$$

so that the component of the discrete momentum map in the ξ direction is preserved by the discrete flow.

Chapter 4

Discretized Lagrangian Mechanics

4.1 Introduction

This chapter specializes the constructions in Chapter 3 so that the discrete tangent bundle is an open neighbourhood of the zero section in TQ and the maps ∂^+ and ∂^- correspond to the endpoints of curves adapted to the constraint distribution. This gives the class of numerical methods that will be called *Lagrange–d'Alembert* integrators.

Local versions of Theorems 3 and 4 appear in Appendix D.

4.2 Discretized Tangent Bundle

A *discretization* of TQ is a triple of smooth mappings $(\psi, \alpha^+, \alpha^-)$ described as follows:

$$\psi\colon U \subset \mathbb{R}^2 \times TQ \to Q \;:\; (h, t, v_q) \mapsto \psi(h, t, v_q),$$

$$\alpha^+\colon [0, a) \to [0, \infty),$$

$$\alpha^-\colon [0, a) \to (-\infty, 0],$$

for some $a > 0$. These mappings satisfy:

$$\alpha^+(h) - \alpha^-(h) = h,\;\; h \in [0, a),$$

$$\psi(h, 0, v_q) = q,$$

$$\frac{\partial \psi}{\partial t}(h, 0, v_q) = v_q.$$

Define

$$\partial_h^+(v_q) = \psi(h, \alpha^+(h), v_q), \tag{4.1}$$

$$\partial_h^-(v_q) = \psi(h, \alpha^-(h), v_q). \tag{4.2}$$

Theorem 3. *Let $(\psi, \alpha^+, \alpha^-)$ be a discretization of TQ and let $Q_0 \subset Q$ be a relatively compact open submanifold of Q. Then there is an $a > 0$ and an open submanifold P_h of TQ such that*

1. *$(P_h, \partial_h^+, \partial_h^-)$ is a discrete tangent bundle of Q_0 for all $h \in (0, a)$,*

2. *$\partial_h^\pm$ is a diffeomorphism of P_h to an open neighbourhood of $\Delta(Q_0 \times Q_0)$.*

Further, for all $v_q \in T_qQ_0$, there is an h and P_h such that $v_q \in P_h$.

Proof. Let E be a normal bundle to $\Delta(Q \times Q)$. Then there is a tubular neighbourhood, $W^{0(E)}$ of the zero section of E, a neighbourhood $W^{\Delta(Q\times Q)}$ of $\Delta(Q\times Q)$ and a diffeomorphism $\zeta\colon W^{0(E)} \to W^{\Delta(Q\times Q)}$ (Lang [16]).

Let $w = T_{0_{(q,q)}}E$. There is a natural decomposition of w into horizontal and vertical subspaces by

$$(v_{(q,q)}, w_{(q,q)}) \mapsto \left.\frac{d}{dt}\right|_{t=0} 0_{(q(t),q(t))} + \left.\frac{d}{dt}\right|_{t=0} tw_{(q,q)} = w, \tag{4.3}$$

where $(q'(0), q'(0)) = v_{(q,q)} \in T\Delta(Q\times Q)$ and $w_{(q,q)} \in E_{(q,q)}$. The first term in Equation (4.3) is denoted $\operatorname{hor} w$ and the second term is $\operatorname{vert} w$ so that $w = \operatorname{hor} w + \operatorname{vert} w$.

For (v_q, h) such that $\partial_h^\pm(v_q) \in W^{\Delta(Q\times Q)}$, define the curve c_{v_q} in $W^{0(E)}$ implicitly by

$$\zeta(hc_{v_q}(h)) = \partial_h^\pm(v_q), \qquad h > 0. \tag{4.4}$$

In fact, c_{v_q} is defined, for $h > 0$ as

$$c_{v_q}(h) = \frac{1}{h}\zeta^{-1} \circ \partial_h^\pm(v_q),$$

so that c_{v_q} is smooth for $h > 0$. c_{v_q} can be extended to a C^1 curve on $h \geq 0$, as will now be shown.

Suppose ζ and $\partial_h^\pm$ are C^k. Let $f(h) = hc_{v_q}(h) = (\zeta^{-1} \circ \partial_h^\pm)(v_q)$. Then f is C^k and $f(0) = 0$. Locally, the Taylor expansion of f at $h = 0$ is, with continuous remainder term $R(h)$,

$$\begin{aligned}\hat{f}(h) &= \hat{f}(0) + \sum_{j=1}^{k} \frac{h^j}{j!} \left.\frac{d^j}{dt^j}\right|_{t=0} \hat{f}(t) + \frac{h^k}{k!} R(h)\\ &= \sum_{j=1}^{k} \frac{h^j}{j!} \left.\frac{d^j}{dt^j}\right|_{t=0} \hat{f}(t) + \frac{h^k}{k!} R(h). \qquad \text{since } f(0) = 0\end{aligned}$$

This gives c_{v_q} locally as

$$\hat{c}_{v_q}(h) = \sum_{j=1}^{k} \frac{h^{j-1}}{j!} \left.\frac{d^j}{dt^j}\right|_{t=0} \hat{f}(t) + \frac{h^{k-1}}{k!} R(h).$$

Then, by Taylor's theorem (Abraham, Marsden and Ratiu [2] Section 2.4), $\hat{c}_{v_q}$ is C^{k-1}.

Differentiating both sides of Equation (4.4) with respect to h at $h = 0$ gives

$$\begin{aligned}\left.\frac{d}{dh}\right|_{h=0} \partial_h^\pm(v_q) &= \left.\frac{d}{dh}\right|_{h=0} \zeta(hc_{v_q}(h))\\ &= T_{0_{(q,q)}}\zeta \left.\frac{d}{dh}\right|_{h=0} (hc_{v_q}(h))\\ &= T_{0_{(q,q)}}\zeta \left.\frac{d}{dh}\right|_{h=0} (hc_{v_q}(0)) + T_{0_{(q,q)}}\zeta \left.\frac{d}{dh}\right|_{h=0} (0c_{v_q}(h))\\ &= T_{0_{(q,q)}}\zeta \left.\frac{d}{dh}\right|_{h=0} (hc_{v_q}(0)). \end{aligned} \tag{4.5}$$

Equation (4.5) is the derivative of ζ on a vertical vector. So,

$$c_{v_q}(0) = \operatorname{vert} T_{(q,q)}\zeta^{-1} \left.\frac{d}{dh}\right|_{h=0} \partial_h^\pm(v_q).$$

For clarity, choose $E = \{(v_q, -v_q) \,|\, v_q \in TQ\}$. Let X be a spray on TQ and exp the corresponding exponential map. Then ζ can be chosen to be $\zeta(v_q, -v_q) = (\exp(v_q), \exp(-v_q))$.

Let $(v_q, v_q) \in T\Delta(Q \times Q)$ and let $q(t)$ be a curve in Q such that $q'(0) = v_q$. At $0_{(q,q)}$, let $w_h = \left(\left.\frac{d}{dt}\right|_{t=0} 0_{q(t)}, \left.\frac{d}{dt}\right|_{t=0} 0_{q(t)} \right)$ be a horizontal vector. Then,

$$\begin{aligned}T_{0_{(q,q)}}\zeta\, w_h &= \left.\frac{d}{dt}\right|_{t=0} \zeta(0_{q(t)}, 0_{q(t)})\\ &= \left(\left.\frac{d}{dt}\right|_{t=0} \exp(0_{q(t)}), \left.\frac{d}{dt}\right|_{t=0} \exp(0_{q(t)}) \right)\end{aligned}$$

$$= \big(T_{0_q} \exp\, v_q, T_{0_q} \exp\, v_q\big) \quad \text{chain rule}$$

$$= (v_q, v_q). \quad \text{since } T_{0_q} \exp = \text{id}$$

Let $(v_q, -v_q) \in E_{(q,q)}$. At $0_{(q,q)}$, let $w_v = \big(\frac{d}{dt}\big|_{t=0} tv_q, -\frac{d}{dt}\big|_{t=0} tv_q\big)$ be a vertical vector. Then,

$$\begin{aligned} T_{0_{(q,q)}}\zeta\, w_v &= \frac{d}{dt}\bigg|_{t=0} \zeta\,(tv_q, -tv_q) \\ &= \left(\frac{d}{dt}\bigg|_{t=0} \exp(tv_q),\, \frac{d}{dt}\bigg|_{t=0} \exp(-tv_q)\right) \\ &= (v_q, -v_q). \end{aligned}$$

Therefore, for $w = w_h + w_v = (u_q, u_q) + (v_q, -v_q) \in T_{0_{q,q}}E$,

$$T_{0_{(q,q)}}\zeta(w_h + w_v) = (u_q, u_q) + (v_q, -v_q), \tag{4.6}$$

which is the decomposition $T(Q \times Q) = T\Delta(Q \times Q) \oplus E_{(q,q)}$.

Let $(v^+, v^-) \in T(Q \times Q)$. Then,

$$(v^+, v^-) = \frac{1}{2}\big((v^+ + v^-), (v^+ + v^-)\big) + \frac{1}{2}\big((v^+ - v^-), (v^- - v^+)\big).$$

Using Equation (4.6),

$$\begin{aligned} w_h + w_v &= T_{(q,q)}\zeta^{-1}\,(v^+, v^-) \\ &= T_{(q,q)}\zeta^{-1}\frac{1}{2}\big((v^+ + v^-), (v^+ + v^-)\big) + T_{(q,q)}\zeta^{-1}\frac{1}{2}\big((v^+ - v^-), (v^- - v^+)\big) \\ &= \frac{1}{2}\big((v^+ + v^-), (v^+ + v^-)\big) + \frac{1}{2}\big((v^+ - v^-), (v^- - v^+)\big), \end{aligned}$$

so that

$$w_v = \text{vert}\, T_{(q,q)}\zeta^{-1}\,(v^+, v^-) = \frac{1}{2}\big((v^+ - v^-), (v^- - v^+)\big)$$

Define $\phi\colon \big\{(v_q, h) \,|\, \partial_h^{\pm}(v_q) \in W^{\Delta(Q \times Q)}\big\} \to W^{0(E)}$ by

$$\phi(v_q, h) = \begin{cases} \big(\frac{1}{h}\zeta^{-1} \circ \partial_h^{\pm}(v_q), h\big), & h > 0, \\[2ex] \big(\text{vert}\, T_{(q,q)}\zeta^{-1}\ \frac{d}{dh}\big|_{h=0}\ \partial_h^{\pm}(v_q), 0\big), & h = 0. \end{cases}$$

To simplify ϕ, compute

$$\begin{aligned}
\left.\frac{d}{dh}\right|_{h=0} \partial_h^+(v_q) &= \left.\frac{d}{dh}\right|_{h=0} \psi(h, \alpha^+(h), v_q) \\
&= \left.\frac{d}{dh}\right|_{h=0} \psi(h, 0, v_q) + \left.\frac{d}{dh}\right|_{h=0} \psi(0, \alpha^+(h), v_q) \\
&= \left.\frac{d}{dh}\right|_{h=0} q + v_q \left.\frac{d}{dh}\right|_{h=0} \alpha^+(h) \\
&= (\alpha^+)'(0) v_q. \qquad (4.7)
\end{aligned}$$

A similar calculation shows that

$$\left.\frac{d}{dh}\right|_{h=0} \partial_h^-(v_q) = (\alpha^-)'(0) v_q. \qquad (4.8)$$

Then,

$$\begin{aligned}
\left.\frac{d}{dh}\right|_{h=0} \partial_h^\pm(v_q) &= \left(\left.\frac{d}{dh}\right|_{h=0} \partial_h^+(v_q), \left.\frac{d}{dh}\right|_{h=0} \partial_h^-(v_q) \right) \\
&= \left((\alpha^+)'(0) v_q, (\alpha^-)'(0) v_q \right).
\end{aligned}$$

And,

$$\phi(v_q, h) = \begin{cases} \left(\frac{1}{h} \zeta^{-1} \circ \partial_h^\pm(v_q), h \right), & h > 0, \\ \left(\frac{1}{2} v_q, -\frac{1}{2} v_q, 0 \right), & h = 0. \end{cases}$$

ϕ is smooth by construction, is a local diffeomorphism at $(v_q, 0)$ for all $v_q \in TQ$ and is a diffeomorphism of $TQ \times \{0\}$ to $E \times \{0\}$. Therefore, by Lemma 2 in Appendix A, there are open neighbourhoods U and V

$$E \times \mathbb{R} \supset U \supset E \times \{0\},$$

$$TQ \times \mathbb{R} \supset V \supset TQ \times \{0\},$$

such that $\phi\colon V \to U$ is a diffeomorphism.

Define $\tilde{\phi}\colon U \to W$ by

$$\tilde{\phi}(e, h) = (\zeta(he), h),$$

so that $(\partial_h^\pm(v_q), h) = \tilde{\phi} \circ \phi(v_q, h)$. Since $\tilde{\phi}$ is a diffeomorphism when $h \neq 0$, $\partial_h^\pm$ is a diffeomorphism from $V \backslash TQ \times \{0\}$ to $W = \zeta(U \backslash E \times \{0\})$.

Let Q_0 be a relatively compact submanifold of Q. Then there is an $a > 0$ such that $Q_0 \times Q_0 \times (0, a) \subset W$. Let $h \in (0, a)$ and define

$$P_h = \{v_q \in TQ \,|\, (v_q, h) \in V \text{ and } \partial_h^{\pm}(v_q) \in Q_0 \times Q_0\}.$$

$\partial_h^{\pm}$ is a diffeomorphism on P_h since $\partial_h^{\pm} \times \mathrm{id}$ is a diffeomorphism on V. In addition, ∂_h^{+} and ∂_h^{-} are onto Q_0 since there is always a v_q satisfying $\partial_h^{\pm}(v_q, h) = (q_0, q_0, h)$ for any $q_0 \in Q_0$. Therefore, P_h is a discrete tangent bundle of Q_0.

Let $v_q \in TQ_0$. Since $\partial^{\pm} \times \mathrm{id}$ is continuous, there is an $h \in (0, a)$ such that P_h contains v_q. □

4.3 Discretized Lagrangian Systems

A discretization of a Lagrangian system $L \colon TQ \to \mathbb{R}$ is a tuple $(L_h, \psi, \alpha^+, \alpha^-)$ where $(\psi, \alpha^+, \alpha^-)$ is a discretization of TQ and L_h is defined by

$$L_h(v_q) = \frac{1}{h} \int_{\alpha^-(h)}^{\alpha^+(h)} L\left(\frac{\partial \psi}{\partial t}(h, t, v_q)\right) dt. \tag{4.9}$$

That is, $L_h(v_q)$ is the average of L along the curve $t \mapsto \frac{\partial}{\partial t}\psi(h, t, v_q)$ for $t \in [\alpha^-(h), \alpha^+(h)]$.

Example 2. Let X be a second order vector field with flow F_t^X. Then $(L_h, \psi, \alpha^+, \alpha^-)$ is a discretization of a Lagrangian system with

$$\psi(h, t, v_q) = (\tau_Q \circ F_t^X)(v_q),$$
$$L_h(v_q) = \frac{1}{h} \int_{\alpha^-(h)}^{\alpha^+(h)} L(F_t^X(v_q))\, dt.$$

The discrete action is given in Equation (3.4) as

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d(p_k).$$

A direct substitution of L_h for L_d is possible, but leaves the discrete action with the same units as the discrete Lagrangian (which has the same units as the continuous Lagrangian).

In fact, the trend in the literature (Marsden and West [18], Cortés and Martínez, [7] and Marsden, Patrick and Shkoller [17] to name a few), use this convention. In order to keep the analogy with the continuous system as consistent as possible, the discretized action will be a scaling of S_d by the time step,

$$S_h(q_d) = \sum_{k=0}^{N-1} L_h(v_k)\, h.$$

4.4 Discretized Constrained Lagrangian Systems

Let $\mathcal{D}$ be a distribution on Q, $(\psi, \alpha^+, \alpha^-)$ a discretization of TQ. The discretization of $\mathcal{D}$ is achieved by requiring the curve $t \mapsto \frac{\partial \psi}{\partial t}(h, t, v_q)$ be in $\mathcal{D}$ so that $\mathcal{D}_d = \mathcal{D}$ (Definition 9).

Write the following for short form notation, where the ordered pairs are with respect to standard bundle coordinates,

$$R_t(v_q) = \psi(h, t, v_q), \quad R_t'(v_q) = \frac{\partial \psi}{\partial t}(h, t, v_q), \tag{4.10}$$

$$\left(R_t(v_q), \frac{d}{dt} R_t(v_q)\right) = (R_t(v_q), \dot{R}_t(v_q)). \tag{4.11}$$

Let $v_q \in TQ$ and $\delta v \in T_{v_q}TQ$. Let $v(\epsilon)$ be a curve in TQ such that $v(0) = v_q$ and $\frac{d}{d\epsilon}\big|_{\epsilon=0} v(\epsilon) = \delta v$. The coordinate calculations are with respect to natural bundle coordinates on TQ for which v_q is $(q, \dot{q})$ and δv is $(\delta q, \delta \dot{q})$. Further, the curve $v(\epsilon)$ is $(q(\epsilon), \dot{q}(\epsilon))$. Then, compute in these local coordinates,

$$\begin{aligned} dL_h(q,\dot{q})\,(\delta q, \delta \dot{q}) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L_h(q(\epsilon), \dot{q}(\epsilon)) \\ &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \frac{1}{h} \int_{\alpha^-(h)}^{\alpha^+(h)} L(R_t'(q(\epsilon), \dot{q}(\epsilon)))\, dt \quad \text{definition of } L_h \\ &= \frac{1}{h} \int_{\alpha^-(h)}^{\alpha^+(h)} dL(R_t'(q, \dot{q}))\, T_{(q,\dot{q})} R_t'\, (\delta q, \delta \dot{q})\, dt. \quad \text{chain rule} \end{aligned}$$

Since $R_t'(q, \dot{q}) = \left(R_t(q, \dot{q}), \dot{R}_t(q, \dot{q})\right)$ and expanding dL,

$$dL_h(q, \dot{q})\,(\delta q, \delta \dot{q}) = \frac{1}{h} \int_{\alpha^-(h)}^{\alpha^+(h)} \left(\frac{\partial L}{\partial q^i}(R_t'(q, \dot{q})) \left(\frac{\partial R_t^i}{\partial q^j} \delta q^j + \frac{\partial R_t^i}{\partial \dot{q}^j} \delta \dot{q}^j\right) + \right.$$

$$+\frac{\partial L}{\partial \dot q^i}(R'_t(q,\dot q))\left(\frac{\partial \dot R^i_t}{\partial q^j}\delta q^j+\frac{\partial \dot R^i_t}{\partial \dot q^j}\delta \dot q^j\right)\Bigg)\,dt.$$

Integration by parts gives

$$dL_h(q,\dot q)\,(\delta q,\delta \dot q)=\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}\left(\frac{\partial L}{\partial q^i}(R'_t(q,\dot q))-\frac{d}{dt}\frac{\partial L}{\partial \dot q^i}(R'_t(q,\dot q))\right)\left(\frac{\partial R^i_t}{\partial q^j}\delta q^j+\frac{\partial R^i_t}{\partial \dot q^j}\delta \dot q^j\right)dt+$$
$$+\frac{1}{h}\frac{\partial L}{\partial \dot q^i}(R'_t(q,\dot q))\left(\frac{\partial R^i_t}{\partial q^j}\delta q^j+\frac{\partial R^i_t}{\partial \dot q^j}\delta \dot q^j\right)\Bigg|_{t=\alpha^-(h)}^{\alpha^+(h)}.$$

Which is equivalent to

$$dL_h(q,\dot q)\,(\delta q,\delta \dot q)=\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}D_{EL}L\left(R'_t(q,\dot q)\right)\,TR_t\left(\delta q,\delta \dot q\right)dt+$$
$$+\frac{1}{h}\Theta_L(R'_t(q,\dot q))\,TR'_t\left(\delta q,\delta \dot q\right)\Big|_{t=\alpha^-(h)}^{\alpha^+(h)}.$$

So that,

$$dL_h(v_q)\,\delta v=\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}D_{EL}L(R'_t(v_q))\,TR_t\,\delta v\,dt+\frac{1}{h}\Theta_L(R'_t(v_q))\,TR'_t(v_q)\,\delta v\Big|_{t=\alpha^-(h)}^{\alpha^+(h)}. \quad (4.12)$$

Using Equation (4.12), $TR_{\alpha^-(h)}(v_k)\,\delta v_k=0$ and $TR_{\alpha^+(h)}(v_{k+1})\,\delta v_{k+1}=0$, as required by the definition of $D_{EL}L_h$, calculate

$$\begin{aligned}D_{EL}L_h(v_k,v_{k+1})\,\delta q&=dL_h(v_k)\,\delta v_k+dL_h(v_{k+1})\,\delta v_{k+1}\\&=\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}D_{EL}L(R'_t(v_k))\,TR_t\,\delta v_k\,dt+\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}D_{EL}L(R'_t(v_{k+1}))\,TR_t\,\delta v_{k+1}\,dt+\\&\quad+\frac{1}{h}\Theta_L(R'_{\alpha^+(h)}(v_k))\,TR'_{\alpha^+(h)}\,\delta v_k-\frac{1}{h}\Theta_L(R'_{\alpha^-(h)}(v_{k+1}))\,TR'_{\alpha^-(h)}\,\delta v_{k+1}. \quad (4.13)\end{aligned}$$

There are two important points to make about Equation (4.13):

1. The discrete Lagrange–d'Alembert principle requires that $D_{EL}L_h\in\mathcal{D}^0$, which is equivalent to requiring that $\delta q\in\mathcal{D}$. However, this does not imply that the vectors δv_k and δv_{k+1} are in $T\mathcal{D}$, nor does it imply that $TR_t\delta v_k\in\mathcal{D}$ so that a discretization of $\mathcal{D}$ is not enough. In order to differentiate L_h in the direction of the variations, a neighbourhood of $\mathcal{D}$ is necessary.

2. If $R'_t = F_t^{X_E}$, the exact flow, then for L regular,

$$F_h^{X_E}(v_k) = v_{k+1}.$$

Since $\alpha^+(h) - \alpha^-(h) = h$,

$$F_{\alpha^+(h)}(v_k) = F_{\alpha^-(h)}(v_{k+1}).$$

L regular implies that the fibre derivative of L, $\mathbb{F}L$, is a local diffeomorphism so that

$$\mathbb{F}L(F_{\alpha^+(h)}(v_k))\,\delta q = \mathbb{F}L(F_{\alpha^-(h)}(v_{k+1}))\,\delta q.$$

Since, locally, $\Theta_L = \frac{\partial L}{\partial \dot{q}^i} dq^i$,

$$\Theta_L(F_{\alpha^+(h)}(v_k))\,TR'_{\alpha^+(h)}\,\delta v_k = \Theta_L(F_{\alpha^-(h)}(v_{k+1}))\,TR'_{\alpha^-(h)}\,\delta v_{k+1}. \tag{4.14}$$

Equation (4.14) forms the last two terms of Equation (4.13). The two integral terms in Equation (4.13) need not be zero if the constraint is nonholonomic. This fact interferes with the requirement that one should be able to use the discrete Lagrange–d'Alembert to construct arbitrary order integrators.

Item 1 is dealt with by discretizing an affine constraint (see Appendix C). Let Y be a vector field on Q such that $\phi Y \neq 0$ for any ϕ in the annihilator of $\mathcal{D}$. Take the curves $\psi(h, t, v_q)$ such that they are in $\mathcal{D}_Y$. The vector field Y is treated as a parameter by setting the discrete Euler–Lagrange equations on $\mathcal{D}_Y$ and solving them for $Y = 0$. In Chapter 5, Y is chosen so that, locally, $Y(q) = a^i X_i(q), i = d+1, \ldots, n$ where the a^i are constant and $\{X_i\}_{i=1}^n$ is a local basis of TQ at q such that $\{X_i\}_{i=1}^d$ is a local basis of $\mathcal{D}$ at q.

Item 2 can only be dealt with in an *ad hoc* manner at this point.

Let $\bar{\alpha}_h : TQ \times TQ \to T^*Q$ and adjust the discretized Lagrange–d'Alembert principle by

$$D_{EL}L_h(v_k, v_{k+1}) = \bar{\alpha}_h(v_k, v_{k+1}),$$

where, for a one form α_h on TQ,

$$\bar{\alpha}_h(v_k, v_{k+1})\,\delta q = \alpha_h(v_k)\,\delta v_k + \alpha_h(v_{k+1})\,\delta v_{k+1},$$

and

$$T_{v_k}\partial_h^- \,\delta v_k = 0,$$

$$T_{v_{k+1}}\partial_h^+ \,\delta v_{k+1} = 0,$$

$$T_{v_k}\partial_h^+ \,\delta v_k = T_{v_{k+1}}\partial_h^- \,\delta v_{k+1},$$

$$\partial_h^+(v_k) = \partial_h^-(v_{k+1}).$$

While there may be many choices for α_h, an obvious one motivated by the definition of L_h and the calculation of dL_h is

$$\alpha_h(v_q)\,\delta v = \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} F_c(R_t'(v_q))\,TR_t\,\delta v\,dt, \tag{4.15}$$

where F_c is the constraint force.

This α_h serves the purpose of isolating the work done by the *numeric force*, that is, the force responsible for the deviation of R_t from the exact trajectory.

In summary, let $L\colon TQ \to \mathbb{R}$ be a Lagrangian and $\mathcal{D}$ a distribution on Q. Let Y be a vector field on Q. The discretization of TQ is given by curves $t \mapsto \psi(h,t,v_q)$ such that $\frac{\partial\psi}{\partial t} - Y \in \mathcal{D}$. The discretized Lagrangian is

$$L_h(v_q) = \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} \left(L \circ \frac{\partial\psi}{\partial t}\right)(v_q)\,dt.$$

The discrete Lagrange–d'Alembert principle is

$$D_{EL}L_h(v_k, v_{k+1}) = \alpha_h(v_k) + \alpha_h(v_{k+1}), \tag{4.16}$$

$$v_k, v_{k+1} \in \mathcal{D}_\gamma,$$

for $Y = 0$. Written out explicitly, this is

$$dL_h(v_k)\,\delta v_k + dL_h(v_{k+1})\,\delta v_{k+1} = \alpha_h(v_k)\,\delta v_k + \alpha_h(v_{k+1})\,\delta v_{k+1}, \tag{4.17}$$

$$T_{v_k}\partial_h^- \,\delta v_k = 0, \tag{4.18}$$

$$T_{v_{k+1}}\partial_h^+ \,\delta v_{k+1} = 0, \tag{4.19}$$

$$T_{v_k}\partial_h^+\,\delta v_k = T_{v_{k+1}}\partial_h^-\,\delta v_{k+1}, \tag{4.20}$$

$$\partial_h^+(v_k) = \partial_h^-(v_{k+1}), \tag{4.21}$$

$$T_{v_k}\partial_h^+\,\delta v_k \in \mathcal{D}, \tag{4.22}$$

$$v_k \text{ and } v_{k+1} \in \mathcal{D}_Y, \tag{4.23}$$

for $Y = 0$. Equations (4.18), (4.19) and (4.20) each provide n equations and Equation (4.22) gives another $n-d$ so that of the $4n$ variations $\delta v_k, \delta v_{k+1}$, only d can be chosen independently.

Choosing d independent variations gives d equations from Equation (4.17). There are n equations in Equation (4.21) and $2(n-d)$ in Equation (4.23). This gives a total of $3n-d$ equations in the $4n$ variables v_k, v_{k+1}.

Theorem 4 gives the conditions for the discrete Euler–Lagrange equations to have solutions.

Theorem 4. *Let $(L_h, \psi, \alpha^+, \alpha^-)$ be a discretization of a $\mathcal{D}$–hyperregular Lagrangian system. Then there are neighbourhoods U, V: $\mathcal{D}\times\mathcal{D}\times\mathbb{R} \supseteq U \supseteq \Delta(\mathcal{D}\times\mathcal{D})\times\{0\}$ and $\mathcal{D}\times\mathbb{R}\supseteq V\supseteq \mathcal{D}\times\{0\}$ such that for all $(v_1, h)\in V$ there is a unique $v_2\in\mathcal{D}$ such that $(v_1, v_2, h)\in U$ and (v_1, v_2) satisfy the discrete Euler–Lagrange equations.*

Proof. A neighbourhood of $h = 0$ is required, however, the discrete Euler–Lagrange equations are singular when $h = 0$, as will now be shown.

First, note that $R_0 = \tau_Q = \partial_0^+ = \partial_0^-$. Both α_h and L_h are singular at $h = 0$, so take limits, letting $T\tau_Q\,\delta v = \delta q$,

$$\begin{aligned}
\lim_{h\to 0}\alpha_h(v_q)\,\delta v &= \lim_{h\to 0}\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} F_c(R_t'(v_q))\,TR_t\,\delta v\,dt \quad \text{definition of } \alpha_h\\
&= \frac{d}{dh}\bigg|_{h=0}\int_{\alpha^-(h)}^{\alpha^+(h)} F_c(R_t'(v_q))\,TR_t\,\delta v\,dt \quad \text{definition of derivative}\\
&= F_c(v_q)\,\delta q\ \frac{d}{dh}\bigg|_{h=0}(\alpha^+(h)-\alpha^-(h)) \quad \text{fundamental theorem of calculus}\\
&= F_c(v_q)\,\delta q. \quad \text{since } \alpha^+(h)-\alpha^-(h) = h
\end{aligned} \tag{4.24}$$

Write α for α_0. Also

$$\begin{aligned}
\lim_{h\to 0} L_h(v_q) &= \lim_{h\to 0} \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} L(R_t'(v_q))\,dt \quad \text{definition of } L_h \\
&= \frac{d}{dh}\bigg|_{h=0} \int_{\alpha^-(h)}^{\alpha^+(h)} L(R_t'(v_q))\,dt \quad \text{definition of derivative} \\
&= L(v_q)\ \frac{d}{dh}\bigg|_{h=0} (\alpha^+(h)-\alpha^-(h)) \quad \text{fundamental theorem of calculus} \\
&= L(v_q). \quad \text{since } \alpha^+(h)-\alpha^-(h) = h
\end{aligned} \tag{4.25}$$

Using v_1 for v_k and v_2 for v_{k+1}, Equations (4.17) to (4.23) become

$$dL(v_1)\,\delta v_1 + dL(v_2)\,\delta v_2 = \alpha(v_1)\,\delta v_1 + \alpha(v_2)\,\delta v_2, \tag{4.26}$$

$$T_{v_1}\tau_Q\,\delta v_1 = 0, \tag{4.27}$$

$$T_{v_2}\tau_Q\,\delta v_2 = 0, \tag{4.28}$$

$$T_{v_1}\tau_Q\,\delta v_1 = T_{v_2}\tau_Q\,\delta v_2, \tag{4.29}$$

$$\tau_Q(v_1) = \tau_Q(v_2), \tag{4.30}$$

$$T_{v_1}\tau_Q\,\delta v_1 \in \mathcal{D}, \tag{4.31}$$

$$v_1 \text{ and } v_2 \in \mathcal{D}_Y, \tag{4.32}$$

solved at $Y = 0$. Let $\tau_Q(v_1) = q_1$, $\tau_Q(v_2) = q_2$, $T\tau_Q\delta v_1 = \delta q_1$ and $T\tau_Q\delta v_2 = \delta q_2$, then Equations (4.27) and (4.28) give $\delta q_1 = 0$ and $\delta q_2 = 0$ respectively. Equations (4.29) and (4.31) are then trivially satisfied, leaving $2n$ degrees of freedom in choosing the variations δv_1 and δv_2.

The $2n$ independent variations give $2n$ equations from Equation (4.26). Equation (4.30) gives another n and Equation (4.32) gives $2(n-d)$.

There is then $5n-d$ equations in the $4n$ variables v_1 and v_2, possibly overdetermining the system.

To desingularize, the normal bundle E and the tubular neighbourhood used in the proof of Theorem 3 will be used.

Since ∂_h^+ and ∂_h^- are submersions on $\mathcal{D}_Y$ at $h = 0$, there is an open neighbourhood, $\hat{U}$,

$$\mathcal{D}_Y \times \{0\} \subset \hat{U} \subset \mathcal{D}_Y \times \mathbb{R},$$

such that $\partial_h^+ \times \mathrm{id}$ and $\partial_h^- \times \mathrm{id}$ are submersions. Therefore,

$$C = \{(v_1, v_2, h) \in \mathcal{D}_Y \times \mathcal{D}_Y \times \mathbb{R} \,|\, \partial_h^+(v_1) = \partial_h^-(v_2), (v_1, h) \in \hat{U}, (v_2, h) \in \hat{U}\}$$

is a submanifold of $\mathcal{D}_Y \times \mathcal{D}_Y \times \mathbb{R}$.

Define $\beta_h^\pm \colon C \to Q \times Q$ by

$$\beta_h^\pm(v_1, v_2) = \left(\partial_h^+(v_2), \partial_h^-(v_1)\right).$$

Then, for $(v_1, v_2, h) \in C$ such that $\beta_h^\pm(v_1, v_2) \in W^{\Delta(Q\times Q)}$ (see the proof of Theorem 3), define the curve c_{v_1,v_2} in E by

$$\zeta(hc_{v_1,v_2}(h)) = \beta_h^\pm(v_1, v_2).$$

Then, for $h > 0$,

$$c_{v_1,v_2}(h) = \frac{1}{h}\zeta^{-1} \circ \beta_h^\pm(v_1, v_2).$$

By a calculation and argument similar to the one used in the proof of Theorem 3, c_{v_1,v_2} is C^{k-1} and

$$c_{v_1,v_2}(0) = \mathrm{vert}T_{(q,q)}\zeta^{-1} \left.\frac{d}{dh}\right|_{h=0} \beta_h^\pm(v_1, v_2).$$

Using Equations (4.7) and (4.8),

$$\begin{aligned} \mathrm{vert}\, T_{(q,q)}\zeta^{-1} \left.\frac{d}{dh}\right|_{h=0} \beta_h^\pm(v_1, v_2) &= \mathrm{vert}\, T_{(q,q)}\zeta^{-1} \Big(\left(\alpha^+\right)'(0)v_2, \left(\alpha^-\right)'(0)v_1\Big) \\ &= \frac{1}{2}\Big(\left(\alpha^+\right)'(0)v_2 - \left(\alpha^-\right)'(0)v_1, \left(\alpha^-\right)'(0)v_1 - \left(\alpha^+\right)'(0)v_2\Big). \end{aligned} \tag{4.33}$$

On C, $\partial_h^+(v_1) = \partial_h^-(v_2)$, so

$$\left.\frac{d}{dh}\right|_{h=0} \partial_h^+(v_1) = \left.\frac{d}{dh}\right|_{h=0} \partial_h^-(v_2).$$

The definitions of ∂_h^+ and ∂_h^- give

$$\left(\alpha^+\right)'(0)\, v_1 = \left(\alpha^-\right)'(0)\, v_2.$$

Adding and subtracting $(\alpha^-)'(0)v_1$ to the left hand side and $(\alpha^+)'(0)v_2$ to the right hand side gives

$$\left(\alpha^+\right)'(0)v_1 - \left(\alpha^-\right)'(0)v_1 + \left(\alpha^-\right)'(0)v_1 = \left(\alpha^-\right)'(0)v_2 - \left(\alpha^+\right)'(0)v_2 + \left(\alpha^+\right)'(0)v_2.$$

Since $(\alpha^+)'(0) - (\alpha^-)'(0) = 1$,

$$\left(\alpha^+\right)'(0)v_2 - \left(\alpha^-\right)'(0)v_1 = v_1 + v_2.$$

So that Equation (4.33) is

$$\mathrm{vert}T_{(q,q)}\zeta^{-1} \left.\frac{d}{dh}\right|_{h=0} \beta_h^{\pm}(v_1, v_2) = \frac{1}{2}\left(v_1 + v_2, -v_1 - v_2\right).$$

Define $\phi\colon \left\{(v_1, v_2, h) \in C \,|\, \beta_h^{\pm}(v_1, v_2) \in W^{\Delta(Q\times Q)}\right\} \to W^{0(E)} \times \mathbb{R}$ by

$$\phi(v_1, v_2, h) = \begin{cases} \left(\frac{1}{h}\zeta^{-1} \circ \beta_h^{\pm}(v_1, v_2), h\right), & h > 0, \\ \left(\frac{1}{2}\left(v_1 + v_2, -v_1 - v_2\right), 0\right), & h = 0. \end{cases}$$

By construction, ϕ is smooth and a submersion.

The derivative of ϕ for h constant is

$$T_{(v_1,v_2,h)}\phi\left(\delta v_1, \delta v_2, 0\right) = \begin{cases} \left(\frac{1}{h}T_{\beta_h^{\pm}(v_1,v_2)}\zeta^{-1}\left(T_{v_2}\partial_h^+\delta v_2, T_{v_1}\partial_h^-\delta v_1\right), 0\right), & h > 0, \\ \left(\frac{1}{2}\left(\delta v_1 + \delta v_2, -\delta v_1 - \delta v_2\right), 0\right), & h = 0. \end{cases}$$

The kernel of $T\phi$ at (v_1, v_2, h) is therefore, for $h > 0$

$$T_{v_1}\partial_h^-\delta v_1 = 0, \qquad T_{v_2}\partial_h^+\delta v_2 = 0, \qquad \delta h = 0.$$

The first two are Equations (4.18) and (4.19) for the infinitesimal variations of the discrete Euler–Lagrange equations.

