UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS
Departamento de Geometría y Topología

TESIS DOCTORAL

New developments and applications of the inverse problem of the calculus of variations

Nuevos desarrollos y aplicaciones del problema inverso del cálculo de variaciones

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Marta Farré Puiggalí

Directores

David Martín de Diego
María Barbero Liñán

Madrid, 2018

Doctoral Thesis

New developments and applications of the Inverse Problem of the Calculus of Variations.
Nuevos desarrollos y aplicaciones del problema inverso del cálculo de variaciones.

UNIVERSIDAD COMPLUTENSE MADRID

Facultad de Ciencias Matemáticas
Departamento de Geometría y Topología

ICMAT INSTITUTO DE CIENCIAS MATEMÁTICAS

Instituto de Ciencias Matemáticas
(CSIC-UAM-UC3M-UCM)

Author:
Marta Farré Puiggalí

Advisor:
David Martín de Diego
Co-advisor:
María Barbero Liñán

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

January 13, 2017

Acknowledgements

En primer lugar quiero agradecer a mi director de tesis David Martín de Diego, mentor, amigo y ejemplo, por todo el tiempo y esfuerzo dedicados a la realización de esta tesis, por todas las oportunidades, enseñanzas y buen humor. Han sido unos años felices, y en gran parte te lo debo a ti.

En segundo lugar a mi co-directora María Barbero Liñán por toda la ayuda, energía e iniciativas mostradas a lo largo de estos años, y por encontrar siempre un hueco a pesar de tantas cosas.

I am deeply thankful to my past and current collaborators Anthony M. Bloch, Elena Celledoni, Eirik Hoel Hoiseth, Sebastián Ferraro and Tom Mestdag. Our discussions were extremely helpful and it has been a privilege to work together.

During these years I had the opportunity to visit many places. Everywhere I found welcoming people who contributed to make my stays abroad pleasant. I would like to thank all of them. I feel particularly grateful to Professors Anthony M. Bloch, Melvin Leok, Tom Mestdag and Geoff Prince for hosting me to their institutions, for the hospitality and all the time devoted to discussions. I learned a lot from them. I would also like to thank Frans Cantrijn and Willy Sarlet for their kindness and hospitality during my visit to Ghent.

I am also grateful to Frans Cantrijn and Miguel Muñoz Lecanda for refereeing this thesis.

Quiero agradecer a la red de Geometría, Mecánica y Control todo el empeño puesto en dar oportunidades a los jóvenes, organizando eventos en los cuales estos son los protagonistas. Especialmente a Juan Carlos Marrero, Edith Padrón, Manuel de León y Eduardo Martínez.

Agradezco al ICMAT el proporcionar a lo largo de estos años un ambiente ideal de trabajo. En particular a todo el grupo de mecánica geométrica y muy especialmente a Leo que desde el principio adoptó el papel de hermano mayor académico. Gracias por todos los consejos, buena disposición y por tu acogida en Michigan. A Cédric, por animar siempre la cuarta planta mientras estuvo aquí. Y a Miguel por seguir haciéndolo hasta ahora. A Javi y David por sus consejos y apoyo cuando llegué. I al Roger pels ànims, l'energia i la decisió. Quiero agradecer también a Ángel las lecciones básicas en campos desconocidos. A los compañeros de despacho y a los de almuerzo, en definitiva a toda la gente que de una forma u otra contribuye a crear un ambiente agradable.

Al departamento de geometría y topología de la Universidad Complutense de Madrid, en particular a Elena Martín Peinador con quien tuve el placer de compartir mi primera experiencia docente. A Marco Castrillón por ponerme en contacto con David, y por todo lo que él mismo me enseñó, con gran entusiasmo, durante mi etapa de máster.

Anant més enrere aprofito per agrair el suport de Pilar Bayer i Carles Casacuberta en moments clau. Y la confianza y consejo de María Jesús Carro y Javier Soria.

Tot comença a la Universitat de Barcelona, on vaig tenir la immensa sort de compartir el camí amb la Celia, el David, l'Elba i la Laia. Gràcies per tantes hores junts.

Finalment a la meva família per ser-hi sempre. Y a Álvaro.

Contents

Resumen

Gran parte del trabajo incluido en esta tesis tiene un tema común que es el problema inverso del cálculo de variaciones. De manera concisa, este problema inverso se refiere a si un sistema de ecuaciones diferenciales ordinarias de segundo orden (SODE para abreviar) es equivalente a un sistema Lagrangiano regular. Este problema se remonta a finales del siglo XIX, momento en el que sólo se comprendía completamente el caso unidimensional. Cuarenta años más tarde, el medallista Fields J. Douglas dio una clasificación para sistemas bidimensionales [55]. Después de esto no se ha clasificado completamente ninguna otra dimensión, pero se ha conseguido una comprensión geométrica más profunda del problema gracias a las contribuciones de varias personas incluyendo G. Prince, W. Sarlet, M. Crampin, I. Anderson y G. Thompson [49, 144], que hicieron posible la extensión de algunos de los casos de la clasificación de Douglas a dimensiones arbitrarias [3, 47].

Las condiciones de Helmholtz son un conjunto bien conocido de ecuaciones algebraicas y ecuaciones en derivadas parciales que son necesarias y suficientes para que una SODE sea variacional, es decir, equivalente a un sistema Lagrangiano regular. Estas condiciones vienen dadas en términos de una matriz de multiplicadores, que corresponde a la matriz Hessiana del Lagrangiano buscado con respecto a las velocidades, y fueron utilizadas por Douglas para describir su clasificación. Hay un teorema de M. Crampin que caracteriza el problema en términos de la existencia de una 2-forma de Poincaré-Cartan. Nos referiremos a este resultado como Teorema de Crampin [43].

El término mecánica geométrica se refiere a una variedad de temas que se encuentran en la intersección de la geometría diferencial, los sistemas dinámicos, tanto discretos como continuos, y la mecánica analítica. Algunas referencias clásicas son los libros de R. Abraham y J. E. Marsden [1], el libro de V. I. Arnold [7], y enfatizando el papel de la simetría, el libro de J. E. Marsden y T. S. Ratiu [111]. El problema inverso es el hilo principal de esta tesis, mientras se consideran algunos temas centrales de la mecánica geométrica, como los sistemas noholónomos y el problema de Hamiltonización, la mecánica Lagrangiana en algebroides de Lie, la estabilización de sistemas mecánicos mediante controles apropiados y la mecánica discreta, en particular integradores geométricos. Ahora daremos una breve introducción a cada uno de estos temas.

Sistemas noholónomos

Los sistemas noholónomos son sistemas dinámicos en el fibrado tangente de un variedad, con restricciones en las velocidades que usualmente se definen por una distribución no integrable. Ejemplos de sistemas noholónomos son, entre otros, un disco que rueda sobre una superficie sin resbalar o un rattleback. La propiedad clave de las ligaduras noholónomas es que no pueden reducirse a ligaduras sobre las posibles configuraciones del sistema, en contraste con las ligaduras holónomas.

Los sistemas noholónomos tienen una amplia bibliografía y una historia de confusión con los sistemas vakónomos [102]. Ahora bien, es bien sabido que los sistemas noholónomos no son variacionales, pero una pregunta abierta es tratar de caracterizar cuándo tales sistemas pueden ser estudiados desde un punto de vista Hamiltoniano. Estos intentos pueden denominarse como Hamiltonización de sistemas noholónomos. Uno de los enfoques consiste en intentar poner las trayectorias del sistema noholónomo como la restricción a la subvariedad de ligaduras de las trayectorias de un sistema Lagrangiano [17]. Otros enfoques incluyen Hamiltonización de Chaplygin [10, 65].

Integradores geométricos

Encontrar integradores numéricos que posean las mismas propiedades geométricas que el sistema continuo original, como por ejemplo simplecticidad para un sistema Hamiltoniano [112], puede ser muy útil para obtener un comportamiento cualitativo adecuado. Algunas propiedades interesantes son la preservación de constantes del movimiento, volumen, simetrías o ligaduras entre otras. Los integradores numéricos que preservan estas propiedades se conocen como integradores geométricos [77, 92]. Hay una amplia literatura sobre integradores geométricos para sistemas Hamiltonianos, por ejemplo el libro de J. M. Sanz-Serna y M. P. Calvo [139], pero no tanto para sistemas noholónomos [39, 63, 117].