On C, define the distribution $\mathcal{F}$ by

$$\mathcal{F}_{(v_1,v_2,h)} = \ker T\phi \cap (T_{v_1}(\partial_h^+)^{-1}\mathcal{D} \times T_{v_2}(\partial_h^-)^{-1}\mathcal{D} \times \mathbb{R}),$$

$$\mathcal{F} = \bigcup_{(v_1,v_2,h)\in C} F_{(v_1,v_2,h)}.$$

Also define $g\colon C \to \mathcal{D}_Y \times \mathbb{R}$ by

$$g(v_1, v_2, h) = (v_1, h).$$

Solutions to the discrete Euler–Lagrange equations are therefore solutions to the skew critical point problem, for $Y = 0$,

$$dL_h(v_1)\,\delta v_1 + dL_h(v_2)\,\delta v_2 = \alpha_h(v_1)\,\delta v_1 + \alpha_h(v_2)\,\delta v_2, \tag{4.34}$$

$$(\delta v_1, \delta v_2, \delta h) \in \mathcal{F}, \tag{4.35}$$

$$g(v_1, v_2, h) = (v_1, h). \tag{4.36}$$

At $h = 0$, $\partial_0^+(v_q) = \partial_0^-(v_q) = \tau_Q(v_q)$ so that $(\delta v_1, \delta v_2, \delta h) \in \ker T\phi$ if

$$\delta v_2 = -\delta v_1, \qquad T\tau_Q\,\delta v_1 = T\tau_Q\,\delta v_2. \tag{4.37}$$

In addition, Equations (4.27) and (4.28) imply that δv_1 and δv_2 are vertical so that the second set of equations in Equation (4.37) are redundant.

Let v_q be fixed in $\mathcal{D}_Y$ and let $\delta q(h)$ be a curve in $\mathcal{D}$ such that $\tau_Q \delta q(h) = \partial_h^+(v_q)$. Let $\delta v(h)$ be the curve in $T_{v_q}\mathcal{D}_\gamma$ satisfying $T_{v_q}\partial_h^+\,\delta v(h) = \delta q(h)$. Then $\delta v(h)$ is vertical and,

$$\begin{aligned}
\left.\frac{d}{dh}\right|_{h=0} \delta q(h) &= \left.\frac{d}{dh}\right|_{h=0} T_{v_q}\partial_h^+\,\delta v(h) \\
&= \left.\frac{d}{dh}\right|_{h=0} T_{v_q}\partial_0^+\,\delta v(h) + \left.\frac{d}{dh}\right|_{h=0} T_{v_q}\partial_h^+\delta v(0) \\
&= T_{v_q}\tau_Q \left.\frac{d}{dh}\right|_{h=0} \delta v(h) + \left(\alpha^+\right)'(0)\delta v(0) \\
&= \left(\alpha^+\right)'(0)\delta v(0). \quad \text{since } \delta v(h) \text{ is vertical}
\end{aligned}$$

Therefore, $\left.\frac{d}{dh}\right|_{h=0} \delta q(h)$ is vertical.

Define the distribution on TQ

$$\mathcal{D}' = \left\{ \delta v \in TTQ \,|\, \delta v = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \epsilon\delta q, \delta q \in \mathcal{D} \right\}.$$

On $g^{-1}(v_1, 0)$, $\mathcal{F}$ is therefore

$$\mathcal{F} = \{(\delta v_1, \delta v_2, 0) \in Tg^{-1}(v_1, 0) \,|\, \delta v_2 = -\delta v_1, \delta v_1 \in \mathcal{D}'\}. \tag{4.38}$$

Equation (4.24) implies that $\alpha(v_1)\,\delta v_1 = \alpha(v_2)\,\delta v_2 = 0$ since δv_1 and δv_2 are vertical. Using Equation (4.25), the discrete Euler–Lagrange equations at $h = 0$ are, for $Y = 0$ on C,

$$dL(v_1)\,\delta v_1 + dL(v_2)\,\delta v_2 = 0, \tag{4.39}$$

$$(\delta v_1, \delta v_2, \delta h) \in \mathcal{F}, \tag{4.40}$$

$$g(v_1, v_2, h) = (v_1, 0). \tag{4.41}$$

Using Equation (4.38) for $\mathcal{F}$ on $g^{-1}(v_1, 0)$, Equations (4.39), (4.40) and (4.41) have solution $v_2 = v_1$ since L is $\mathcal{D}$–hyperregular.

Lemma 5 then furnishes the open neighbourhoods U and V of the theorem.

□

Theorem 5. *Let L be a $\mathcal{D}$–hyperregular Lagrangian and $F_t^{X_E}$ the flow of the Euler–Lagrange vector field. Let $R_t = \tau_Q \circ F_t^{X_E}$ and V be the neighbourhood of Theorem 4. Then $(v_1, v_2, h) \in V$ satisfies the discrete Euler–Lagrange equations if and only if*

$F_h^{X_E}(v_1) = v_2.$

Proof. First, recall Equation (4.12),

$$dL_h(v_q)\,\delta v = \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} D_{EL}L\left(F_t^{X_E}(v_q)\right)\, T_{v_q}\left(\tau_Q \circ F_t^{X_E}\right)\,\delta v\, dt + \\ + \frac{1}{h}\Theta_L\left(F_t^{X_E}(v_q)\right)\, T_{v_q} F_t^{X_E}\,\delta v\Big|_{t=\alpha^-(h)}^{\alpha^+(h)}.$$

The continuous Lagrange–d'Alembert principle implies

$$dL_h(v_q)\,\delta v = \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} F_c\left(F_t^{X_E}(v_q)\right)\, T_{v_q}\left(\tau_Q \circ F_t^{X_E}\right)\,\delta v\, dt + \\ + \frac{1}{h}\Theta_L\left(F_t^{X_E}(v_q)\right)\, T_{v_q} F_t^{X_E}\,\delta v\Big|_{t=\alpha^-(h)}^{\alpha^+(h)}.$$

Using the definition of α_h gives

$$dL_h(v_q)\,\delta v = \alpha_h(v_q)\,\delta v + \frac{1}{h}\Theta_L\left(F_t^{X_E}(v_q)\right)\, T_{v_q} F_t^{X_E}\,\delta v\Big|_{t=\alpha^-(h)}^{\alpha^+(h)},$$

so that equation (4.17) becomes

$$\Theta_L\left(F^{X_E}_{\alpha^+(h)}(v_1)\right)\,T_{v_1}F^{X_E}_{\alpha^+(h)}\,\delta v_1-\Theta_L\left(F^{X_E}_{\alpha^-(h)}(v_2)\right)\,T_{v_2}F^{X_E}_{\alpha^-(h)}\,\delta v_2=0. \tag{4.42}$$

Equations (4.20) and (4.22) give

$$T_{v_1}F^{X_E}_{\alpha^+(h)}\,\delta v_1=\delta q=T_{v_2}F^{X_E}_{\alpha^-(h)}\,\delta v_2$$

for $\delta q\in\mathcal{D}$, so that the equations implied by (4.42) are,

$$\mathbb{F}L\left(F^{X_E}_{\alpha^+(h)}(v_1)\right)\,\delta q=\mathbb{F}L\left(F^{X_E}_{\alpha^-(h)}(v_2)\right)\,\delta q,$$

$$\delta q\in\mathcal{D},$$

$$v_1,v_2\in\mathcal{D}.$$

Since L is $\mathcal{D}$–hyperregular, these equations have solution

$$F^{X_E}_{\alpha^+(h)}(v_1)=F^{X_E}_{\alpha^-(h)}(v_2),$$

which is

$$F^{X_E}_h(v_1)=v_2.$$

□

Proposition 16. *Let* $(\psi^2,\alpha^+,\alpha^-)$ *and* $(\psi^1,\alpha^+,\alpha^-)$ *be discretizations of* TQ *such that* $\psi^2(h,t,v_q)=\psi^1(h,t,v_q)+O(t^r)$. *Let* G *be a one form on* TQ *and define, for* $i=1,2$, *the following one forms on* TQ:

$$G^i_h(v_q)w_{v_q}=\frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}G\left(\frac{\partial\psi^i}{\partial t}(h,t,v_q)\right)\,T_{v_q}\left(\frac{\partial}{\partial t}\psi^i\right)\,w_{v_q}\,dt.$$

Then $G^2_h=G^1_h+O(h^{r-1})$.

Proof. Use the notation of Equations (4.10) and (4.11). The hypothesis of the theorem is then $R^2_t(v_q)=R^1_t(v_q)+O(t^r)$. In local coordinates $(q,\dot q)$ for TQ,

$$\hat R^2_t(q,\dot q)=\hat R^1_t(q,\dot q)+t^r\delta R_t(q,\dot q),$$

for some continuous δR_t. In fact, since R_t^2 and R_t^1 are assumed smooth, δR_t will also be smooth. Then

$$\begin{aligned}(\hat{R}_t^2)'(q,\dot{q}) &= (\hat{R}_t^1)'(q,\dot{q}) + (t^r \delta R_t(q,\dot{q}))' \\ &= (\hat{R}_t^1)'(q,\dot{q}) + t^{r-1}\delta\tilde{R}_t(q,\dot{q}),\end{aligned}$$

for some $\delta\tilde{R}_t(q,\dot{q})$. Theorem 9 of Appendix B gives

$$\hat{G}\circ(\hat{R}_t^2)'(q,\dot{q}) = \hat{G}\circ(\hat{R}_t^1)'(q,\dot{q}) + t^{r-1}\delta G(q,\dot{q},t).$$

Also,

$$\begin{aligned}T_{(q,\dot{q})}(\hat{R}_t^2)'\,w &= T_{(q,\dot{q})}((\hat{R}_t^1)' + t^{r-1}\delta\tilde{R}_t)\,w \\ &= T_{(q,\dot{q})}((\hat{R}_t^1)'\,w + t^{r-1}T_{(q,\dot{q})}\delta\tilde{R}_t\,w,\end{aligned}$$

so that

$$\begin{aligned}\hat{G}\circ(\hat{R}_t^2)'(q,\dot{q})\,T_{(q,\dot{q})}(\hat{R}_t^2)'\,w &= \left(\hat{G}\circ(\hat{R}_t^1)'(q,\dot{q}) + t^{r-1}\delta G(q,\dot{q},t)\right)T_{(q,\dot{q})}\left((\hat{R}_t^1)' + t^{r-1}\delta\tilde{R}_t\right)\,w \\ &= \hat{G}\circ(\hat{R}_t^1)'(q,\dot{q})\,T_{(q,\dot{q})}(\hat{R}_t^1)'\,w + t^{r-1}\delta\tilde{G}(q,\dot{q},t)\,w,\end{aligned}$$

for some continuous $\delta\tilde{G}$. Then,

$$\begin{aligned}\hat{G}_h^2(q,\dot{q})\,w &= \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}\hat{G}\circ(\hat{R}_t^2)'\,T_{(q,\dot{q})}(\hat{R}_t^2)'\,w\,dt \\ &= \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)}\left(\hat{G}\circ(\hat{R}_t^1)'(q,\dot{q})\,T_{(q,\dot{q})}(\hat{R}_t^1)'\,w + t^{r-1}\delta\tilde{G}(q,\dot{q},t)\,w\right)\,dt \\ &= \hat{G}_h^1(q,\dot{q})\,w + \frac{1}{h}\int_{\alpha^-(h)}^{\alpha^+(h)} t^{r-1}\delta\tilde{G}(q,\dot{q},t)\,w\,dt. \qquad (4.43)\end{aligned}$$

Assume $|\alpha^-(h)| \leq |\alpha^+(h)|$ (the other case is similar). Then the integral term in Equation (4.43) satisfies

$$h(\alpha^-(h))^{r-1}\delta\tilde{G}_{\min}(q,\dot{q}) \leq \int_{\alpha^-(h)}^{\alpha^+(h)} t^{r-1}\delta\tilde{G}(q,\dot{q},t)\,w\,dt \leq h(\alpha^+(h))^{r-1}\delta\tilde{G}_{\max}(q,\dot{q}),$$

where

$$\delta\tilde{G}_{\min}(q,\dot{q}) = \min_{t\in[\alpha^-(h),\alpha^+(h)]}\delta\tilde{G}(q,\dot{q},t),$$

$$\delta\tilde{G}_{\max}(q,\dot{q}) = \max_{t\in[\alpha^-(h),\alpha^+(h)]} \delta\tilde{G}(q,\dot{q},t).$$

Since $\alpha^-(h) = O(h)$ and $\alpha^+(h) = O(h)$, for some constants $\tilde{G}_{\min}$ and $\tilde{G}_{\max}$,

$$\lim_{h\to 0} \frac{|h(\alpha^-(h))^{r-1}\delta\tilde{G}_{\min}(q,\dot{q})|}{|h|^r} = \tilde{G}_{\min},$$
$$\lim_{h\to 0} \frac{|h(\alpha^-(h))^{r-1}\delta\tilde{G}_{\max}(q,\dot{q})|}{|h|^r} = \tilde{G}_{\max}.$$

Therefore,

$$\lim_{h\to 0} \frac{|\int_{\alpha^-(h)}^{\alpha^+(h)} t^{r-1}\delta\tilde{G}(q,\dot{q},t)\, w\, dt|}{|h|^r} \leq \delta\tilde{G}_{\min},$$

and

$$\int_{\alpha^-(h)}^{\alpha^+(h)} t^{r-1}\delta\tilde{G}(q,\dot{q},t)\, w\, dt = h^r\delta G_h(q,\dot{q}),$$

for some continuous δG_h. And finally,

$$\hat{G}_h^2(q,\dot{q}) = \hat{G}_h^1(q,\dot{q}) + h^{r-1}\delta G_h(q,\dot{q}),$$

so that $G_h^2 = G_h^1 + O(h^{r-1})$. □

The discrete Lagrange–d'Alembert principle gives the discrete evolution equations as a fixed endpoint problem. Lemma 14 in Appendix B implies that the problem can also be set as an initial value problem. Given a solution to the discrete Lagrange–d'Alembert equations, $(v_1, v_2, h) \in TQ \times TQ \times \mathbb{R}$, there is a mapping $F_h\colon TQ \to TQ$ such that $F_h(v_1) = v_2$. The mapping F_h is the *discrete flow.*

Theorem 6. *Let $(L_h^2, \psi^2, \alpha^+, \alpha^-)$ and $(L_h^1, \psi^1, \alpha^+, \alpha^-)$ be discretizations of a $\mathcal{D}$–hyper-regular Lagrangian system with discrete flows F_h^2 and F_h^1 respectively. If*
$\psi^2 = \psi^1 + O(h^r)$, then $F_h^2 = F_h^1 + O(h^r)$.

Proof. Define, for $i = 1, 2$, the following one forms on $TQ \times TQ \times \mathbb{R}$:

$$\mathfrak{S}^i(v_1, v_2, h)\,(\delta v_1, \delta v_2, \delta h) = dL_h^i(v_1)\,\delta v_1 + dL_h^i(v_2)\,\delta v_2 - \alpha_h^i(v_1)\,\delta v_1 - \alpha_h^i(v_2)\,\delta v_2.$$

For $i = 1, 2$, use $G^i = dL$ for the dL_h^i terms and $G^i = F_c^i$ for the α_h terms. Proposition 16 implies that $\mathfrak{S}^2 = \mathfrak{S}^1 + O(h^{r-1})$. Order r is confirmed by a direct calculation using the desingularization maps in Theorem 22 in Appendix D. For $i = 1, 2$, these are

$$
\begin{aligned}
\eta\colon & \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}, \\
& \eta(q^+, q^-, h) = (\alpha^+(h)q^- - \alpha^-(h)q^+, q^+ - q^-, h), \\
\rho\colon & \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{2n} \times \mathbb{R}, \\
& \rho(q^+, q^-, h) = \left(\frac{1}{h}(q^+, q^-), h\right), \\
(\beta^\pm)^i \times \mathrm{id}\colon & TQ \times TQ \times \mathbb{R} \to Q \times Q \times \mathbb{R}, \\
& (\beta^\pm)^i(v_1, v_2, h) = ((\partial_h^+)^i(v_2), (\partial_h^-)^i(v_1), h), \\
\phi^i\colon & \mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^{2n} \times \mathbb{R}, \\
& \phi^i((q_1, \dot{q}_1), (q_2, \dot{q}_2), h) = (\rho \circ \eta \circ \widehat{(\beta^\pm)^i \times} \,\mathrm{id})((q_1, \dot{q}_1), (q_2, \dot{q}_2), h), \\
g^i\colon & \mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}, \\
& g^i((q_1, \dot{q}_1), (q_2, \dot{q}_2), h) = ((\hat{\partial}_h^+)^i(q_1, \dot{q}_1) - (\hat{\partial}_h^-)^i(q_2, \dot{q}_2), h).
\end{aligned}
$$

Working in local coordinates $(q, \dot{q})$ for TQ, Define the distributions $\hat{D}_i$, $i = 1, 2$ as in Equations (D.10) to (D.13).

Let $((\delta q_1^2, \delta\dot{q}_1^2), (\delta q_2^2, \delta\dot{q}_2^2), \delta h^2)$ be a vector field in a neighbourhood of $((q, \dot{q}), (q, \dot{q}), 0)$ and taking its values in $\mathcal{D}_2$. Then, by Lemma 11 in Appendix B there is a vector field $((\delta q_1^1, \delta\dot{q}_1^1), (\delta q_2^1, \delta\dot{q}_2^1), \delta h^1)$ in $\mathcal{D}_1$ such that

$$((\delta q_1^2, \delta\dot{q}_1^2), (\delta q_2^2, \delta\dot{q}_2^2), \delta h^2) = ((\delta q_1^1, \delta\dot{q}_1^1), (\delta q_2^1, \delta\dot{q}_2^1), \delta h^1) + O(h^r).$$

Then, using the definition of $\mathfrak{S}^i$,

$$
\begin{aligned}
&\mathfrak{S}^2((q_1, \dot{q}_1), (q_2, \dot{q}_2), h)\,((\delta q_1^2, \delta\dot{q}_1^2), (\delta q_2^2, \delta\dot{q}_2^2), \delta h^2) \\
&\qquad - \mathfrak{S}^1((q_1, \dot{q}_1), (q_2, \dot{q}_2), h)\,((\delta q_1^1, \delta\dot{q}_1^1), (\delta q_2^1, \delta\dot{q}_2^1), \delta h^1) \\
&= dL_h^2(q_1, \dot{q}_1)\,(\delta q_1^2, \delta\dot{q}_1^2) - dL_h^1(q_1, \dot{q}_1)\,(\delta q_1^1, \delta\dot{q}_1^1) + dL_h^2(q_2, \dot{q}_2)\,(\delta q_2^2, \delta\dot{q}_2^2) - dL_h^1(q_2, \dot{q}_2)\,(\delta q_2^1, \delta\dot{q}_2^1) +
\end{aligned}
$$

$$+\alpha_h^1(q_1,\dot q_1)\,(\delta q_1^1,\delta\dot q_1^1)-\alpha_h^2(q_1,\dot q_1)\,(\delta q_1^2,\delta\dot q_1^2)+\alpha_h^1(q_2,\dot q_2)\,(\delta q_2^1,\delta\dot q_2^1)-\alpha_h^2(q_2,\dot q_2)\,(\delta q_2^2,\delta\dot q_2^2)$$

$$=h^{r-1}\left(\delta L_h(q_1,\dot q_1)\,(\delta q_1^1,\delta\dot q_1^1)+\delta L_h(q_2,\dot q_2)\,(\delta q_2^1,\delta\dot q_2^1)\;+\delta\alpha_h(q_1,\dot q_1)\,(\delta q_1^1,\dot q_1^1)+\delta\alpha_h(q_2,\dot q_2)\,(\delta q_2^1,\delta\dot q_2^1)\right).$$

The proof of Theorem 4 shows that at $h=0$, the solution of the discrete Euler–Lagrange equations is $(q_2,\dot q_2)=(q_1,\dot q_1)$ and $((\delta q_1,\delta\dot q_1),(\delta q_2,\delta\dot q_2))=((0,\delta\dot q_1),(0,-\delta\dot q_1))$, for $\delta\dot q_1\in\hat{\mathcal{D}}$. Then

$$\begin{aligned}\operatorname{res}^{r-1}\left(\mathfrak{S}^2,\mathfrak{S}^1\right)((q_1,\dot q_1),(q_2,\dot q_2),0)&=\delta L_h\left(q_1,\dot q_1\right)\,(0,\delta\dot q_1)+\\&+\delta L_h\left(q_1,\dot q_1\right)\,(0,-\delta\dot q_1)+\delta\alpha_h\left(q_1,\dot q_1\right)\,(0,\delta\dot q_1)+\delta\alpha_h\left(q_1,\dot q_1\right)\,(0,-\delta\dot q)=0.\end{aligned}$$

Therefore $\mathfrak{S}^2=\mathfrak{S}^1+O(h^r)$. The definitions of g^i give $g^2=g^1+O(h^r)$, so Lemma 13 in Appendix B gives $\gamma^2=\gamma^1+O(h^r)$ and $F_h^2=F_h^1+O(h^r)$ follows from Lemma 15.

□

Theorem 7. *Let the flow of the Euler–Lagrange vector field be $F_t^{X_E}$ and suppose $\psi(h,t,v_q)=\tau_Q\circ F_t^{X_E}(v_q)+O(h^r)$. Then $F_h(v_q)=F_h^{X_E}+O(h^r)$.*

Proof. In Theorems 5 and 6, let $\psi^1=\psi$ and $\psi^2=\tau_Q\circ F_t^{X_E}$. □

4.5 Construction of a Lagrange–d'Alembert Integrator

The construction of a Lagrange–d'Alembert integrator requires a relatively compact submanifold, Q_0, of Q that is suitable to the problem at hand. To construct an integrator that is r order accurate, choose ψ such that $\psi(h,t,v_q)=\tau_Q\circ F_t^{X_E}(v_q)+O(h^{r+1})$ and exactly satisfies the constraints.

The time step, h, must be chosen so that:

1. The discrete tangent bundle, P_h, guaranteed by Theorem 3 contains the region of interest.

2. The neighbourhood V of Theorem 4 contains the region of interest.

Chapter 5 contains details on the construction of a specific type of Lagrange–d'Alembert integrator.

4.6 Symmetry

The terminology in this section refers to Section 3.7. The condition that $\partial^\pm \circ \Phi_g = \phi_g \circ \partial^\pm$ is strengthened somewhat by requiring that ϕ_g take the curve $t \mapsto \psi(h, t, v_q)$ to the curve $t \mapsto \psi(h, t, \Phi_g(v_q))$, where Φ_g is the discrete lifted action. The symmetry of the continuous system is related to the symmetry of the discrete system in the following theorem.

Proposition 17. *Let $\mathcal{G}$ be a Lie group acting on Q by ϕ_g and on TQ by Φ_g (not necessarily the tangent lift). Denote the tangent lift of ϕ_g by Φ_g^T. Let $R_t(v_q) = \psi(h, t, v_q)$ and suppose that $\phi_g \circ R_t = R_t \circ \Phi_g$. Let L be a Lagrangian and L_h the discretized Lagrangian. Then, L_h is invariant under the Φ action if L is invariant under the Φ^T action.*

Proof.

$$\begin{aligned}
(L_h \circ \Phi_g)(v_q) &= \frac{1}{h} \int_{\alpha_-(h)}^{\alpha^+(h)} L\left((R_t \circ \Phi_g)'(v_q) \right) dt \quad \text{definition of } L_h \\
&= \frac{1}{h} \int_{\alpha_-(h)}^{\alpha^+(h)} L\left((\phi_g \circ R_t)'(v_q) \right) dt \quad \text{definition of } \Phi_g \\
&= \frac{1}{h} \int_{\alpha_-(h)}^{\alpha^+(h)} L\left(T_{R_t(v_q)}\phi_g \, R_t'(v_q) \right) dt \quad \text{chain rule} \\
&= \frac{1}{h} \int_{\alpha_-(h)}^{\alpha^+(h)} \left(L \circ \Phi_g^T \right) \left(R_t'(v_q) \right) dt \quad \text{definition of } \Phi_g^T \\
&= \frac{1}{h} \int_{\alpha_-(h)}^{\alpha^+(h)} L\left(R_t'(v_q) \right) dt \quad \text{since } L \text{ is invariant} \\
&= L_h(v_q). \quad \text{definition of } L_h
\end{aligned}$$

□

The discrete momentum equation developed in Section 3.7 needs an adjustment because of the addition of α_h to the Discrete Lagrange–d'Alembert principle. Rewriting Equation (4.16), this is

$$D_{EL}L_h(v_k, v_{k+1}), \delta q_k = \alpha_h(v_k)\,\delta v_k + \alpha_h(v_{k+1})\,\delta v_{k+1}. \tag{4.44}$$

Then, as in Section 3.7, using Equation (4.44) and employing the identical steps as in Section 3.7, Equation (3.34) is

$$\begin{aligned} &J^{L_d}(F^N_{L_h}(v_q))\,(\xi^N - \xi^{N-1}) + F^*_{L_h}\alpha_h(F^{N-1}_{L_h}(v_q))\,\xi^{N-1}_P(F^{N-1}_{L_h}(v_q)) \\ &\qquad = J^{L_d}_{\xi^N}(F^N_{L_h}(v_q)) - J^{L_d}_{\xi^{N-1}}(F^{N-1}_{L_h}(v_q)) - \alpha_h(F^{N-1}_{L_h}(v_q))\,\xi^{N-1}_P(F^{N-1}_{L_h}(v_q)). \end{aligned} \tag{4.45}$$

For $N = 1$, Equation (4.45) is

$$\begin{aligned} &J^{L_d}(F_{L_h}(v_q))\,(\xi^1 - \xi^0) + F^*_{L_h}\alpha_h(v_q)\,\xi^0_P(v_q) \\ &\qquad = J^{L_d}_{\xi^1}(F_{L_h}(v_q)) - J^{L_d}_{\xi^0}(v_q) - \alpha_h(v_q)\,\xi^0_P(v_q). \end{aligned} \tag{4.46}$$

If ξ^i can be chosen to be constant, then $\xi^1 = \xi^0 = \xi$ and Equation (4.46) is, with some rearranging,

$$J^{L_h}_\xi(F_{L_h}(v_q)) - F^*_{L_h}\alpha_h(v_q)\,\xi_P(v_q) = J^{L_h}_\xi(v_q) + \alpha_h(v_q)\,\xi_P(v_q).$$

Chapter 5

Implementation

5.1 Introduction

This chapter provides a description of the steps required to construct the Lagrange–d'Alembert integrators used to numerically evolve the examples in Chapter 6. Each aspect of the necessary computations is listed in Section 5.2 with the details in appropriate subsections. This chapter includes an explicit construction of a first order Lagrange–d'Alembert integrator in Section 5.2.2 for a constrained Lagrangian system. The computations will be repeated for each example in Chapter 6.

There are various summations over local quantities that require different start and stop indices. Throughout this chapter, the following convention will be applied:

$$a = 1, \ldots, d, \tag{5.1}$$

$$b = d + 1, \ldots, n, \tag{5.2}$$

$$c = d + 1, \ldots, n, \tag{5.3}$$

$$i = 1, \ldots, n. \tag{5.4}$$

Primed indices, ie. a', run over the same numbers as the unprimed indices.

5.2 Methodology

The examples in Chapter 6 are separated into several parts.

Physical and Mathematical analysis.

Construction of a first order Lagrange–d'Alembert integrator.

Numerical Results.

Remarks.

Each part is further subdivided

Physical and Mathematical analysis.

Description of the physical system and the mathematical model. This includes the configuration manifold, Q, the phase space, TQ, the Lagrangian, L and the constraint distribution, $\mathcal{D}$.

Local description, in bundle coordinates for TQ, of L and $\mathcal{D}$. The description of $\mathcal{D}$ takes the form of a local basis of $\mathcal{D}$ and the extension to a local basis of TQ. In addition, a local basis for the annihilator, $\mathcal{D}^0$ of $\mathcal{D}$ is given as well as the extension to a local basis of T^*Q. Suitable local vector fields Y are chosen for the affine constraints introduced in Section 4.4. Using the basis elements of $\mathcal{D}^0$, the local constraint functions, c, are defined.

The local equations of motion are derived and the Euler–Lagrange vector field, X_E, of Theorem 3 in Chapter 2 is obtained.

Conserved quantities are described and their invariance along solutions of the Euler–Lagrange equations are proved.

Constraint adapted description. The mapping G that gives c–adapted local coordinates of TQ is described. The local description of L, $\mathcal{D}$, $\mathcal{D}^0$, Y and c in c–adapted coordinates are given. The equations of motion in c–adapted coordinates are derived.

Construction of a first order Lagrange–d'Alembert integrator.

Construct a discrete tangent bundle. Refer to Section 4.2. Choose α^+, α^- and ψ such that ψ exactly satisfies the affine constraints. This involves choosing an initial ψ and then making a suitable modification. ∂_h^+ and ∂_h^- are defined. In practice, if an order r integrator is required, ψ will be chosen so that it matches $F_t^{X_E}$ to order r at $t = 0$. See Theorem 6 in Chapter 4.

Compute the infinitesimal variations. Compute the exact derivatives of ∂_h^+ and ∂_h^-. Solve Equations (4.18), (4.19), (4.20) and (4.22) for the δv_k and δv_{k+1}. The system of equations are underdetermined and provide a d parameter solution set for δv_k, δv_{k+1}.

Construct L_h and α_h using L, ψ and the constraint force.

The local discrete Euler–Lagrange equations are derived. This involves the exact derivative of L_h.

Numerical Results

Short time evolution by a Lagrange–d'Alembert integrator to compare with exact solutions of the Euler–Lagrange equations.

Long time evolution using the constructed first order Lagrange–d'Alembert integrator.

Evolution of the conserved quantities along the computed first order long time evolution and a comparison with Matlab's ode45 time adaptive method.

Order matching demonstration. The method order is estimated for various orders.

Remarks

Comments and further analysis of the numerical results that are specific to the system.

5.2.1 Physical and Mathematical Analysis

This section is concerned with an analysis of the continuous Lagrangian system. It includes computing the equations of motion in local bundle coordinates as well as in coordinates adapted to the constraint. The main purpose of working in coordinates adapted to the constraint is that the role of the vector field Y for the affine constraint is replaced by the off constraint coordinates in the adapted system.

In addition, conserved quantities are defined and proved to be preserved by solutions of the Euler–Lagrange equations.

Description of the Physical System and Mathematical Model

In practice, many physical systems are described by applied forces and constraints. The systems that are described in this thesis are such that they can be characterized by a Lagrangian, L, and a constraint distribution. The mathematical model is therefore a suitable configuration manifold, Q, the phase space, TQ, and the Lagrangian, L, that characterize the *unconstrained* motion along with a distribution, $\mathcal{D}$, on Q. Chapter 2 gives a thorough description of this model.

Local Description

Numerical computations are always done in coordinates, so an atlas for Q is chosen. Let $\dim Q = n$ and $\dim \mathcal{D} = d$. Further, let (U, μ) be a coordinate chart and assume that U is such that there is a set, $\mathcal{X}$, of smooth nonsingular vector fields $\{X_i(q)\}_{i=1}^n$ such that $\mathcal{X}$ is a basis for T_qU for each $q \in U$. In addition, it is convenient to construct $\mathcal{X}$ so that $\{X_a(q)\}_{a=1}^d$ is a basis for $\mathcal{D}_q$. Typically, the basis is constructed for $\mathcal{D}_q$ and extended to a basis for all of T_qU.

The dual basis, $\mathcal{Y}$, is computed. The elements of $\mathcal{Y}$, $\{\phi^i(q)\}_{i=1}^n$ have the property that $\{\phi^b(q)\}_{b=d+1}^n$ span the annihilator, $\mathcal{D}_q^0$ of $\mathcal{D}_q$.

Write $(q, \dot{q}) = T\mu(v_q)$, where both q and $\dot{q}$ are in $\mathbb{R}^n$. The local representatives, $\hat{X}_i$ and $\hat{\phi}^i$ of X_i and ϕ^i are computed, as are the local constraint functions, $\hat{c}^b(q, \dot{q}) = \hat{\phi}^b_i(q)\,\dot{q}^i$. Also, define $\hat{c}(q, \dot{q}) = (c^{d+1}(q, \dot{q}), \ldots, c^n(q, \dot{q}))$.

The affine vector fields, Y are chosen so that their local representatives are $\hat{Y}(q) = \hat{Y}^b X_b(q)$, where each $\hat{Y}^b$ is constant. Then, $\hat{c}^b(q, \dot{q}) = Y^b$ and the constraint manifold, $\mathcal{D}_Y$ is locally given as the intersection of the level sets of $\hat{c}^b$. Appendix C contains the necessary details regarding affine constraints.

Local Equations of Motion

A direct computation, using Lagrange multipliers, λ_b, is performed to obtain the Euler–Lagrange equations. In local coordinates these are

$$\frac{\partial \hat{L}}{\partial q^i} - \frac{d}{dt}\frac{\partial \hat{L}}{\partial \dot{q}^i} = \lambda_b\,\hat{\phi}^b_i(q). \tag{5.5}$$

A complete set of equations consists of the Euler–Lagrange Equations (5.5) and

$$\frac{d}{dt}q = \dot{q}, \tag{5.6}$$

$$\hat{c}^b(q, \dot{q}) = Y^b. \tag{5.7}$$

If L is $\mathcal{D}$ regular on TU, then the vector field, X_E can be found, as in Theorem 3 of Chapter 2. In practice, this amounts to expanding the time derivative in Equation (5.5). Using Equation (5.6), this is

$$\frac{\partial \hat{L}}{\partial q^i} - \frac{\partial^2 \hat{L}}{\partial q^j \partial \dot{q}^i}\dot{q}^j - \frac{\partial^2 \hat{L}}{\partial \dot{q}^j \partial \dot{q}^i}\ddot{q}^j = \lambda_b\,\hat{\phi}^b_i(q), \tag{5.8}$$

where $\ddot{q}^i = \frac{d}{dt}\dot{q}^i$. Equation (5.7) is then differentiated along a solution and the λ_b in Equation (5.8) are eliminated. Finally, the $\ddot{q}^j$ are solved for. The local Euler–Lagrange vector field $\hat{X}_E$ is

$$\hat{X}_E(q, \dot{q}) = (\dot{q}, \ddot{q}).$$

Solutions of Equations (5.5), (5.6) and (5.7) are integral curves of $\hat{X}_E$. The local flow of $\hat{X}_E$ on $T\mu TU$ will be written $F^{\hat{X}_E}_t$.

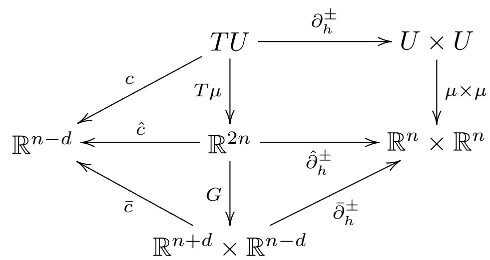

Figure 5.1: Mappings

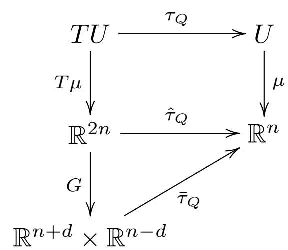

Figure 5.2: Mappings

Conserved Quantities

Proposition 6 in Chapter 2 shows that the energy,

$$\hat{E}(q, \dot{q}) = \frac{\partial \hat{L}}{\partial \dot{q}^i}\, \dot{q}^i - \hat{L}(q, \dot{q})$$

is preserved. Other conserved quantities are system specific and will be described for each example.

Constraint Adapted Description

It is convenient to work in coordinates adapted to c (see Lemma 1 in Appendix A). Figures 5.1 and 5.2 are diagrams showing the relevant mappings for this chapter. Let $(q, \dot{s}, r) = (q^1, \dots q^n, \dot{s}^1, \dots, \dot{s}^d, r^{d+1}, \dots, r^n)$ be coordinates for $\mathbb{R}^{n+d} \times \mathbb{R}^{n-d}$ and define

$$G\colon T\hat{U} \to \mathbb{R}^{n+d} \times \mathbb{R}^{n-d},$$

$$G(q, \dot{q}) = (q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^d, \hat{c}^{d+1}(q, \dot{q}), \dots \hat{c}^n(q, \dot{q})),$$

so that $r^b = \hat{\phi}^b(q)\, \dot{q}$.