Mecánica Lagrangiana en algebroides de Lie

Se puede pensar en un algebroide de Lie como una generalización de un fibrado tangente. El uso de funciones Lagrangianas definidas en algebroides de Lie permite incluir bajo un mismo esquema la dinámica de sistemas mecánicos para una variedad de situaciones, como los sistemas mecánicos reducidos por la simetría de un grupo de Lie. Por ejemplo las ecuaciones de Euler-Poincaré y las ecuaciones de Lagrange-Poincaré se obtienen como ecuaciones de Euler-Lagrange en una álgebra de Lie y en un algebroide de Atiyah, respectivamente. El estudio de la mecánica Lagrangiana en algebroides de Lie tiene su precursor en A. Weinstein [165]. Otras referencias son [51, 113]. Incluso los sistemas noholónomos se pueden estudiar en este contexto si eliminamos la identidad de Jacobi y nos restringimos a algebroides anti-simétricos [37, 50].

Control geométrico

Los sistemas dinámicos se pueden modelar incluyendo fuerzas externas (los controles) que se pueden escoger para conseguir algún objetivo, por ejemplo la estabilización de un equilibrio inestable o llegar a una posición deseada con el mínimo coste posible. La geometría diferencial ha proporcionado muchas herramientas para el diseño y estudio de controles para sistemas controlados no lineales. Muchos de estos desarrollos fueron impulsados por R. W. Brockett [25], como también por A. M. Bloch [14], V. Jurdjevic [95], H. Nijmeijer y A. J. Van der Schaft [127], E. D. Sontag [149] y H. J. Sussmann [151]. Nosotros nos centraremos en el problema de estabilización de un equilibrio inestable para una cierta clase de sistemas mecánicos con simetría.

Daremos una introducción más completa a estos temas a lo largo del texto, según vayan apareciendo. A continuación detallamos una lista de las publicaciones y preprints en los que se basa esta tesis, una breve descripción de los resultados obtenidos en cada uno, y también un esquema de esta tesis relacionando cada capítulo con los contenidos de la lista.

Publicaciones y preprints

P1-[11]. María Barbero-Liñán, Marta Farré Puiggalí, David Martín de Diego: *Isotropic submanifolds and the inverse problem for mechanical constrained systems.* J. Phys. A: Math. Theor. 48 (2015) 045203 (35pp).

P2-[12]. María Barbero Liñán, Marta Farré Puiggalí, David Martín de Diego: *Inverse problem for Lagrangian systems on Lie algebroids and applications to reduction by symmetries.* Monatsh. Math. 180 (2016), no. 4, 665-691.

P3-[57]. Marta Farré Puiggalí, Tom Mestdag: *The inverse problem of the calculus of variations and the stabilization of controlled Lagrangian systems.* SIAM J. Control Optim. 54-6 (2016), pp. 3297-3318.

P4-[32]. Elena Celledoni, Marta Farré Puiggalí, Eirik Hoel Høiseth, David Martín de Diego: *Energy-preserving integrators applied to nonholonomic systems.* arXiv:1605.02845.

P5. María Barbero Liñán, Marta Farré Puiggalí, Sebastián Ferraro, David Martín de Diego: *Inverse problem of the calculus of variations for discrete systems.*

P6. Anthony M. Bloch, Marta Farré Puiggalí: *New matching conditions from Helmholtz conditions.*

Breve descripción de los resultados

Problema inverso para sistemas con ligaduras

En el primer artículo P1-[11] presentamos una nueva caracterización del problema inverso en términos de subvariedades Lagrangianas, tanto para el caso autónomo como el no autónomo. La caracterización se da en términos de la existencia de la transformación de Legendre, en lugar de la existencia de una matriz de multplicadores. Este enfoque nuevo permite una formulación geométrica del problema inverso en el contexto de sistemas con ligaduras. Damos un noción de variacionalidad para una SODE en una subvariedad del fibrado tangente simplemente reemplazando subvariedad Lagrangiana por subvariedad isótropa.

Esta noción nos permite demostrar un análogo del Teorema de Crampin para SODEs con ligaduras. También nos permite comprobar si un sistema noholónomo es un subsistema de un sistema Lagrangiano, para una función Lagrangiana que podría ser degenerada. Esto está relacionado con una de las aproximaciones al problema de Hamiltonización de sistemas noholónomos [17, 120].

Problema inverso en algebroides de Lie

En el segundo artículo P2-[12] somos capaces de extender la definición del problema inverso al contexto de SODEs en algebroides de Lie regulares usando prolongados de algebroides de Lie y subvariedades Lagrangianas de algebroides de Lie simplécticos. La diferencia clave con P1-[11] es que para un algebroide de Lie general el Lema de Poincaré no se cumple. Reobtenemos las condiciones de Helmholtz para algebroides de Lie presentes en la literatura pero además identificamos su insuficiencia para garantizar la existencia de un Lagrangiano local como la ausencia del Lema de Poincaré. Entonces podemos encontrar una condición adicional que proporciona un conjunto de condiciones necesarias y suficientes, todas en función de las componentes de la transformación de Legendre como incógnita. También estudiamos el comportamiento de la propiedad variacional bajo morfismos de algebroides de Lie y finalmente mostramos algunos ejemplos y comparamos con anteriores aproximaciones.

Problema inverso y estabilización de sistemas mecánicos

En [21] A. M. Bloch, N. E. Leonard y J. E. Marsden presentaron una estrategia para construir controles explícitos de forma que estabilicen ciertos sistemas mecánicos con simetría. La idea consiste en modificar el Lagrangiano del sistema mecánico dado de un modo predeterminado para conseguir que el sistema Lagrangiano controlado original sea equivalente al sistema Lagrangiano correspondiente a la nueva función Lagrangiana, que dependerá de algunos parámetros. Las *matching conditions* son condiciones suficientes para lograr este objetivo, y se dispone de una variedad de ellas para atacar distintas situaciones, incluyendo una rotura de simetría para la energía potencial [15].

Subyacente a este proceso se encuentra el problema inverso del cálculo de variaciones. Observemos que, en lugar de una SODE, tenemos un familia de SODEs con los controles desconocidos como parámetros, y el objetivo consiste en encontrar parámetros adecuados para que la SODE controlada sea variacional. Por ello en el artículo P3-[57] usamos la clasificación de Douglas para conseguir nuevos controles estabilizadores para una clase de sistemas mecánicos de dimensión dos. La diferencia principal en la estrategia respecto a los resultados anteriores es que ahora no imponemos ninguna forma específica para el nuevo Lagrangiano, sino que trabajamos con una clase concreta de controles. De momento esto se restringe a dos dimensiones ya que se basa en los resultados de Douglas.

Integradores nonholónomos energía-preservantes

En contraste con los sistemas Hamiltonianos, los sistemas noholónomos no poseen necesariamente ninguna medida invariante. Por otro lado, la energía siempre es una cantidad conservada para los sistemas noholónomos. Por eso, en lugar de intentar adaptar integradores variacionales al caso noholónomo, nos hemos centrado en P4-[32] en el estudio de integradores energía-preservantes para sistemas noholónomos.

Diremos que un sistema de ODEs $\dot{x} = \Pi(x)\nabla H(x)$, con $\Pi(x)$ una matriz anti-simétrica, está dado en forma de *skew gradient*. Un gradiente discreto $\bar{\nabla} H : \mathbb{R}^{2N} \longrightarrow \mathbb{R}^N$ es una aproximación al gradiente de H, que cumple $\bar{\nabla} H(x,x')^T(x'-x) = H(x') - H(x)$, y $\bar{\nabla} H(x,x) = \nabla H(x)$, para todo $x, x' \in \mathbb{R}^N$. Entonces se puede definir un integrador energía-preservante como

$$\frac{x'-x}{h} = \tilde{\Pi}(x,x',h)\bar{\nabla} H(x,x')\,, \tag{1}$$

donde $\tilde{\Pi}(x,x',h)$ es una aproximación anti-simétrica de $\Pi(x)$.

En P4-[32] usamos el hecho de que las ecuaciones del movimiento para un sistema noholónomo se pueden escribir en el dual de la distribución de ligaduras, en forma *skew gradient*. Por lo tanto podemos aplicar integradores del tipo (1), que automáticamente preservarán tanto las ligaduras como la energía.

Además, para facilitar la aplicación de estos integradores a sistemas complejos que no se puedan poner fácilmente en la forma *skew gradient* mencionada, conseguimos reescribir los integradores usando sólo la información inicial del sistema noholónomo. Finalmente testeamos los integradores resultantes en distintos ejemplos.

Problema inverso discreto

P5 es un proyecto en proceso con María Barbero Liñán, Sebastián Ferraro y David Martín de Diego. Usando otra vez subvariedades Lagrangianas e isótropas conseguimos extender el problema inverso al caso de mecánica discreta libre y con ligaduras. Ahora el espacio de estados es dos copias de la variedad de configuraciones, que sustituye al fibrado tangente. En este contexto será clave tratar con sistemas implícitos. Nuestro interés principal en esta extensión son las posibles aplicaciones al estudio de integradores geométricos, ya que tendremos una noción de variacionalidad para sistemas con ligaduras tanto continuos como discretos. Si el sistema original es variacional con ligaduras entonces podemos estudiar las posibles ventajas de usar un integrador variacional con ligaduras discreto, y compararlos con otros integradores geométricos para sistemas noholónomos.

Matching usando el problema inverso I

P6 es un proyecto en proceso con Anthony M. Bloch. Usando las condiciones de Helmholtz conseguimos recuperar las *matching conditions* de [21], ya que las condiciones de Helmholtz son necesarias y suficientes para la existencia de un Lagrangiano regular. Además podemos derivar nuevas condiciones de *matching* para una clase particular de sistemas mecánicos. Resulta que para esta clase de sistemas obtenemos controles que sólo dependen de las variables de configuración.

Esquema de la tesis

- Introducción
- Capítulo 1: Primero damos una introducción al problema inverso del cálculo de variaciones, recordando algunos de los principales resultados, con especial énfasis en la aproximación geométrica. En la Sección 1.5 presentamos una nueva caracterización del problema inverso, introducida en P1-[11]. También en esta sección incluimos la versión implícita del problema dada en P5.

- Capítulo 2: Presentamos algunas aplicaciones del problema inverso a la teoría de control, más concretamente a la estabilización de sistemas mecánicos. El capítulo se divide en dos partes, la primera se basa en P3-[57] y proporciona una aplicación de la clasificación de Douglas a la estabilización de una clase de sistemas mecánicos de dimensión dos. La segunda parte se basa en P6 y describe cómo reobtener explícitamente las *matching conditions* a partir de las condiciones de Helmholtz y cómo obtener condiciones de *matching* nuevas para una clase particular de sistemas mecánicos.

- Capítulo 3: Este capítulo se basa en P2-[12]. Primero damos una introducción a la teoría de algebroides de Lie y a la mecánica Lagrangiana en algebroides de Lie para luego extender el problema inverso a este contexto.

- Capítulo 4: Este capítulo contiene parte del trabajo desarrollado en P5. Primero damos una introducción a la mecánica discreta y a los integradores variacionales. Luego presentamos una extensión del problema inverso a sistemas discretos, tanto en el caso explícito como implícito, usando otra vez subvariedades Lagrangianas. Finalmente probamos un análogo del Teorema de Crampin en este contexto.

- Capítulo 5: Los resultados presentados en este capítulo se basan en P1-[11] y P5. Primero damos una introducción a los sistemas con ligaduras, con especial énfasis en los sistemas noholónomos. Luego presentamos una extensión del problema inverso a sistemas con ligaduras tanto en el caso continuo (P1-[11]) como en el discreto (P5). También demostramos las correspondientes versiones del Teorema de Crampin para cada caso.

- Capítulo 6: Este capítulo se basa en P4-[32]. Los sistemas noholónomos no son variacionales en el sentido habitual de la palabra (no confundir con la noción de SODE restringida variacional que se introduce en el Capítulo 5). Por lo tanto, en general no deberíamos utilizar un integrador variacional para integrar un sistema noholónomo. Hay diversos métodos diseñados específicamente para integrar sistemas noholónomos. En este capítulo proponemos usar integradores energía-preservantes para sistemas noholónomos. Comparamos su comportamiento con otros métodos estándar en mecánica noholónoma y comprobamos sus ventajas en distintos ejemplos.

- Capítulo 7: Damos un esquema de posibles direcciones para trabajo futuro, que surgen de los resultados obtenidos en esta tesis.

- Apéndice A: Introducimos dos elementos clave de este trabajo, las subvariedades Lagrangianas y las subvariedades isótropas. También recordamos una construcción clave que permite extender una subvariedad isótropa a una subvariedad Lagrangiana.

- Apéndice B: Por completitud incluimos el triple de Tulczyjew, que se usa en los Capítulos 1 y 5.

- Apéndice C: Proporcionamos cálculos directos que muestran la equivalencia entre las condiciones clásicas de Helmholtz escritas en términos de los multiplicadores y las condiciones de Helmholtz introducidas en el Capítulo 1 dadas en términos de la transformación de Legendre.
- Apéndice D: Mostramos un algoritmo que permite usar los integradores descritos en el Capítulo 6 sin tener que calcular la formulación casi-Poisson del sistema noholónomo.

Introduction

Much of the work included in this dissertation has a gluing topic which is the inverse problem of the calculus of variations. In brief, this inverse problem addresses the question of whether a system of second order ordinary differential equations (SODE for short) is equivalent to a regular Lagrangian system. This problem dates back to the end of the 19th century, at which time only the one-dimensional case was understood. Forty years latter, the Fields medalist J. Douglas gave a classification for two-dimensional systems [55]. After that no other dimension has been classified, but deeper geometric understanding of the problem was gained by various researchers including G. Prince, W. Sarlet, M. Crampin, I. Anderson and G. Thompson [49, 144], which allowed the extension of some of the cases in Douglas' classification to arbitrary dimension [3, 47].

The Helmholtz conditions are a well-known set of algebraic and PDE equations which are necessary and sufficient for a SODE to be variational, that is, equivalent to a regular Lagrangian system. These conditions are written in terms of a multiplier matrix which corresponds to the Hessian of the sought Lagrangian with respect to the velocities, and they were used by Douglas in order to derive his classification. There is a theorem by M. Crampin which characterizes the problem in terms of the existence of a Poincaré-Cartan two-form. We will refer to this result as Crampin's Theorem [43].

Geometric mechanics refers to a variety of topics that lie at the intersection of differential geometry, dynamical systems, both discrete and continuous, and analytical mechanics. Some classical references are the books by R. Abraham and J. E. Marsden [1] and V. I. Arnold [7], and emphasizing the role of symmetry, the book by J. E. Marsden and T. S. Ratiu [111]. The inverse problem is the leading thread of this thesis, but it runs through some central issues in Geometric Mechanics, namely nonholonomic systems and the Hamiltonization problem, Lagrangian mechanics on Lie algebroids, stabilization of mechanical systems using appropriate controls and discrete mechanics, in particular geometric integrators. We will now give a brief introduction to each one of these topics.

Nonholonomic systems

Nonholonomic systems are dynamical systems on the tangent bundle of a manifold with constraints in the velocities that are usually defined by a non-integrable distribution. Examples of nonholonomic systems are for instance a disk rolling on a surface without slipping or a rattleback. The key property of nonholonomic constraints is that they cannot be reduced to constraints on the possible configurations of the system, in contrast to holonomic constraints.

Nonholonomic systems have a wide literature [14, 18, 40, 126] and a history of confusion with vakonomic systems, see [102] for a discussion. It is now well-known that nonholonomic systems are not variational, but one open question is trying to characterize when such systems can be studied from a Hamiltonian point of view. Such attempts can be referred to as Hamiltonization of nonholonomic systems. One of the approaches consists in trying to put the trajectories of the nonholonomic system as the restriction to the constraint submanifold of the trajectories of a Lagrangian system [17]. Other approaches include Chaplygin Hamiltonization [10, 65].

Geometric integration

Finding numerical integrators that possess the same geometric properties as the original system, for instance symplecticity for a Hamiltonian system [112], can be very useful in order to obtain appropriate qualitative behaviour. Some interesting properties are preservation of constants of motion, volume, symmetries or constraints among others. Numerical integrators that preserve such properties are known as geometric integrators [77, 92]. There is a wide literature on geometric integration for Hamiltonian systems, for instance the book by J. M. Sanz-Serna and M. P. Calvo [139], but not so much for nonholonomic systems, although there are also some interesting contributions [39, 63, 117].

Lagrangian mechanics on Lie algebroids

Lie algebroids can be thought of as a generalization of tangent bundles. The use of Lagrangian functions on them allows to include under the same scheme the dynamics of mechanical systems for a variety of situations, including mechanical systems reduced by some symmetry Lie group. For instance Euler-Poincaré equations and Lagrange-Poincaré equations are obtained as Euler-Lagrange equations on a Lie algebra and Atiyah algebroid respectively. The study of Lagrangian mechanics on Lie algebroids has a precursor in A. Weinstein [165]. Further references are [51, 113]. Even nonholonomic mechanics can be studied in this context if we drop the Jacobi identity and stick to skew-symmetric Lie algebroids [37, 50].