The constraint functions are

$$\hat{c}^b(q,\dot{q}) = \hat{\phi}^b_i(q)\,\dot{q}^i = \hat{\phi}^b_a(q)\,\dot{q}^a + \hat{\phi}^b_c(q)\,\dot{q}^c.$$

Since the ϕ^b are assumed pointwise linearly independent, it can be assumed that $\hat{\phi}^b_c$ is invertible with inverse θ^c_b. Then, $\dot{q}^c = \theta^c_b(q)\,(r^b - \hat{\phi}^b_a(q)\,\dot{q}^a)$. This gives

$$G^{-1}(q,\dot{s},r) = (q,\dot{s},\theta^{d+1}_b(q)\,(r^b - \hat{\phi}^b_a(q)\,\dot{s}^a),\ldots,\theta^n_b(q)\,(r^b - \hat{\phi}^b_a(q)\,\dot{s}^a)).$$

Proposition 18. *Let the local representation of the Euler–Lagrange vector field, X_E in $(TU,T\mu)$ coordinates be given by $\hat{X}_E(q,\dot{q}) = \dot{q}^i\partial_{q^i} + \hat{X}^i(q,\dot{q})\partial_{\dot{q}^i}$. Then in the c adapted coordinates, $(G(T\mu U), G\circ T\mu)$, the local representation of X_E is*

$$\bar{X}_E(q,\dot{s},r) = \dot{s}^a\partial_{q^a} + \theta^c_b(q)\,(r^b - \hat{\phi}^b_a(q)\,\dot{s}^a)\partial_{q^c} + \hat{X}^a_E(G^{-1}(q,\dot{s},r))\partial_{\dot{s}^a}.$$

Proof. The derivative of G has the block form

$$DG = \begin{bmatrix} I & 0 \\ \frac{\partial\hat{c}^b}{\partial q^i} & \frac{\partial\hat{c}^b}{\partial\dot{q}^i} \end{bmatrix},$$

where I and 0 are $(n+d)\times n$ and $I^i_j = 1$ if $i=j$ and 0 otherwise and 0 is a matrix of 0s.

Write the principle part of the local representation of the Euler–Lagrange vector field as

$$\hat{X}_E(q,\dot{q}) = (\dot{q}^1,\ldots,\dot{q}^n,\hat{X}^1(q,\dot{q})\;\ldots,\hat{X}^n(q,\dot{q})). \tag{5.9}$$

Let $(q(t),\dot{q}(t))$ be an integral curve of $\hat{X}_E$. Then, by hypothesis,

$$\hat{c}^b(q(t),\dot{q}(t)) = \hat{Y}^b, \quad \text{for } Y^b \text{ constant}$$

so that differentiating both sides with respect to t gives

$$\begin{aligned} 0 &= \frac{d}{dt}\hat{c}^b(q(t),\dot{q}(t)) \\ &= \frac{\partial\hat{c}^b}{\partial q^i}(q(t),\dot{q}(t))\,\frac{d}{dt}q^i(t) + \frac{\partial\hat{c}^b}{\partial\dot{q}^i}(q(t),\dot{q}(t))\,\frac{d}{dt}\dot{q}^i(t) \quad \text{chain rule} \\ &= \frac{\partial\hat{c}^b}{\partial q^i}(q(t),\dot{q}(t))\,\dot{q}^i(t) + \frac{\partial\hat{c}^b}{\partial\dot{q}^i}(q(t),\dot{q}(t))\,\hat{X}^i(q(t),\dot{q}(t)). \quad \text{Equation (5.9)} \end{aligned}$$

Then

$$DG(q,\dot{q})\,\hat{X}_E = (\dot{q}^1,\ldots,\dot{q}^n,\hat{X}^1(q,\dot{q}),\ldots\hat{X}^d(q,\dot{q}),0,\ldots,0),$$

where the last 0s are repeated $n-d$ times. The proposition follows from the fact that

$$\bar{X}_E(q,\dot{s},r) = T_{G^{-1}(q,\dot{s},r)}G\,\hat{X}_E \circ G^{-1}(q,\dot{s},r).$$

□

Referring to Figure 5.1, the constraint function on $\mathbb{R}^{n+d}\times\mathbb{R}^{n-d}$ is, by design,

$$\bar{c}(q,\dot{s},r) = (\hat{c}\circ G^{-1})(q,\dot{s},r) = r.$$

Therefore, an integral curve $(q(t),\dot{q}(t))$ of $\hat{X}_E$ in $\mathbb{R}^{2n}$ through the point $(q,\dot{q})$ must satisfy the constraint equations $\hat{c}^b(q(t),\dot{q}(t)) = \hat{Y}^b$ and consequently, for $G(q,\dot{q}) = (q,\dot{s},r)$, the curve $G(q(t),\dot{q}(t)) = (q(t),\dot{s}(t),r(t))$ must satisfy $r^b(t) = \hat{Y}^b$, where Y^b is chosen to be constant. The affine constraints are thus imposed by keeping r^b constant thereby replacing the abstract Y with concrete coordinates.

5.2.2 A First Order Integrator

To illustrate the construction of a Lagrange–d'Alembert integrator and to provide a blueprint for the Examples in Chapter 6, a first order integrator will be developed.

Let $Q=\mathbb{R}^n$ and $TQ=\mathbb{R}^n\times\mathbb{R}^n$. Since $(\mathbb{R}^n,\mathrm{id})$ forms a suitable coordinate chart for Q, all local objects will be written without the $\hat{}$. For example, $\hat{L}$ will be written L. The notation for c adapted coordinates will still be used.

Let $\mathcal{D}$ be a distribution on Q. Let U be open in Q and $\{X_i(q)\}_{i=1}^n$ be a local basis for T_qU such that $\{X_q(q)\}_{a=1}^d$ is a local basis of $\mathcal{D}_q$ for each $q\in U$. Let $\{\phi^i(q)\}_{i=1}^n$ be a local basis of T^*U dual to $\{X_i(q)\}_{i=1}^n$.

Let $L\colon TQ\to\mathbb{R}$ be a $\mathcal{D}$–regular Lagrangian on TU with Euler–Lagrange vector field $X_E(q,\dot{q}) = (\dot{q},\ddot{q})$.

Discrete Tangent Bundle

See Section 4.2 in Chapter (4) for the discretization of the tangent bundle.

Define $\psi(h, t, (q, \dot{q})) = q + t\dot{q}, \alpha^+(h) = h, \alpha^-(h) = 0$. Then $\partial_h^+(q, \dot{q}) = q + h\dot{q}$ and $\partial_h^-(q, \dot{q}) = q$. $\partial_h^{\pm}(q, \dot{q}) = (q + h\dot{q}, q) = (q^+, q^-)$ has inverse $\left(\partial_h^{\pm}\right)^{-1}(q^+, q^-) = (q^-, \frac{q^+ - q^-}{h}) = (q, \dot{q})$. A suitable discrete tangent bundle is then $P_h = TU \cap \left(\partial_h^{\pm}\right)^{-1}(U \times U)$, for $h > 0$.

Define R_t on $\bar{P}_h = G(P_h)$ by (recall the indexing from Equations (5.1) to (5.4))

$$R_t^a(q, \dot{s}, r) = \psi^a(h, t, G^{-1}(q, \dot{s}, r)) = q^a + t\dot{s}^a \tag{5.10}$$

and $R_t^c(q, \dot{s}, r)$, to satisfy the (possibly implicit) set of ordinary differential equations coming from the constraints:

$$c^b\left(R_t'(q, \dot{s}, r)\right) = r^b, \tag{5.11}$$

where,

$$R_t'(q, \ldots s, r) = \Big(R_t(q, \dot{s}, r), \dot{R}_t(q, \dot{s}, r)\Big), \tag{5.12}$$

$$\dot{R}_t(q, \dot{s}, r) = \frac{d}{dt} R_t(q, \dot{s}, r). \tag{5.13}$$

Note that c is used in Equation (5.11) since the curve $t \mapsto R_t(q, \dot{s}, r)$ is in $\mathbb{R}^n$ and therefore $t \mapsto R_t'(q, \dot{s}, r)$ is in $\mathbb{R}^n \times \mathbb{R}^n$. The equations in 5.11 are

$$\begin{aligned}
r^b &= c^b((R_t'(q, \ldots s, r)) \\
&= \phi_i^b(R_t(q, \dot{s}, r))\, \dot{R}_t^i(q, \dot{s}, r) \quad \text{definition of } c^b \\
&= \phi_a^b(R_t(q, \dot{s}, r))\, \dot{R}_t^a(q, \dot{s}, r) + \phi_c^b(R_t(q, \dot{s}, r))\, \dot{R}_t^c(q, \dot{s}, r) \quad \text{splitting the sum} \\
&= \phi_a^b(R_t(q, \dot{s}, r))\, \dot{s}^a + \phi_c^b(R_t(q, \dot{s}, r))\, \dot{R}_t^c(q, \dot{s}, r). \quad \text{Equation (5.10)}
\end{aligned}$$

Assuming that the matrix ϕ_c^b is invertible, with inverse θ_b^c, the above constraint equations can be solved for $\dot{R}_t^c$,

$$\dot{R}_t^c(q, \dot{s}, r) = \theta_b^c\left(R_t(q, \dot{s}, r)\right)(r^b - \phi_a^b(R_t(q, \dot{s}, r))\, \dot{s}^a).$$

If neither θ^c_b nor ϕ^b_a depend on q^b, then a solution to the equations is

$$R^c_t(q,\dot{s},r) = q^c + \int_0^t \theta^c_b(R_\tau(q,\dot{s},r))\left(r^b - \phi^b_a(R_\tau(q,\dot{s},r))\,\dot{s}^a\right)\,d\tau. \tag{5.14}$$

The curve segments $t \mapsto R_t(q,\dot{s},r), 0 \le t \le h$ take the place of $\psi(h,t,(q,\dot{q}))$ so that the discrete tangent bundle that is actually used is $\bar{P}_h = G(P_h)$, along with $\bar{\partial}^+_h(q,\dot{s},r) = R_h(q,\dot{s},r)$ and $\bar{\partial}^-_h(q,\dot{s},r) = q$.

Remark–Higher Order Integrators

The generalization to higher order integrators is achieved by taking ψ to have higher order contact with the exact flow, $F^{X_E}_t$. The constructions of the previous section are generalized by not explicitly computing R^a_t and $\dot{R}^a_t$.

For example, let Φ_h be any order r numerical method for approximating the flow of X_E with constant step size h. Define $\psi(h,t,(q,\dot{q})) = \tau_Q \circ \Phi_t(q,\dot{q}),\ \alpha^-(h) \le t \le \alpha^+(h)$. Φ_h might, for example, be an order r Runge–Kutta method or the Taylor polynomial for an integral curve of X_E truncated after the rth term.

Infinitesimal Variations

The derivatives of $\bar{\partial}^+_h(q,\dot{q}) = R_h(q,\dot{q})$ are computed directly from Equations (5.10) and (5.14), assuming that R^c_t can be explicitly calculated. Recall that this assumption is based on the fact that θ^c_b and ϕ^b_a do not depend on q^b. The arguments of θ^c_b and ϕ^b_a are $R_\tau(q,\dot{s},r)$ and are omitted for compactness. Also note that the primed indices run over the same numbers as the unprimed versions. See Equations (5.1) to (5.4) for the indexing.

$$\begin{aligned}
\frac{\partial R^a_h}{\partial q^i} &= \delta^a_i, \\
\frac{\partial R^a_h}{\partial \dot{s}^{a'}} &= h\delta^a_{a'}, \\
\frac{\partial R^a_h}{\partial r^b} &= 0, \\
\frac{\partial R^c_h}{\partial q^i} &= \delta^c_i + \int_0^h \left(\frac{\partial \theta^c_b}{\partial q^i}(r^b - \phi^b_a\,\dot{s}^a) - \theta^c_b\frac{\partial \phi^b_a}{\partial q^i}\,\dot{s}^a\right)\,d\tau,
\end{aligned}$$

$$\frac{\partial R_h^c}{\partial \dot{s}^{a'}} = \int_0^h \left(\frac{\partial \theta_b^c}{\partial \dot{s}^{a'}} (r^b - \phi_a^b \dot{s}^a) - \theta_b^c \frac{\partial \phi_a^b}{\partial \dot{s}^{a'}} \, \dot{s}^a - \theta_b^c \phi_{a'}^b \right) d\tau,$$
$$\frac{\partial R_h^c}{\partial r^{b'}} = \int_o^h \theta_{b'}^c \, d\tau.$$

The derivative of $\bar{\partial}_h^-(q, \dot{s}, r) = q$ is

$$\frac{\partial \left(\bar{\partial}_h^-\right)^i}{\partial q^{i'}} = \delta_{i'}^i,$$

and zero for all other partial derivatives.

Write $\delta v_k = (\delta q_k, \delta \dot{s}_k, \delta r_k)$. Then Equations (4.18), (4.19) and (4.20) are

$$D\bar{\partial}_h^-(q_k, \dot{s}_k, r_k)\,(\delta q_k, \delta \dot{s}_k, \delta r_k) = 0, \tag{5.15}$$

$$D\bar{\partial}_h^+(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\,(\delta q_{k+1}, \delta \dot{s}_{k+1}, \delta r_{k+1}) = 0, \tag{5.16}$$

$$D\bar{\partial}_h^+(q_k, \dot{s}_k, r_k)\,(\delta q_k, \delta \dot{s}_k, \delta r_k) = \\ D\bar{\partial}_h^-(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\,(\delta q_{k+1}, \delta \dot{s}_{k+1}, \delta r_{k+1}). \tag{5.17}$$

Using the computed derivatives, Equation (5.15) is

$$\delta q_k = 0. \tag{5.18}$$

Equation (5.16) is the following set of equations:

$$\delta_i^a \, \delta q_{k+1}^i + h \delta_{a'}^a \, \delta \dot{s}_{k+1}^{a'} = 0, \tag{5.19}$$

$$\frac{\partial R_h^c}{\partial q^i}(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\, \delta q_{k+1}^i + \frac{\partial R_h^c}{\partial \dot{s}^{a'}}(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\, \delta \dot{s}_{k+1}^{a'} + \\ + \frac{R_h^c}{\partial r^b}(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\, \delta r^b = 0. \tag{5.20}$$

Equation (5.17) is the following set of equations:

$$\delta_i^a \, \delta q_k^i + h \delta_{a'}^a \, \delta \dot{s}_k^{a'} = \delta q_{k+1}^a, \tag{5.21}$$

$$\frac{\partial R_h^c}{\partial q^i}(q_k, \dot{s}_k, r_k)\, \delta q_k^i + \frac{\partial R_h^c}{\partial \dot{s}^{a'}}(q_k, \dot{s}_k, r_k)\, \delta \dot{s}_k^{a'} + \frac{R_h^c}{\partial r^b}(q_k, \dot{s}_k, r_k)\, \delta r^b = \delta q_{k+1}^c. \tag{5.22}$$

In addition, there is the requirement of Equation (4.22)

$$D\bar{\partial}_h^+(q_k, \dot{s}, r)\,(\delta q_k, \delta \dot{s}_k, \delta r_k) \in \mathcal{D}.$$

From Equation (5.17) this is equivalent to

$$\delta q_{k+1} \in \mathcal{D}. \tag{5.23}$$

Explicit solutions for $(\delta q_k, \delta \dot{s}_k, \delta r_k)$ and $(\delta q_{k+1}, \delta \dot{s}_{k+1}, \delta r_{k+1})$ will be computed for each example in Chapter 6. It is enough to note, for now, that Equations (5.18) to (5.23) give a d parameter solution set.

Construction of L_h and α_h

The discrete Lagrangian, L_h and α_h are computed directly from their definitions in Equations (4.9) and (4.15) respectively. That is,

$$L_h(q, \dot{s}, r) = \frac{1}{h} \int_0^h L\left(R'_\tau(q, \dot{s}, r)\right) \, d\tau, \tag{5.24}$$

$$\alpha_h(q, \dot{s}, r)\,(\delta q, \delta \dot{s}, \delta r) = \frac{1}{h} \int_0^h F_c\left(R'_\tau(q, \dot{s}, r)\right) \, T_{(q,\dot{s},r)} R_h\,(\delta q, \delta \dot{s}, \delta r)\, d\tau, \tag{5.25}$$

where F_c is the constraint force and is usually calculated as

$$F_c(q, \dot{q}) = \lambda_b(q, \dot{q})\, \phi^b(q).$$

For an order r integrator it is only necessary to compute L_h and α_h to $O(h^{r-1})$ accuracy, as that is the best that can be expected for the order matching of the method. See Proposition 16 in Chapter 4.

Discrete Euler–Lagrange Equations

The discrete Euler–Lagrange Equations (4.17) are, recalling $\delta q_k = 0$,

$$\begin{aligned}
&\frac{\partial L_h}{\partial \dot{s}^a}(q_k, \dot{s}_k, r_k)\, \delta \dot{s}_k^a + \frac{\partial L_h}{\partial r^b}(q_k, \dot{s}_k, r_k)\, \delta r_k^b + \frac{\partial L_h}{\partial q^i}(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\, \delta q_{k+1}^i \\
&\quad + \frac{\partial L_h}{\partial \dot{s}^a}(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\, \delta \dot{s}_{k+1}^a + \frac{\partial L_h}{\partial r^b}(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\, \delta r_{k+1}^b \\
&\quad = \alpha_h(q_k, \dot{s}_k, r_k)\,(\delta q_k, \delta \dot{s}_k, \delta r_k) + \alpha_h(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\,(\delta q_{k+1}, \delta \dot{s}_{k+1}, \delta r_{k+1}).
\end{aligned}$$

Equations (4.21) and (4.23) are, respectively,

$$R_h(q_k, \dot{s}_k, r_k) = (q_{k+1}, \dot{s}_{k+1}, r_{k+1}),$$

$$c(q_k, \dot{s}_k, r_k) = 0, \qquad c(q_{k+1}, \dot{s}_{k+1}, r_{k+1}) = 0.$$

These equations are explicitly computed for the examples in Chapter 6.

5.2.3 Numerical Results

Short Time

Analytic solutions are compared to the evolution of a Lagrange–d'Alembert integrator. The term *short time* means just long enough to capture the behaviour of the solution. For example, if the exact solution is oscillatory, then *short time* means a few oscillations.

Long Time

In this section, the qualitative behaviour of the first order integrator developed above is investigated. This is specific to each system and will be demonstrated in Chapter 6.

Evolution of the Conserved Quantities

Constants of the motion are evaluated at each time step and their absolute error is computed and plotted. Since these quantities are preserved by the exact flow, it is desirable for the numerical method to also preserve them. Exact conservation is not observed in practice, but *near* conservation often is. That is, the error is bounded and depends on the time step.

Included in this section are holonomic constraints. The Lagrange–d'Alembert integrators developed in this thesis exactly preserve the constraint manifold, $\mathcal{D} \subset TQ$. If $\mathcal{D}$ is integrable, then the leaves of $\mathcal{D}$ are invariant under the exact flow. This invariance is tested for the Lagrange–d'Alembert integration of holonomically constrained systems.

Order Matching Demonstration

By choosing ψ to have various contact order with the exact solution, the theoretical method order predicted by Proposition 6 in Chapter 4 is verified. The order is estimated by the step–doubling method described in Section B.4.

5.2.4 Remarks

Every system has its own points of interest. Such elements of interest are collected here.

Chapter 6

Examples

6.1 Introduction

This chapter provides examples as a *proof of concept* for the theory in Chapters 3, and 4. A nonholonomic oscillator, a disk rolling on a circle and the two dimensional Kepler problem represent, respectively, nonholonomic, holonomic and unconstrained systems. The systems are treated identically as systems with linear velocity constraints. That is, the holonomic constraint is not used to eliminate variables in the disk on a circle example.

The results of the simulations substantiate the theory of this thesis by showing that Lagrange–d'Alembert integrators produce numerical approximations of similar quality to standard methods.

In addition, when compared to a standard method, Matlab's ode45, it is apparent that Lagrange–d'Alembert integrators exhibit far superior long term behaviour when computing geometric quantities associated to the physical system. For instance, energy remains bounded for the Lagrange–d'Alembert integrator whereas a linear in time error appears in the ode45 solution.

Each example is presented in the format introduced in Chapter 5.

The computer programs were written in C++ and used the ADOL-C algorithmic differentiation library [24].

6.2 Nonholonomic Oscillator

The nonholonomic oscillator appears in Cushman and Sniatycki [10] as a toy problem for the purpose of comparing various ways of generating the equations of motion. It serves as a good example for a Lagrange–d'Alembert integrator since it contains a conserved momentum.

6.2.1 Physical and Mathematical Analysis

Description of the Physical System and Mathematical Model

This example is not based on a physical system. It serves as a test example for the integrators developed in this thesis.

Configuration space is $Q = \mathbb{R}^4$ and phase space is $TQ = \mathbb{R}^4 \times \mathbb{R}^4$. Points in Q are denoted by $q = (x, y, z, w)$ and in TQ by $(q, \dot{q}) = (x, y, z, w, \dot{x}, \dot{y}, \dot{z}, \dot{w})$.

The Lagrangian for this system is

$$L(q, \dot{q}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2 + \dot{w}^2) - \frac{1}{2}(x^2 + y^2 + w^2).$$

The constraint distribution is

$$\mathcal{D} = \text{span}\{\partial_x, \partial_y, \partial_z + (1 + x^2 + y^2)\partial_w\}. \tag{6.1}$$

Local Description

Since $Q = \mathbb{R}^4$, a suitable coordinate chart is $(\mathbb{R}^4, \text{id})$. The basis of $\mathcal{D}$ given in Equation (6.1) is extended to a basis of TQ by

$$T_q Q = \text{span}\{\partial_x, \partial_y, \partial_z + (1 + x^2 + y^2)\partial_w, \partial_z\}.$$

T^*Q is spanned by the dual basis

$$T_q^* Q = \text{span}\{dx, dy, \frac{1}{1 + x^2 + y^2} dw, dz - \frac{1}{1 + x^2 + y^2} dw\}.$$

Finally the constraint codistribution is

$$\mathcal{D}^0 = \text{span}\{dz - \frac{1}{1+x^2+y^2} dw\}, \tag{6.2}$$

which induces the constraint function

$$c(q,\dot{q}) = \dot{z} - \frac{1}{1+x^2+y^2}\dot{w}. \tag{6.3}$$

Local Equations of Motion

The Euler–Lagrange equations, with Lagrange multiplier λ are

$$\ddot{x} = -x,$$

$$\ddot{y} = -y,$$

$$\ddot{z} = -\lambda,$$

$$\ddot{w} = \frac{\lambda}{1+x^2+y^2} - w.$$

Using $\frac{d}{dt}c(q,\dot{q}) = 0$, λ is computed as

$$\lambda = \frac{(1+x^2+y^2)w + 2\dot{w}(x\dot{x}+y\dot{y})}{1+(1+x^2+y^2)^2}, \tag{6.4}$$

so that the differential equations are

$$\ddot{x} = -x, \tag{6.5}$$

$$\ddot{y} = -y, \tag{6.6}$$

$$\ddot{z} = -\frac{(1+x^2+y^2)w + 2\dot{w}(x\dot{x}+y\dot{y})}{1+(1+x^2+y^2)^2},$$

$$\ddot{w} = \frac{-(1+x^2+y^2)^3 w + 2\dot{w}(x\dot{x}+y\dot{y})}{(1+(1+x^2+y^2)^2)(1+x^2+y^2)},$$

which gives the Euler–Lagrange vector field $X_E(q,\dot{q}) = (\dot{q},\ddot{q})$.

Conserved Quantities

The energy function,

$$E(q,\dot{q}) = \frac{1}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2+\dot{w}^2) + \frac{1}{2}(x^2+y^2+w^2) \tag{6.7}$$

is conserved. See Theorem 6 in Chapter 2.

The momentum,

$$J(q,\dot{q}) = x\dot{y} - y\dot{x} \tag{6.8}$$

is also conserved. Using Equations (6.5) and (6.6),

$$\begin{aligned}\frac{d}{dt}J\left(F_t^{X_E}(q,\dot{q})\right) &= dJ\left(F_t^{X_E}(q,\dot{q})\right)\, X_E\left((F_t^{X_E}(q,\dot{q})\right)\\ &= \dot{y}\dot{x} - \dot{x}\dot{y} + x\ddot{y} - y\ddot{x}\\ &= 0.\end{aligned}$$

Constraint Adapted Description

Make the coordinate change given by $G\colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^7 \times \mathbb{R}$,

$$G(x,y,z,w,\dot{x},\dot{y},\dot{z},\dot{w}) = \left(x,y,z,w,\dot{x},\dot{y},\dot{z},\dot{z} - \frac{1}{1+x^2+y^2}\dot{w}\right).$$

Writing coordinates for $\mathbb{R}^7 \times \mathbb{R}$ as $(q,\dot{s},r) = (x,y,z,w,\dot{x},\dot{y},\dot{z},r)$ gives the formula for r as

$$r = \dot{z} - \frac{1}{1+x^2+y^2}\dot{w},$$

and the inverse formula for $\dot{w}$ as

$$\dot{w} = (1+x^2+y^2)(\dot{z}-r).$$

The constraint function on $\mathbb{R}^7 \times \mathbb{R}$ is, by design,

$$\bar{c}(q,\dot{s},r) = r.$$

The equations of motion on $\mathbb{R}^7 \times \mathbb{R}$ are, using Proposition 18 of Chapter 5,

$$\bar{X}_E = TG\,X_E = \left(\dot{x},\dot{y},\dot{z},\frac{dw}{dt},-x,-y,\frac{d\dot{z}}{dt},\frac{dr}{dt}\right), \tag{6.9}$$

where,

$$\frac{dw}{dt} = (1+x^2+y^2)(\dot{z}-r), \tag{6.10}$$

$$\frac{d\dot{z}}{dt} = -\frac{(1+x^2+y^2)(w+2(\dot{z}-r)(x\dot{x}+y\dot{y}))}{1+(1+x^2+y^2)^2}, \tag{6.11}$$

$$\frac{dr}{dt} = 0. \tag{6.12}$$

6.2.2 A First Order Integrator

In this section, a first order Lagrange–d'Alembert integrator is developed.

Discrete Tangent Bundle

For $(q, \dot{q}) \in \mathbb{R}^4 \times \mathbb{R}^4$, define the curve segments in $\mathbb{R}^4$, $t \mapsto \psi(h, t, (q, \dot{q}))$ by $\psi(h, t, (q, \dot{q})) = q + t\dot{q}$, $0 \leq t \leq h$, for $h > 0$. In Section 4.2, this is taking $\alpha^+(h) = h$ and $\alpha^-(h) = 0$. The discrete tangent bundle will be in $\mathbb{R}^7 \times \mathbb{R}$ using the constraint adapted coordinates given by the mapping G.

For $(q, \dot{s}, r) \in \mathbb{R}^7 \times \mathbb{R}$, define the curve segments in $\mathbb{R}^4$, $t \mapsto R_t(q, \dot{s}, r)$ by

$$R_t(q, \dot{s}, r) = \left(R_t^1(q, \dot{s}, r), R_t^2(q, \dot{s}, r), R_t^3(q, \dot{s}, r), R_t^4(q, \dot{s}, r)\right).$$

where, for $a = 1, 2, 3$,

$$R_t^a(q, \dot{s}, r) = q^a + t\dot{s}^a, \tag{6.13}$$

and $R_t^4(q, \dot{s}, r)$ will be chosen so that that $c\left(R_t'(q, \dot{s}, r)\right) = r$ for $0 \leq t \leq h$, where the notation of Equations (5.12) and (5.13) in Chapter 5 is used. Explicitly,

$$\dot{R}_t^1(q, \dot{s}, r) = \dot{x}, \tag{6.14}$$

$$\dot{R}_t^2(q, \dot{s}, r) = \dot{y}, \tag{6.15}$$

$$\dot{R}_t^3(q, \dot{s}, r) = \dot{z}. \tag{6.16}$$

Using Equation (6.3), $c\left(R_t'(q, \dot{s}, r)\right) = r$ is, suppressing the arguments of R_t and $\dot{R}_t$,

$$\dot{R}_t^3 - \frac{1}{1 + (R_t^1)^2 + (R_t^2)^2} \frac{dR_t^4}{dt} = r,$$
$$\frac{dR_t^4}{dt} = \left(1 + (R_t^1)^2 + (R_t^2)^2\right)\left(\dot{R}_t^3 - r\right). \tag{6.17}$$

Using the definitions of $R_t^a(q, \dot{s}, r)$, for $a = 1, 2, 3$, and their derivatives, the following differential equation for $R_t^4(q, \dot{s}, r)$ is obtained:

$$\frac{dR_t^4}{dt} = (1 + (x + t\dot{x})^2 + (y + t\dot{y})^2)(\dot{z} - r). \tag{6.18}$$

In Equation (6.18), the $x, y, \dot{z}, r$ are held constant so that the differential equation is separable with solution

$$R_t^4(q, \dot{s}, r) = w + t(\dot{z} - r)(1 + x^2 + y^2) + t^2(\dot{z} - r)(x\dot{x} + y\dot{y}) + \frac{t^3}{3}(\dot{z} - r)(\dot{x}^2 + \dot{y}^2). \quad (6.19)$$

The discretization of $TQ = \mathbb{R}^4 \times \mathbb{R}^4$ of Section 4.2 is given by the tuple $(R_t, \alpha^+(h), \alpha^-(h))$, where $\alpha^+(h) = h$ and $\alpha^-(h) = 0$. R_t is, in fact, a mapping from $\mathbb{R}^7 \times \mathbb{R}$ to $Q = \mathbb{R}^4$ but $t \mapsto R_t'(q, \dot{s}, r)$ is a curve in TQ, ensuring that $t \mapsto R_t(q, \dot{s}, r)$ are suitable curve segments for the discretization of TQ.

On $\mathbb{R}^7 \times \mathbb{R}$, define

$$\partial_h^+(q, \dot{s}, r) = R_h(q, \dot{s}, r),$$

$$\partial_h^-(q, \dot{s}, r) = R_0(q, \dot{s}, r).$$

These are the mappings defined by Equations (4.1) and (4.2) in Chapter 4. An explicit calculation gives

$$\begin{aligned}\partial_h^+(q, \dot{s}, r) =&(x + h\dot{x}, y + h\dot{y}, z + h\dot{z}, \\ & w + h(\dot{z} - r)(1 + x^2 + y^2) + h^2(\dot{z} - r)(x\dot{x} + y\dot{y}) + \frac{h^3}{3}(\dot{z} - r)(\dot{x}^2 + \dot{y}^2), \end{aligned} \quad (6.20)$$

$$\partial_h^-(q, \dot{s}, r) = q. \quad (6.21)$$

Proposition 3 of Chapter 4 guarantees that there is a $P_h \subset \mathbb{R}^7 \times \mathbb{R}$ such that $\partial_h^\pm(q, \dot{s}, r) = (\partial_h^+(q, \dot{s}, r), \partial_h^-(q, \dot{s}, r))$ is a diffeomorphism of P_h to an open neighbourhood of $\Delta(\mathbb{R}^4 \times \mathbb{R}^4)$.

Infinitesimal Variations

The infinitesimal variations δv_k and δv_{k+1} are computed from Equations (4.18), (4.19) and (4.20). Writing $\delta v_k = (\delta q_k, \delta \dot{s}_k, \delta r_k)$, these are

$$T_{(q_k, \dot{s}_k, r_k)} \partial_h^- (\delta q_k, \delta \dot{s}_k, \delta r_k) = 0, \quad (6.22)$$

$$T_{(q_{k+1},\dot{s}_{k+1},r_{k+1})}\partial_h^+(\delta q_{k+1},\delta\dot{s}_{k+1},\delta r_{k+1}) = 0, \tag{6.23}$$

$$T_{(q_k,\dot{s}_k,r_k)}\partial_h^+(\delta q_k,\delta\dot{s}_k,\delta r_k) = T_{(q_{k+1},\dot{s}_{k+1},r_{k+1})}\partial_h^-(\delta q_{k+1},\delta\dot{s}_{k+1},\delta r_{k+1}). \tag{6.24}$$

In addition,

$$T_{(q_{k+1},\dot{s}_{k+1},r_{k+1})}\partial_h^-(\delta q_{k+1},\delta\dot{s}_{k+1},\delta r_{k+1}) \in \mathcal{D}. \tag{6.25}$$

Let $\alpha = \dot{z} - r$ and $\beta = h(1+x^2+y^2)+h^2(x\dot{x}+y\dot{y})+\frac{1}{3}h^3(\dot{x}^2+\dot{y}^2)$. The spacial derivative of $R_h(q,\dot{s},r)$ is

$$DR_h(q,\dot{s},r) =$$
$$\begin{bmatrix} 1 & 0 & 0 & 0 & h & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & h & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & h & 0 \\ \alpha h(2x+h\dot{x}) & \alpha h(2y+t\dot{y}) & 0 & 1 & \alpha h^2(x+\frac{2}{3}h\dot{x}) & \alpha h^2(y+\frac{2}{3}h\dot{y}) & \beta & -\beta \end{bmatrix}.$$

Compress notation further by defining a,b,c,d so that

$$DR_h(q,\dot{s},r) = \begin{bmatrix} 1 & 0 & 0 & 0 & h & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & h & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & h & 0 \\ a & b & 0 & 1 & c & d & \beta & -\beta \end{bmatrix}.$$

Expanding other notation, let $\delta q_k = (\delta x_k, \delta y_k, \delta z_k, \delta w_k)$ and $\delta\dot{s}_k = (\delta\dot{x}_k, \delta\dot{y}_k, \delta\dot{z}_k)$. Using Equation (6.21), Equation (6.22) implies

$$\delta q_k = 0. \tag{6.26}$$

Using Equation (6.20) and the derivative of $R_h(q,\dot{s},r)$, Equation (6.23) implies

$$\delta x_{k+1} + h\delta\dot{x}_{k+1} = 0, \tag{6.27}$$

$$\delta y_{k+1} + h\delta\dot{y}_{k+1} = 0, \tag{6.28}$$

$$\delta z_{k+1} + h\delta\dot{z}_{k+1} = 0, \tag{6.29}$$

$$a_{k+1}\delta x_{k+1} + b_{k+1}\delta y_{k+1} + \tag{6.30}$$

$$\delta w_{k+1} + c_{k+1}\delta \dot{x}_{k+1} + d_{k+1}\delta \dot{y}_{k+1} + \beta_{k+1}\delta \dot{z}_{k+1} - \beta_{k+1}\delta r_{k+1} = 0.$$

Equation (6.24) implies, using Equations (6.20), (6.21) and $\delta q_k = 0$,

$$\delta x_{k+1} = h\delta \dot{x}_k, \tag{6.31}$$

$$\delta y_{k+1} = h\delta \dot{y}_k, \tag{6.32}$$

$$\delta z_{k+1} = h\delta \dot{z}_k, \tag{6.33}$$

$$\delta w_{k+1} = c_k\delta \dot{x}_k + d_k\delta \dot{y}_k + \beta_k\delta \dot{z}_k - \beta_k\delta r_k. \tag{6.34}$$

The condition of Equation (6.25) is, using Equation (6.22), $\delta q_{k+1} \in \mathcal{D}$. This is realized by applying an element of $\mathcal{D}^0$, defined in Equation (6.2), to δq_{k+1}. Namely, let $\phi(q) = dz - dw/(1 + x^2 + y^2)$. Then $\phi(q)\,\delta q_{k+1} = 0$ is equivalent to $\delta q_{k+1} \in \mathcal{D}$. This is

$$(1 + x_{k+1}^2 + y_{k+1}^2)\delta z_{k+1} - \delta w_{k+1} = 0. \tag{6.35}$$

Equations (6.26) to (6.35) are 13 equations (Equation (6.26) counts for 4) in the 16 unknowns

$$(\delta x_k, \delta y_k, \delta z_k, \delta w_k, \delta \dot{x}_k, \delta \dot{y}_k, \delta \dot{z}_k, \delta r_k),$$

$$(\delta x_{k+1}, \delta y_{k+1}, \delta z_{k+1}, \delta w_{k+1}, \delta \dot{x}_{k+1}, \delta \dot{y}_{k+1}, \delta \dot{z}_{k+1}, \delta r_{k+1}).$$

The solution set is, using $\delta \dot{x}_k, \delta \dot{y}_k, \delta \dot{z}_k$ as the free parameters

$$\delta x_k = 0, \tag{6.36}$$

$$\delta y_k = 0, \tag{6.37}$$

$$\delta z_k = 0, \tag{6.38}$$

$$\delta w_k = 0, \tag{6.39}$$

$$\delta r_k = \frac{1}{\beta_k}\left(c_k\delta \dot{x}_k + d_k\delta \dot{y}_k + \beta_k\delta \dot{z}_k - (1 + x_{k+1}^2 + y_{k+1}^2)h\delta \dot{y}_k\right), \tag{6.40}$$

$$\delta x_{k+1} = h\delta \dot{x}_k, \tag{6.41}$$

$$\delta y_{k+1} = h\delta\dot{y}_k, \tag{6.42}$$

$$\delta z_{k+1} = h\delta\dot{z}_k, \tag{6.43}$$

$$\delta w_{k+1} = (1 + x_{k+1}^2 + y_{k+1}^2)h\delta\dot{y}_k, \tag{6.44}$$

$$\delta\dot{x}_{k+1} = -\delta\dot{x}_k, \tag{6.45}$$

$$\delta\dot{y}_{k+1} = -\delta\dot{y}_k, \tag{6.46}$$

$$\delta\dot{z}_{k+1} = -\delta\dot{z}_k, \tag{6.47}$$

$$\delta r_{k+1} = \frac{1}{\beta_{k+1}}\left((h\alpha_{k+1} - c_{k+1})\delta\dot{x}_k + (hb_{k+1} - d_{k+1} + h(1 + x^2 + y^2))\delta\dot{y}_k - \beta_{k+1}\delta\dot{z}_k\right). \tag{6.48}$$

Construction of L_h and α_h

The discrete Lagrangian, L_h and α_h are computed directly from their definitions in Equations (4.9) and (4.15) respectively. That is,

$$L_h(q,\dot{s},r) = \frac{1}{h}\int_0^h L\left(R'_\tau(q,\dot{s},r)\right)\,d\tau, \tag{6.49}$$

$$\alpha_h(q,\dot{s},r)\left(\delta q,\delta\dot{s},\delta r\right) = \frac{1}{h}\int_0^h F_c\left(R'_\tau(q,\dot{s},r)\right)\,T_{(q,\dot{s},r)}R_h\left(\delta q,\delta\dot{s},\delta r\right)d\tau, \tag{6.50}$$

where

$$L(q,\dot{q}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2 + \dot{w}^2) - \frac{1}{2}(x^2 + y^2 + w^2)$$

and

$$F_c(q,\dot{q}) = \lambda(q,\dot{q})\left(dz - \frac{1}{1 + x^2 + y^2}dw\right),$$

for the Lagrange multiplier, λ from Equation (6.4).

Since the integrator is to be first order, the exact integrals in Equations (6.49) and (6.50) need not be computed. A left endpoint approximation suffices. That is, take

$$L_h(q,\dot{s},r) = L(G^{-1}(q,\dot{s},r)) \tag{6.51}$$

$$= \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2 + (1 + x^2 + y^2)^2(\dot{z} - r)^2) - \frac{1}{2}(x^2 + y^2 + w^2) \tag{6.52}$$

and, using Equation (6.4) for λ,

$$\begin{aligned}\alpha_h(q,\dot{s},r)\,(\delta q,\delta\dot{s},\delta r) &= F_c(G^{-1}(q,\dot{s},r))\,(\delta q,\delta\dot{s},\delta r)\\ &= \lambda(G^{-1}(q,\dot{s},r))\left(dz - \frac{1}{1+x^2+y^2}dw\right)(\delta q,\delta\dot{s},\delta r)\\ &= \frac{(1+x^2+y^2)w + 2(1+x^2+y^2)(\dot{z}-r)(x\dot{x}+y\dot{y})}{1+(1+x^2+y^2)^2}\left(\delta z - \frac{1}{1+x^2+y^2}\delta w\right).\end{aligned} \tag{6.53}$$

Discrete Euler–Lagrange Equations

The discrete Euler–Lagrange Equations (4.17) in Chapter 4 are

$$dL_h(q_k,\dot{s}_k,r_k)\,(\delta q_k,\delta\dot{s}_k,\delta r_k) + dL_h(q_{k+1},\dot{s}_{k+1},r_{k+1})\,(\delta q_{k+1},\delta\dot{s}_{k+1},\delta r_{k+1}) =$$
$$\alpha_h(q_k,\dot{s}_k,r_k)\,(\delta q_k,\delta\dot{s}_k,\delta r_k) + \alpha_h(q_{k+1},\dot{s}_{k+1},r_{k+1})\,(\delta q_{k+1},\delta\dot{s}_{k+1},\delta r_{k+1}).$$

To obtain the explicit equations, each of $\delta\dot{x}_k$, $\delta\dot{y}_k$ and $\delta\dot{z}_k$ are taken to be 1 and the rest 0, in turn, so that three equations are generated. In addition, the constraint is $c(q,\dot{s},r) = r$ for r constant. To realize the constraint given in the original problem, set $r = 0$.

Using the L_h of Equation (6.52), the α_h of Equation (6.53) and the infinitesimal variations of Equations (6.36) to (6.48), the discrete Euler–Lagrange equations are

$$\dot{x}_k - (1+x_k^2+y_k^2)\frac{c_k\dot{z}_k}{\beta_k} + h(2(1+x_k^2+y_k^2)\dot{z}_{k+1}^2 - x_{k+1}) - \tag{6.54}$$
$$\dot{x}_{k+1} - (1+x_{k+1}^2+y_{k+1}^2)\dot{z}_{k+1}\frac{h\alpha_{k+1}-c_{k+1}}{\beta_{k+1}} = 0,$$

$$\dot{y}_k - \frac{h}{\beta_k}(1+x_k^2+y_k^2)^2\dot{z}_k(d_k - (1+x_{k+1}^2+y_{k+1}^2)) + \tag{6.55}$$
$$h(2(1+x_{k+1}^2+y_{k+1}^2)\dot{z}_{k+1}^2 y_{k+1} - y_{k+1}) - \dot{y}_{k+1} -$$
$$\frac{\dot{z}_k}{\beta_{k+1}}(1+x_{k+1}^2+y_{k+1}^2)(hb_{k+1} - d_{k+1} + h(1+x_{k+1}^2+y_{k+1}^2)\cdot$$
$$hw_{k+1}(1+x_{k+1}^2+y_{k+1}^2) = h\lambda_{k+1},$$

$$(1+x_k^2+y_k^2)^2\dot{z}_k + \dot{z}_{k+1} = -h\lambda_{k+1}. \tag{6.56}$$

There are also the Equations (4.21) from Chapter 4. These are $q_{k+1} = R_h(q_k, \dot{s}_k, r_k)$, which are, from Equations (6.13) and (6.19)

$$x_{k+1} = x_k + h\dot{x}_k, \tag{6.57}$$

$$y_{k+1} = y_k + h\dot{y}_k, \tag{6.58}$$

$$z_{k+1} = z_k + h\dot{z}_k, \tag{6.59}$$

$$w_{k+1} = w_k + h\dot{z}_k(1 + x_k^2 + y_k^2) + h^2\dot{z}_k(x_k\dot{x}_k + y_k\dot{y}_k) + \frac{h^3}{3}\dot{z}_k(\dot{x}_k^2 + \dot{y}_k^2). \tag{6.60}$$

To evolve the system, choose initial conditions for $k = 0$ and set $(q_k, \dot{s}_k, r_k) = (q_0, \dot{s}_0, r_0)$. Equations (6.54) to (6.60) are then 7 nonlinear equations in the 7 unknowns $(q_{k+1}, \dot{s}_{k+1}, r_{k+1}) = (q_1, \dot{s}_1, r_1)$. Of course, to satisfy the constraint, $r_0 = 0$ necessarily and $r_1 = 0$ is an 8th omitted equation. If $(q_0, \dot{s}_0, r_0)$ and the time step, h, are suitably chosen, then a solution for $(q_1, \dot{s}_1, r_1)$ is guaranteed by Proposition 4 in Chapter 4. If $(q_1, \dot{s}_1, r_1)$ and h are still suitable, then set $(q_k, \dot{s}_k, r_k) = (q_1, \dot{s}_1, r_1)$ and solve the equations again.

6.2.3 Numerical Results

The differential equations (6.9) to (6.12) have solution

$$x(t) = \cos t, \tag{6.61}$$

$$y(t) = \sin t, \tag{6.62}$$

$$z(t) = \frac{\sqrt{5}}{2}\sin\left(\frac{2}{\sqrt{5}}t\right), \tag{6.63}$$

$$w(t) = \sqrt{5}\sin\left(\frac{2}{\sqrt{5}}t\right), \tag{6.64}$$

$$\dot{x}(t) = -\sin t, \tag{6.65}$$

$$\dot{y}(t) = \cos t, \tag{6.66}$$

$$\dot{z}(t) = \cos\left(\frac{2}{\sqrt{5}}t\right), \tag{6.67}$$

$$r = 0, \tag{6.68}$$

through $(1, 0, 0, 0, 0, 1, 1, 0)$. This solution is bounded and oscillatory.

Short Time

A fourth order Lagrange–d'Alembert Method is used to produce a short term simulation of the nonholonomic oscillator with the initial conditions $(q, \dot{s}, r) = (1, 0, 0, 0, 0, 1, 1, 0)$. The simulation is run for a little longer than two complete oscillations and plotted with the exact solution given in Equations (6.61) to (6.68). The fourth order method is generated from a ψ that is the order 4 Taylor polynomial of the exact solution. That is,

$$\psi(h, h, (q, \dot{q})) = (q, \dot{q}) + h \left.\frac{d}{dt}\right|_{t=0} F_t^{X_E}(q, \dot{q}) + \frac{h^2}{2} \left.\frac{d^2}{dt^2}\right|_{t=0} F_t^{X_E}(q, \dot{q}) + \\ + \frac{h^3}{6} \left.\frac{d^3}{dt^3}\right|_{t=0} F_t^{X_E}(q, \dot{q}) + \frac{h^4}{24} \left.\frac{d^4}{dt^4}\right|_{t=0} F_t^{X_E}(q, \dot{q}).$$

The plot for r is omitted since it is exactly zero by design. See Figure 6.1

Long Time

Long time integration of the first order Lagrange–d'Alembert method given by Equations (6.54) to (6.60) is shown in Figure 6.2. The simulation was run long enough to obtain 20 000 oscillations of the x component. The figures show that the evolution remains bounded.

Figure 6.3 shows the last computed oscillation of the x component of displacement and a sine curve to show that after a long integration, the character of the numerical solution remains.

Evolution of the Conserved Quantities

Energy and momentum in Equations (6.7) and (6.8) are conserved by the exact flow. Figure 6.4 shows the long term energy and momentum error of the order one Lagrange–d'Alembert integrator. Errors in both quantities are bounded.

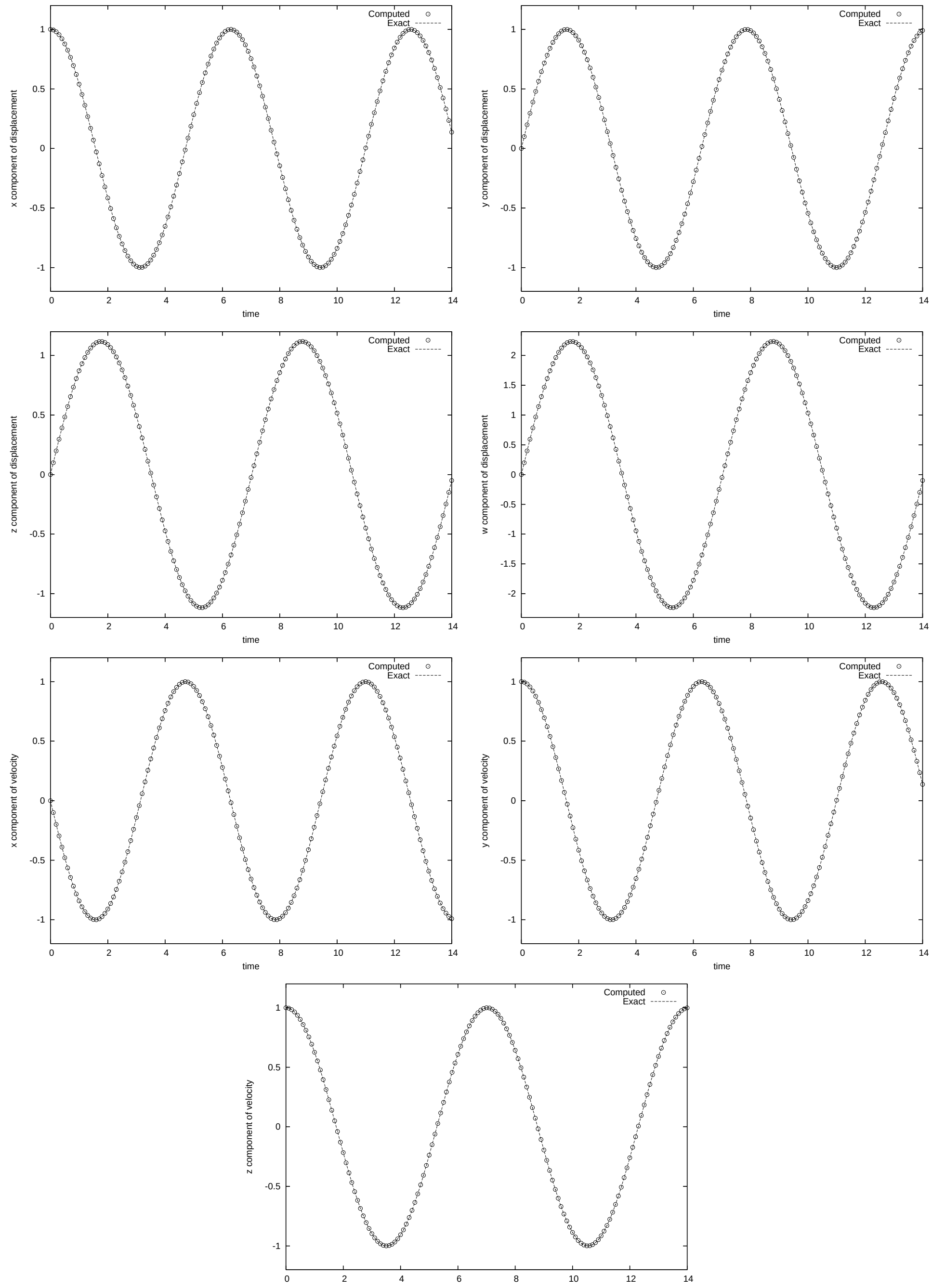

Figure 6.1: Lagrange–d'Alembert integration and exact solution of the components of the nonholonomic oscillator. Order 4 method. $h = 0.1$

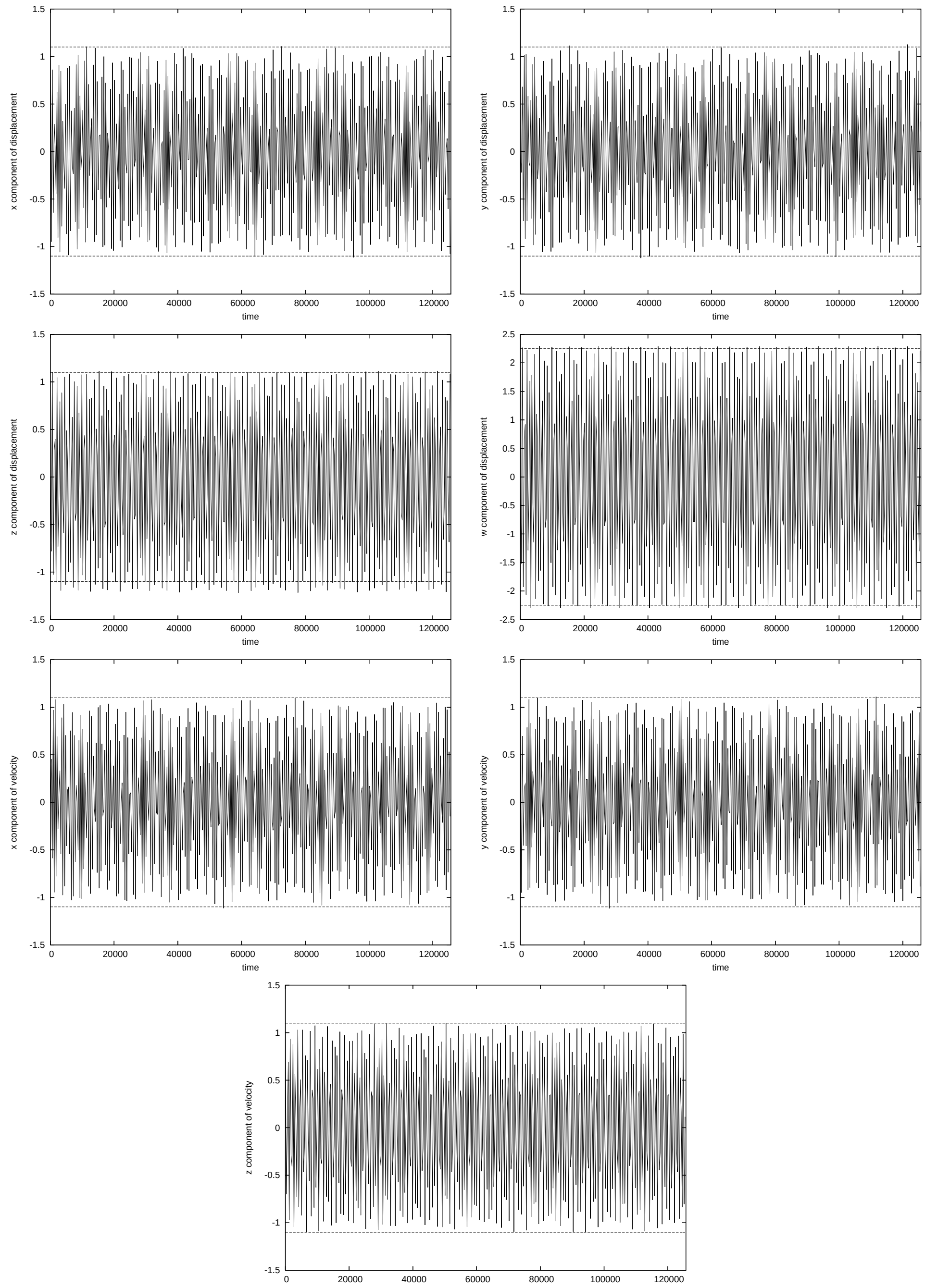

Figure 6.2: Long time Lagrange–d'Alembert integration of the components of the nonholonomic oscillator. Horizontal lines illustrate the upper and lower bounds of the oscillations. Order one method. $h = 0.1$. Every 1000 point plotted.

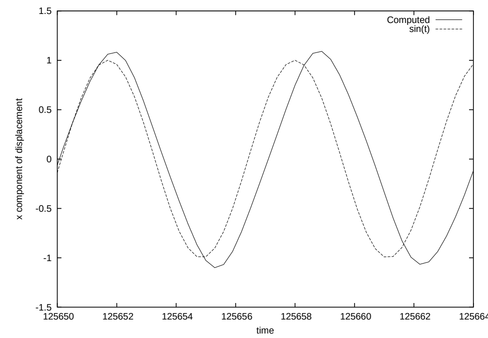

Figure 6.3: Tail end of the x component of the long time, first order Lagrange–d'Alembert integration of the nonholonomic oscillator. The sine curve is plotted in order to give a sense of scale to the amplitude and period of the oscillation.

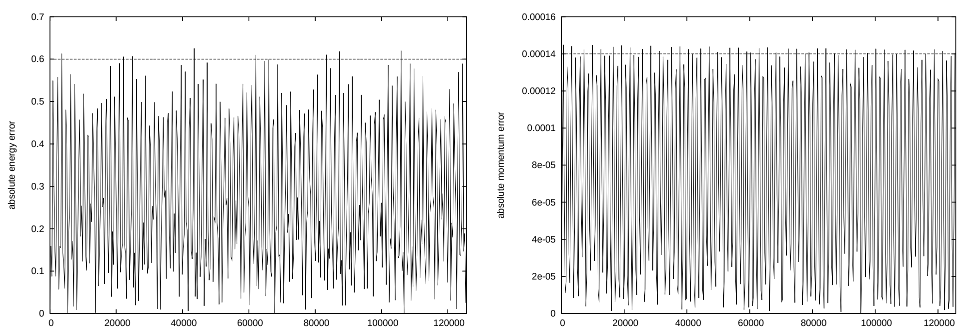

Figure 6.4: Energy and momentum error for the order one Lagrange–d'Alembert integrator. Horizontal lines illustrate the upper bound of the error. $h = 0.1$. Every 1000 point plotted

Order Matching

To generate examples showing that the residual order of Proposition 6 in Chapter 4 is obtained, the curve segments $t \mapsto \psi(h, t, (q, \dot{q}))$ of Chapter 5 is taken to be the Taylor series of the exact solution truncated at various orders, p. The method order is estimated using step doubling as outlined in Section (B.4). Table 6.1 gives the results of using first to fourth order approximations by Taylor series.

Step h	Error 1	Rate	Error 2	Rate	Error 3	Rate	Error 4	Rate
0.1								
0.05	2.50e-03		3.65e-06		3.65e-06		1.35e-09	
0.025	1.25e-03	1.00	4.54e-07	3.01	4.55e-07	3.00	4.26e-11	4.98
0.0125	6.24e-04	1.00	5.68e-08	3.00	5.69e-08	3.00	1.35e-12	4.98
0.00625	3.12e-04	1.00	7.10e-09	3.00	7.11e-09	3.00	4.00e-14	5.08

Table 6.1: Method order estimates for the x component of nonholonomic oscillator simulations. The column headings refer to the truncation order of the Taylor series. The exact solution has even symmetry, which is reflected in this table.

6.2.4 Remarks

The energy behaviour for the symplectic integration of holonomic systems is well understood. See Hairer, Lubich and Wanner [14], particularly Theorem 8.1 of Chapter IX. Roughly, this theorem states that, for $(q_n, \dot{q}_n)$ the nth iterate of an order r symplectic integration of a Hamiltonian system, $E(q_n, \dot{q}_n) = E(q_0, \dot{q}_o) + O(h^r)$. An integrator exhibiting this behaviour will have a bounded energy error. As shown in Chapter 3, a Lagrange–d'Alembert integrator is symplectic whenever the constraint is holonomic, so the theorem holds.

There are no similar theorems yet for Lagrange–d'Alembert integrators for nonholonomic systems, however, the graph of Energy error in Figure 6.4 provides some evidence that such a theorem is possible. The order of the energy error is presented in Table 6.2, further indicating that there may be a theorem.

Matlab's ode45 routine was used to evolve the same system in order to demonstrate the

Step h	Rate 1	Rate 2	Rate 3	Rate 4
0.1	0.98	2.17	3.00	4.11
0.05	0.99	2.09	3.00	4.06
0.025	1.00	2.02	3.00	4.03

Table 6.2: Energy error order estimates for nonholonomic oscillator Lagrange–d'Alembert simulations. The column headings refer to the truncation order of the Taylor series.

long term energy error for standard methods. Figure 6.5 shows the linear dependence on time of the absolute energy error. The absolute error is still small, however, the linear dependence indicates that the error will grow without bound. This is is contrast with the bounded energy error of the Lagrange–d'Alembert integrator. Even though the error is larger, the results in Table 6.2 indicate that using a high order method coupled with a suitably small step size will decrease the energy error.

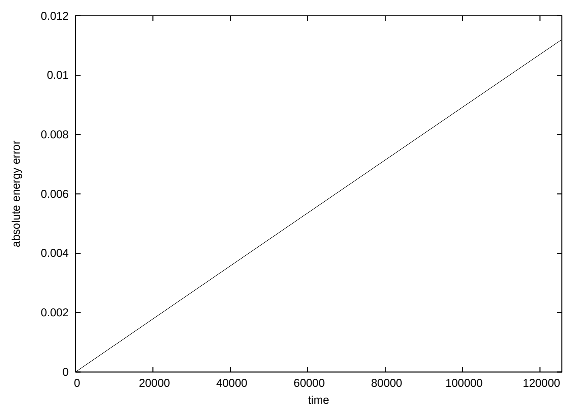

Figure 6.5: Matlab's ode45 long term energy error. Relative tolerance 10^{-12}. The magnitude of the error shows a linear dependence on time.

6.3 Disk Rolling on a Circle

This example is of a disk rolling on a circle without slipping. The no–slip rolling constraint for this system turns out to be holonomic. This example shows that the Lagrange–d'Alembert integrators naturally preserve the configuration constraints.

The configuration space and constraint are adaptations to two dimensional rigid bodies of the treatment of three dimensional rolling rigid bodies in Hermans [15]. This example is integrated on the complex unit circle and provides a future path to the integration of rolling rigid body systems on the set of unit quaternions. Both the complex circle and the unit quaternions are convenient configuration manifolds, as they require only one holonomic constraint to realize, whereas three and six constraints are necessary to realize the matrix Lie groups $SO(2)$ and $SO(3)$ respectively.

6.3.1 Facts About $SO(2)$

This section follows the treatment of $SO(3)$ in Bates and Cushman [4], Chapter 3.

On $\mathbb{R}^2$ with Euclidean inner product $(\cdot\,,\,\cdot)$, the special orthogonal group is the group of linear transformations that preserve the inner product and orientation. That is, $A \in SO(2)$ if and only if $(Ax, Ay) = (x, y)$ for all $x, y \in \mathbb{R}^2$.

The lie algebra, $so(2)$ is

$$T_e SO(2) = so(2) = \{X \in gl(2, \mathbb{R}) \,|\, X + X^t = 0\}.$$

The elements of $so(2)$ are 2×2 antisymmetric matrices. A basis for $so(2)$ is

$$E_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

where the subscript 3 on E_3 is a carryover from $SO(3)$. The lie bracket on $so(2)$ is the trivial one defined by $[E_3, E_3] = 0$.

Define the linear map

$$i \colon so(2) \to \mathbb{R},$$

$$i(X) = x,$$

where $X = xE_3$.

There is a positive definite metric, k, on $so(2)$ defined by

$$k(X,Y) = -\frac{1}{2}tr(XY).$$

That is, for $X = xE_3$ and $Y = yE_3$, $k(X,Y) = xy$.

The exponential map is defined

$$\exp\colon so(2) \to SO(2),$$
$$\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

Proposition 19. *Let* $X = \theta E_3 \in so(2)$*. Then* $\exp(X) = I\cos\theta + E_3 \sin\theta$*.*

Proof. First, note that

$$E_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, E_3^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, E_3^3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, E_3^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then,

$$\begin{aligned} \exp(X) &= \exp(\theta E_3) \\ &= \sum_k \frac{1}{k!}\theta^k E_3^k \\ &= I + \theta E_3 + \frac{1}{2}\theta^2 E_3^2 + \frac{1}{3!}\theta^3 E_3^3 + \frac{1}{4!}\theta^4 E_3^4 + \dots \\ &= I(1 - \frac{1}{2}\theta^2 + \frac{1}{4!}\theta^4 + \dots) + E_3(\theta - \frac{1}{3!}\theta^3 + \dots) \\ &= I\cos\theta + E_3\sin\theta. \end{aligned}$$

□

Proposition 20. *The mapping* $\exp\colon so(2) \to SO(2)$ *has the following properties:*

1. *If* $\exp(X) = A$*, then* $\exp(-X) = A^{-1}$*,*

2. $\exp$ *is a diffeomorphism from the open ball/interval* $B_\pi = \{X = \theta E_3 \in so(2) \mid \theta \in (-\pi, \pi)\}$ *onto its image,*

3. $\exp$ *is continuous.*

4. exp *is surjective from the closed ball* $\bar{B}_\pi = \{X = \theta E_3 \in so(2) \,|\, \theta \in [-\pi, \pi]\}$ *to* $SO(2)$*,*

Proof. 1. Since $SO(2)$ is abelian,

$$I = \exp(0) = \exp(X - X) = \exp(X)\exp(-X).$$

2. This follows directly from Proposition 19.

3. Proposition 19 implies that the mapping $\theta \mapsto \exp(\theta E_3)$ is differentiable and therefore continuous.

4. To be painfully explicit, write $A \in SO(2)$ as

$$A = \begin{bmatrix} A^1_1 & A^1_2 \\ A^2_1 & A^2_2 \end{bmatrix}.$$

The conditions on A are

$$A^1_1 A^1_2 + A^2_1 A^2_2 = 0,$$

$$(A^1_1)^2 + (A^2_1)^2 = 1,$$

$$(A^1_2)^2 + (A^2_2)^2 = 1.$$

These equations are satisfied for

$$A^1_1 = A^2_2 = \cos\theta, \quad -A^1_2 = A^2_1 = \sin\theta,$$

for some θ.

□

Let $\dot{A} \in T_A SO(2)$. Then there is a curve $A(t)$ in $SO(2)$ such that $A(0) = A$ and $\left.\frac{d}{dt}\right|_{t=0} A(t) = \dot{A}$. Then,

$$A^{-1}\dot{A} = \exp(-\theta E_3) \left.\frac{d}{dt}\right|_{t=0} \exp(\theta(t) E_3)$$

$$= \exp(-\theta E_3)\dot{\theta} E_3 \exp(\theta E_3)$$

$$= \dot{\theta} E_3,$$

which is in $so(2)$.

On $TSO(2)$, there is the left invariant metric κ,

$$\kappa_A(\dot{A}, \dot{B}) = k(A^{-1}\dot{A}, A^{-1}\dot{B})$$

6.3.2 Mapping $SO(2)$ into $\mathbb{D}$

Let

$$\mathbb{D} = \{z \in \mathbb{C} \,|\, |z| = 1\},$$

and

$$\eta\colon so(2) \to \mathbb{D},$$

$$\eta(\theta E_3) = e^{i\theta}.$$

Then, for θ such that $\exp(\theta E_3) = A$,

$$(\eta \circ \exp^{-1})(A) = e^{i\theta},$$

and, for $A(t)$ such that $A(0) = A$ and $\dot{A}(0) = \dot{A}$,

$$\begin{aligned} T_A(\eta \circ \exp^{-1})\, \dot{A} &= \left.\frac{d}{dt}\right|_{t=0} (\eta \circ \exp^{-1})A(t) \\ &= \left.\frac{d}{dt}\right|_{t=0} e^{i\theta(t)} \\ &= i\dot{\theta} e^{i\theta}, \end{aligned}$$

where $\exp(\theta(t)E_3) = A(t)$ and $\theta(0) = \theta$, $\dot{\theta}(0) = \dot{\theta}$. Further,

$$\begin{aligned} |T_A(\eta \circ \exp^{-1})\, \dot{A}| &= |i\dot{\theta} e^{i\theta}| \\ &= |\dot{\theta}|, \end{aligned}$$

and

$$\kappa_A(\dot{A}, \dot{A}) = k(A^{-1}\dot{A}, A^{-1}\dot{A})$$

$$= k(\dot{\theta}E_3, \dot{\theta}E_3)$$

$$= \dot{\theta}^2,$$

so

$$\sqrt{\kappa_A(\dot{A}, \dot{A})} = |T_A(\eta \circ \exp^{-1})\dot{A}|.$$

6.3.3 Description of the Physical System and Mathematical Model

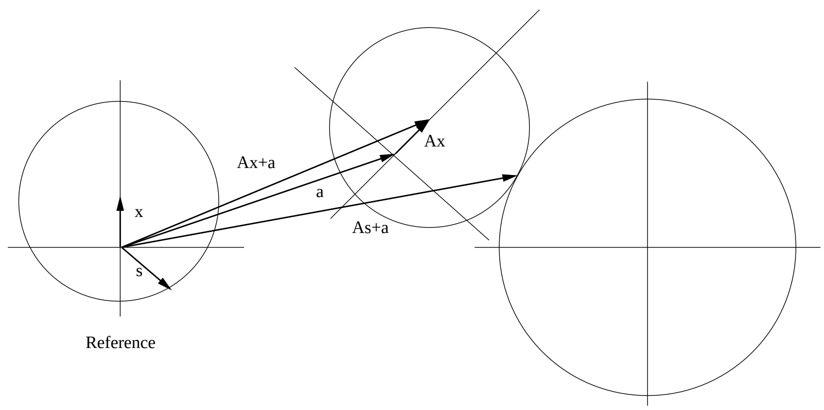

Figure 6.6: Configuration of a disk rolling on a circle

The configuration space is the Lie group of special Euclidean transformations of $\mathbb{R}^2$, $SE(2) = SO(2) \times \mathbb{R}^2$. $SE(2)$ acts on $\mathbb{R}^2$ on the left by rigid translations and rotations by $(A, \mathbf{a})\mathbf{q} = A\mathbf{q} + \mathbf{a}$.

See Figure 6.6. The *reference* position of the disk in $\mathbb{R}^2$ has the center of mass at the origin and the geometric center at the position $\mathbf{x} = x\mathbf{e}_2$. A configuration of the disk is identified with the element $(A, \mathbf{a}) \in SE(2)$ that takes the reference disk to the current configuration.

Phase space is $TSE(2) = TSO(2) \times T\mathbb{R}^2$. Let $(A(t), \mathbf{a}(t))$ be a curve in $SE(2)$ and $(\dot{A}(t), \dot{\mathbf{a}}(t))$ be the tangent to the curve. A point $\mathbf{q}$ in $\mathbb{R}^2$ traverses the curve $(A(t), \mathbf{a}(t))\mathbf{q} = A(t)\mathbf{q} + \mathbf{a}(t)$ with velocity $\dot{A}(t)\mathbf{q} + \dot{\mathbf{a}}(t)$.

Let I be the moment of inertia of the disk and m its mass and R_1 its radius. The

Lagrangian is

$$L(A, \mathbf{a}, \dot{A}, \dot{\mathbf{a}}) = \frac{I}{2}\kappa_A(\dot{A}, \dot{A}) + \frac{m}{2}(\dot{\mathbf{a}}, \dot{\mathbf{a}})$$
$$= \frac{I}{2}\dot{\theta}^2 + \frac{m}{2}(\dot{\mathbf{a}}, \dot{\mathbf{a}}).$$

The point, $\mathbf{s}$, on the reference disk goes to the contact point $A\mathbf{s} + \mathbf{a}$. The no slip rolling constraint requires that the velocity of the contact point relative to the circle be zero. In terms of the point $\mathbf{s}$, this is $\dot{A}\mathbf{s} + \dot{\mathbf{a}} = 0$. The outward normal to the disk at $\mathbf{s}$ is mapped to the inward normal of the circle by A, thereby supplying an equation for $\mathbf{s}$

$$\frac{1}{R_1}A(\mathbf{s} - \mathbf{x}) = -\frac{1}{R_2}(A\mathbf{s} + \mathbf{a}),$$

where R_2 is the radius of the circle. Solving for $\mathbf{s}$ gives

$$\mathbf{s} = \rho_2\mathbf{x} - \rho_1 A^{-1}\mathbf{a}, \quad \text{where } \rho_1 = \frac{R_1}{R_1 + R_2},\ \rho_2 = \frac{R_2}{R_1 + R_2}.$$

The no slip rolling constraint is then

$$\rho_2\dot{A}\mathbf{x} - \rho_1\dot{A}A^{-1}\mathbf{a} + \dot{\mathbf{a}} = 0.$$

Therefore, the constraint distribution is given by

$$\mathcal{D}_{(A,\mathbf{a})} = \left\{(\dot{A}, \dot{\mathbf{a}}) \,|\, \rho_2\dot{A}\mathbf{x} - \rho_1\dot{A}A^{-1}\mathbf{a} + \dot{\mathbf{a}} = 0\right\}.$$

This will be shown to be a holonomic constraint in the next section.

Local Description

The complex unit circle provides a suitable configuration manifold, with coordinates provided by $\mathbb{R}$ via the complex exponential, $\theta \mapsto e^{i\theta}$, $\theta \in [0, 2\pi)$. Coordinates for $Q = SE(2)$ are (θ, a^1, a^2) and for $TQ = TSE(2)$ are $(\theta, a^1, a^2, \dot{\theta}, \dot{a}_1, \dot{a}_2)$.