Geometric control

Dynamical systems may be modeled including some external forces (the controls) which may be chosen in order to achieve a desired goal, which could be for instance stabilization of an unstable equilibrium or attainment of a desired position while minimizing some cost. Differential geometry has provided many tools for the design and study of control laws for nonlinear control systems. Many of such developments were driven by R. W. Brockett [25] as well as A. M. Bloch [14], V. Jurdjevic [95], H. Nijmeijer and A. J. Van der Schaft [127], E. D. Sontag [149] and H. J. Sussmann [151]. We will focus on the problem of stabilization of an unstable equilibrium for a certain class of mechanical systems with symmetry.

These topics will be given a proper introduction along the manuscript as they appear. We will next give a list of the publications and preprints included in this thesis, a brief description of the results obtained in each one and also an outline of the present manuscript relating each chapter to the contents of the list. Finally we will fix some notation that will be used throughout.

Publications and preprints

P1-[11]. María Barbero-Liñán, Marta Farré Puiggalí, David Martín de Diego: *Isotropic submanifolds and the inverse problem for mechanical constrained systems.* J. Phys. A: Math. Theor. 48 (2015) 045203 (35pp).

P2-[12]. María Barbero Liñán, Marta Farré Puiggalí, David Martín de Diego: *Inverse problem for Lagrangian systems on Lie algebroids and applications to reduction by symmetries.* Monatsh. Math. 180 (2016), no. 4, 665-691.

P3-[57]. Marta Farré Puiggalí, Tom Mestdag: *The inverse problem of the calculus of variations and the stabilization of controlled Lagrangian systems.* SIAM J. Control Optim. 54-6 (2016), pp. 3297-3318.

P4-[32]. Elena Celledoni, Marta Farré Puiggalí, Eirik Hoel Høiseth, David Martín de Diego: *Energy-preserving integrators applied to nonholonomic systems.* arXiv:1605.02845.

P5. María Barbero Liñán, Marta Farré Puiggalí, Sebastián Ferraro, David Martín de Diego: *Inverse problem of the calculus of variations for discrete systems.*

P6. Anthony M. Bloch, Marta Farré Puiggalí: *New matching conditions from Helmholtz conditions.*

Brief description of results

Inverse problem for constrained systems

In the first paper P1-[11] we give a new characterization of the inverse problem in terms of Lagrangian submanifolds, both for the autonomous and non-autonomous cases. The characterization is given in terms of the existence of a Legendre transformation instead of the existence of a multiplier matrix. This new approach allows a geometric formulation of the inverse problem in the context of constrained systems. We give a notion of variationality for a SODE on a submanifold of the tangent bundle just by changing Lagrangian submanifold into isotropic submanifold.

This notion allows us to prove an analog of Crampin's Theorem for a constrained SODE. It also allows us to check whether a nonholonomic system is a subsystem of a Lagrangian one, for a Lagrangian function which might be degenerate. This is related to one of the approaches to Hamiltonization of nonholonomic systems [17, 120].

Inverse problem on Lie algebroids

In the second paper P2-[12] we were able to further extend the definition of the inverse problem to the context of SODEs on regular Lie algebroids by using prolongations of Lie algebroids and Lagrangian submanifolds of symplectic Lie algebroids. The key difference with P1-[11] is that for a general Lie algebroid the Poincaré Lemma does not hold. We recover Helmholtz conditions for Lie algebroids already known in the literature but further identify their insufficiency to ensure the existence of a local Lagrangian as the lack of the Poincaré Lemma. We can then find an extra condition which provides a set of necessary and sufficient conditions, all of them with the components of the Legendre transformation as unknown. Finally we also study the behaviour of the variational property under morphisms of Lie algebroids and provide some examples and comparison with previous approaches in the literature.

Inverse problem and stabilization of mechanical systems

In [21] A. M. Bloch, N. E. Leonard and J. E. Marsden presented a strategy to explicitly construct feedback controls that would stabilize certain mechanical systems with symmetry. The idea is to modify the Lagrangian of the given mechanical system in a prescribed way in order to obtain that the original controlled Lagrangian system is equivalent to the Lagrangian system corresponding to the new Lagrangian, which will depend on some parameters. Sufficient conditions to achieve this goal are referred to as matching conditions and there is a variety of them that allow to tackle different situations, including a symmetry break in the potential energy [15].

Underlying this process there is the inverse problem of the calculus of variations. Notice that instead of one SODE we have a family of SODEs with the unknown controls as parameters and the goal is to find appropriate parameters such that the controlled SODE becomes variational. Therefore in paper P3-[57] we use the classification of Douglas in order to obtain new stabilizing controls for a class of two-dimensional mechanical systems. The main difference in the strategy is that now we do not impose any specific form for the new Lagrangian, but instead we work among a certain class of controls. This is so far restricted to dimension two since it relies on Douglas' results.

Energy-preserving nonholonomic integrators

In contrast with Hamiltonian systems, nonholonomic systems do not necessarily possess an invariant measure. On the other hand, energy is always a conserved quantity for nonholonomic systems. This is why, instead of trying to adapt variational integrators to the nonholonomic case we have focused in P4-[32] on the study of energy-preserving integrators for nonholonomic systems.

We will say that an ODE system $\dot{x} = \Pi(x)\nabla H(x)$ with $\Pi(x)$ a skew-symmetric matrix is given in skew gradient form. A discrete gradient $\bar{\nabla} H : \mathbb{R}^{2N} \longrightarrow \mathbb{R}^N$ is an approximation of the gradient of H. It satisfies $\bar{\nabla} H(x, x')^T(x' - x) = H(x') - H(x)$, and $\bar{\nabla} H(x, x) = \nabla H(x)$, for all $x, x' \in \mathbb{R}^N$. Then an energy-preserving integrator can be defined as

$$\frac{x' - x}{h} = \tilde{\Pi}(x, x', h)\bar{\nabla} H(x, x'), \tag{2}$$

where $\tilde{\Pi}(x, x', h)$ is a skew-symmetric approximation of $\Pi(x)$. In P4-[32] we use the fact that the equations of motion for a nonholonomic system can be written on the dual of the constraint distribution, in skew gradient form. Therefore we can apply integrators of type (2), which will automatically preserve both the constraints and the energy.

Moreover, in order to facilitate their applicability to complex systems which cannot be easily transformed into the aforementioned almost-Poisson form, we manage to rewrite our integrators using just the initial information of the nonholonomic system. The derived procedures are tested on several examples.

Discrete inverse problem

P5 is an ongoing project with María Barbero Liñán, Sebastián Ferraro and David Martín de Diego. Using again Lagrangian and isotropic submanifolds we can extend the inverse problem to both discrete mechanics and constrained discrete mechanics. Now the evolution space is two copies of the configuration manifold, which replaces the tangent bundle. In this context it will be key to deal with implicit systems. Our main interest for this extension is its possible applicability to the study of geometric integrators. Now we have a notion of variationality for both continuous and discrete constrained systems. If the original system is constrained variational we intend to study the possible advantages of using a discrete constrained variational integrator, and compare them to other existing geometric integrators for nonholonomic systems.

Matching via the inverse problem I

P6 is an ongoing project with Anthony M. Bloch. By using the Helmholtz conditions we are able to recover the matching conditions from [21], since the Helmholtz conditions are necessary and sufficient conditions for existence of a regular Lagrangian, but further we can derive new matching conditions for a particular class of mechanical systems. It turns out that for this class of systems we obtain feedback controls that only depend on the configuration variables.

Outline of the thesis

- Introduction
- Chapter 1: We first give an overview of the inverse problem of the calculus of variations, recalling some of the main results with particular emphasis in the geometric approach. Then in Section 1.5 we introduce a new characterization for the inverse problem which was given in P1-[11]. We also introduce in this section the implicit version provided in P5.
- Chapter 2: We give some applications of the inverse problem to control theory, more precisely to the stabilization of mechanical systems. The chapter is divided into two parts, the first one is based on P3-[57] and provides an application of Douglas' classification to the stabilization of a class of two-dimensional mechanical systems. The second one is based on P6 and describes how to explicitly recover some of the matching conditions from the Helmholtz conditions and how to derive additional ones for a particular class of mechanical systems.
- Chapter 3: This chapter is based on P2-[12]. We first give an introduction to Lie algebroids and mechanics on Lie algebroids. Then we extend the inverse problem to this context.

- Chapter 4: This chapter is based on some of the work in P5. We first introduce discrete mechanics and variational integrators. Then we provide an extension of the inverse problem to discrete systems, both implicit and explicit, using again Lagrangian submanifolds. Finally we prove an analog of Crampin's Theorem in this context.