Using the coordinate change $\eta \circ \exp^{-1} \colon SO(2) \to \mathbb{D}^2$, the Lagrangian looks the same, so abusing notation, the Lagrangian on $T\mathbb{D} \times T\mathbb{R}^2$ is

$$L(\theta, \mathbf{a}, \dot{\theta}, \dot{\mathbf{a}}) = \frac{I}{2}\dot{\theta}^2 + \frac{m}{2}(\dot{\mathbf{a}}, \dot{\mathbf{a}}). \tag{6.69}$$

The constraint becomes

$$\rho_2 i\dot{\theta}e^{i\theta}\mathbf{x} - \rho_1 i\dot{\theta}\mathbf{a} + \dot{\mathbf{a}} = 0.$$

Since $\rho_1 + \rho_2 = 1$,

$$\begin{aligned}
&\rho_2 i\dot{\theta}e^{i\rho_2\theta}\mathbf{x} - \rho_1 i\dot{\theta}e^{-i\rho_1\theta}\mathbf{a} + e^{-i\rho_1\theta}\dot{\mathbf{a}} = 0,\\
&\frac{d}{dt}\left(e^{i\rho_2\theta}\mathbf{x} + e^{-i\rho_1\theta}\mathbf{a}\right) = 0,\\
&e^{i\rho_2\theta}\mathbf{x} + e^{-i\rho_1\theta}\mathbf{a} = e^{i\rho_2\theta_0}\mathbf{x} + e^{-i\rho_1\theta_0}\mathbf{a}_0,
\end{aligned}$$

where $\theta_0 = \theta(0)$ and $\mathbf{a}_0 = \mathbf{a}(0)$. The constraint is therefore holonomic. Solving for $\mathbf{a}$,

$$\mathbf{a} = -e^{i\theta}\mathbf{x} + e^{i\rho_1\theta}\boldsymbol{\alpha}_0, \qquad \text{where} \quad \boldsymbol{\alpha}_0 = e^{i\rho_2\theta_0}\mathbf{x} + e^{-i\rho_1\theta_0}\mathbf{a}_0 \tag{6.70}$$

and

$$\dot{\mathbf{a}} = -i\dot{\theta}e^{i\theta}\mathbf{x} + i\rho_1\dot{\theta}e^{i\rho_1\theta}\boldsymbol{\alpha}_0. \tag{6.71}$$

Let $\boldsymbol{\beta} = -ie^{i\theta}\mathbf{x} + i\rho_1 e^{i\rho_1\theta}\boldsymbol{\alpha}_0$. Then, $\dot{\mathbf{a}} = \dot{\theta}\boldsymbol{\beta}$. This is simply to compact the notation. Write β^1 and β^2 for the components of $\boldsymbol{\beta}$ and

$$\dot{\boldsymbol{\beta}} = \frac{d}{dt}\boldsymbol{\beta}, \qquad \boldsymbol{\beta}' = \frac{d}{d\theta}\boldsymbol{\beta}.$$

For explicit computations, use

$$e^{i\theta}\mathbf{y} = (y_1\cos\theta - y_2\sin\theta, y_1\sin\theta + y_2\cos\theta).$$

The constraint distribution is

$$\mathcal{D} = \operatorname{span}\{\partial_\theta + \beta^1\partial_{a^1} + \beta^2\partial_{a^2}\}.$$

The codistribution is

$$\mathcal{D}^0 = \operatorname{span}\{da^1 - \beta^1 d\theta, da^2 - \beta^2 d\theta\}.$$

A basis for TQ is

$$TQ = \operatorname{span}\{\partial_\theta + \beta^1\partial_{a^1} + \beta^2\partial_{a^2}, \partial a^1, \partial a^2\},$$

with dual basis for T^*Q

$$T^*Q = \text{span}\{d\theta, da^1 - \beta^1 d\theta, da^2 - \beta^2 d\theta\}.$$

The constraint functions are

$$c^1(\theta, a^1, a^2, \dot{\theta}, \dot{a}_1, \dot{a}_2) = \dot{a}_1 - \dot{\theta}\beta^1,$$

$$c^2(\theta, a^1, a^2, \dot{\theta}, \dot{a}_1, \dot{a}_2) = \dot{a}_2 - \dot{\theta}\beta^2.$$

The Euler–Lagrange equations, with Lagrange multipliers λ_1, λ_2, are

$$\ddot{\theta} = \frac{1}{I}\lambda_1\beta^1 + \frac{1}{I}\lambda_2\beta^2, \tag{6.72}$$

$$\ddot{a}_1 = -\frac{1}{m}\lambda_1, \tag{6.73}$$

$$\ddot{a}_2 = -\frac{1}{m}\lambda_2. \tag{6.74}$$

$$\tag{6.75}$$

Using the derivative of c^1 and c^2 and Equations (6.72) to (6.74),

$$\lambda_1 = -m\dot{\theta}\frac{I\dot{\beta}^1 + m(\beta^2)^2\dot{\beta}^1 - m\beta^1\beta^2\dot{\beta}^2}{I + m(\beta^1)^2 + m(\beta^2)^2}, \tag{6.76}$$

$$\lambda_2 = -m\dot{\theta}\frac{I\dot{\beta}^2 + m(\beta^1)^2\dot{\beta}^2 - m\beta^1\beta^2\dot{\beta}^1}{I + m(\beta^1)^2 + m(\beta^2)^2}. \tag{6.77}$$

The Euler–Lagrange equations are

$$\frac{d\theta}{dt} = \dot{\theta},$$

$$\frac{da^1}{dt} = \dot{a}_1,$$

$$\frac{da^2}{dt} = \dot{a}_2,$$

$$\frac{d\dot{\theta}}{dt} = -m\dot{\theta}\frac{\beta^1\dot{\beta}^1 + \beta^2\dot{\beta}^2}{I + m(\beta^1)^2 + m(\beta^2)^2},$$

$$\frac{d\dot{a}^1}{dt} = \dot{\theta}\frac{I\dot{\beta}^1 + m(\beta^2)^2\dot{\beta}^1 - m\beta^1\beta^2\dot{\beta}^2}{I + m(\beta^1)^2 + m(\beta^2)^2},$$

$$\frac{d\dot{a}^2}{dt} = \dot{\theta}\frac{I\dot{\beta}^2 + m(\beta^1)^2\dot{\beta}^2 - m\beta^1\beta^2\dot{\beta}^1}{I + m(\beta^1)^2 + m(\beta^2)^2}.$$

Alternatively, since the constraints are integrable, substitute Equation (6.71) into the Lagrangian (6.69) to eliminate the constraint. This gives

$$L(\theta,\dot\theta) = \frac{I}{2}\dot\theta^2 + \frac{m}{2}\dot\theta^2(\beta,\beta).$$

Using this Lagrangian, the Euler–Lagrange equations are

$$\begin{aligned}
\frac{d\theta}{dt} &= \dot\theta,\\
\frac{d\dot\theta}{dt} &= \frac{m\dot\theta^2(\boldsymbol{\beta}',\boldsymbol{\beta}) - 2m\dot\theta(\dot{\boldsymbol{\beta}},\boldsymbol{\beta})}{I + m(\boldsymbol{\beta},\boldsymbol{\beta})}.
\end{aligned}$$

This version of the Euler–Lagrange equations will not be used, as the purpose of this example is to treat the disk on a circle as a constrained system in order to demonstrate that a Lagrange–d'Alembert integrator can preserve the holonomic constraints automatically.

Conserved Quantities

Energy is the only conserved quantity. From Equation (2.37),

$$\begin{aligned}
E(\theta,\dot\theta) &= \Theta_L(\theta,\dot\theta)\, X_E(\theta,\dot\theta) - L(\theta,\dot\theta)\\
&= \frac{\partial L}{\partial\dot\theta}\,\dot\theta - L(\theta,\dot\theta)\\
&= \frac{I}{2}\dot\theta^2 + \frac{m}{2}(\dot a,\dot a).
\end{aligned} \tag{6.78}$$

Constraint Adapted Description

Define $G\colon \mathbb{R}^3\times\mathbb{R}^3 \to \mathbb{R}^4\times\mathbb{R}^2$ by

$$G(\theta,a^1,a^2,\dot\theta,\dot a_1,\dot a_2) = (\theta,a^1,a^2,\dot\theta,\dot a_1 - \dot\theta\beta^1,\dot a_2 - \dot\theta\beta^2).$$

Writing coordinates for $\mathbb{R}^4\times\mathbb{R}^2$ as $(q,\dot s,r) = (\theta,a^1,a^2,\dot\theta,r^1,r^2)$ gives the formula for r as

$$r^1 = \dot a_1 - \dot\theta\beta^1,$$

$$r^2 = \dot a_2 - \dot\theta\beta^2.$$

The equations of motion on $\mathbb{R}^4 \times \mathbb{R}^2$, using Theorem 18 of Chapter 5,

$$\bar{X}_E = TG\, X_E,$$

to give

$$\frac{d\theta}{dt} = \dot{\theta}, \tag{6.79}$$
$$\frac{da^1}{dt} = \dot{a}^1, \tag{6.80}$$
$$\frac{da^2}{dt} = \dot{a}^2, \tag{6.81}$$
$$\frac{d\dot{\theta}}{dt} = -m\dot{\theta}\frac{\beta^1\dot{\beta}^1 + \beta^2\dot{\beta}^2}{I + m(\beta^1)^2 + m(\beta^2)^2}, \tag{6.82}$$
$$\frac{dr^1}{dt} = 0, \tag{6.83}$$
$$\frac{dr^2}{dt} = 0. \tag{6.84}$$

6.3.4 A First Order Integrator

In this section, a first order Lagrange–d'Alembert integrator is developed.

Discrete Tangent Bundle

For $(q, \dot{q}) \in \mathbb{R}^3 \times \mathbb{R}^3$, define the curve segments in $\mathbb{R}^3$, $t \mapsto \psi(h, t, (q, \dot{q}))$ by $\psi(h, t, (q, \dot{q})) = q + t\dot{q}$, $0 \le t \le h$, for $h > 0$. In Section 4.2, this is taking $\alpha^+(h) = h$ and $\alpha^-(h) = 0$. The discrete tangent bundle will be in $\mathbb{R}^4 \times \mathbb{R}^2$ using the constraint adapted coordinates given by the mapping G.

For $(q, \dot{s}, r) \in \mathbb{R}^4 \times \mathbb{R}^2$, define the curve segments in $\mathbb{R}^3$, $t \mapsto R_t(q, \dot{s}, r)$ by

$$R_t(q, \dot{s}, r) = (\theta + t\dot{\theta}, R_t^3(q, \dot{s}, r), R_t^2(q, \dot{s}, r)),$$

where R_t^2 and R_t^3 will be chosen so that $c(R_t') = r$ for $0 \le t \le h$ where the notation of Equations (5.12) and (5.13) in Chapter 5 is used. The arguments of R_t are being suppressed in order to compress notation. Explicitly,

$$\dot{R}_t^2 - \dot{R}_t^1\beta^1 = r^1, \qquad \dot{R}_t^3 - \dot{R}_t^1\beta^2 = r^2,$$

where $R_t^1(q,\dot{s},r) = \theta + t\dot{\theta}$. The differential equations to solve are then

$$\frac{dR_t^2}{dt} = r^1 + \dot{\theta}\beta^1,$$
$$\frac{dR_t^3}{dt} = r^2 + \dot{\theta}\beta^2.$$

These equations are separable, so

$$R_t^2(q,\dot{s},r) = a^1 + r^1 t + \dot{\theta}\int_0^t \beta^1(\theta + \tau\dot{\theta})\, d\tau,$$
$$R_t^3(q,\dot{s},r) = a^2 + r^2 t + \dot{\theta}\int_0^t \beta^2(\theta + \tau\dot{\theta})\, d\tau.$$

For more notation compression, define $\bar{\boldsymbol{\beta}} = 1/\dot{\theta}(-e^{i(\theta+h\dot{\theta})}\mathbf{x} + ie^{i\rho_1(\theta+h\dot{\theta})}\boldsymbol{\alpha}_0)$. Then,

$$R_h^2(q,\dot{s},r) = a^1 + r^1 h + \dot{\theta}\bar{\beta}^1,$$
$$R_h^3(q,\dot{s},r) = a^2 + r^2 h + \dot{\theta}\bar{\beta}^2.$$

The discretization of $TQ = \mathbb{R}^3 \times \mathbb{R}^3$ of Section 4.2 is given by the tuple $(R_t, \alpha^+(h), \alpha^-(h))$, where $\alpha^+(h) = h$ and $\alpha^-(h) = 0$. R_t is, in fact, a mapping from $\mathbb{R}^4 \times \mathbb{R}^2$ to $Q = \mathbb{R}^3$ but $t \mapsto R_t'(q,\dot{s},r)$ is a curve in TQ, ensuring that $t \mapsto R_t(q,\dot{s},r)$ are suitable curve segments for the discretization of TQ.

On $\mathbb{R}^4 \times \mathbb{R}^2$, define

$$\partial_h^+(q,\dot{s},r) = R_h(q,\dot{s},r),$$
$$\partial_h^-(q,\dot{s},r) = R_0(q,\dot{s},r).$$

These are the mappings defined by Equations (4.1) and (4.2) in Chapter 4. An explicit calculation gives

$$\partial_h^+(q,\dot{s},r) = (\theta + h\dot{\theta}, a^1 + r^1 h + \dot{\theta}\bar{\beta}^1, a^2 + r^2 h + \dot{\theta}\bar{\beta}^2), \tag{6.85}$$

$$\partial_h^-(q,\dot{s},r) = q. \tag{6.86}$$

Proposition 3 of Chapter 4 guarantees that there is a $P_h \subset \mathbb{R}^4 \times \mathbb{R}^2$ such that $\partial_h^\pm(q,\dot{s},r) = (\partial_h^+(q,\dot{s},r), \partial_h^-(q,\dot{s},r))$ is a diffeomorphism of P_h to an open neighbourhood of $\Delta(\mathbb{R}^3 \times \mathbb{R}^3)$.

Infinitesimal Variations

The infinitesimal variations δv_k and δv_{k+1} are computed from Equations (4.18), (4.19) and (4.20). Writing $\delta v_k = (\delta q_k, \delta \dot{s}_k, \delta r_k)$, these are

$$T_{(q_k,\dot{s}_k,r_k)}\partial_h^-\,(\delta q_k, \delta\dot{s}_k, \delta r_k) = 0, \tag{6.87}$$

$$T_{(q_{k+1},\dot{s}_{k+1},r_{k+1})}\partial_h^+\,(\delta q_{k+1}, \delta\dot{s}_{k+1}, \delta r_{k+1}) = 0, \tag{6.88}$$

$$T_{(q_k,\dot{s}_k,r_k)}\partial_h^+\,(\delta q_k, \delta\dot{s}_k, \delta r_k) = T_{(q_{k+1},\dot{s}_{k+1},r_{k+1})}\partial_h^-\,(\delta q_{k+1}, \delta\dot{s}_{k+1}, \delta r_{k+1}). \tag{6.89}$$

In addition,

$$T_{(q_{k+1},\dot{s}_{k+1},r_{k+1})}\partial_h^-\,(\delta q_{k+1}, \delta\dot{s}_{k+1}, \delta r_{k+1}) \in \mathcal{D}. \tag{6.90}$$

The spatial derivative of $R_h(q, \dot{s}, r)$ is

$$DR_h(q,\dot{s},r) = \begin{bmatrix} 1 & 0 & 0 & h & 0 & 0 \\ \frac{\partial\bar{\beta}^1}{\partial\theta} & 1 & 0 & \frac{\partial\bar{\beta}^1}{\partial\dot{\theta}} + \bar{\beta}^1 & h & 0 \\ \frac{\partial\bar{\beta}^2}{\partial\theta} & 1 & 0 & \frac{\partial\bar{\beta}^2}{\partial\dot{\theta}} + \bar{\beta}^2 & 0 & h \end{bmatrix}.$$

Expanding some notation, let $\delta q_k = (\delta\theta_k, \delta a_k^1, \delta a_k^2)$, $\delta\dot{s}_k = \delta\dot{\theta}_k$ and $\delta r_k = (\delta r_k^1, \delta r_k^2)$. Using Equation (6.86), Equation (6.87) implies

$$\delta q_k = 0. \tag{6.91}$$

Using Equation (6.85), Equation (6.88) implies

$$\delta\theta_{k+1} + h\delta\dot{\theta}_{k+1} = 0, \tag{6.92}$$

$$\frac{\partial\bar{\beta}^1_{k+1}}{\partial\theta}\delta\theta_{k+1} + \delta a^1_{k+1} + \left(\frac{\partial\bar{\beta}^1_{k+1}}{\partial\dot{\theta}} + \bar{\beta}^1_{k+1}\right)\delta\dot{\theta}_{k+1} + h\delta r^1_{k+1} = 0, \tag{6.93}$$

$$\frac{\partial\bar{\beta}^2_{k+1}}{\partial\theta}\delta\theta_{k+1} + \delta a^2_{k+1} + \left(\frac{\partial\bar{\beta}^2_{k+1}}{\partial\dot{\theta}} + \bar{\beta}^2_{k+1}\right)\delta\dot{\theta}_{k+1} + h\delta r^2_{k+1} = 0. \tag{6.94}$$

Equation (6.89) implies, using Equations (6.85), (6.86) and $\delta q_k = 0$,

$$\delta\theta_{k+1} = h\delta\dot{\theta}_k, \tag{6.95}$$

$$\delta a^1_{k+1} = \left(\frac{\partial\bar{\beta}^1_k}{\partial\dot{\theta}} + \bar{\beta}^1_k\right)\delta\dot{\theta}_k + h\delta r^1_k, \tag{6.96}$$

$$\delta a^2_{k+1} = \left(\frac{\partial \bar{\beta}^2_k}{\partial \dot{\theta}} + \bar{\beta}^2_k \right) \delta \dot{\theta}_k + h \delta r^2_k. \tag{6.97}$$

Let $\phi^1(q) = da^1 - \beta^1 d\theta$ and $\phi^2(q) = da^2 - \beta^2 d\theta$. These form a spanning set for $\mathcal{D}^0$. Then, the condition in Equation (6.90) is, using Equation (6.89),

$$\delta a^1_{k+1} - \beta^1_{k+1} \delta \theta_{k+1} = 0, \tag{6.98}$$

$$\delta a^2_{k+1} - \beta^2_{k+1} \delta \theta_{k+1} = 0. \tag{6.99}$$

Equations (6.91) to (6.99) are 11 equations (Equation (6.91) counts for 3) in the 12 unknowns

$$\begin{gathered} \left(\delta \theta_k, \delta a^1_k, \delta a^2_k, \delta \dot{\theta}_k, \delta r^1_k, \delta r^2_k \right), \\ \left(\delta \theta_{k+1}, \delta a^1_{k+1}, \delta a^2_{k+1}, \delta \dot{\theta}_{k+1}, \delta r^1_{k+1}, \delta r^2_{k+1} \right). \end{gathered}$$

The solution set is, using $\delta \dot{\theta}_k$ as the free parameter,

$$\delta \theta_k = 0, \tag{6.100}$$

$$\delta a^1_k = 0, \tag{6.101}$$

$$\delta a^2_k = 0, \tag{6.102}$$

$$\delta r^1_k = \beta^1_{k+1} \delta \dot{\theta}_k - \frac{1}{h} \left(\frac{\partial \bar{\beta}^1_k}{\partial \dot{\theta}} + \bar{\beta}^1_k \right) \delta \dot{\theta}_k, \tag{6.103}$$

$$\delta r^2_k = \beta^2_{k+1} \delta \dot{\theta}_k - \frac{1}{h} \left(\frac{\partial \bar{\beta}^2_k}{\partial \dot{\theta}} + \bar{\beta}^2_k \right) \delta \dot{\theta}_k, \tag{6.104}$$

$$\delta \theta_{k+1} = h \delta \dot{\theta}_k, \tag{6.105}$$

$$\delta a^1_{k+1} = h \beta^1_{k+1} \delta \dot{\theta}_k, \tag{6.106}$$

$$\delta a^2_{k+1} = h \beta^2_{k+1} \delta \dot{\theta}_k, \tag{6.107}$$

$$\delta \dot{\theta}_{k+1} = -\delta \dot{\theta}_k, \tag{6.108}$$

$$\delta r^1_{k+1} = -\beta^1_{k+1} \delta \dot{\theta}_k - \frac{\partial \bar{\beta}^1_{k+1}}{\partial \theta} \delta \dot{\theta}_k + \frac{1}{h} \left(\frac{\partial \bar{\beta}^1_k}{\partial \dot{\theta}} + \bar{\beta}^1_k \right) \delta \dot{\theta}_k, \tag{6.109}$$

$$\delta r^2_{k+1} = -\beta^2_{k+1} \delta \dot{\theta}_k - \frac{\partial \bar{\beta}^2_{k+1}}{\partial \theta} \delta \dot{\theta}_k + \frac{1}{h} \left(\frac{\partial \bar{\beta}^2_k}{\partial \dot{\theta}} + \bar{\beta}^2_k \right) \delta \dot{\theta}_k. \tag{6.110}$$

Construction of L_h and α_h

The discrete Lagrangian, L_h and α_h are computed directly from their definitions in Equations (4.9) and (4.15) respectively. That is,

$$L_h(q, \dot{s}, r) = \frac{1}{h}\int_0^h L\left(R'_\tau(q, \dot{s}, r)\right)\, d\tau, \tag{6.111}$$

$$\alpha_h(q, \dot{s}, r)\,(\delta q, \delta\dot{s}, \delta r) = \frac{1}{h}\int_0^h F_c\left(R'_\tau(q, \dot{s}, r)\right)\, T_{(q,\dot{s},r)}R_h\,(\delta q, \delta\dot{s}, \delta r)\, d\tau, \tag{6.112}$$

where

$$L(q, \dot{q}) = \frac{I}{2}\dot{\theta}^2 + \frac{m}{2}(\dot{\mathbf{a}}, \dot{\mathbf{a}})$$

and

$$F_c(q, \dot{q}) = \lambda_1(q, \dot{q})(da^1 - \beta^1 d\theta) + \lambda_2(q, \dot{q})(da^2 - \beta^2 d\theta),$$

with λ_1 and λ_2 the Lagrange multipliers from Equations (6.76) and (6.77).

Since the integrator is to be first order, the exact integrals in Equations (6.111) and (6.112) need not be computed. A left endpoint approximation suffices. That is,

$$\begin{aligned} L_h(q, \dot{s}, r) &= L(G^{-1}(q, \dot{s}, r)) \\ &= \frac{I}{2}\dot{\theta}^2 + \frac{m}{2}\dot{\theta}^2(\boldsymbol{\beta}, \boldsymbol{\beta}) \end{aligned} \tag{6.113}$$

and, using Equations (6.76) and (6.77) for λ_1 and λ_2,

$$\begin{aligned} \alpha_h(q, \dot{s}, r)(\delta q, \delta\dot{s}, \delta r) &= F_c(G^{-1}(q, \dot{s}, r))\,(\delta q, \delta\dot{s}, \delta r) \\ &= \lambda_1(G^{-1}(q, \dot{s}, r))(da^1 - \beta^1 d\theta) + \lambda_2(G^{-1}(q, \dot{s}, r))(da^2 - \beta^2 d\theta) \\ &= -m\dot{\theta}\frac{I\dot{\beta}^1 + m(\beta^2)^2\dot{\beta}^1 - m\beta^1\beta^2\dot{\beta}^2}{I + m(\beta^1)^2 + m(\beta^2)^2}(da^1 - \beta^1 d\theta) + \\ &\qquad - m\dot{\theta}\frac{I\dot{\beta}^2 + m(\beta^1)^2\dot{\beta}^2 - m\beta^1\beta^2\dot{\beta}^1}{I + m(\beta^1)^2 + m(\beta^2)^2}(da^2 - \beta^2 d\theta). \end{aligned} \tag{6.114}$$

Discrete Euler–Lagrange Equations

The discrete Euler–Lagrange Equations (4.17) in Chapter 4 are

$$dL_h(q_k, \dot{s}_k, r_k)\,(\delta q_k, \delta\dot{s}_k, \delta r_k) + dL_h(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\,(\delta q_{k+1}, \delta\dot{s}_{k+1}, \delta r_{k+1}) =$$

$$\alpha_h(q_k, \dot{s}_k, r_k)\,(\delta q_k, \delta \dot{s}_k, \delta r_k) + \alpha_h(q_{k+1}, \dot{s}_{k+1}, r_{k+1})\,(\delta q_{k+1}, \delta \dot{s}_{k+1}, \delta r_{k+1}).$$

To obtain the explicit equations, $\delta\dot{\theta}_k$ is set to 1. In addition, the constraint $c(q, \dot{s}, r) = r$ for r constant. To realize the constraint given in the original problem, set $r = 0$.

Using the L_h of Equation (6.113), the α_h of Equation (6.114) and the infinitesimal variations of Equations (6.100) to (6.110), the discrete Euler–Lagrange equation is

$$\begin{aligned} I\dot{\theta}_k + m(\boldsymbol{\beta}_k, \boldsymbol{\beta}_k) + m\dot{\theta}_{k+1}^2(\boldsymbol{\beta}'_{k+1}, \boldsymbol{\beta}_{k+1})h - I\dot{\theta}_{k+1} + m(\boldsymbol{\beta}_{k+1}, \boldsymbol{\beta}_{k+1}) = \\ - m\dot{\theta}_{k+1}\frac{I\dot{\beta}^1_{k+1} + m(\beta^2_{k+1})^2\dot{\beta}^1_{k+1} - m\beta^1_{k+1}\beta^2_{k+1}\dot{\beta}^2_{k+1}}{I + m(\beta^1_{k+1})^2 + m(\beta^2_{k+1})^2}(h\beta^1_{k+1} - \beta^1_{k+1}h) + \\ - m\dot{\theta}_{k+1}\frac{I\dot{\beta}^2_{k+1} + m(\beta^1_{k+1})^2\dot{\beta}^2_{k+1} - m\beta^1_{k+1}\beta^2_{k+1}\dot{\beta}^1_{k+1}}{I + m(\beta^1_{k+1})^2 + m(\beta^2_{k+1})^2}(h\beta^2_{k+1} - \beta^2_{k+1}h). \end{aligned} \tag{6.115}$$

There are also the Equations (4.21) from Chapter 4. These are $q_{k+1} = R_h(q_k, \dot{s}_k, r_k)$, which are, from Equation (6.85)

$$\theta_{k+1} = \theta_k + h\dot{\theta}_k, \tag{6.116}$$

$$a^1_{k+1} = a^1_k + r^1_k h + \dot{\theta}_k\bar{\beta}^1_k, \tag{6.117}$$

$$a^2_{k+1} = a^2_k + r^2_k h + \dot{\theta}_k\bar{\beta}^2_k. \tag{6.118}$$

To evolve the system, choose initial conditions for $k = 0$ and set $(q_k, \dot{s}_k, r_k) = (q_0, \dot{s}_0, r_0)$. Equations (6.54) to (6.60) are then 4 nonlinear equations in the 4 unknowns $(q_{k+1}, \dot{s}_{k+1}, r_{k+1}) = (q_1, \dot{s}_1, r_1)$. Of course, to satisfy the constraint, $r_0 = 0$ necessarily and $r_1 = 0$ contains a 5th and 6th omitted equation. If $(q_0, \dot{s}_0, r_0)$ and the time step, h, are suitably chosen, then a solution for $(q_1, \dot{s}_1, r_1)$ is guaranteed by Theorem 4 in Chapter 4. If $(q_1, \dot{s}_1, r_1)$ and h are still suitable, then set $(q_k, \dot{s}_k, r_k) = (q_1, \dot{s}_1, r_1)$ and solve the equations again.

6.3.5 Numerical Results

For a sample solution of Equations (6.79) to (6.84), take $\mathbf{x} = 0.9$ and the disk and circle radius to be 1. See Figure 6.6. Matlab' ode45 routine is used with a relative error tolerance of

10^{-12} to produce a trajectory very close to the exact solution for the short time simulation.

Short Time

Initial conditions are $(\theta, a^1, a^2, \dot{\theta}, r^1, r^2) = (0, 0, 1.1, -1, 0, 0)$. The simulation is run for a little longer than two complete oscillations and plotted with the Matlab ode45 solution. As can be seen in Figure 6.7, the two solutions are indistinguishable. The plots for r^1 and r^2 are omitted since they are exactly zero by design.

The time step for the Lagrange–d'Alembert integrator is $h = 0.05$.

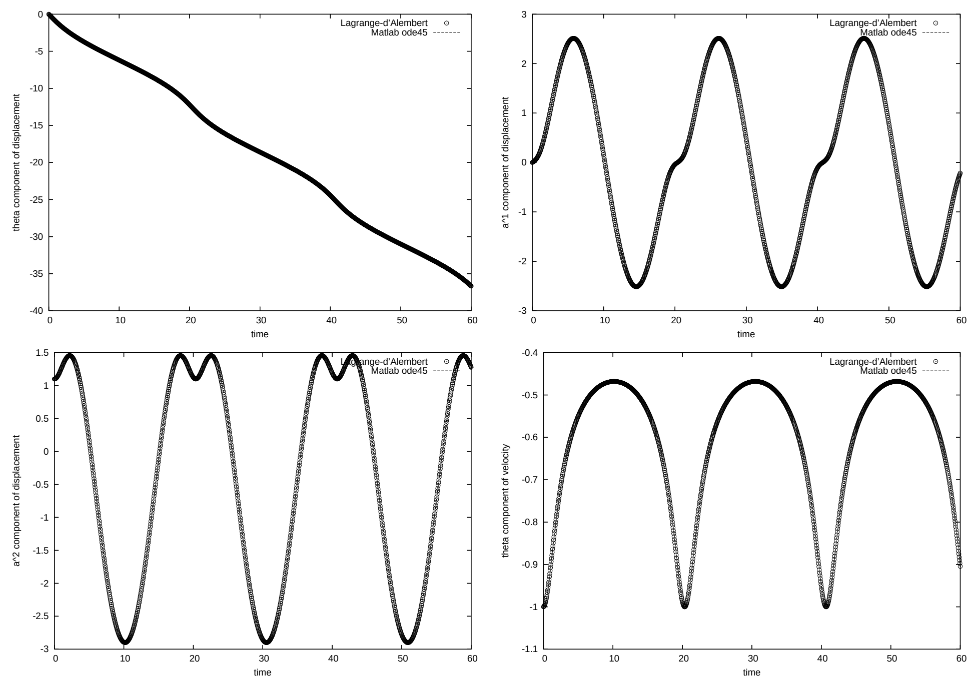

Figure 6.7: Lagrange–d'Alembert integration and Matlab ode45 solution components of the disk on a circle.

Long Time

Long time integration of the first order Lagrange–d'Alembert method given by Equations (6.79) to (6.82) is shown in Figure 6.8. The figures show that all but the θ component of displacement remains bounded.

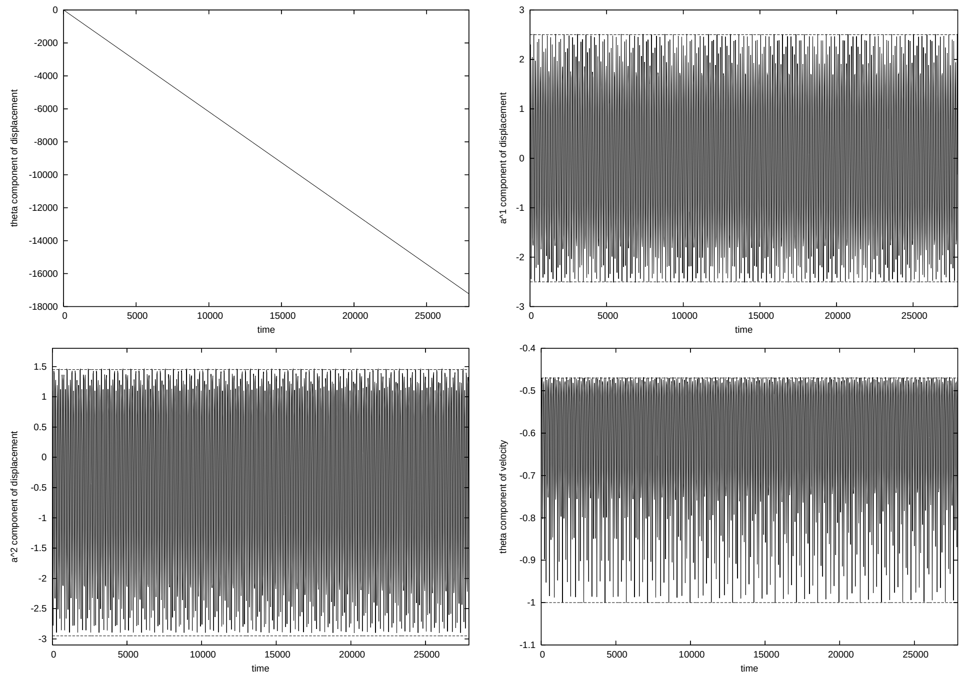

Figure 6.8: Long time Lagrange–d'Alembert integration of the components of the disk on a circle. Horizontal lines represent the upper and lower bounds of the plots. Order one method. $h = 0.05$. Every 500 point plotted.

Figure 6.9 shows the last computed oscillation of the x component of displacement to show that even after a long integration the character of the numerical solution remains.

Evolution of the Conserved Quantities

Energy, Equation (6.78), is the only conserved quantity. Figure 6.10 shows the long term energy error. The error is bounded.

The holonomic constraints, Equations (6.71) are also conserved by the exact flow. Figure 6.11 shows the long term error in the constraints produced by the order one Lagrange–d'Alembert method. The error is within the tolerance of the nonlinear equation solver used to solve the discrete Euler–Lagrange equations.

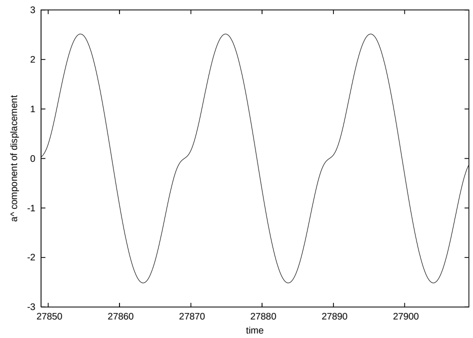

Figure 6.9: Tail end of the x component of the long time, first order Lagrange–d'Alembert integration of the disk on a circle.

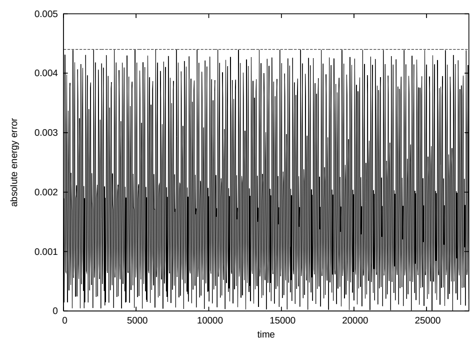

Figure 6.10: Energy error for the order one Lagrange–d'Alembert integrator. The horizontal line is the upper bound of the error. $h = 0.05$. Every 500 point plotted.

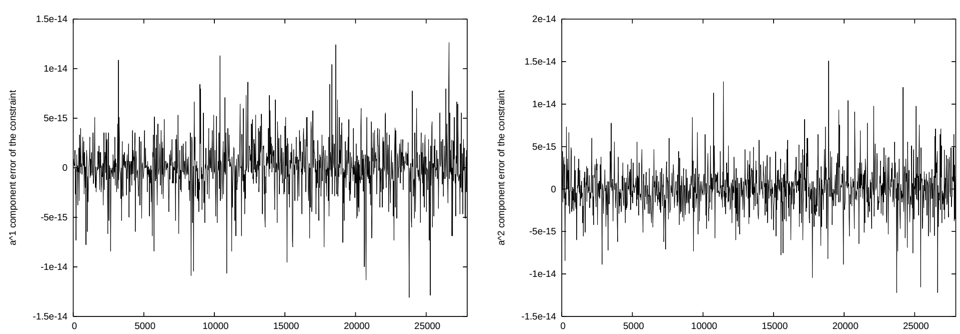

Figure 6.11: Holonomic constraint error for the order 1 Lagrange–d'Alembert method. Every 500 point plotted.

6.4 Kepler System

The evolution of the Kepler system is unconstrained in phase space $\mathbb{R}^4$ and has analytic solutions.

6.4.1 Physical and Mathematical Analysis

Description of the Physical System and Mathematical Model

The Kepler system is a two body problem where the bodies attract each other according to an inverse square force law.

The Kepler problem is treated as the equivalent one body problem subject to the potential $V(q) = -\frac{k}{|q|}$. See Goldstein, Poole and Safko [13] for the details.

The configuration space is $Q = \mathbb{R}^2$ with points represented by $q = (x, y)$. Phase space is $TQ = \mathbb{R}^2 \times \mathbb{R}^2$ with points represented by $(q, \dot{q}) = (x, y, \dot{x}, \dot{y})$.

The Lagrangian for this system is

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{\sqrt{x^2 + y^2}}.$$

Local Description

Since $Q = \mathbb{R}^2$, a suitable coordinate chart is $(\mathbb{R}^2, \mathrm{id})$.

Local Equations of Motion

The Euler–Lagrange equations are

$$\frac{dx}{dt} = \dot{x}, \tag{6.119}$$

$$\frac{dy}{dt} = \dot{y}, \tag{6.120}$$

$$\frac{d\dot{x}}{dt} = -\frac{x}{(x^2 + y^2)^{\frac{3}{2}}}, \tag{6.121}$$

$$\frac{d\dot{y}}{dt} = -\frac{y}{(x^2+y^2)^{\frac{3}{2}}}. \tag{6.122}$$

The Euler–Lagrange vector field is therefore

$$X_E(q,\dot{q}) = \left(\dot{x}, \dot{y}, -\frac{x}{(x^2+y^2)^{\frac{3}{2}}}, -\frac{y}{(x^2+y^2)^{\frac{3}{2}}}\right). \tag{6.123}$$

Conserved Quantities

The energy function,

$$E(q,\dot{q}) = \frac{1}{2}(\dot{x}^2+\dot{y}^2) - \frac{1}{\sqrt{x^2+y^2}} \tag{6.124}$$

is conserved. See Theorem 6 in Chapter 2.

There is a momentum

$$J(x,y,\dot{x},\dot{y}) = x\dot{y} - y\dot{x}. \tag{6.125}$$

Write the flow of the Euler–Lagrange vector field X_E (Equation (6.123)) as $F_t^{X_E}$. Then $F_t^{X_E}$ preserves J since

$$\begin{aligned}\frac{d}{dt}J(F_t^{X_E}(q,\dot{q}) =& dJ(F_t^{X_E}(q,\dot{q})\, X_E(F_t^{X_E}(q,\dot{q})\\ &= \dot{y}\dot{x} - \dot{x}\dot{y} - y\ddot{x} + x\ddot{y}\\ &= 0,\end{aligned}$$

where Equations (6.121) and (6.122) were used and $\ddot{q} = \frac{d\dot{q}}{dt}$.

There is another conserved quantity, the *eccentricity* or *Laplace–Runge–Lenz* vector

$$\mathbf{A} = (\dot{y}, -\dot{x})J - \frac{1}{\sqrt{x^2+y^2}}(x,y). \tag{6.126}$$

Differentiating along $F_t^{X_E}(q,\dot{q})$ and using the fact that $\frac{d}{dt}J = 0$ gives

$$\begin{aligned}\frac{d}{dt}\mathbf{A}(F_t^{X_E}(q,\dot{q})) &= (\ddot{y}, -\ddot{x})J(F_t^{X_E}(q,\dot{q})) + (\dot{y}, -\dot{x})\frac{d}{dt}J(F_t^{X_E}(q,\dot{q})) +\\ &\qquad + \frac{x\dot{x}+y\dot{y}}{(x^2+y^2)^{3/2}}(x,y) - \frac{1}{\sqrt{x^2+y^2}}(\dot{x},\dot{y})\\ &= \frac{x\dot{y}-y\dot{x}}{(x^2+y^2)^{3/2}}(-y,x) + \frac{x\dot{x}+y\dot{y}}{(x^2+y^2)^{3/2}}(x,y) - \frac{x^2+y^2}{(x^2+y^2)^{3/2}}(\dot{x},\dot{y})\\ &= 0.\end{aligned}$$

6.4.2 A First Order Integrator

In this section, a first order Lagrange–d'Alembert integrator is developed.

Discrete Tangent Bundle

For $(q, \dot{q}) \in \mathbb{R}^2 \times \mathbb{R}^2$, define the curve segments in $\mathbb{R}^2$, $t \mapsto \psi(h, t, (q, \dot{q}))$ by $\psi(h, t, (q, \dot{q})) = q + t\dot{q}$, $0 \le t \le h$, for $h > 0$. In Section 4.2, this is taking $\alpha^+(h) = h$ and $\alpha^-(h) = 0$. The discrete tangent bundle will be in $\mathbb{R}^2 \times \mathbb{R}^2$.

To compress notation, write $R_t(q, \dot{q}) = \psi(h, t, (q, \dot{q}))$. On $\mathbb{R}^2 \times \mathbb{R}^2$, define

$$\partial_h^+(q, \dot{q}) = R_h(q, \dot{q}),$$

$$\partial_h^-(q, \dot{q}) = R_0(q, \dot{q}).$$

These are the mappings defined by Equations (4.1) and (4.2) in Chapter 4. An explicit calculation gives

$$\partial_h^+(q, \dot{q}) = (x + h\dot{x}, y + h\dot{y}), \tag{6.127}$$

$$\partial_h^-(q, \dot{q}) = q. \tag{6.128}$$

Infinitesimal Variations

The infinitesimal variations δv_k and δv_{k+1} are computed from Equations (4.18), (4.19) and (4.20). Writing $\delta v_k = (\delta q_k, \delta \dot{q}_k)$, these are

$$T_{(q_k, \dot{q}_k)} \partial_h^- (\delta q_k, \delta \dot{q}_k) = 0, \tag{6.129}$$

$$T_{(q_{k+1}, \dot{q}_{k+1})} \partial_h^+ (\delta q_{k+1}, \delta \dot{q}_{k+1}) = 0, \tag{6.130}$$

$$T_{(q_k, \dot{q}_k)} \partial_h^+ (\delta q_k, \delta \dot{q}_k) = T_{(q_{k+1}, \dot{q}_{k+1})} \partial_h^- (\delta q_{k+1}, \delta \dot{q}_{k+1}). \tag{6.131}$$

The spatial derivative of $R_h(q, \dot{q})$ is

$$DR_h(q, \dot{q}) = \begin{bmatrix} 1 & 0 & h & 0 \\ 0 & 1 & 0 & h \end{bmatrix}.$$

Expanding notation, let $\delta q_k = (\delta x_k, \delta y_k)$ and $\delta \dot{q}_k = (\delta \dot{x}_k, \delta \dot{y}_k)$. Using Equation (6.128), Equation (6.129) implies

$$\delta q_k = 0. \tag{6.132}$$

Using Equation (6.127), Equation (6.130) implies

$$\delta x_{k+1} + h\delta \dot{x}_{k+1} = 0, \tag{6.133}$$

$$\delta y_{k+1} + h\delta \dot{y}_{k+1} = 0. \tag{6.134}$$

Equation (6.131) implies, using Equations (6.127) and (6.128) and $\delta q_k = 0$,

$$\delta x_{k+1} = h\delta \dot{x}_k, \tag{6.135}$$

$$\delta y_{k+1} = h\delta \dot{y}_k. \tag{6.136}$$

Equations (6.132) to (6.136) are 6 equations (Equation (6.132) counts for 2) in the 8 unknowns

$$(\delta x_k, \delta y_k, \delta \dot{x}_k, \delta \dot{y}_k),$$

$$(\delta x_{k+1}, \delta y_{k+1}, \delta \dot{x}_{k+1}, \delta \dot{y}_{k+1}).$$

The solution set is, using $\delta \dot{x}_k$ and $\delta \dot{y}_k$ as the free parameters,

$$\delta x_k = 0, \tag{6.137}$$

$$\delta y_k = 0, \tag{6.138}$$

$$\delta x_{k+1} = h\delta \dot{x}_k, \tag{6.139}$$

$$\delta y_{k+1} = h\delta \dot{y}_k, \tag{6.140}$$

$$\delta \dot{x}_{k+1} = -\delta \dot{x}_k, \tag{6.141}$$

$$\delta \dot{y}_{k+1} = -\delta \dot{y}_k. \tag{6.142}$$

Construction of L_h

The discrete Lagrangian, L_h is computed directly from its definition in Equations (4.9), That is,

$$L_h(q, \dot{s}, r) = \frac{1}{h} \int_0^h L\left(R'_\tau(q, \dot{s}, r)\right) \, d\tau, \tag{6.143}$$

where

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{\sqrt{x^2 + y^2}}.$$

Since the integrator is to be first order, the exact integral in Equation (6.143) need not be computed. A left endpoint approximation suffices. That is, take

$$L_h(q, \dot{q}) = L(q, \dot{q}) \tag{6.144}$$

$$= \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{\sqrt{x^2 + y^2}}. \tag{6.145}$$

Discrete Euler–Lagrange Equations

The discrete Euler–Lagrange Equations (4.17) in Chapter 4 are

$$dL_h(q_k, \dot{q}_k)\,(\delta q_k, \delta \dot{q}_k) + dL_h(q_{k+1}, \dot{q}_{k+1})\,(\delta q_{k+1}, \delta \dot{q}_{k+1}) = 0.$$

To obtain the explicit equations, $\delta \dot{x}_k$ and $\delta \dot{y}_k$ are each set to 1 in turn, and the other to 0. Using the L_h of Equation (6.145) and the infinitesimal variations of Equations (6.137) to (6.142), the discrete Euler–Lagrange equations are

$$\dot{x}_k + h\frac{x_{k+1}}{(x_{k+1}^2 + y_{k+1}^2)^{3/2}} - \dot{x}_{k+1} = 0, \tag{6.146}$$

$$\dot{y}_k + h\frac{y_{k+1}}{(x_{k+1}^2 + y_{k+1}^2)^{3/2}} - \dot{y}_{k+1} = 0. \tag{6.147}$$

There are also the Equations (4.21) from Chapter 4. These are $q_{k+1} = R_h(q_k, \dot{q}_k)$, which are, from Equation (6.127),

$$x_{k+1} = x_k + h\dot{x}_k, \tag{6.148}$$

$$y_{k+1} = y_k + h\dot{y}_k. \qquad (6.149)$$

To evolve the system, choose initial conditions for $k = 0$ and set $(q_k, \dot{q}_k) = (q_0, \dot{q}_0)$. Equations (6.146) to (6.149) are then 4 nonlinear equations in the 4 unknowns $(q_{k+1}, \dot{q}_{k+1}) = (q_1, \dot{q}_1)$. If $(q_0, \dot{q}_0)$ and the time step, h, are suitably chosen, then a solution for $(q_1, \dot{q}_1)$ is guaranteed by Proposition 4 in Chapter 4. If $(q_1, \dot{q}_1)$ and h are still suitable, then set $(q_k, \dot{q}_k) = (q_1, \dot{q}_1)$ and solve the equations again.

6.4.3 Numerical Results

As shown in Goldstein, Poole and Safko [13] and Hairer, Lubich and Wanner [14], solutions of the Kepler problem are ellipses. The exact solutions can be found in Goldstein, Poole and Safko [13].

Short Time

A fourth order Lagrange–d'Alembert Method is used to produce a short term simulation of the nonholonomic oscillator with the initial conditions $(q, \dot{q}) = (0.4, 0, 0, 2)$. The simulation is run for a little longer than two complete oscillations and plotted with the exact solution given by Matlab's ode45 solver with a relative error tolerance of 10^{-12}. The fourth order method is generated from a ψ that is the order 4 Taylor polynomial of the exact solution. That is,

$$\psi(h, h, (q, \dot{q})) = (q, \dot{q}) + h\, \frac{d}{dt}\bigg|_{t=0} F_t^{X_E}(q, \dot{q}) + \frac{h^2}{2}\, \frac{d^2}{dt^2}\bigg|_{t=0} F_t^{X_E}(q, \dot{q}) + \\ + \frac{h^3}{6}\, \frac{d^3}{dt^3}\bigg|_{t=0} F_t^{X_E}(q, \dot{q}) + \frac{h^4}{24}\, \frac{d^4}{dt^4}\bigg|_{t=0} F_t^{X_E}(q, \dot{q}).$$

See Figure 6.12. The Lagrange–d'Alembert integrator is showing an error in all the plots. This will be commented on in the remarks at the end of this chapter.

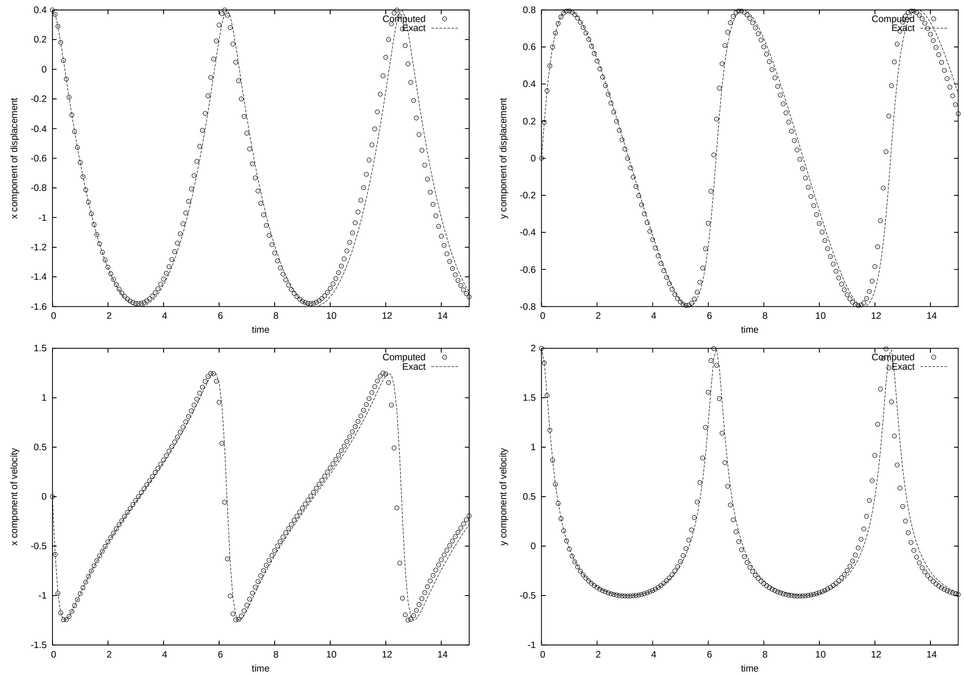

Figure 6.12: Lagrange–d'Alembert integration and exact solution of the components of the Kepler problem. Order 4 method. $h = 0.1$

Long Time

Long time integration of the first order Lagrange–d'Alembert method given by Equations (6.146) to Equation (6.149) is shown in Figure 6.13. The figure shows that the evolution remains bounded. The figures represent 20 000 oscillations of the system.

Figures 6.14 to 6.17 show the first and last computed oscillations after a long integration. The character of the numerical solution remains but the magnitudes have changed significantly, and the period has shortened.

Evolution of the Conserved Quantities

Energy, momentum and the Laplace–Runge–Lenz vector in Equations (6.124), (6.125) and (6.126) are conserved by the exact flow. Figure 6.18 shows the absolute error in the energy and momentum for the order one Lagrange–d'Alembert integrator. The momentum error is due entirely to numerical error, since the order one Lagrange–d'Alembert method exactly pre-

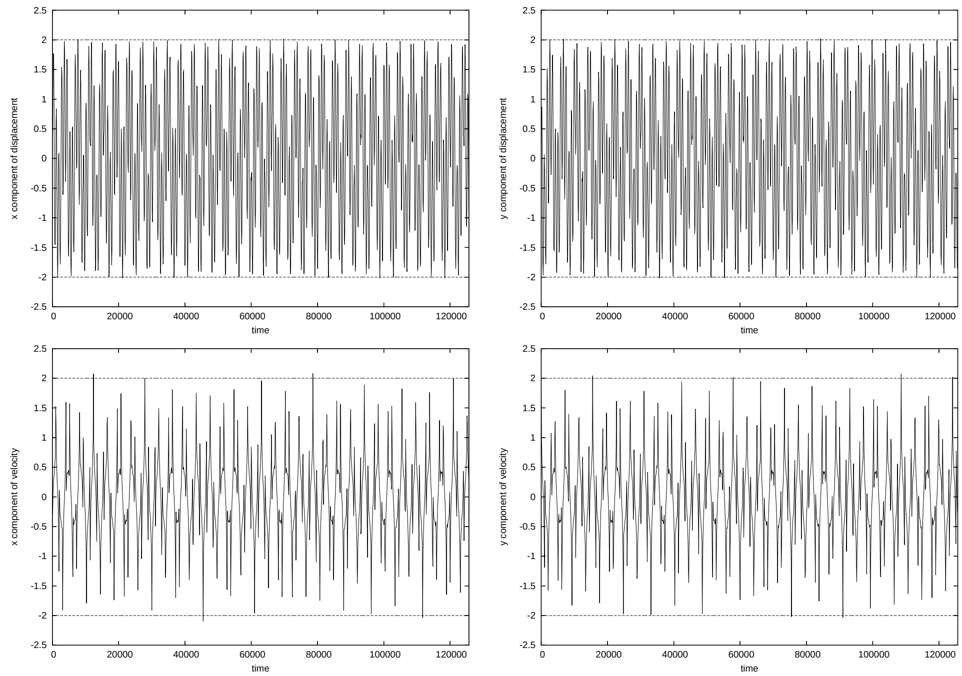

Figure 6.13: Long time Lagrange–d'Alembert integration of the components of the Kepler problem. Horizontal lines illustrate the upper and lower bounds of the oscillations. Order one method. $h = 0.1$. Every 1000 data point plotted.

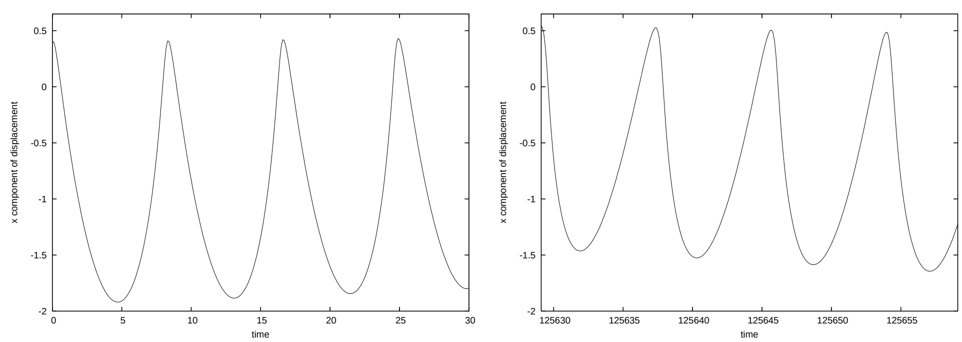

Figure 6.14: First and last oscillations of the x component of the Kepler problem.

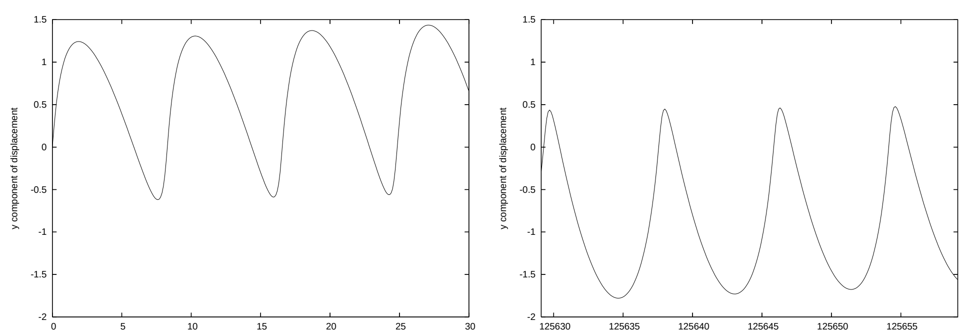

Figure 6.15: First and last oscillations of the y component of the Kepler problem.

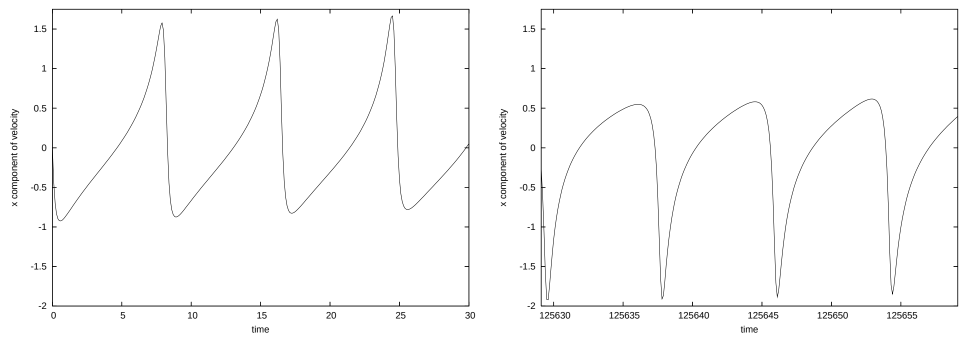

Figure 6.16: First and last oscillations of the $\dot{x}$ component of the Kepler problem.

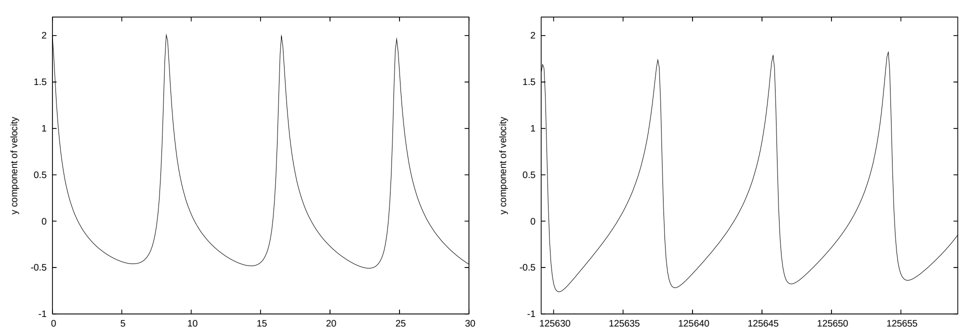

Figure 6.17: First and last oscillations of the $\dot{y}$ component of the Kepler problem.

serves the momentum, as the following computation, using Equations (6.146) to (6.149) shows. Write $r_{k+1} = (x_{k+1}^2 + y_{k+1}^2)^{3/2}$,

$$
\begin{aligned}
J(q_{k+1}, \dot{q}_{k+1}) &= x_{k+1}\dot{y}_{k+1} - y_{k+1}\dot{x}_{k+1} \\
&= (x_k + h\dot{x}_k)\left(\dot{y}_k + \frac{h}{r_{k+1}}(y_k + h\dot{y}_k)\right) - \\
&\qquad (y_k + h\dot{y}_k)\left(\dot{x}_k + \frac{h}{r_{k+1}}(x_k + h\dot{x}_{k+1})\right) \\
&= x_k\dot{y}_k - y_k\dot{x}_k \\
&= J(q_k, \dot{q}_k).
\end{aligned}
$$

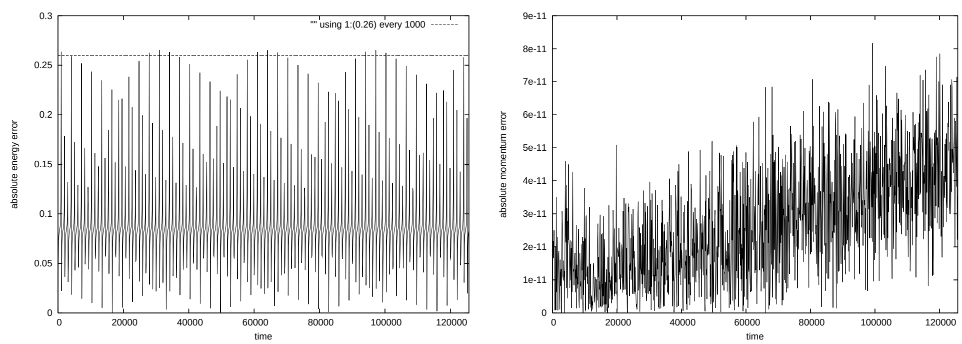

Figure 6.18: Energy and momentum error for the order one Lagrange–d'Alembert integrator. Horizontal lines illustrate the upper bound of the error. $h = 0.1$. Every 1000 point plotted.

Figure 6.19 shows the x and y components of the Laplace–Runge–Lenz vector for a short time and its magnitude over a long time. The vector is *not* preserved by the integrator and its magnitude remains bounded. See the remarks at the end of this chapter.

Order Matching

To generate examples showing that the residual order of Proposition 6 in Chapter 4 is obtained, the curve segments $t \mapsto \psi(h, t, (q, \dot{q}))$ of Chapter 5 are taken to be the Taylor series of the exact solution truncated at various orders, p. The method order is estimated using step doubling as outlined in Section B.4. Table 6.1 gives the results of using first to fourth order approximations by Taylor series.

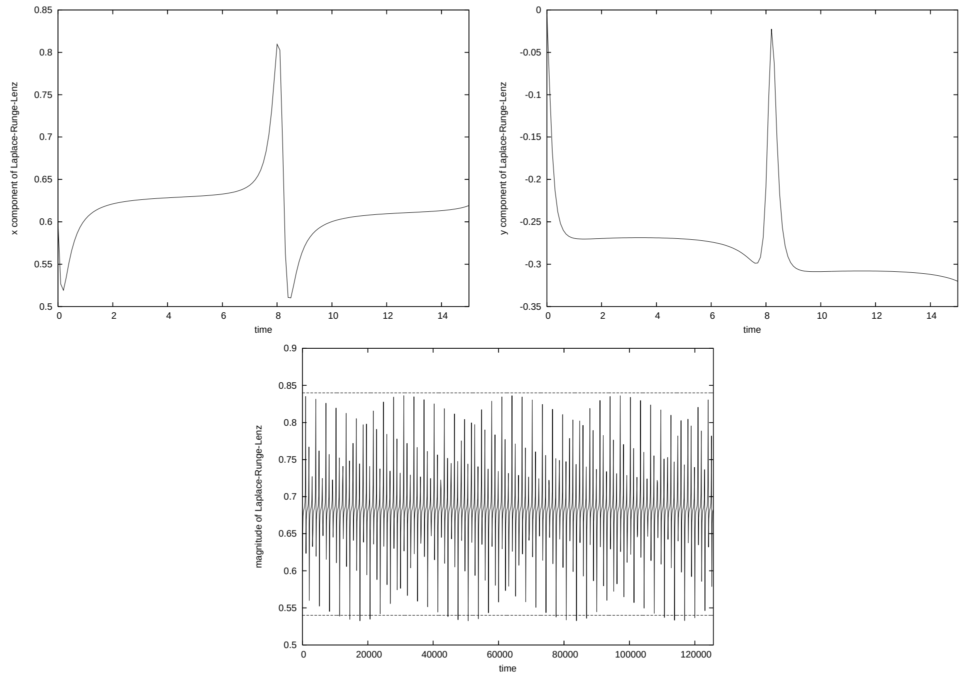

Figure 6.19: Components of the Laplace–Runge–Lenz and its magnitude as evolved by the order one Lagrange–d'Alembert integrator. $h = 0.1$.

6.4.4 Remarks

Figure 6.19 shows that the Laplace–Runge–Lenz vector is not preserved. Figure 6.20 confirms this by showing a precession that is not evident in the physical system.

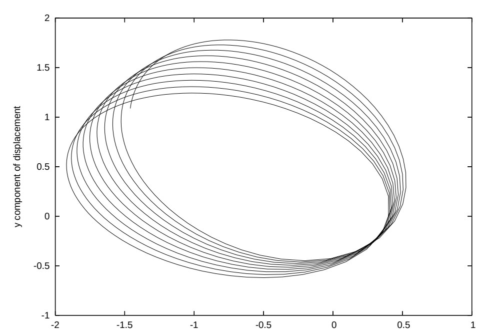

Figure 6.20: y vs. x of the order one Lagrange–d'Alembert integrator exhibiting a precession in the orbit.

Hairer, Lubich and Wanner [14] show this same precession in their demonstration of the symplectic Euler method. The inverse square force law arising from the potential $V(q) =$

Step h	Error 1	Rate	Error 2	Rate	Error 3	Rate	Error 4	Rate
0.1								
0.05	1.43e-02		9.48e-04		9.32e-04		6.07e-05	
0.025	7.8e-03	0.87	1.15e-04	3.04	1.27e-04	2.89	1.83e-06	5.04
0.0125	4.02e-03	0.96	1.23e-05	3.23	1.62e-05	2.97	4.97e-08	5.21
0.00625	2.03e-03	1.00	1.01e-06	3.61	2.03e-06	2.99	1.05e-09	5.56

Table 6.3: Method order estimates for the x component of nonholonomic oscillator simulations. The column headings refer to the truncation order of the Taylor series.

$-1/|q|$ is quite special in that no other centrally attracting potential has the eccentricity vector as a constant of the motion. Given this, and the fact that a symplectic integrator is exactly integrating the equations of motion for a nearby Hamiltonian system (Hairer, Lubich, Wanner [14], Chapter IX), a reasonable explanation for the precession is that the numerical method is giving the exact solution for a system under the influence of a potential that is not exactly given by an inverse square law.

To verify this, the order 3 modified equations, see Hairer, Lubich, Wanner [14] were solved using Matlab's ode45 routine. The results are illustrated in Figure (6.21). As can be seen, the agreement is quite good.

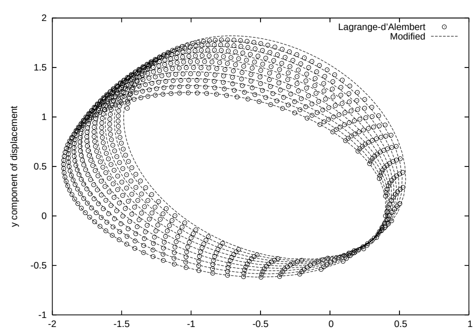

Figure 6.21: y vs. x of the order one Lagrange–d'Alembert integrator exhibiting and the Matlab ode45 solution trajectory for the order 3 modified equation.

The long term energy and momentum error for Matlab's ode45 routine is plotted in Figure 6.22. The components of the Laplace–Runge–Lenz vector and its magnitude are plotted in Figure 6.23. In all figures,a linear error in time is observed.

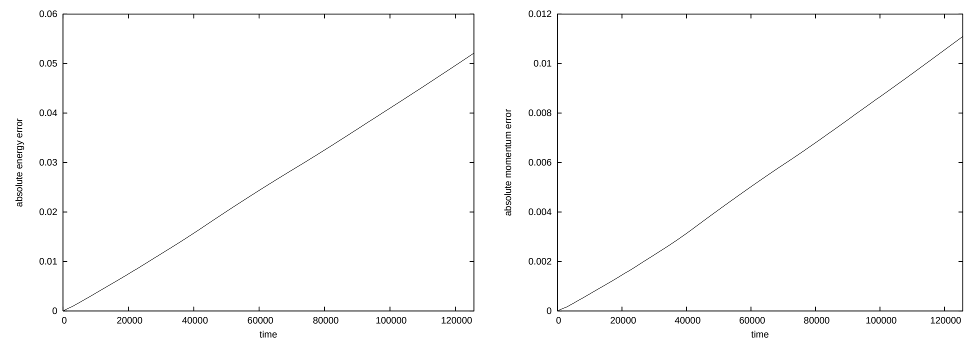

Figure 6.22: Energy and momentum error for Matlab's ode45 routine. Every 1000 point plotted.

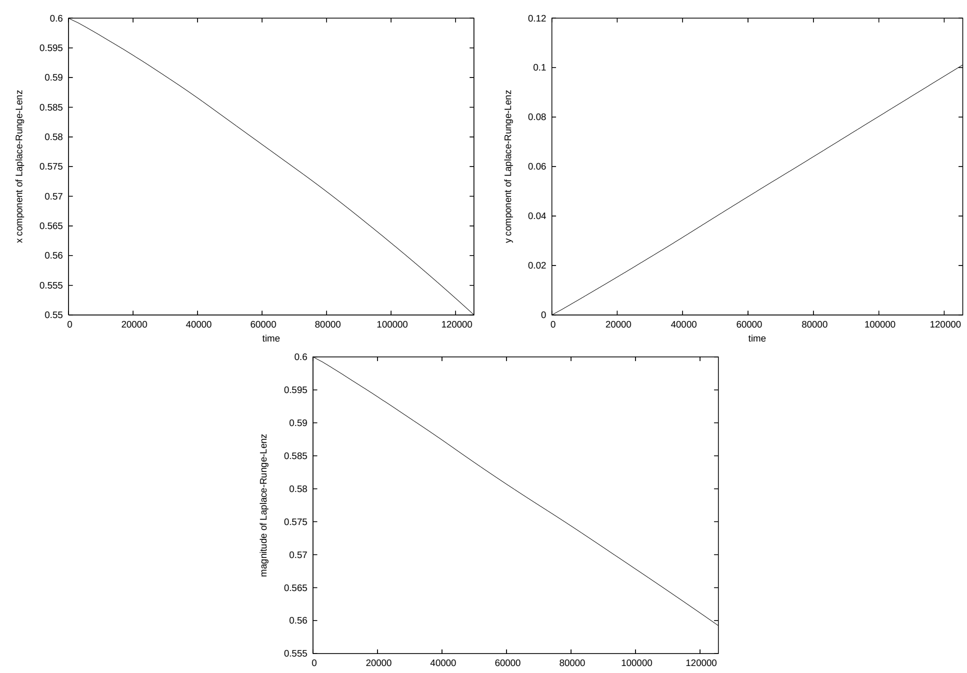

Figure 6.23: Components of the Laplace–Runge–Lenz and its magnitude as evolved by Matlab's ode45 routine

Chapter 7

Conclusion and Further Work

In this thesis, a geometric theory of Lagrange–d'Alembert integrators has been introduced and applied to mechanical problems. As a result, it is possible to generate numerical simulations for linearly constrained mechanical systems that exactly preserve the constraints and naturally reflect the geometry of the continuous system.

The discrete Euler–Lagrange equations are a skew critical point problem that is generated by a discrete Lagrange–d'Alembert principal.

Of course, there are other investigations to be made and further work to be done.

Applied Forces. Constraints can be realized as the application of a force that causes the constraints to be satisfied. The introduction of α_h in Chapter 4 suggests that any force may be applied, including dissipative forces.

Methods. The methods used to demonstrate the theory in Chapters 5 and 6 are *proof of concept* only. They are neither easy to implement, nor efficient. Faster and better methods, perhaps based directly on the exact Lagrangian in the calculation of dL_h in Equation (4.13), could be developed.

Numerical Analysis. The numerics of the discrete Euler–Lagrange equations still need to be investigated. They need to be classified and questions such as stability need to be examined.

Numerical Mechanics. A strong motivation for using geometric integrators is their (possible) applications in geometric mechanics. The link between numerical and theoretical results should be strengthened with theorems of a more quantitative nature.

Discrete Vakonomic Mechanics. The discrete Lagrange–d'Alembert principle can be replaced with a Vakonomic Principle. This could be done, if it was thought to be useful.

Appendix A

Theorems, Lemmas and Definitions

The Lemmas in this appendix provide results that are necessary for the theory in this thesis.

Lemma 1. *(Wasserman [25]) Let M and N be manifolds of dimension m and n respectively. Let $g\colon M \to N$ be a C^1 mapping and suppose g is a submersion at $m_0 \in M$. Then there are charts (U, μ) at m_0 and (V, ν) at $n_0 = g(m_0)$ such that $\hat{g}(x, y) = (\nu \circ g \circ \mu^{-1})(x, y) = x$ where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^{m-n}$.*

Definition 12. The coordinate system (U, μ) given in Lemma 1 is called a g *adapted* coordinate system.

A g adapted coordinate system is useful for submanifolds generated by the level sets of g, a construct used throughout this thesis.

The following Lemma is a generalization of the inverse function theorem for a neighbourhood of a closed submanifold.

Lemma 2. *(Cuell and Patrick [9]) Let M and N be manifolds and $f\colon M \to N$ be C^k, $k \geq 1$. Suppose that*

1. *M_0 is a closed submanifold of M, N_0 is a closed submanifold of N and $f|_{M_0} \to N_0$ is a diffeomorphism,*
2. *f is a local diffeomorphism at every $m \in M_0$.*

Then f is a C^k diffeomorphism from some open neighbourhood $U \supset M_0$ to some open neighbourhood $V \supset N_0$.

The next Lemma is similar to the previous one, but for submanifolds arising from a constrained critical point problem.

Lemma 3. *(Cuell and Patrick [9]) Let M and N be manifolds and let $f\colon M \to \mathbb{R}$ and $g\colon M \to N$ be C^k, $k \geq 2$. Suppose that*

1. *M_0 is a closed submanifold of M, N_0 is a closed manifold of N and $\gamma_0\colon N_0 \to M_0$ is a C^{k-1} diffeomorphism,*

2. *every $n_0 \in N_0$ is a regular value of g,*

3. *for all $n_0 \in N_0$, $\gamma(n_0)$ is a nondegenerate critical point of the function $f|_{g^{-1}(n_0)}$.*

Then there are open neighbourhoods $U \subseteq M_0$ and $V \subseteq N_0$ and a C^{k-1} extension $\gamma\colon V \to U$ such that

1. *for all $n \in V$, $\gamma(n)$ is a critical point of $f|_{g^{-1}(n)}$,*

2. *$\gamma(n)$ is the unique critical point of $f|_{g^{-1}(n)}$ in U.*

Appendix B

Skew Critical Points, Order Notation and Residuals

B.1 Introduction

Given a function $f\colon M \to \mathbb{R}$ and a mapping $g\colon M \to N$, a point $m_0 \in M$ is a critical point of f with respect to g if $df(m_0)\,X = 0$, $g(m_0) = n_0$ for all X such that $T_{m_0}gX = 0$. This is a standard constrained optimization problem that appears in undergraduate calculus courses. The key elements of this problem are df, g and $\ker Tg = Tg^{-1}(n_0)$. The theory is not more difficult if the tangent space to $g^{-1}(n_0)$ is replaced by a general distribution, $\mathcal{D}$. In fact, it allows for a slightly more general class of optimization problems, including one of the key problems of this thesis: finding critical points for a discrete variational principle.

Also collected in this appendix are a number of results concerning contact order for mappings on manifolds. That is, there is defined the notion of $f = g + O(h^r)$ when f, g are mappings from a manifold M to a manifold N with a little extra structure included. This reduces to the standard definition when $M = U \times \mathbb{R}$, and $N = V$ where U and V are Euclidean spaces.

The theory in this appendix provides the necessary tools for the development of Lagrange–D'Alembert integrators directly on finite dimensional manifolds. Cuell and Patrick [9] includes the infinite dimensional presentation.

B.2 Skew Critical Points

Let M and N be manifolds of dimension m and n respectively. Let $\mathcal{D}$ be a C^1, d dimensional distribution on M, α a C^1 one form on M and $g\colon M \to N$ a mapping.

Definition 13. The point $m_0 \in M$ is a *skew critical point of α with respect to g and $\mathcal{D}$* if

$$\alpha(m_0)\,\mathcal{D}_{m_0} = 0, \tag{B.1}$$

$$g(m_0) = n_0, \tag{B.2}$$

where $\alpha(m_0)\,\mathcal{D}_{m_0} = 0$ means that $\alpha(m_0)$ annihilates all vectors in $\mathcal{D}_{m_0}$.