- Chapter 5: The results presented in this chapter are based on P1-[11] and P5. We will first give an introduction to constrained systems, especially to nonholonomic systems. Then we will present the extension of the inverse problem to constrained systems both in the continuos (P1-[11]) and discrete cases (P5). We will also see the corresponding versions of Crampin's Theorem for each case.

- Chapter 6: This chapter is based on P4-[32]. Nonholonomic systems are not variational in the usual sense of the word (not to be confused with the notion of constrained variational SODE introduced in Chapter 5). We should then in general not use a variational integrator for integrating a nonholonomic system. There are various methods specifically designed for integrating nonholonomic systems. In this chapter we propose to use energy-preserving integrators for nonholonomic systems. Their performance is compared with other standard methods in nonholonomic dynamics, and their merits verified in practice.

- Chapter 7: We outline some possible directions for future research from the results obtained in this thesis.

- Appendix A: We introduce two key elements of this work, namely Lagrangian and isotropic submanifolds. We also review a key construction that allows to extend an isotropic submanifold into a Lagrangian one.

- Appendix B: We review the Tulczyjew triple, which is used in Chapters 1 and 5.

- Appendix C: We provide direct computations showing the equivalence between the classical Helmholtz conditions in terms of the multipliers and the Helmholtz conditions introduced in Chapter 1 in terms of the Legendre transformation.

- Appendix D: We show an algorithm that allows to use the integrators presented in Chapter 6 without having to compute the almost-Poisson formulation of the nonholonomic system.

Notation

All manifolds and maps are assumed to be smooth unless otherwise stated. Given a manifold Q, TQ and T^*Q denote respectively the tangent and the cotangent bundle of Q, and $\tau_Q : TQ \longrightarrow Q$, $\pi_Q : T^*Q \longrightarrow Q$ are the corresponding canonical projections. $\mathcal{C}^\infty(Q)$ denotes the set of smooth functions on a manifold Q. The set of vector fields on Q is denoted by $\mathfrak{X}(Q)$ and the set of k-forms on Q is denoted by $\Lambda^k(Q)$. In general $T^r_s(Q)$ denotes the set r-covariant and s-contravariant tensor fields on Q. If $R \in T^r_s(Q)$ then we say that R is an (s,r)-tensor field on Q. $\Gamma(E)$ will denote the space of sections of a fiber bundle $E \to M$. If $(q^1, \ldots, q^n)$ are local coordinates on Q then we will in general use the shorthand notation $q = (q^1, \ldots, q^n)$, $\dot{q} = (\dot{q}^1, \ldots, \dot{q}^n)$ and so on. The Einstein summation convention is also used throughout.

Chapter 1

The inverse problem of the calculus of variations

The calculus of variations is mainly concerned with finding extrema of functionals of various kinds acting on curves or functions with different regularity requirements and with fixed or varying boundary conditions. In this manuscript the expression "inverse problem of the calculus of variations" will appear frequently and it refers to the opposite direction of one of the simplest cases. More specifically we will be concerned with functionals of the type

$$c(t) \longmapsto \int_a^b L(t, c(t), \dot{c}(t))\, dt\,, \tag{1.1}$$

where $c(t)$ are C^2 curves on a manifold with fixed end points. Additionally assume that we only look for critical points and are not concerned with them being actual extrema [24, 66, 137]. We regard this problem in the calculus of variations as the direct problem. The main topic of this chapter is the opposite question, that is, given a family of curves on the manifold, which are the solutions of a second order differential system, can we find a functional of type (1.1) such that its critical values are the original curves?

In this chapter we will first give in Section 1.1 an introduction to Lagrangian mechanics. Then in Section 1.2 we review the main versions of the inverse problem of the calculus of variations. Next we give an overview of some of the main contributions, including the results for the two-dimensional case in Section 1.3 and a characterization of the problem in terms of the Poincaré-Cartan two-form in Section 1.4. Finally Section 1.5 is based on part of [11] and provides a new characterization of the inverse problem, both in the autonomous and non-autonomous cases, which will be very useful in the next chapters. Section 1.5 also includes an implicit version of the problem, presented in [32].

1.1 Lagrangian mechanics

The Euler-Lagrange equations are a system of second order ordinary differential equations which model many physical phenomena in classical mechanics, see [1, 7, 29, 85, 86, 87, 111]. In this section we will show how to derive them in two alternative ways, first using a variational principle, namely Hamilton's principle, and second using the geometry of the tangent bundle.

An autonomous Lagrangian system is defined by a smooth manifold Q, known as the configuration manifold, since it models all possible positions or configurations of the system, and a smooth function

$L : TQ \to \mathbb{R}$, known as the Lagrangian function, which encodes the dynamics. We will say that the Lagrangian system is of mechanical type if there is a Riemannian metric g on Q and a smooth map $V : Q \longrightarrow \mathbb{R}$ such that $L(v_q) = \frac{1}{2}g(v_q, v_q) - V(q)$, where $q \in Q$ and $v_q \in T_qQ$. From now on, unless otherwise stated, all manifolds and maps are assumed to be smooth.

Let us consider a configuration manifold Q of dimension n with local coordinates (q^i), $i = 1, \ldots, n$. The corresponding local fibered coordinates on TQ will be denoted by $(q^i, \dot{q}^i)$ and the canonical projection by $\tau_Q : TQ \to Q$. Consider a curve $c : [a, b] \to Q$ of class C^2 connecting two fixed points q_0 and q_1 in the configuration manifold Q. The set of all these curves is denoted by

$$\mathcal{C}^2(q_0, q_1, [a, b]) = \left\{ c : [a, b] \subseteq \mathbb{R} \longrightarrow Q \;\middle|\; c \in C^2([a, b]),\ c(a) = q_0,\ c(b) = q_1 \right\}.$$

This set is a smooth infinite dimensional manifold [1]. Its tangent space at a curve $c \in \mathcal{C}^2(q_0, q_1, [a, b])$ is given by

$$T_c\mathcal{C}^2(q_0, q_1, [a, b]) = \left\{ X : [a, b] \longrightarrow TQ \;\middle|\; X \in C^2([a, b]),\ \tau_Q \circ X = c \ \text{ and } X(a) = X(b) = 0 \right\}.$$

Now if we introduce a Lagrangian function $L : TQ \to \mathbb{R}$ we can define on $\mathcal{C}^2(q_0, q_1, [a, b])$ the action functional

$$\begin{array}{rccl} \mathcal{J} : & \mathcal{C}^2(q_0, q_1, [a, b]) & \longrightarrow & \mathbb{R} \\ & c & \longmapsto & \displaystyle\int_a^b L(c(t), \dot{c}(t))\, dt\,, \end{array}$$

whose critical curves are the trajectories of the Lagrangian system.

Definition 1.1.1. [Hamilton's principle] *A curve $c \in \mathcal{C}^2(q_0, q_1, [a, b])$ is a solution of the Lagrangian system defined by $L : TQ \to \mathbb{R}$ if and only if c is a critical point of $\mathcal{J}$, that is,*

$$d\mathcal{J}(c) = 0\,. \tag{1.2}$$

Using standard arguments from variational calculus, it is easy to show that the solutions of the Lagrangian system given by (1.2) are the solutions of the system of second order differential equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0, \quad 1 \leq i \leq n = \dim Q\,,$$

known as the Euler-Lagrange equations for $L : TQ \to \mathbb{R}$. Notice that when the Lagrangian is regular, that is, when the n by n Hessian matrix $(\partial^2 L/\partial \dot{q}^i \partial \dot{q}^j)$ is regular, the Euler-Lagrange equation can be rewritten in explicit or normal form as

$$\ddot{q}^k = g^{ki}\left(-\frac{\partial^2 L}{\partial q^j \partial \dot{q}^i}\dot{q}^j + \frac{\partial L}{\partial q^i}\right)\,, \tag{1.3}$$

where $g_{ki} = \frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^i}$ and g^{ki} denote the components of the inverse matrix. Such systems can be represented as vector fields on TQ of a special type, which will be referred to as SODEs (Second Order Differential Equations). In order to define them we first need to introduce two geometric objects associated with TQ, namely the **Liouville** or **dilation vector field** $\Delta \in \mathfrak{X}(TQ)$ and a $(1,1)$-tensor field S on TQ called the **vertical endomorphism**. They are intrinsically given by

$$\begin{array}{rccc} \Delta : & TQ & \longrightarrow & TTQ \\ & v_q & \longmapsto & \frac{d}{dt}\big|_{t=0}(e^t v_q) \end{array} \quad \text{and} \quad \begin{array}{rccc} S : & TTQ & \longrightarrow & TTQ \\ & w_{v_q} & \longmapsto & \frac{d}{dt}\big|_{t=0}(v_q + tT\tau_Q(w_{v_q}))\,, \end{array}$$

where $v_q \in T_qQ$, $w_{v_q} \in T_{v_q}TQ$, and the corresponding expressions in canonical coordinates are

$$\Delta = \dot{q}^i \frac{\partial}{\partial \dot{q}^i} \quad \text{and} \quad S = dq^i \otimes \frac{\partial}{\partial \dot{q}^i}\,.$$

Definition 1.1.2. *A SODE Γ is a vector field on TQ satisfying $S(\Gamma) = \Delta$. In fibered coordinates it is written as*

$$\Gamma = \dot{q}^i \frac{\partial}{\partial q^i} + \Gamma^i(q, \dot{q}) \frac{\partial}{\partial \dot{q}^i}\,.$$

The solutions of the SODE Γ are precisely the solutions of the system of second order differential equations $\ddot{q}^i = \Gamma^i(q, \dot{q})$.

Now using the Liouville vector field and the vertical endomorphism we can derive the Euler-Lagrange equations intrinsically, which will have a SODE as solution. Given a Lagrangian function $L : TQ \to \mathbb{R}$ we define the **Poincaré-Cartan one-form** $\Theta_L = S^*(dL)$, the associated **Poincaré-Cartan two-form** $\Omega_L = -d\Theta_L$ and the **energy function** $E_L : TQ \to \mathbb{R}$ by $E_L = \Delta(L) - L$. Locally we have

$$\Theta_L = \frac{\partial L}{\partial \dot{q}^i}\, dq^i \quad \text{and} \quad E_L = \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L\,.$$

When the Lagrangian L is regular, that is, Ω_L is a symplectic two-form, then there exists a unique SODE Γ_L solution of the equation

$$i_\Gamma \Omega_L = dE_L\,, \tag{1.4}$$

or alternatively of

$$\mathcal{L}_\Gamma \Theta_L = dL\,, \tag{1.5}$$

where $\mathcal{L}_\Gamma \Theta_L$ is the Lie derivative of Θ_L along Γ. The solutions of the SODE Γ_L are precisely the solutions to the Euler-Lagrange equations for L.

If we now want to include time-dependent systems we will consider a time-dependent Lagrangian function, that is $L : \mathbb{R} \times TQ \longrightarrow \mathbb{R}$, or more generally a Lagrangian function on the first jet bundle of a bundle $\pi : E \longrightarrow \mathbb{R}$, see [146]. We can again use Hamilton's principle to derive the Euler-Lagrange equations, which can now be seen as a vector field on $\mathbb{R} \times TQ$.

Definition 1.1.3. *A vector field Γ on $\mathbb{R} \times TQ$ is a SODE if $\langle \Gamma, \theta^i \rangle = 0$ and $\langle \Gamma, dt \rangle = 1$, where $\theta^i = dq^i - \dot{q}^i dt$ are the usual contact one-forms. In local coordinates $(t, q, \dot{q})$ for $\mathbb{R} \times TQ$ we get*

$$\Gamma = \frac{\partial}{\partial t} + \dot{q}^i \frac{\partial}{\partial q^i} + \Gamma^i(t, q, \dot{q}) \frac{\partial}{\partial \dot{q}^i}\,.$$

The integral curves of the SODE Γ are the solutions of the system of explicit second order differential equations $\ddot{q}^i = \Gamma^i(t, q, \dot{q})$.

Remark 1.1.4. An example of a SODE on $\mathbb{R} \times TQ$ is precisely the Euler-Lagrange vector field associated to a regular Lagrangian function $L : \mathbb{R} \times TQ \longrightarrow \mathbb{R}$, which is defined as the unique vector field Γ_L on $\mathbb{R} \times TQ$ satisfying

$$i_\Gamma \Omega_L = 0 \quad \text{and} \quad \langle \Gamma, dt \rangle = 1\,, \tag{1.6}$$

where $\Omega_L = -d\theta_L$ is the Poincaré-Cartan two-form, $\theta_L = Ldt + dL \circ S$ is the Poincaré-Cartan one-form and $S = \frac{\partial}{\partial \dot{q}^i} \otimes \theta^i$ is the vertical endomorphism. Note that (Ω_L, dt) provides $\mathbb{R} \times TQ$ with a cosymplectic structure if L is regular [31]. Notice also that we have used the same notation for the Poincaré-Cartan two-form and for the vertical endomorphism as in the autonomous case and will keep doing so since it will be clear from the context to which case we refer.

1.2 The inverse problem of the calculus of variations

In the previous section we have seen that given a regular Lagrangian, either autonomous $L : TQ \to \mathbb{R}$ or time-dependent $L : \mathbb{R} \times TQ \longrightarrow \mathbb{R}$, we can always find a unique SODE, denoted by Γ_L, satisfying equations (1.4) or (1.6). In this section we will focus on the time-dependent case $L : \mathbb{R} \times TQ \longrightarrow \mathbb{R}$.

The inverse problem of the calculus of variations studies whether or not it is possible to go the other way around, that is, given a system of second order differential equations on $\mathbb{R} \times TQ$, either implicit or explicit, decide whether or not it is related to a regular Lagrangian function. There are two main versions of the inverse problem, which depend on what we mean by related. Some references clarifying their difference and providing a historical review are [101, 147].

1.2.1 The covariant version

In this case related means to have exactly the same expression. More precisely, given an implicit system of second order differential equations

$$\Phi_i(t, q, \dot{q}, \ddot{q}) = 0\,, \quad i = 1, \ldots, n\,,$$

determine whether or not it is possible to find a regular Lagrangian $L(t, q, \dot{q})$ such that

$$\Phi_i(t, q, \dot{q}, \ddot{q}) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\,, \quad i = 1, \ldots, n\,. \tag{1.7}$$

This question was first raised by Helmholtz [161] in 1887, who provided a set of necessary and sufficient conditions for (1.7) to hold, namely,

$$\frac{\partial \Phi_i}{\partial \ddot{q}^j} - \frac{\partial \Phi_j}{\partial \ddot{q}^i} = 0\,, \tag{1.8}$$

$$\frac{\partial \Phi_i}{\partial q^j} - \frac{\partial \Phi_j}{\partial q^i} - \frac{1}{2}\frac{d}{dt}\left(\frac{\partial \Phi_i}{\partial \dot{q}^j} - \frac{\partial \Phi_j}{\partial \dot{q}^i}\right) = 0\,, \tag{1.9}$$

$$\frac{\partial \Phi_i}{\partial \dot{q}^j} + \frac{\partial \Phi_j}{\partial \dot{q}^i} - \frac{d}{dt}\left(\frac{\partial \Phi_i}{\partial \ddot{q}^j} + \frac{\partial \Phi_j}{\partial \ddot{q}^i}\right) = 0\,, \tag{1.10}$$

although the sufficiency part was proved by Mayer [116] later in 1896. The conditions (1.8)-(1.10) are sometimes referred to as Helmholtz conditions, but in the present manuscript this could lead to confusion. Therefore we will refer to them as **covariant Helmholtz conditions**. They can be obtained in differents ways, see for instance [100]. Santilli calls them self-adjointness conditions in [138].

The necessity part of Helmholtz conditions (1.8)-(1.10) can be proved by a straightforward computation. The sufficiency part can be proved using the existence of the Vainberg-Tonti Lagrangian, which is a second order Lagrangian $K(t, q, \dot{q}, \ddot{q})$ such that

$$\Phi_i = \frac{\partial K}{\partial x^i} - \frac{d}{dt}\frac{\partial K}{\partial \dot{x}^i} + \frac{d^2}{dt^2}\frac{\partial K}{\partial \ddot{x}^i}\,. \tag{1.11}$$

Then it can be seen that the equality (1.11) implies $K = K_0 + \frac{dC}{dt}$ for some first order Lagrangian $K_0(t, q, \dot{q})$ and some function C on the base space, see [100] for details. The right-hand side of Equation (1.11) is the Euler-Lagrange operator corresponding to a second order Lagrangian and will be clear after Section 1.2.3.