Definition 14. Let m_0 be a skew critical point of α with respect to g and $\mathcal{D}$. Let $u \in \ker T_{m_0} g$, $v \in \mathcal{D}_{m_0}$ and V a vector field defined in a neighbourhood of m_0 taking its values in $\mathcal{D}$ such that $V(m_0) = v$. The *skew–Hessian* of α with respect to g and $\mathcal{D}$ is the *asymmetric* bilinear form

$$d_{\mathcal{D}}(\alpha, g)(m_0)(u, v) = u(\alpha\, V)(m_0).$$

The skew–Hessian is asymmetric since u and v belong to different vector subspaces and hence cannot be exchanged in the argument. It needs to be shown that the definition of $d_{\mathcal{D}}(\alpha, g)$ depends only on v and not on the extension to V. Let $\tilde{V}$ be another vector field such that $\tilde{V}(m_0) = v$ and let $\{X_b\}_{b=1}^{d}$ be a local basis of $\mathcal{D}$ in a neighbourhood of m_0. Then there are functions, $\bar{V}^b$ such that $\tilde{V} - V = \bar{V}^b X_b$, and

$$\begin{aligned} u(\alpha\,(\tilde{V} - V))(m_0) &= u(\alpha\,(\bar{V}^b X_b))(m_0) \\ &= u(\bar{V}^b)(m_0)(\alpha\, X_b)(m_0) + \bar{V}^b(m_0) u(\alpha\, X_b)(m_0). \end{aligned}$$

The first term is zero since $(\alpha\, X_b)(m_0) = 0$. The second term is zero since

$$\bar{V}^b(m_0) = \tilde{V}^b(m_0) - V^b(m_0) = v - v = 0.$$

Assume now that g is C^1 in a neighbourhood of m_0 and that n_0 is a regular value. Then there are g adapted coordinates (U, μ) at m_0 (see Definition (12) in Appendix A). Write $\mu(m_0) = (x_0, y_0)$. Let $\{X_a\}_{a=1}^{m}$ be a local basis of TM in a neighbourhood of m_0 such that $\{X_b\}_{b=1}^{d}$ is a basis of $\mathcal{D}$ and let $\{\phi^a\}_{a=1}^{m}$ be a basis of $T_m^* M$ dual to $\{X_a\}_{a=1}^{m}$. Write $\hat{X}_a$ and $\hat{\phi}^a$ as the local representations of X_a and ϕ^a with respect to μ. Let the indices j run from

1 to $m-n$ (y coordinates for $g^{-1}(n_0)$), k from 1 to m (x and y coordinates for M) and b from 1 to d (index for the basis of $\mathcal{D}$). Write $v^b = \phi^b(v)$. The skew–Hessian is then given in coordinates as

$$
\begin{aligned}
\hat{d}_{\hat{\mathcal{D}}}(\hat{\alpha}, \hat{g})(x_0, y_0)(\hat{u}, \hat{v}) &= \hat{u}(\hat{\alpha}\,\hat{V})(x_0, y_0) \\
&= \hat{u}^j \partial_{y^j}(\hat{\alpha}_k \hat{V}^b \hat{X}_b^k)(x_0, y_0) \\
&= \hat{u}^j \frac{\partial \hat{\alpha}_k}{\partial y^j}(x_0, y_0)\hat{v}^b \hat{X}_b^k(x_0, y_0) + \hat{u}^j \hat{\alpha}_k \frac{\partial \hat{V}^b}{\partial y^j} \hat{X}_b^k(x_0, y_0) + \\
&\qquad + \hat{u}^j \hat{\alpha}_k(x_0, y_0)\hat{v}^b \frac{\partial \hat{X}_b^k}{\partial y^j}(x_0, y_0) \qquad \text{Leibniz rule} \\
&= \hat{u}^j \frac{\partial \hat{\alpha}_k}{\partial y^j}(x_0, y_0)\hat{v}^b \hat{X}_b^k(x_0, y_0) + \\
&\qquad + \hat{u}^j \hat{\alpha}_k(x_0, y_0)\hat{v}^b \frac{\partial \hat{X}_b^k}{\partial y^j}(x_0, y_0) \qquad \text{since } \hat{\alpha}_k \hat{X}_b^k = 0 \\
&= \frac{\partial}{\partial y^j}\left(\hat{\alpha}\,\hat{X}_b\right)(x_0, y_0)\hat{u}^j \hat{v}^b \qquad \text{(B.3)}
\end{aligned}
$$

Definition 15. Suppose n_0 is a regular value of $g, \dim M = d + \dim N$ and that the following are equivalent and satisfied:

1. $d_{\mathcal{D}}(\alpha, g)(m_0)(u, v) = 0 \;\forall\; v \in \mathcal{D}_{m_0} \Rightarrow u = 0$,

2. $d_{\mathcal{D}}(\alpha, g)(m_0)(u, v) = 0 \;\forall\; u \in \ker T_{m_0} g \Rightarrow v = 0$.

Then m_0 is said to be a *nondegenerate skew critical point* of α with respect to g and $\mathcal{D}$.

When m_0 is a nondegenerate skew critical point, the matrix $H_{jb} = \frac{\partial}{\partial y^j}(\hat{\alpha}\hat{X}_b)$ in Equation (B.3) is square and invertible when restricted to $\ker T_{m_0} g \times \mathcal{D}_{m_0}$.

Lemma 4. *Let m_0 be a nondegenerate skew critical point of α with respect to g and $\mathcal{D}$ and let n_0 be a regular value of g. Then there is a neighbourhood, W_0, of n_0 such that the skew critical equations have a unique solution for each $n \in W_0$.*

Proof. Let X and Y be vector fields in a neighbourhood of m_0 such that $Y(m_0) \in T_{m_0} g^{-1}(n_0)$ and $T_{m_0} g X(m_0) \neq 0$.

Let $v \in \mathcal{K}_{m_0}$ and V a vector field in a neighbourhood of m_0 taking its values in $\mathcal{K}$ such that $V(m_0) = v$. Differentiating Equations (B.1) and (B.2) in the direction of $X + Y$ at m_0 and setting equal to zero gives,

$$X(\alpha\, V)(m_0) + Y(\alpha\, V)(m_0) = 0, \tag{B.4}$$

$$(X\, g)(m_0) = 0. \tag{B.5}$$

Since $T_{m_0} g\, X \neq 0$, Equation (B.5) implies $X(m_0) = 0$ which further implies that Equation (B.4) becomes

$$Y(\alpha\, V)(m_0) = 0.$$

But this is $d_{g,\mathcal{K}}\alpha(m_0)(Y(m_0), v) = 0$ for any $v \in \mathcal{K}_{m_0}$. Since m_0 is a nondegenerate skew critical point, this implies that $Y(m_0) = 0$. Therefore the derivative of the skew critical Equations (B.1) and (B.2) is nonsingular and the inverse function theorem implies that the equations are uniquely solvable in a neighbourhood, U, of m_0. □

Continuing, shrink U, if necessary, so that it is a coordinate neighbourhood containing m_0 and $V \supset g(U)$ is a coordinate neighbourhood of n_0. Let $\nu\colon V \to W$ be a local coordinate map. Since g is a submersion at m_0 there is a local coordinate map $\mu\colon U \to W \times S$ such that the local representative of g is $g(x, y) = x$. Then $D_1 g = \mathrm{Id}$ so that the implicit function theorem gives neighbourhoods W_0, V_0 and a unique mapping $f\colon W_0 \times V_0$ such that $g(f(w, y), y) = w$ for all $(w, y) \in W_0 \times S_0$.

The skew critical equations in this chart are then

$$\alpha(f(w, y), y)\, X_b(f(w, y), y) = 0,$$

$$g(f(w, y), y) = w,$$

and are uniquely solvable since $(w, y) \in \mu(U)$.

There is a skew version of Lemma 3 in Appendix A .

Lemma 5. *Let M and N be manifolds. Let α be a C^k one form on M, $\mathcal{D}$ a C^k distribution on M and $g\colon M \to N$ be C^k, $k \geq 1$. Suppose that*

1. *M_0 is a closed submanifold of M, N_0 is a closed manifold of N and $\gamma_0\colon N_0 \to M_0$ is a C^k diffeomorphism,*

2. *every $n_0 \in N_0$ is a regular value of g,*

3. *for all $n_0 \in N_0$, $\gamma(n_0)$ is a nondegenerate skew–critical point of α with respect to g and $\mathcal{D}$.*

Then there are open neighbourhoods $U \subseteq M_0$ and $V \subseteq N_0$ and a C^k extension $\gamma\colon V \to U$ such that

1. *for all $n \in V$, $\gamma(n)$ is a skew–critical point of α with respect to g and $\mathcal{D}$,*

2. *$\gamma(n)$ is the unique skew–critical point of α with respect to g and $\mathcal{D}$ in U.*

Proof. Using g adapted coordinates, the skew–critical equations are

$$\hat{\alpha}(x_0, y_0)\, \hat{X}(x_0, y_0) = 0, \tag{B.6}$$
$$\hat{g}(x_0, y_0) = x_0.$$

Since $\gamma_0(n_0)$ is a nondegenerate skew–critical point, the implicit function theorem states that Equation (B.6) can be solved for y in terms of x in a neighbourhood of (x_0, y_0). The rest of the proof then follows Cuell and Patrick [9]. □

B.3 Order Notation and Residuals

Definition 16. Let M be a manifold and $h_M\colon M \to \mathbb{R}$ be a C^∞ function such that 0 is a regular value. The pair (M, h_M) will be called a *manifold*.

Let (M, h_M) and (N, h_N) be manifolds. A C^k function $f\colon (M, h_M) \to (N, h_N)$ is a mapping $f\colon M \to N$ if $h_N \circ f = h_M$. If $f\colon (M, h_M) \to N$ or $f\colon M \to (N, h_N)$, then f is a mapping $f\colon M \to N$ without any conditions involving h_M or h_N.

Definition 17. Let (M, h_M) and N be manifolds and $f_i\colon (M, h_M) \to N$, $i = 1, 2$ such that $f_1 = f_2$ on $h_M^{-1}(0)$. For each $m_0 \in h_M^{-1}(0)$, let $f_i(m_0) = n_0$ and let (V_{n_0}, ν_{n_0}) be a coordinate chart for N at n_0. Define $f_2 = f_1 + O(h_M^r)$, $r \geq 1$ if, for all $m_0 \in h_M^{-1}(0)$,

$$\nu_{n_0}(f_2(m)) - \nu_{n_0}(f_1(m)) = h_M(m)^r (\delta f)_{\nu_{n_0}}(m),$$

for all m in some neighbourhood, U, of m_0 and $(\delta f)_{\nu_{n_0}}$ continuous at m_0.

Lemma 6. *The definition of $f_2 = f_1 + O(h_M^r)$ does not depend on the coordinate chart.*

Proof. Let (V, ν) and $(\tilde{V}, \tilde{\nu})$ be coordinate charts at $f(m_0)$ for $m_0 \in h_M^{-1}(0)$. Then on $V \cap \tilde{V}$, for m in the neighbourhood U of Definition 17,

$$\begin{aligned}
\tilde{\nu}(f_2(m)) - \tilde{\nu}(f_1(m)) &= ((\tilde{\nu} \circ \nu^{-1}) \circ \nu)(f_2(m)) - ((\tilde{\nu} \circ \nu^{-1}) \circ \nu)(f_1(m)) \qquad \text{inserting } \nu^{-1} \circ \nu \\
&= (\tilde{\nu} \circ \nu^{-1})(\nu(f_2(m))) - (\tilde{\nu} \circ \nu^{-1})(\nu(f_1(m))) \\
&= (\tilde{\nu} \circ \nu^{-1})(\nu(f_1(m)) + h_M(m)^r (\delta f)_\nu(m)) - (\tilde{\nu} \circ \nu^{-1})(\nu(f_1(m))) \qquad \text{since } f_2 = f_1 + O(h^r) \\
&= \int_0^1 \frac{d}{du}(\tilde{\nu} \circ \nu^{-1})(\nu(f_1(m)) + u\, h_M(m)^r (\delta f)_\nu(m))\, du \qquad \text{fundamental theorem of calculus} \\
&= \left(\int_0^1 D(\tilde{\nu} \circ \nu^{-1})(\nu(f_1(m)) + t\, h_M(m)^r (\delta f)_\nu(m))\, dt \right) h_M(m)^r (\delta f)_\nu(m), \qquad \text{applying } \frac{d}{du}
\end{aligned}$$

which proves the claim. □

Since $h_M(m)^r (\delta f)_{\tilde{\nu}}(m) = \tilde{\nu}(f_2(m)) - \tilde{\nu}(f_1(m))$, the calculation in Lemma 6 gives

$$(\delta f)_{\tilde{\nu}}(m) = \left(\int_0^1 D(\tilde{\nu} \circ \nu^{-1})(\nu(f_1(m)) + t\, h_M(m)^r (\delta f)_\nu(m))\, dt \right) (\delta f)_\nu(m),$$

so that at $m_0 \in h_M^{-1}(0)$ and $f_i(m_0) = n_0$,

$$(\delta f)_{\tilde{\nu}} = D(\tilde{\nu} \circ \nu^{-1})(\nu(n_0))(\delta f)_\nu(m_0),$$

indicating that $(\delta f)_\nu(m_0)$ transforms as a tangent vector.

Definition 18. The vector $\operatorname{res}^r(f_2, f_1)(m_0) \in T_{n_0}N$ with coordinate representation $(\delta f)_\nu(m_0)$, for any chart (V, ν) is called the *r–residual of f_1 and f_2*.

Note that $\operatorname{res}^r(f_2, f_1)$ is defined only on $h_M^{-1}(0) \subset M$ and takes its values in TN.

Since 0 is a regular value of h_M, there are h_M adapted coordinates in a neighbourhood of each $m_0 \in h_M^{-1}(0)$. See Definition 12 in Appendix A. Suppose $f\colon (M, h_M) \to (N, h_N)$. Let (U, μ) be an h_M adapted coordinate chart at $m_0 \in M$ and (V, ν) an h_N adapted coordinate chart at $n_0 = f(m_0) \in N$. Denote the coordinate representations of h_M, h_N and f by $\hat{h}_M$, $\hat{h}_N$ and $\hat{f}$ respectively. Then

$$h_N \circ f = h_M$$

is, after inserting $\nu^{-1} \circ \nu$ and $\mu^{-1} \circ \mu$,

$$h_N \circ \nu^{-1} \circ \nu \circ f \circ \mu^{-1} \circ \mu = h_M,$$

which is the local statement

$$\hat{h}_N \circ \hat{f} = \hat{h}_M.$$

Let $m \in U$ and let $\mu(m) = (x, t)$. Then $(\hat{h}_N \circ \hat{f})(x, t) = \hat{h}_M(x, t) = t$, so that there is a mapping $\tilde{f}$ such that $\hat{f}(x, t) = (\tilde{f}(x, t), t)$.

Coordinate calculations will often produce the result that $\tilde{f}_2(x, h) = \tilde{f}_1(x, h) + O(h^r)$. That is, $\tilde{f}_2(x, h) = \tilde{f}_1(x, h) + h^r \delta\tilde{f}(x, h)$ for some $\delta\tilde{f}(x, h)$. The following calculation shows that this implies $f_2 = f_1 + O(h_M^r)$:

$$\begin{aligned}\hat{f}_2(x, h) - \hat{f}_1(x, h) &= (\tilde{f}_2(x, h), h) - (\tilde{f}_1(x, h), h) \\ &= (\tilde{f}_2(x, h) - \tilde{f}_1(x, h), 0) \\ &= h^r(\delta\tilde{f}(x, h), 0).\end{aligned}$$

This gives $(\delta f)_\nu(x, h) = (\delta\tilde{f}(x, h), 0)$. In addition,

$$(\delta f)_\nu(x, 0) = \widehat{\operatorname{res}}^r(f_2, f_1)(x, 0) \in \ker d\hat{h}_N(x_0, 0),$$

$$
\begin{array}{ccccccc}
M & \xleftarrow{\ \iota\ } & h_M^{-1}(0) & \xrightarrow{\operatorname{res}^r(f_2,f_1)} & \ker\{dh_N\} & \xrightarrow{\ \iota\ } & TN \\
\Big\downarrow{\scriptstyle \mu} & & \Big\downarrow{\scriptstyle \mu} & & \Big\downarrow{\scriptstyle T\nu} & & \Big\downarrow{\scriptstyle T\nu} \\
\mathbb{R}^{m-1} \times \mathbb{R} & \xleftarrow[\iota]{} & \mathbb{R}^{m-1} \times \{0\} & \xrightarrow[(\delta f)_\nu(\cdot,0)]{} & \mathbb{R}^{2(n-1)} \times \{0\} \times \{0\} & \xrightarrow[\iota]{} & \mathbb{R}^{2(n-1)} \times \mathbb{R} \times \mathbb{R}
\end{array}
$$

Figure B.1: Residuals

and therefore $\operatorname{res}^r(f_2, f_1) \in Th_N^{-1}(0)$. Figure B.1 is a diagram of these relationships, with $\dim M = m$, $\dim N = n$ and ι indicating the natural inclusion, wherever it is present.

Let $f_i\colon (M, h_M) \to N$, $i = 1, 2$ be C^r and let (U, μ) be an h_M adapted coordinate chart about $m_0 \in h_M^{-1}(0)$. Let $f_i(m_0) = n_0$ and let (V, ν) be a coordinate chart for N containing n_0. With x fixed, the Taylor expansion of $\hat{f}_i$ about $t = 0$ is

$$\hat{f}_i(x,h) = \hat{f}_i(x,0) + h\frac{\partial \hat{f}_i}{\partial t}(x,0) + \frac{h^2}{2}\frac{\partial^2 \hat{f}_i}{\partial t^2}(x,0) + \cdots + \frac{h^r}{r!}\frac{\partial^r \hat{f}_i}{\partial t^r}(x,0) + (R_r)_i(x,h)\, h^r,$$

where (Abraham, Marsden, Ratiu [2])

$$(R_r)_i(x,h) = \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!}\left(\frac{\partial^r \hat{f}_i}{\partial t^r}(x, t\,h) - \frac{\partial^r \hat{f}_i}{\partial t^r}(x,0)\right) dt.$$

The condition that $f_2 = f_1 + O(h_M^r)$ implies that the Taylor expansions for $\hat{f}_2$ and $\hat{f}_1$ match up to and including the $r-1$ term. Therefore,

$$\begin{aligned}
&\hat{f}_2(x,h) - \hat{f}_1(x,h) \\
&\qquad = \frac{h^r}{r!}\left(\frac{\partial^r}{\partial t^r}\hat{f}_2(x,0) - \frac{\partial^r}{\partial t^r}\hat{f}_1(x,0)\right) + (R_r)_2(x,h)\, h^r - (R_r)_1(x,h)\, h^r.
\end{aligned}$$

This identifies $(\delta f)_\nu(m)$ in ν coordinates, with respect to the h_M adapted coordinate chart as

$$(\delta \hat{f})_\nu(x,h) = \frac{1}{r!}\left(\frac{\partial^r \hat{f}_2}{\partial t^r}(x,0) - \frac{\partial^r \hat{f}_1}{\partial t^r}(x,0)\right) + (R_r)_2(x,h) - (R_r)_1(x,h).$$

At $h = 0$, $(R_r)_i(x,0) = 0$ so that $\operatorname{res}^r(f_2, f_1)(m_0)$ in coordinates is

$$\begin{aligned}
(\delta \hat{f})_\nu(x,0) &= \frac{1}{r!}\left(\frac{\partial^r \hat{f}_2}{\partial t^r}(x,0) - \frac{\partial^r \hat{f}_1}{\partial t^r}(x,0)\right) \\
&= \frac{1}{r!}\left.\frac{\partial^r}{\partial t^r}\right|_{t=0}\left(\hat{f}_2(x,t) - \hat{f}_1(x,t)\right).
\end{aligned}$$

If $\mathrm{res}^r(f_2, f_1) = 0$, then the order of agreement in the Taylor series is at least one greater than r. The notation $(\delta f)_\nu(m)$ will often be shortened to $\delta f(m)$.

If N is a vector bundle over M with α^1 and α^2 sections, the definition of $\alpha^2 = \alpha^1 + O(h_M^r)$ can be written coordinate free as

$$\alpha^2(m) - \alpha^1(m) = h_M(m)^r \delta\alpha(m).$$

For local results, take $M = \mathbb{R}^{m-1} \times \mathbb{R}$ and $h_M(m,t) = h(m,t) = t$, as in the following

Lemma 7. *Let $f_i \colon \mathbb{R}^{m-1} \times \mathbb{R} \to \mathbb{R}^{m-1}$ and $x_i \colon \mathbb{R} \to \mathbb{R}^{m-1}$, $i = 1, 2$. Suppose*

1. *$f_2 = f_1 + O(h^r)$ and f_i are C^1,*
2. *$x_1(0) = x_2(0)$,*
3. *$f_2(x_2(h), h) = f_1(x_1(h), h) = 0$,*
4. *$\dfrac{\partial f_i}{\partial x}(x_i(0), 0)$ are nonsingular.*

Then $x_2(h) = x_1(h) + O(h^r)$.

Proof. Write $x_2(h) = x_1(h) + \delta x(h)$. Then,

$$\begin{aligned}
0 &= f_2(x_2(h), h) \\
&= f_1(x_2(h), h) + h^r \delta f(x_2(h), h) \\
&= f_1(x_1(h) + \delta x(h), h) + h^r \delta f(x_2(h), h) \\
&= \int_0^1 \frac{d}{du} f_1(x_1(h) + u\,\delta x(h), h)\,du + h^r \delta f(x_2(h), h) \qquad \text{fundamental theorem of calculus} \\
&= \left(\int_0^1 \frac{\partial f_1}{\partial x}(x_1(h) + u\,\delta x(h), h)\,du \right) \delta x(h) + h^r \delta f(x_2(h), h). \qquad \text{apply } \frac{d}{du}
\end{aligned}$$

Solving for $\delta x(h)$ gives

$$\delta x(h) = -h^r \left(\int_0^1 \frac{\partial f_1}{\partial x}(x_1(h) + u\,\delta x(h), h)\,du \right)^{-1} \delta f(x_2(h), h),$$

where the inverse exists for h small enough. □

Note that Item 3 can be replaced by $f_i(x_i(h), h) = y$, y constant without affecting the result.

Lemma 8. *Let (M, h_M) and (N, h_N) be manifolds and suppose that $f_i\colon (M, h_M) \to (N, h_N)$ are C^k diffeomorphisms, $k \geq 1$, $i = 1, 2$. If $f_2 = f_1 + O(h_M^r)$ then $f_2^{-1} = f_1^{-1} + O(h_N^r)$.*

Proof. Let (x, t) be h_M–adapted coordinates for M and (y, s) be h_N adapted coordinates for N. Fix y and let $x_i(s) = \tilde{f}_i^{-1}(y, s)$, where
$\hat{f}_i(\tilde{f}_i^{-1}(y, s), s) = (y, s)$. Then

$$x_2(0) = \tilde{f}_2^{-1}(y, 0) = \tilde{f}_1^{-1}(y, 0) = x_1(0),$$

$\hat{f}_2 = \hat{f}_1$ for all $(x, 0)$. Also,

$$\hat{f}_2(x_2(s), s) = \hat{f}_2(\tilde{f}_2^{-1}(y, s), s) = (y, s) = \hat{f}_1(\tilde{f}_1^{-1}(y, s), s) = \hat{f}_1(x_1(s), s).$$

Lemma 7 then gives $\tilde{f}_2^{-1} = \tilde{f}_1^{-1} + O(h^r)$ so that $f_2^{-1} = f_1^{-1} + O(h_N^r)$. ☐

Lemma 9. *Let (M, h_M), (N, h_N) and P be manifolds. Let $f_i\colon (M, h_M) \to (N, h_N)$ and $g_i\colon (N, h_N) \to P$, $i = 1, 2$. Further, assume $f_2 = f_1 + O(h_M^r)$ and $g_2 = g_1 + O(h_N^r)$. Then $g_2 \circ f_2 = g_1 \circ f_1 + O(h_M^r)$ and*

$$res^r(g_2 \circ f_2, g_1 \circ f_1)(m) = res^r(g_2, g_1)(f_1(m)) + T_{f_1(m)} g_2\, res^r(f_2, f_1)(m). \tag{B.7}$$

Proof. Let h_M–adapted coordinates on M be (x, t) and h_N–adapted coordinates on N be (y, s). Then,

$$\begin{aligned}
(\hat{g}_2 \circ \hat{f}_2)(x, t) &= \hat{g}_2(\tilde{f}_2(x, t), t) \\
&= \hat{g}_2(\tilde{f}_1(x, t) + t^r \delta \tilde{f}(x, t), t) \quad \text{since } f_2 = f_1 + O(h_M^r) \\
&= \hat{g}_2(\tilde{f}_1(x, t)) + \int_0^1 \frac{d}{du} \hat{g}_2(\tilde{f}_1(x, t) + u\, t^r \delta \tilde{f}(x, t), t)\, du \\
&\qquad\qquad\qquad\qquad \text{fundamental theorem of calculus} \\
&= \hat{g}_1(\hat{f}_1(x, t)) + t^r \delta g(\hat{f}_1(x, t)) +
\end{aligned}$$

$$+\, t^r \left(\int_0^1 \frac{\partial \hat{g}_2}{\partial y}(\tilde{f}_1(x,t) + u\, t^r \delta \tilde{f}(x,t), t)\, du\, \delta \tilde{f}(x,t), 0 \right),$$

applying $\dfrac{d}{du}$

so that $\delta(g \circ f)$ can be identified from

$$(\hat{g}_2 \circ \hat{f}_2)(x,t) - (\hat{g}_1 \circ \hat{f}_1)(x,t) = t^r \delta(g \circ f)(x,t),$$

as

$$\delta(g \circ f)(x,t) = \delta g(\hat{f}_1(x,t)) + \left(\int_0^1 \frac{\partial \hat{g}_2}{\partial y}(\tilde{f}_1(x,t) + u\, t^r \delta \tilde{f}(x,t), t)\, du\, \delta \tilde{f}(x,t), 0 \right). \tag{B.8}$$

Note that

$$\frac{\partial \hat{g}_2}{\partial y}(\hat{f}_i(x,t))\, \delta \tilde{f}(x,t) = D\hat{g}_2(\hat{f}_i(x,t))\, (\delta \tilde{f}(x,t), 0) = D\hat{g}_2(\hat{f}_i(x,t))\, \delta f(x,t),$$

and setting $t = 0$ in Equation (B.8) gives the component form of Equation (B.7). □

Since $\delta f(x,t) = (\delta \tilde{f}(x,t), 0)$ and $\hat{g}_2(y,0) = \hat{g}_1(y,0)$,

$$D\hat{g}_2(\hat{f}_i(x,0))\, \delta f(x,0) = D\hat{g}_1(\hat{f}_i(x,0))\, \delta f(x,0),$$

so that Tg_2 may be replaced by Tg_1 in Equation (B.7).

For n a nonnegative integer, and $f\colon M \to M$ a mapping, recursively define

$$f^0(m) = m, \qquad f^n(m) = (f^{n-1} \circ f)(m).$$

Lemma 10. *Let $n \geq 1$ be an integer and $f_i\colon (M, h_M) \to (M, h_M)$, $i = 1, 2$ be differentiable and $f_2 = f_1 + O(h_M^r)$. Then*

$$res^r(f_2^n, f_1^n)(m) = res^r(f_2, f_1)(f_1^{n-1}(m)) + \sum_{i=1}^{n-1} T_{f_1(m)} f_2^i\, res^r(f_2, f_1)(m).$$

Proof. The proof is by induction. For $n = 1$, the conclusion is trivially true. Assume the result for $n = k$. Then, using Lemma 9 with $g_2 = f_2^k$ and $g_1 = f_1^k$,

$$\text{res}^r(f_2^{k+1}, f_1^{k+1})(m) = \text{res}^r(f_2^k \circ f_2, f_1^k \circ f_1)(m)$$

$$= \text{res}^r(f_2^k, f_1^k)(f_1(m)) + T_{f_1(m)} f_2^k \text{res}^r(f_2, f_1)(m) \quad \text{Lemma 9}$$

$$= \text{res}^r(f_2, f_1)(f_1^{k-1}(f_1(m))) + \sum_{i=1}^{k-1} T_{f_1(m)} f_2^i \text{res}^r(f_2, f_1)(m) +$$

$$+ T_{f_1(m)} f_2^k \text{res}^r(f_2, f_1)(m) \quad \text{by the induction hypothesis}$$

$$= \text{res}^r(f_2, f_1)(f_1^k(m)) + \sum_{i=1}^{k} T_{f_1(m)} f_2^i \text{res}^r(f_2, f_1)(m),$$

thus proving the Lemma. □

Definition 19. Let (M, h_M) be a manifold of dimension m. Let $\mathcal{D}_2$ and $\mathcal{D}_1$ be d dimensional distributions on (M, h_M). Define

$$\mathcal{D}_2 = \mathcal{D}_1 + O(h_M^r)$$

if, for every $m_0 \in h_M^{-1}(0)$, there is a neighbourhood U_{m_0} of m_0 and one forms $\{\beta_i^b\}_{b=d+1}^m$, $i = 1, 2$ defined on U_{m_0} such that $\mathcal{D}_i = \ker\{\beta_i^b\}_{b=d+1}^m$ and $\beta_2^b = \beta_1^b + O(h_M^r)$ for $b = d+1, \ldots, m$.

Lemma 11. *Let $\mathcal{D}_2$ and $\mathcal{D}_1$ be d dimensional distributions on (M, h_M) such that $\mathcal{D}_2 = \mathcal{D}_1 + O(h_M^r)$. Then for any vector field Y_2 on (M, h_M) taking its values in $\mathcal{D}_2$ there is a vector field Y_1 taking its values in $\mathcal{D}_1$ such that $Y_2 = Y_1 + O(h_M^r)$.*

Proof. As in Definition 19, let U_{m_0} be a neighbourhood of $m_0 \in h_M^{-1}(0)$ and $\{\beta_i^j\}_{j=1}^m, i = 1, 2$ be a basis for $T^*U_{m_0}$ such that $\{\beta_i^b\}_{b=d+1}^m$ annihilates $\mathcal{D}_i$ and $\beta_2^b = \beta_1^b + O(h_M^r)$. Let $\{X_j^i\}_{j=1}^m$, be the local basis dual to $\{\beta_i^j\}_{j=1}^m$ and let $Y_2 = Y_2^a X_a^2, a = 1 \ldots d$. Define on U_{m_0},

$$Y_1^{m_0} = Y_2 - \langle \beta_1^b, Y_2 \rangle X_b^1. \tag{B.9}$$

For $m \in h_M^{-1}(0) \cap U_{m_0}$, $\beta_2(m) = \beta_1(m)$ so that

$$Y_1^{m_0}(m) = Y_2(m) - \langle \beta_1^b(m), Y_2(m) \rangle X_b^1(m)$$

$$= Y_2(m) - \langle \beta_2^b(m), Y_2(m) \rangle X_b^1(m)$$

$$= Y_2(m).$$

$Y_1^{m_0}$ is in $\mathcal{D}_1$ since, for $c = d+1, \ldots, m$

$$\begin{aligned}
\langle \beta_1^c, Y_1^{m_0} \rangle &= \langle \beta_1^c, Y_2 - \langle \beta_1^b, Y_2 \rangle X_b^1 \rangle \quad \text{Equation (B.9)} \\
&= \langle \beta_1^c, Y_2 \rangle - \langle \beta_1^b, Y_2 \rangle \langle \beta_1^c, X_b^1 \rangle \quad \text{since } X_b^1 \in \mathcal{D}_1 \\
&= \langle \beta_1^c, Y_2 \rangle - \langle \beta_1^b, Y_2 \rangle \delta_b^c \\
&= \langle \beta_1^c, Y_2 \rangle - \langle \beta_1^c, Y_2 \rangle \\
&= 0.
\end{aligned}$$

Rewriting Equation (B.9), for $m \in U_{m_0}$,

$$\begin{aligned}
Y_2(m) &= Y_1^{m_0}(m) + \langle \beta_1^b(m), Y_2(m) \rangle X_b^1(m) \\
&= Y_1^{m_0}(m) + \langle \beta_2^b(m) + h_M^r(m)\delta\beta^b(m), Y_2(m) \rangle X_b^1(m) \quad \text{since } \beta_1^b = \beta_2^b + O(h_M^r) \\
&= Y_1^{m_0}(m) + h_M^r(m) \langle \delta\beta^b(m), Y_2(m) \rangle X_b^1(m). \quad \text{since } \beta_2^b Y_2 = 0
\end{aligned}$$

This establishes the Lemma for every U_{m_0}.

Let $\{\psi_{m_0}\}_{m_0 \in h_M^{-1}(0)} \cup \{\psi_0\}$ be a partition of unity subordinate to $\{U_{m_0}\}_{m_0 \in h_M^{-1}(0)} \cup \{M \backslash h_M^{-1}(0)\}$ such that the support of ψ_{m_0} is in U_{m_0} and the support of ψ_0 is in $M \backslash h_M^{-1}(0)$.

The vector field $\psi_{m_0} Y^{m_0}$ on U_{m_0} can be extended to all of M by setting it to be the zero vector field on $M \backslash \text{supp } \psi_{m_0}$. Then, for $m \in M$, define

$$Y_1(m) = \sum_{m_0 \in h_M^{-1}(0)} \psi_{m_0}(m) Y_1^{m_0}(m).$$

For $m \in h_M^{-1}(0)$, $\psi_0(m) = 0$ and

$$\begin{aligned}
Y_1(m) &= \sum_{m_0 \in h_M^{-1}(0)} \psi_{m_0}(m) Y_1^{m_0}(m) \\
&= \left(\psi_0(m) + \sum_{m_0 \in h_M^{-1}(0)} \psi_{m_0}(m) \right) Y_2(m) \\
&= Y_2(m).
\end{aligned}$$

And, for $m \in \cup_{m_0 \in h_M^{-1}(0)} U_{m_0}$,

$$\begin{aligned} Y_1(m) &= \sum_{m_0 \in h_M^{-1}(0)} \psi_{m_0}(m) Y_1^{m_0}(m) \\ &= \sum_{m_0 \in h_M^{-1}(0)} \psi_{m_0}(m) \left(Y_2(m) + h_M^r(m) \delta Y^{m_0}(m)\right) \\ &= Y_2(m) + h_M^r(m) \sum_{m_0 \in h_M^{-1}(0)} \psi_{m_0}(m) \delta Y^{m_0}(m). \end{aligned}$$

Thereby proving the Lemma. □

Lemma 12. *Let $\mathcal{D}_2$ and $\mathcal{D}_1$ be d dimensional distributions on the manifold (M, h_M) such that $\mathcal{D}_2 = \mathcal{D}_1 + O(h_M^r)$. Let $\{X_a^2\}_{a=1}^d$ be a local basis for $\mathcal{D}_2$ in a neighbourhood, U of $m \in h_M^{-1}(0)$ and let X_a^1 be the vector field on U given by Lemma (11) for each $a = 1, \ldots, d$. Then $\{X_a^1\}_{a=1}^d$ is a local basis for $\mathcal{D}_1$.*

Proof. Suppose $\{X_a^1\}_{a=1}^d$ is not a basis. Then, without loss of generality, assume

$$X_1^1 = Z^c X_c^1, \; c = 2, \ldots, d,$$

for some functions Z^c. By hypothesis

$$X_1^2 = X_1^1 + O(h_M^r) = Z^c X_c^1 + O(h_M^r)$$

so that at m,

$$X_1^2(m) = Z^c(m) X_c^1(m) = Z^c(m) X_c^2(m),$$

contradicting the hypothesis that $\{X_a^2\}_{a=1}^d$ is a basis. □

Lemma 13. *Let (M, h_M) and (N, h_N) be manifolds and suppose α^i, g_i, γ_i and $\mathcal{D}_i$, $i = 1, 2$ are as in Lemma 5 and $M_0 \subseteq h_M^{-1}(0)$, $N_0 \subseteq h_N^{-1}(0)$. If $\alpha^2 = \alpha^1 + O(h_M^r)$, $g_2 = g_1 + O(h_M^r)$ and $D_2 = D_1 + O(h_M^r)$ then $\gamma_2 = \gamma_1 + O(h_N^r)$.*

Proof. Let (x, t) be h_M–adapted coordinates for M and (y, s) be h_N–adapted coordinates for N. For fixed y, $\hat{\gamma}_i(y, h)$ is a curve of skew–critical points of $\hat{\alpha}^i$ with respect to $\hat{g}_i$ and

$\hat{\mathcal{D}}_i$, $i = 1, 2$. Let $\{(X_a)_2\}_{a=1}^d$ be a local basis for $\hat{\mathcal{D}}_2$ and let $\{(X_a)_1\}_{a=1}^d$ be the local basis guaranteed by Lemma 12. Then, for $a = 1, \ldots, d$,

$$\begin{aligned}\hat{\alpha}^2(x,t)\,(\hat{X}_a)_2 &= \left(\hat{\alpha}^1(x,t) + t^r\delta\alpha(x,t)\right)\left((\hat{X}_a)_1 + t^r\delta X_a(x,t)\right)\\ &= \hat{\alpha}^1(x,t)\,(\hat{X}_a)_1(x,t) + t^r\delta\tilde{\alpha}_a(x,t)\end{aligned}$$

for some continuous $\delta\tilde{\alpha}_a$. For $i = 1, 2$ define

$$f_i(x,t) = \left(\hat{g}_i(x,t), \hat{\alpha}^i\,(\hat{X}_1)_i\,\ldots,(\hat{X}_d)_i\right).$$

Then $f_i(\hat{\gamma}_i(y,t)) = 0$ and

$$\begin{aligned}f_2(x,t) &= \left(\hat{g}_2(x,t), \hat{\alpha}^2(x,t)\,(\hat{X}_1)_2\,\ldots, \alpha^2(\hat{x},t)(\hat{X}_d)_2\right)\\ &= \Big(\hat{g}_1(x,t) + t^r\delta g(x,t), \hat{\alpha}^1(x,t)(\hat{X}_1)_1 + t^r\delta\tilde{\alpha}_1(x,t)\,\ldots,\\ &\qquad\qquad \alpha^1(\hat{x},t)(\hat{X}_d)_1 + t^r\delta\tilde{\alpha}_d(x,t)\Big)\\ &= f_1(x,t) + t^r\delta f(x,t).\end{aligned}$$

Then, by Lemma 7, $\hat{\gamma}_2(x,t) = \hat{\gamma}_1(x,t) + t^r\delta\gamma(x,t)$ which gives $\gamma_2 = \gamma_1 + O(h^r)$. □

Lemma 14. *Let M and (N, h_N) be manifolds. Let $\gamma\colon (U, h_N) \subseteq (N, h_N) \to M \times M$ be C^k, $k \geq 1$. Suppose that $\gamma|_{h_N^{-1}(0)}$ is a diffeomorphism to $\Delta(M \times M)$. Then there are neighbourhoods $\tilde{U} \subseteq U$ and $V \subseteq M \times \mathbb{R}$ of $M \times \{0\}$ such that, for all $(m, h) \in V$ there is a unique $\tilde{m} \in M$ so that for some $n \in \tilde{U}$, $\gamma(n) = (m, \tilde{m})$ and $h_N(n) = h$. The map $f\colon V \to M$ defined by $(m, h) \mapsto \tilde{m}$ is C^k.*

Proof. Let π_1 and π_2 be the projections on $M \times M$, $\pi_i(m_1, m_2) = m_i$, $i = 1, 2$. Define $\psi\colon U \to M \times \mathbb{R}$ by $n \mapsto ((\pi_1 \circ \gamma)(n), h_N(n))$. ψ is a diffeomorphism from $h_N^{-1}(0)$ to $\Delta(M \times M) \times \{0\}$ and is a local diffeomorphism at each point of $h_N^{-1}(0)$. Therefore, by Lemma 2, ψ is a diffeomorphism from a neighbourhood $\tilde{U} \subseteq U$ of $h_N^{-1}(0)$ to a neighbourhood V of $M \times \{0\}$.