1.2.2 The contravariant or multiplier version

In this case related means to have the same solutions. More precisely, given an explicit SODE

$$\ddot{q}^i = \Gamma^i(t, q, \dot{q}), \quad i = 1, \ldots, n\,, \tag{1.12}$$

the problem consists in determining whether or not there is a non-singular multiplier matrix $(g_{ij}(t, q, \dot{q}))$ such that

$$g_{ij}\left(\ddot{q}^j - \Gamma^j(t, q, \dot{q})\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}, \quad i, j = 1, \ldots, n \tag{1.13}$$

holds for some regular Lagrangian $L(t, q, \dot{q})$. Notice that in the affirmative case we necessarily have

$$g_{ij} = \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\,,$$

and the solutions to Γ are exactly the same as the solutions to the Euler-Lagrange equations for L, due to the regularity of $(g_{ij}(t, q, \dot{q}))$. In such case the system of second order ordinary differential equations (1.12) is called **variational**. Geometrically, condition (1.13) can be captured into the requirement of the existence of a function $L : TQ \to \mathbb{R}$ such that $\mathcal{L}_\Gamma \Theta_L = dL$, see (1.5).

The existence of a regular Lagrangian for Γ is equivalent to the existence of multipliers $(g_{ij}(t, q, \dot{q}))$ satisfying the **Helmholtz conditions**

$$\det(g_{ij}) \neq 0, \qquad g_{ji} = g_{ij}, \qquad \frac{\partial g_{ij}}{\partial \dot{q}^k} = \frac{\partial g_{ik}}{\partial \dot{q}^j}\,, \tag{1.14}$$

$$\Gamma(g_{ij}) - \nabla^k_j g_{ik} - \nabla^k_i g_{kj} = 0\,, \tag{1.15}$$

$$g_{ik}\Phi^k_j = g_{jk}\Phi^k_i\,, \tag{1.16}$$

where $\Gamma = \frac{\partial}{\partial t} + \dot{q}^i \frac{\partial}{\partial q^i} + \Gamma^i(q, \dot{q})\frac{\partial}{\partial \dot{q}^i}$, $\nabla^i_j = -\frac{1}{2}\frac{\partial \Gamma^i}{\partial \dot{q}^j}$ and $\Phi^k_j = \Gamma\left(\frac{\partial \Gamma^k}{\partial \dot{q}^j}\right) - 2\frac{\partial \Gamma^k}{\partial q^j} - \frac{1}{2}\frac{\partial \Gamma^i}{\partial \dot{q}^j}\frac{\partial \Gamma^k}{\partial \dot{q}^i}$. A proof of the above statement can be found in [55], although the conditions were first given in this geometric form by Sarlet in [140].

Conversely, if (1.12) has the same solutions as the Euler-Lagrange equations for some regular Lagrangian L, then we can write the Euler-Lagrange system in the form (1.3) and therefore obtain (1.13) with $g_{ij} = \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$.

This version of the inverse problem was formulated by Hirsh [84] in 1898, although it had previously been solved for the case $n = 1$ by Sonin [148], showing that one equation is always variational. There are many characterizations of the inverse problem in the literature, but not much is known about the complete solution, since the Helmholtz conditions are a difficult mixed set of algebraic and partial differential equations for the multipliers (g_{ij}). The next case $n = 2$ was solved more than 50 years later by Douglas in [55], but for $n > 2$ no complete classification exists. Some partial results exist, more precisely, some cases in Douglas' classification have been generalized to arbitrary dimension n, see for instance [47, 49, 143]. Some recent progress includes a new classification scheme which was proposed in [54] and turned out to be useful in order to identify some more variational cases, which also generalize some of Douglas' cases.

From now on, the expression "the inverse problem" will always refer to the multiplier version unless otherwise stated.

Remark 1.2.1 (The one-dimensional case)**.** If we are given just one second order differential equation $\ddot{q} = \Gamma(t, q, \dot{q})$, the question consists in deciding whether or not we can find a nonvanishing function $g(t, q, \dot{q})$ and a Lagrangian function $L(t, q, \dot{q})$ such that

$$g(\ddot{q} - \Gamma) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\,.$$

In this case we only need to study one Helmholtz condition, namely (1.15), which becomes

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial q}\dot{q} + \frac{\partial g}{\partial \dot{q}}\Gamma + \frac{\partial \Gamma}{\partial \dot{q}} g = 0\,.$$

Since $g \neq 0$, we may write $g = \pm e^f$ for some function $f(t, q, \dot{q})$. Then the above equation becomes

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial q}\dot{q} + \frac{\partial f}{\partial \dot{q}}\Gamma + \frac{\partial \Gamma}{\partial \dot{q}} = 0\,,$$

which is a first order linear PDE and can be solved using the standard method of characteristics, see [56, 130].

1.2.3 Other versions

In Section 1.1 we have derived one version of the Euler-Lagrange equations, those for classical mechanics, which are a system of ODEs and the solutions are curves on the configuration manifold Q, but variational principles are also valid in order to derive the equations that model many physical theories such as continuum mechanics, electromagnetism and general relativity, see for instance [123].

These are known as field theories. Now curves on Q are replaced by sections of a bundle $\pi : E \longrightarrow M$, where $\dim(M) > 1$ and we assume M is oriented. We will look for solutions among these sections, known as fields. The case of classical mechanics is recovered if we take $E = \mathbb{R} \times Q \longrightarrow \mathbb{R}$.

For first order field theories, see for instance [64, 70], the Lagrangian is a function $L : J^1\pi \longrightarrow \mathbb{R}$ and for adapted coordinates $(x^i, u^\alpha, u^\alpha_i)$ on $J^1\pi$ the Euler-Lagrange equations become the system of PDEs

$$\frac{\partial L}{\partial u^\alpha} - \frac{d}{dx^i}\frac{\partial L}{\partial u^\alpha_i} = 0, \quad \alpha = 1, \dots, n = \dim(E) - \dim(M)\,, \tag{1.17}$$

where $\frac{d}{dx^i} = \frac{\partial}{\partial x^i} + u^\alpha_i \frac{\partial}{\partial u^\alpha}$. Notice that in the case $\pi : E = \mathbb{R} \times Q \longrightarrow \mathbb{R}$ we have $J^1\pi \cong \mathbb{R} \times TQ$ and therefore we recover the Euler-Lagrange equations from Section 1.1.

For higher order field theories the Lagrangian is a function $L : J^k\pi \longrightarrow \mathbb{R}$, $k > 1$. Denote by (x^i, u^α_J) adapted coordinates on $J^k\pi$, where J is a multi-index, see [105, 146]. Then the Euler-Lagrange equations are

$$\sum_{|J|=0}^{k} (-1)^{|J|} \frac{d^{|J|}}{dx^J}\frac{\partial L}{\partial u^\alpha_J} = \frac{\partial L}{\partial u^\alpha} - \frac{d}{dx^i}\frac{\partial L}{\partial u^\alpha_i} + \frac{d^2}{dx^{ij}}\frac{\partial L}{\partial u^\alpha_{1_i+1_j}} - \dots + (-1)^k \frac{d^k}{dx^J}\frac{\partial L}{\partial u^\alpha_J} = 0\,, \tag{1.18}$$

where $\frac{d}{dx^i} = \frac{\partial}{\partial x^i} + \sum_{|I|=0}^{k-1} u^\alpha_{I+1_i} \frac{\partial}{\partial u^\alpha_I}$ and $|I|$ denotes the lenght of the multi-index.

Both (1.17) and (1.18) can be derived from a variational principle for an action functional defined on the space of sections of π.

Now we have a wider class of Euler-Lagrange equations for which the inverse problem can also be addressed. So we can formulate both versions of the inverse problem for PDEs as well as for higher order systems. More precisely, given a system of PDEs of arbitrary order or a higher order ODE system, we may ask whether or not it can be related to a system Euler-Lagrange equations of type (1.17) or (1.18).

For the inverse multiplier problem I. Anderson and G. Thompson [3] studied explicit systems of $2k$-th order differential equations

$$u^b_{(2k)} = f^b(t, u^c, u^c_{(1)}, u^c_{(2)}, \dots, u^c_{(2k-1)})\,,$$

where $(t, u^c, u^c_{(1)}, u^c_{(2)}, \dots, u^c_{(2k-1)})$ denote coordinates on the jet bundle $J^{2k-1}\pi$. The SODE in this case is replaced by a vector field on $J^{2k-1}\pi$ of the form

$$\Gamma = \frac{\partial}{\partial t} + \sum_{r=0}^{2k-2} u^b_{(r+1)} \frac{\partial}{\partial u^b_{(r)}} + f^b \frac{\partial}{\partial u^b_{(2k-1)}}\,.$$

Regarding field theories, if we restrict to second order quasi-linear PDEs, then Henneaux in [81] derived some Helmholtz type conditions in terms of multipliers, but some extra assumptions are needed for the system. The one-dimensional case, that is, just one quasi-linear equation, was considered in [6] providing conditions for variationality. This case is not always variational, in contrast with the ODE case.