Define $f\colon V \to M$ by $f = \pi_2 \circ \gamma \circ \psi^{-1}$. Then f is C^k. Let $(m, h) \in V$ and n such that $\psi(n) = (m, h)$. Then

$$f(m,h) = (\pi_2 \circ \gamma \circ \psi^{-1})(m,h) = \pi_2(\gamma(n)) = \pi_2(m, \tilde{m}) = \tilde{m},$$

for some $\tilde{m}$. From $\psi(n) = (m, h)$, $h_N(n) = h$. $\tilde{m}$ is unique, since if there is another such $\tilde{m}'$, there would have to be an n' such that

$$\tilde{m}' = \pi_2(m, \tilde{m}') = \pi_2(\gamma(n')) = \pi_2(\gamma(\psi^{-1}(m, h))),$$

where $\psi(n') = (m, h) = \psi(n)$. Since ψ is a diffeomorphism, $n = n'$ implying that $\tilde{m}' = \tilde{m}$. □

Lemma 15. *In the notation and setup of Lemma 14, if* $\gamma^2 = \gamma^1 + O(h^r)$*, then* $f^2 = f^1 + O(h^r)$.

Proof. Since $f^i = \pi_2 \circ \gamma^i \circ (\psi^i)^{-1}$, Lemmas 8 and 9 give the result. □

B.4 Method Order

Definition 20. Let X be a vector field on a manifold Q. The flow of X on Q is a map $F^X : Q \times \mathbb{R} \to Q$ such that

1. $F^X(q, 0) = q$ for all $q \in Q$,
2. $\dfrac{\partial}{\partial t} F(q, t) = X(F(q, t))$.

Definition 21. A *one–step method* for a vector field X on a manifold Q is a map $F : Q \times \mathbb{R} \to Q$ such that

1. $F(q, 0) = q$ for all $q \in Q$,
2. $\left.\dfrac{\partial}{\partial t}\right|_{t=0} F(q, t) = X(q)$.

Definition 22. (Hairer, Lubich, Wanner [14]) The one–step method F for X is said to have *method order p* if the *local error* $\hat{F}^X(q, h) - \hat{F}(q, h)$ satisfies

$$\hat{F}^X(q, h) - \hat{F}(q, h) = O(h^{p+1}),$$

where it is implied that calculations take place in local coordinates.

Lemma 16. *Let (M, h_M) be a manifold such that $M = Q \times \mathbb{R}$ and $h_M(q,h) = h$. Let X be a vector field on Q with flow F^X defined on M and F a one–step method for X also defined on M. Then $F^X(q,h) = F(q,h) + O(h^{p+1})$ if and only if F has method order p.*

Proof. The definition of method order is the local version of the definition of $F^X(q,h) = F(q,h) + O(h^{p+1})$. □

Lemma 17. *Let X be a vector field on Q, F_t^X the flow of X and F_t a one step method for X of method order r. Fix $\tau \in \mathbb{R}$ and let $n \geq 1$ be an integer. Let $h = \tau/n$. Then $F^n_{\tau/n} = F^X_\tau + O(h^r)$.*

Proof. Using Lemma 10, the comment preceding it and for δF_t continuous,

$$
\begin{aligned}
F^n_{\tau/n}(q) = \left(F^X_{\tau/n}\right)^n(q) + \\
&+ h^{r+1}\left(\delta F_{\tau/n}\left(\left(F^X_{\tau/n}\right)^{n-1}(q)\right) + \sum_{i=1}^{n-1} T_{F^X_{\tau/n}(q)}\left(F^X_{\tau/n}\right)^i \delta F_{\tau/n}(q)\right) \\
= F^X_\tau(q) + \\
&+ \tau h^r\left(\frac{1}{n}\delta F_{\tau/n}\left(F^X_{(n-1)\tau/n}(q)\right) + \frac{1}{n}\sum_{i=1}^{n-1} T^X_{F_{\tau/n}(q)}\left(F^X_{\tau/n}\right)^i \delta F_{\tau/n}(q)\right). \quad \text{(B.10)}
\end{aligned}
$$

□

Lemma 17 can be used to verify the method order of a one step method by the procedure of *step doubling* as described in Ascher and Petzold [3]. Let $\mathbf{e}$ be a unit vector and define $y^n = \left(F^n_{\tau/n}(q) - F^{2n}_{\tau/2n}(q)\right) \cdot \mathbf{e}$ and δy^n be the error terms in the parenthesis of Equation (B.10) dotted with $\mathbf{e}$. That is, $y^n = F^X_\tau(q) \cdot \mathbf{e} + \tau h^r \delta y^n$ is the error in the $\mathbf{e}$ component of $F_h(q)$ after n iterations. Compute

$$
\begin{aligned}
y^n - y^{2n} &= F^X_\tau(q) \cdot \mathbf{e} + \tau h^r \delta y^n - F^X_\tau(q) \cdot \mathbf{e} - \tau\left(\frac{h}{2}\right)^r \delta y^{2n} \\
&= \tau h^r\left(\delta y^n - \frac{1}{2^r}\delta y^{2n}\right),
\end{aligned}
$$

and

$$
y^{2n} - y^{4n} = F^X_\tau(q) \cdot \mathbf{e} + \tau\left(\frac{h}{2}\right)^r \delta y^{2n} - F^X_\tau(q) \cdot \mathbf{e} - \tau\left(\frac{h}{4}\right)^r \delta y^{4n}
$$

$$= \tau \left(\frac{h}{2}\right)^r \left(\delta y^{2n} - \frac{1}{2^r}\delta y^{4n}\right).$$

Then

$$\log_2\left(\frac{|y^n - y^{2n}|}{|y^{2n} - y^{4n}|}\right) = \log_2\left(\frac{\tau h^r|\delta y^n - \frac{1}{2^r}\delta y^{2n}|}{\tau \frac{h^r}{2^r}|\delta y^{2n} - \delta y^{4n}|}\right) = \log_2(2^r \epsilon) = r + \log_2 \epsilon,$$

where ϵ is the ratio of the two error differences in the argument of $\log_2$ and is assumed to tend to 1 as $n \to \infty$.

Appendix C

Constraints

This appendix contains a brief review of affine constraints and sets out the relevant notation.

Let Q be a manifold, $\mathcal{D}$ a distribution on Q and Y a vector field on Q. An affine constraint is the tuple $(\mathcal{D}, Y)$. A smooth curve, $\gamma(t)$, in Q is said to satisfy the affine constraint if

$$\frac{d\gamma(t)}{dt} - Y(\gamma(t)) \in \mathcal{D}. \tag{C.1}$$

The affine constraint manifold, $\mathcal{D}_Y$, is the submanifold of TQ defined by

$$\mathcal{D}_Y = \{v \in TQ \,|\, v - Y(\tau_Q(v)) \in \mathcal{D}\}. \tag{C.2}$$

If Y is the zero vector field, then the constraint is linear and $\mathcal{D}_Y = \mathcal{D}$.

The condition in Equation (C.1) is equivalent to

$$\phi(\gamma(t))\,\frac{d\gamma(t)}{dt} - \phi(\gamma(t))\,Y(\gamma(t)) = 0, \qquad \forall \phi \in \mathcal{D}^0.$$

Similarly,

$$v \in \mathcal{D}_Y \Leftrightarrow \phi(\tau_Q(v))\,v - \phi(\tau_Q(v))\,Y(\tau_Q(v)) = 0, \qquad \forall \phi \in \mathcal{D}^0.$$

Let $\{X_i(p)\}_{i=1}^n$ be a set of smooth vector fields forming a local basis for T_pQ such that $\{X_a(p)\}_{a=1}^d$ is a local basis for $\mathcal{D}$ and let $\{\phi^i(p)\}_{i=1}^n$ be the dual basis of T_p^*Q. Then $\{\phi^b(p)\}_{b=d+1}^n$ is a local basis for $\mathcal{D}^0$. Define the functions $c^b \colon TU \to \mathbb{R}$

$$c^b(v) = \phi^b(\tau_Q(v))\,v, \qquad b = d+1, \ldots, n \tag{C.3}$$

and $c \colon TU \to \mathbb{R}^{n-d}$,

$$c(v) = (c^{d+1}(v)\ \ldots, c^n(v)). \tag{C.4}$$

Then, for $Y = Y^i X_i$,

$$v \in \mathcal{D}_Y \Leftrightarrow c^b(v) = Y^b(\tau_Q(v)), \qquad b = d+1, \ldots n.$$

Let (U, μ) be a coordinate chart on Q such that the X_i and ϕ^i are smooth on U. Further, let $(TU, T\mu)$ be the natural coordinate chart of Q with respect to (U, μ) and define the following sets,

$$\begin{aligned} \hat{U} &= \mu(U), \\ T\hat{U} &= T\mu(TU), \\ \hat{\mathcal{D}} &= T\mu(\mathcal{D} \cap TU), \\ \hat{\mathcal{D}_Y} &= T\mu(\mathcal{D}_Y \cap TU) \end{aligned}$$

Let $T\mu(v) = (q, \dot{q})$ and $\tau_Q(v) = p$ so that $\mu(p) = q$. Write the local representations of c^b and c as

$$\begin{aligned} \hat{c}^b(q, \dot{q}) &= (c^b \circ (T\mu)^{-1})(q, \dot{q}), \\ \hat{c}(q, \dot{q}) &= (c \circ (T\mu)^{-1})(q, \dot{q}). \end{aligned}$$

The local representatives of X_i and ϕ^i are defined by

$$\begin{aligned} \hat{X}_i(q) &= \mu_* X_i(\mu^{-1}(q)), \\ \mu^* \hat{\phi}^i(p) &= \phi^i(p), \end{aligned}$$

so that the local representative of Y is

$$\begin{aligned} \hat{Y}(q) &= \mu_*(Y^i(\mu^{-1}(q)) X_i(\mu^{-1}(q))) \\ &= Y^i(\mu^{-1}(q)) \mu_* X_i(\mu^{-1}(q)) \\ &= Y^i(\mu^{-1}(q)) \hat{X}_i(q). \end{aligned}$$

Therefore, $\hat{Y}^i(q) = (Y^i \circ \mu^{-1})(q)$ and

$$(q, \dot{q}) \in \hat{\mathcal{D}}_Y \Leftrightarrow \hat{c}^b(q, \dot{q}) = \hat{Y}^b(q).$$

Appendix D

Local Theorems

Proposition 21 gives a local result for establishing $\partial_h^{\pm}$ as a diffeomorphism to a neighbourhood of the diagonal $\Delta(Q \times Q)$, as in Proposition 3.

Proposition 21. *Let $(\psi, \alpha^+, \alpha^-)$ be a discretization of TQ. For each $q \in Q$ there is a relatively compact neighbourhood U_0 of q, an $a > 0$ and an open submanifold P_h of TU_0 such that*

1. *$(P_h, \partial_h^+, \partial_h^-)$ is a discrete tangent bundle of U_0 for all $h \in (0, a)$,*
2. *$\partial_h^{\pm}$ is a diffeomorphism of P_h to an open neighbourhood of $\Delta(U_0 \times U_0)$.*

Further, for all $v_q \in T_qU_0$, there is an h and P_h such that $v_q \in P_h$.

Proof. Let $\{(U_i, \mu_i)\}$ be an atlas for Q and $\{(TU_i, T\mu_i)\}$ the corresponding natural atlas on TQ. Then $\{(U_i \times U_j \times \mathbb{R}, \mu_i \times \mu_j \times \mathrm{id})\}$ is an atlas for $Q \times Q \times \mathbb{R}$ and $\{(TU_i \times \mathbb{R}, T\mu_i \times \mathrm{id})\}$ is an atlas for $TQ \times \mathbb{R}$. It will be convenient to assume that $\mu_i(U_i) = \mathbb{R}^n$ where $n = \dim Q$. Define the following maps (see Figure D.1):

$$\eta\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R},$$

$$\eta(q^+, q^-, h) = (\alpha^+(h)q^- - \alpha^-(h)q^+, q^+ - q^-, h),$$

$$\rho\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{2n} \times \mathbb{R},$$

$$\rho(q^+, q^-, h) = \left(\frac{1}{h}(q^+, q^-), h\right),$$

$$\partial^{\pm} \times \mathrm{id}\colon TQ \times \mathbb{R} \to Q \times Q \times \mathbb{R},$$

$$(\partial^{\pm} \times \mathrm{id})(v_q, h) = (\partial_h^+(v_q), \partial_h^-(v_q), h)$$

Both ρ and η are nonsingular for $h \neq 0$.

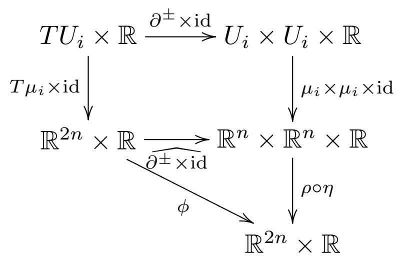

Figure D.1: Mappings for Proposition 21

Let $(v_q, h) \in TU_i \times \mathbb{R}$ such that $(\partial^{\pm} \times \mathrm{id})(v_q, h) \in U_i \times U_i \times \mathbb{R}$ and $T\mu_i v_q = (q, \dot{q})$,

$$(\mu_i \circ \partial_h^+)(v_q) = q + \alpha^+(h)\dot{q} + O(h^2), \tag{D.1}$$

$$(\mu_i \circ \partial_h^-)(v_q) = q + \alpha^-(h)\dot{q} + O(h^2). \tag{D.2}$$

Define

$$\phi \colon \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^{2n} \times \mathbb{R},$$

$$\phi = \rho \circ \eta \circ (\mu_i \times \mu_j \times \mathrm{id}) \circ (\partial_h^{\pm} \times \mathrm{id}) \circ (T\mu_k \times \mathrm{id})^{-1}. \tag{D.3}$$

The mapping ϕ is just $\rho \circ \eta$ composed with a local representative of $\partial^{\pm} \times \mathrm{id}$.

By keeping h small enough, i, j and k may be taken to be the same so that $U_i = U_j = U_k$. Let $q \in Q$ and (U_q, μ_q) a chart for q. Then,

$$\begin{aligned}
\phi((q, \dot{q}), h) &= (\rho \circ \eta \circ (\mu_q \times \mu_q \times \mathrm{id}) \circ (\partial^{\pm} \times \mathrm{id}) \circ (T\mu_q \times \mathrm{id})^{-1})((q, \dot{q}), h) \\
&= (\rho \circ \eta \circ (\mu_q \times \mu_q \times \mathrm{id}) \circ (\partial^{\pm} \times \mathrm{id}))(v_q, h) \\
&= (\rho \circ \eta \circ (\mu_q \times \mu_q \times \mathrm{id}))(\partial_h^+(v_q), \partial_h^-(v_q), h) \quad \text{definition of } \partial_h^{\pm} \\
&= (\rho \circ \eta)(\mu_q(\partial_h^+(v_q)), \mu_q(\partial_h^-(v_q)), h) \\
&= (\rho \circ \eta)(q + \alpha^+(h)\dot{q} + O(h^2), q + \alpha^-(h)\dot{q} + O(h^2), h) \\
&\qquad \text{using Equations (D.1) and (D.2)} \\
&= \rho(hq + O(h^3), h\dot{q} + O(h^2), h) \quad \text{definition of } \eta \\
&= ((q + O(h^2), \dot{q} + O(h)), h). \quad \text{definition of } \rho
\end{aligned}$$

This shows that ϕ is a local diffeomorphism at $((q,\dot{q}),0)$ for every $((q,\dot{q}),0) \in \mathbb{R}^{2n} \times \{0\}$ and a diffeomorphism of $\mathbb{R}^{2n} \times \{0\}$ to $\mathbb{R}^{2n} \times \{0\}$. By Lemma 2, ϕ is therefore a diffeomorphism of a neighbourhood W of $\mathbb{R}^{2n} \times \{0\}$ to a neighbourhood W' of $\mathbb{R}^{2n} \times \{0\}$.

By the definition of ϕ in Equation (D.3), this gives the local representative $(\mu_q \times \mu_q \times \mathrm{id}) \circ (\partial^{\pm} \times \mathrm{id}) \circ (T\mu_q \times \mathrm{id})^{-1}$ as a diffeomorphism of $W \backslash \mathbb{R}^{2n} \times \{0\}$ to $(\rho \circ \eta)^{-1}(W' \backslash \mathbb{R}^{2n} \times \{0\})$ which, in turn, gives open sets $\tilde{V} \subset TU_q \times \mathbb{R}$ and $\tilde{W} \subset U_q \times U_q \times \mathbb{R}$ such that $\partial^{\pm} \times \mathrm{id}$ is a diffeomorphism of $V = \tilde{V} \backslash TU_q \times \{0\}$ to $W = \tilde{W} \backslash U_q \times U_q \times \{0\}$.

Let $U_0 \subset U_q$ be relatively compact. Then there is an $a > 0$ such that $U_0 \times U_0 \times (0,a) \subset W$. Let $h \in (0,a)$ and define

$$P_h = \{v_q \in TU_0 \,|\, (v_q,h) \in V \text{ and } \partial_h^{\pm}(v_q) \in U_0 \times U_0\}.$$

$\partial_h^{\pm}$ is a diffeomorphism on P_h since $\partial^{\pm} \times \mathrm{id}$ is a diffeomorphism on V. In addition, ∂_h^{+} and ∂_h^{-} are onto U_0 since there is always a $v_q \in TU_0$ satisfying $\partial_h^{\pm}(v_q) = (q',q')$ for any $q' \in U_0$. Therefore, P_h is a discrete tangent bundle of U_0.

Let $v_q \in TU_0$. Since $\partial_h^{\pm} \times \mathrm{id}$ is continuous, there is an $h \in (0,a)$ such that P_h contains v_q. □

Proposition 22 is a local version of Proposition 4.

Proposition 22. *Let $(L_h, \psi, \alpha^+, \alpha^-)$ be a discretization of a $\mathcal{D}$–hyperregular Lagrangian system. Then, for each $v_q \in \mathcal{D}$ there is a neighbourhood U_{v_q} of v_q in $\mathcal{D}$ and neighbourhoods U, V satisfying:*

1. *$U_{v_q} \times U_{v_q} \times \mathbb{R} \supseteq U \supseteq \Delta(U_{v_q} \times U_{v_q}) \times \{0\}$ and $U_{v_q} \times \mathbb{R} \supseteq V \supseteq U_{v_q} \times \{0\}$,*
2. *For every $(v_1,h) \in V$ there is a unique $v_2 \in U_{v_q}$ such that $(v_1,v_2,h) \in U$ and (v_1,v_2) satisfy the discrete Euler–Lagrange equations.*

Proof. See the initial comments in the proof of Proposition 4 regarding the singular nature of the discrete Euler–Lagrange equations when $h = 0$.

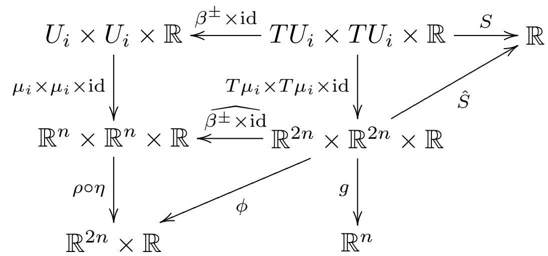

Figure D.2: Diagram of Mappings

See Figure D.2 for a mapping diagram. Let $\{(U_i, \mu_i)\}$ be an atlas for Q. Define the following mappings:

$$\eta \colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R},$$

$$\eta(q^+, q^-, h) = (\alpha^+(h)q^- - \alpha^-(h)q^+, q^+ - q^-, h),$$

$$\rho \colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{2n} \times \mathbb{R},$$

$$\rho(q^+, q^-, h) = \left(\frac{1}{h}(q^+, q^-), h\right),$$

$$\beta^\pm \times \mathrm{id} \colon TQ \times TQ \times \mathbb{R} \to Q \times Q \times \mathbb{R},$$

$$\beta^\pm(v_1, v_2, h) = (\partial_h^+(v_2), \partial_h^-(v_1), h),$$

$$\phi \colon \mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^{2n} \times \mathbb{R},$$

$$\phi((q_1, \dot{q}_1), (q_2, \dot{q}_2), h) = (\rho \circ \eta \circ \widehat{\beta^\pm \times \mathrm{id}})((q_1, \dot{q}_1), (q_2, \dot{q}_2), h),$$

$$g \colon \mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R},$$

$$g((q_1, \dot{q}_1), (q_2, \dot{q}_2), h) = (\hat{\partial}_h^+(q_1, \dot{q}_1) - \hat{\partial}_h^-(q_2, \dot{q}_2), h).$$

The goal is to write down a description of the submanifold $C = \phi^{-1}((q, \dot{q}), 0) \cap g^{-1}(0)$ where $(q, \dot{q})$ is a fixed point in $\mathbb{R}^{2n}$. Recall that the local representatives of ∂_h^+ and ∂_h^- are

$$\hat{\partial}_h^+(q, \dot{q}) = q + \alpha^+(h)\dot{q} + O(h^2),$$

$$\hat{\partial}_h^-(q, \dot{q}) = q + \alpha^-(h)\dot{q} + O(h^2).$$

Then

$$g((q_1,\dot{q}_1),(q_2,\dot{q}_2),h) = (\hat{\partial}_h^+(q_1,\dot{q}_1) - \hat{\partial}_h^-(q_2,\dot{q}_2),h)$$

$$= (q_2 + \alpha^+(h)\dot{q}_2 - q_1 - \alpha^-(h)\dot{q}_1 + O(h^2),h).$$

And

$$\phi((q_1,\dot{q}_1),(q_2,\dot{q}_2),h) = (\rho\circ\eta\circ\widehat{\beta^\pm\times \mathrm{id}})((q_1,\dot{q}_1),(q_2,\dot{q}_2),h)$$

$$= (\rho\circ\eta)(\hat{\partial}_h^+(q_2,\dot{q}_2),\hat{\partial}_h^-(q_1,\dot{q}_1),h) \quad \text{definition of } \beta^\pm$$

$$= \rho(\alpha^+(h)\hat{\partial}_h^-(q_1,\dot{q}_1) - \alpha^-(h)\hat{\partial}_h^+(q_2,\dot{q}_2),\hat{\partial}_h^+(q_2,\dot{q}_2) - \hat{\partial}_h^-(q_1,\dot{q}_1),h)$$

$$\text{definition of } \eta$$

$$= \left(\frac{1}{h}\left(\alpha^+(h)\hat{\partial}_h^-(q_1,\dot{q}_1) - \alpha^-(h)\hat{\partial}_h^+(q_2,\dot{q}_2),\hat{\partial}_h^+(q_2,\dot{q}_2) - \hat{\partial}_h^-(q_1,\dot{q}_1)\right),h\right). \tag{D.4}$$

$$\text{definition of } \rho$$

Then $g((q_1,\dot{q}_1),(q_2,\dot{q}_2),h) = (0,h)$ gives $q_2 = q_1 + \alpha^+(h)\dot{q}_1 - \alpha^-(h)\dot{q}_2 + O(h^2)$. Use this and the local definitions of $\hat{\partial}_h^+$ and $\hat{\partial}_h^-$ in Equation (D.4) to obtain, on $g^{-1}(0,h)$,

$$\phi((q_1,\dot{q}_1),(q_2,\dot{q}_2)),h) = ((q_1 - \alpha^-(h)\dot{q}_2 + O(h^2),\dot{q}_1 + \dot{q}_2 + O(h)),h). \tag{D.5}$$

Write $\mathcal{D}_i = U_{v_q}\cap TU_i$ and $\hat{\mathcal{D}} = T\mu_i(\mathcal{D}_i)$. The infinitesimal variations must be in $\ker\{D\phi, Dg\}$ on $g^{-1}(0,h)$ and $D\hat{\partial}_h^+(q_1,\dot{q}_1)\,(\delta q_1,\delta\dot{q}_1) \in \hat{\mathcal{D}}$. This is

$$\delta q_1 - \alpha^-(h)\delta\dot{q}_2 + O(h^2) = 0, \tag{D.6}$$

$$\delta\dot{q}_1 + \delta\dot{q}_2 + O(h) = 0, \tag{D.7}$$

$$\delta q_2 - \delta q_1 - \alpha^+(h)\delta\dot{q}_1 + \alpha^-(h)\delta\dot{q}_2 + O(h^2) = 0, \tag{D.8}$$

$$\delta q_1 + \alpha^+(h)\delta\dot{q}_1 + O(h^2) \in \hat{\mathcal{D}}. \tag{D.9}$$

Equations (D.6), (D.7) and (D.8) give

$$\delta q_1 = -\alpha^-(h)\delta\dot{q}_1 + O(h^2), \tag{D.10}$$

$$\delta q_2 = \alpha^+(h)\delta\dot{q}_1 + O(h^2), \tag{D.11}$$

$$\delta\dot{q}_2 = -\delta\dot{q}_1 + O(h). \tag{D.12}$$

Using Equation (D.10) in Equation (D.9) gives

$$\delta\dot{q}_1 + O(h) \in \hat{\mathcal{D}}. \tag{D.13}$$

At $h = 0$, the infinitesimal variations are therefore

$$\delta q_1 = \delta q_2 = 0, \qquad \delta\dot{q}_2 = -\delta\dot{q}_1, \qquad \delta\dot{q}_1 \in \hat{\mathcal{D}}. \tag{D.14}$$

Define, for $(q, \dot{q}) \in \mathbb{R}^{2n}$,

$$\mathcal{B}_{(q,\dot{q})} = \phi^{-1}((q,\dot{q}),0) \cap g^{-1}(0) = \{((q_1,\dot{q}_1),(q_2,\dot{q}_2)) \,|\, q_1 = q_2 = q, \dot{q}_1 + \dot{q}_2 = \dot{q}\} \times \{0\}$$

$\mathcal{B}_{(q,\dot{q})}$ is an n dimensional affine subspace of $\mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R}$ and the infinitesimal variations are constrained to be in $T\mathcal{B}_{(q,\dot{q})} \cap \hat{\mathcal{D}}$.

Choose the vector field Y such that $\phi Y \neq 0$ for all $\phi \in \mathcal{D}^0$, the annihilator of $\mathcal{D}$ and write $\hat{Y}$ for the local representation of Y. Write $(\mathcal{D}_Y)_i = \mathcal{D}_Y \cap TU_i$. Then,

$$\hat{\mathcal{D}}_Y \times \hat{\mathcal{D}}_Y \times \mathbb{R} = (T\mu_i \times T\mu_i \times \mathrm{id})((\mathcal{D}_Y)_i \times (\mathcal{D}_Y)_i \times \mathbb{R}).$$

Then,

$$\begin{aligned}
\mathcal{C}_{(q,\dot{q})} &= \hat{\mathcal{D}}_Y \times \hat{\mathcal{D}}_Y \times \mathbb{R} \cap C_{(q,\dot{q})} \\
&= \{((q_1,\dot{q}_1),(q_2,\dot{q}_2)) \in \mathcal{B}_{(q,\dot{q})} \,|\, (q_1,\dot{q}_1) \in \hat{\mathcal{D}}_Y, (q_2,\dot{q}_2) \in \hat{\mathcal{D}}_Y\} \times \{0\} \\
&= \{((q,\dot{q}_1),(q,\dot{q}-\dot{q}_1)) \,|\, \dot{q}_1 - \hat{Y}(q) \in \hat{\mathcal{D}}_q, \dot{q} - \dot{q}_1 - \hat{Y}(q) \in \hat{\mathcal{D}}_q\} \times \{0\} \\
&= \{((q,\dot{q}_1),(q,\dot{q}-\dot{q}_1)) \,|\, \dot{q}_1 - \hat{Y}(q) \in \hat{\mathcal{D}}_q, \dot{q} \in \hat{\mathcal{D}}_q\} \times \{0\}.
\end{aligned}$$

This gives $\mathcal{C}_{(q,\dot{q})}$ as a d dimensional affine subspace of $\mathbb{R}^{2n} \times \mathbb{R}^{2n} \times \mathbb{R}$. Solutions to the discrete Euler–Lagrange equations must be in $\mathcal{C}_{(q,\dot{q})}$.

The local version of the discrete Lagrange–D'Alembert principle is then, for $Y = 0$ and on $g^{-1}(0,0)$,

$$d\hat{L}(q_1,\dot{q}_1)\,(0,\delta\dot{q}_1) + d\hat{L}(q_1,\dot{q}_2)\,(0,-\delta\dot{q}_1) = 0,$$

$$(q_1, \dot{q}_1), (q_1, \dot{q}_2) \in \hat{\mathcal{D}},$$

$$\delta\dot{q}_1 \in \hat{\mathcal{D}},$$

Since L is assumed to be $\mathcal{D}$ hyperregular, these equations have solution $\dot{q}_2 = \dot{q}_1$.

On $g^{-1}(0,0)$, $\phi((q,\dot{q}),(q,\dot{q}),0) = (q, 2\dot{q}, 0)$, so there is a diffeomorphism $f_0 : \hat{\mathcal{D}} \times \{0\} \to \Delta(\hat{\mathcal{D}} \times \hat{\mathcal{D}}) \times \{0\}$ given by $f_0(q,\dot{q},0) = \left(\left(q, \tfrac{1}{2}\dot{q}\right), \left(q, \tfrac{1}{2}\dot{q}\right), 0\right)$. Every $(q,\dot{q},0)$ is a regular value of ϕ since

$$\phi^{-1}(q,\dot{q},0) = \{((q_1,\dot{q}_1),(q_2,\dot{q}_2),0) \,|\, q_1 = q_2 = q, \dot{q}_1 + \dot{q}_2 = \dot{q}\}.$$

And since L is $\mathcal{D}$–hyperregular, Lemma 5 gives open neighbourhoods $\hat{U}$ and $\hat{V}$ such that $\hat{\mathcal{D}} \times \hat{\mathcal{D}} \times \mathbb{R} \supset \hat{U} \supset \Delta(\hat{\mathcal{D}} \times \hat{\mathcal{D}}) \times \{0\}$ and $\hat{\mathcal{D}} \times \mathbb{R} \supset \hat{V} \supset \hat{\mathcal{D}} \times \{0\}$ and for every $((q,\dot{q}),h) \in \hat{V}$ there is a unique $((q_1,\dot{q}_1),(q_2,\dot{q}_2),h) \in \hat{U}$ that satisfies the local discrete Euler–Lagrange equations.

The sets U and V of the theorem are then

$$U = (T\mu_i^{-1} \times T\mu_i^{-1} \times \mathrm{id})(\hat{U}), \qquad V = (T\mu_i^{-1} \times \mathrm{id})(\hat{V}).$$

□

Since $\phi((q_1,\dot{q}_1),(q_2,\dot{q}_2),h) = ((q,\dot{q}),h)$, the definition of ϕ in Equation (D.4) gives, on $g^{-1}(0,h)$,

$$q^+ = \hat{\partial}_h^+(q_2,\dot{q}_2) = q + \frac{\alpha^+(h)}{h}\dot{q}, \qquad q^- = \hat{\partial}_h^-(q_1,\dot{q}_1) = q + \frac{\alpha^-(h)}{h}\dot{q}.$$

This gives a local picture of what the admissible endpoints are.

References

[1] R. Abraham and J. Marsden. *Foundations of Mechanics*. Benjamin/Cummings, 1978.

[2] R. Abraham, J. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer-Verlag, 1988.

[3] U. Ascher and R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential–Algebraic Equations*. Society for Industrial and Applied Mathematics, 1998.

[4] L. Bates and R. Cushman. *Global Aspects of Classical Integrable Systems*. Birkhäuser Verlag, 1997.

[5] L. Bates and J. Sniatycki. Nonholonomic reduction. *Rep. Math. Phys.*, 32:99–115, 1999.

[6] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, and R.M. Murray. Nonholonomic systems with symmetry. *Arch. Rational Mech. Anal.*, 136:21–99, 1996.

[7] J. Cortés and S. Martínez. Non-holonomic integrators. *Nonlinearity*, 14:1365–1392, 2001.

[8] C. Cuell. An electric circuit analogue of a constrained mechanical system. *IEEE Transactions on Circuits and Systems I*, 48:1114–1118, 2000.

[9] C. Cuell and G. Patrick. Skew critical problems. *Submitted Reg. Chaotic Dyn.*, 2007.

[10] R. Cushman and J. Sniatycki. A nonholonomic oscillator. *Rep. Math. Phys.*, 50:85–98, 2002.

[11] M. de León, D. Martín de Diego, and A. Santamaría-Merino. Geometric integrators and nonholonomic mechanics. *Journal of Mathematical Physics*, 45(3):1042–1064, 2004.

[12] I. Gelfand and S. Fomin. *Calculus of Variations*. Prentice-Hall, 1963.

[13] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics*. Addison–Wesley, 2002.

[14] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer–Verlag, 2002.

[15] Hermans Joost. *Rolling Rigid Bodies with and without Symmetries*. PhD thesis, University of Utrecht, 1995.

[16] S. Lang. *Differential and Riemannian Manifolds*. Springer–Verlag, 1995.

[17] J. Marsden, G. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators and nonlinear PDEs. *Comm. Math. Phys.*, 199:351–395, 1998.

[18] J. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, pages 357–514, 2001.

[19] R. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. *J. Nonlin. Sci.*, 16(4):1432–1467, 2006.

[20] G. Patrick. Variational development of the semi–symplectic geometry of nonholonomic mechanics. *Accepted Rep. Math. Phys.*, 2006.

[21] T. Ratiu and J. Marsden. *Introduction to Mechanics and Symmetry*. Springer–Verlag, 1999.

[22] R. Rosenberg. *Analytical Dynamics of Discrete Systems*. Plenum Press, 1977.

[23] H. Sussmann. Orbits of families of vector fields and integrability of distributions. *Transactions of the AMS*, 180, 1973.

[24] A. Walther, A. Griewank, and J. Utke. *ADOL-C. A package for automatic differentiation of algorithms written in C++*.

[25] Robert H. Wasserman. *Tensors and Manifolds*. Oxford University Press, 1992.