For the covariant version there is much work regarding PDEs [5, 101], higher order systems [101] and global issues [5, 4, 99, 152, 160]. Finally there are also other extensions such as a covariant version for nonholonomic mechanics [135] and also a version of the problem for driven SODEs [89].

1.3 Classification for two-dimensional systems

History has shown the inverse problem is extremely difficult since only the full solution for at most two-dimensional systems is known [55]. In this section we will review the available results for two-dimensional systems. In the first part we will recall the original classification of Douglas following [55], but we will only make an emphasis on certain cases that will be interesting in the next chapter. In the second part we will recall the geometric version of Douglas' classification given in [49]. Along the way we will introduce some useful tensor fields and derivations along projections.

1.3.1 Original Douglas' classification

In this section we will recall the results from Douglas' classification that will be relevant in Chapter 2. Douglas' solution consists of an exhaustive classification in different cases using Riquier-Janet theory [55]. A thorough study of the Helmholtz conditions leads to a classification for two-dimensional systems, that is, a classification for

$$y'' = F(t, y, z, y', z'), \quad z'' = G(t, y, z, y', z')$$

into variational and nonvariational systems, where F and G are assumed to be analytic. Both variational and nonvariational examples are included in his work. We keep in this section the original notation used by Douglas, with the exception of t which replaces his x.

The original classification and results are far more involved, complex and detailed, providing for instance the generality of the multipliers, which we will omit. For the full classification, in particular of Case II, as well as for detailed statements including the generality of the solutions, see [55, Section 3].

The first level of Douglas' classification involves the rank of a matrix

$$M = \begin{pmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix},$$

whose components are computed from the given SODE components, that is, F and G, as

$$\begin{aligned}
A &= \frac{d}{dt}F_{z'} - 2F_z - \frac{1}{2}F_{z'}(F_{y'} + G_{z'}), \\
B &= -\frac{d}{dt}F_{y'} + \frac{d}{dt}G_{z'} + 2(F_y - G_z) + \frac{1}{2}(F_{y'}^2 - G_{z'}^2), \\
C &= -\frac{d}{dt}G_{y'} + 2G_y + \frac{1}{2}G_{y'}(F_{y'} + G_{z'}), \\
A_1 &= \frac{dA}{dt} - F_{y'}A - \frac{1}{2}F_{z'}B, \\
B_1 &= \frac{dB}{dt} - G_{y'}A - \frac{1}{2}(F_{y'} + G_{z'})B - F_{z'}C, \\
C_1 &= \frac{dC}{dt} - \frac{1}{2}G_{y'}B - G_{z'}C,
\end{aligned}$$

and A_2, B_2, C_2 are defined using the same formulas that define A_1, B_1, C_1 but replacing A, B, C by A_1, B_1, C_1 respectively. According to the rank of the matrix M we have four main cases, namely

- Case I if $\text{rank}(M) = 0$,
- Case II if $\text{rank}(M) = 1$,
- Case III if $\text{rank}(M) = 2$,
- Case IV if $\text{rank}(M) = 3$.

Case III is further subdivided into

- Case IIIa if $D \neq 0$,
- Case IIIb if $D = 0$,

where $D = \Delta_1\Delta_3 - \Delta_2^2$ with $\Delta_1 = BC_1 - B_1C$, $\Delta_2 = CA_1 - C_1A$ and $\Delta_3 = AB_1 - A_1B$.

Case II is further subdivided into six cases. The subclassification of Case II depends first on the roots of $A\xi^2 + B\xi + C = 0$ being coincident (Case IIb) or different (Case IIa) and later on these roots satisfying certain conditions involving their derivatives. We will only describe two of these cases:

- **Case IIa1**: $A\xi^2 + B\xi + C = 0$ has two different roots $\lambda \neq \mu$ which satisfy $\lambda\lambda_{z'} - \lambda_{y'} = 0$ and $\mu\mu_{z'} - \mu_{y'} = 0$.
- **Case IIb1'**: $A\xi^2 + B\xi + C = 0$ has a double root λ which satisfies $\lambda\lambda_{z'} - \lambda_{y'} = 0$ and $(IX') = 0$, where

$$\begin{aligned}
(IX') &= \lambda(VI)_{z'} - (VI)_{y'} - 2\lambda_{z'z'}(II) + 2\lambda_{yz'} - 2\lambda\lambda_{zz'}\,, \\
(VI) &= 2\lambda_z + 2\lambda_{z'}(IV) - \lambda(IV)_{z'} + (IV)_{y'}\,, \\
(IV) &= F_{z'}\lambda - \frac{1}{2}(F_{y'} - G_{z'})\,, \\
(II) &= \frac{1}{2}F_{z'}\lambda^2 - \frac{1}{2}(F_{y'} - G_{z'})\lambda - \frac{1}{2}G_{y'}\,.
\end{aligned}$$

Remark 1.3.1. If $A = 0$ footnote (23) in [55] describes how to proceed. If $C \neq 0$, then we should interchange coordinates y and z which results in interchanging F for G, A for $-C$ and B for $-B$ and consider the roots of $C\xi^2 + B\xi = 0$ for further subdivision. The case $A = 0$, $C = 0$ and $B \neq 0$ is described in Example (11.5) of [55] and belongs to Case IIa1.

Theorem 1.3.2. *The following statements are simplified versions of some of the results that appear in [55], namely the ones that we will need in Chapter 2 for an application to control theory :*

- *[55, Theorem I] Case I is always variational,*
- *[55, Theorem II] Case IIa1 is always variational,*
- *[55, Theorem IV] Case IIb1' is always variational,*
- *[55, Theorem VII] Case IIIb is never variational,*
- *[55, Theorem VIII] Case IV is never variational.*

1.3.2 Geometric version of Douglas' classification

The techniques used by Douglas turned out to be very difficult to generalize to higher dimensions. In this section we will give a geometric version of the Helmholtz conditions as well as a geometric reformulation of Douglas' classification following [49]. To this end we will use the calculus of sections along a map [114], more precisely of tensor fields along the tangent bundle projection $\tau_Q : TQ \longrightarrow Q$ for the autonomous case and tensor fields along the projection $\pi : \mathbb{R} \times TQ \longrightarrow \mathbb{R} \times Q$ for the non-autonomous case. This theory has also proved useful in the study of separability of SODEs [115, 142].

Douglas' analysis has led the authors of [49] to propose a generalization of the first broad classification of Douglas to arbitrary dimension n, based on properties of the so-called Jacobi endomorphism Φ and the dynamical covariant derivative ∇. Both operators are essentially defined by the geometry of the SODE Γ that is generated by the system (1.12). In the approach of [47, 49, 144], the system (1.12) is represented by the vector field

$$\Gamma = \frac{\partial}{\partial t} + \dot{q}^i \frac{\partial}{\partial q^i} + \Gamma^i(t, q, \dot{q}) \frac{\partial}{\partial \dot{q}^i}$$

on the first jet bundle $J^1\pi$ of a bundle $\pi : E = \mathbb{R} \times Q \to \mathbb{R}$. We will use the notation π_1 for the projection $J^1\pi \to E$. We will refer to sections of the pullback bundle $\pi_1^*(TE) \to J^1\pi$ as vector fields along π_1 and denote the set of such sections by $\mathfrak{X}(\pi_1)$. For most of our purposes in the next chapter one may think of Q as being $\mathbb{R}^n$, and of π and π_1 as the projections $\mathbb{R}^{n+1} \to \mathbb{R}, (t, q) \mapsto t$ and $\mathbb{R}^{2n+1} \to \mathbb{R}^{n+1}, (t, q, \dot{q}) \mapsto (t, q)$, respectively.

We will first give the necessary ingredients to deal with autonomous SODEs and then modify them for the time-dependent case.

Calculus along the projection $\tau_Q : TQ \to Q$. A vector field along τ_Q is a section of the pullback bundle $\tau_Q^*(TQ) \longrightarrow TQ$. It can also be thought of as a map $X : TQ \longrightarrow TQ$ that makes the diagram in the left of Figure 1.1 commutative. We denote the set of vector fields along τ_Q by $\mathfrak{X}(\tau_Q)$. Analogous definitions can be given for other tensor fields, for instance a one-form along τ_Q will be a map $\theta : TQ \longrightarrow T^*Q$ making the diagram in the middle of Figure 1.1 commutative. The last type we will need in this section is a (0,2)-tensor field along τ_Q, shown in the right of Figure 1.1.

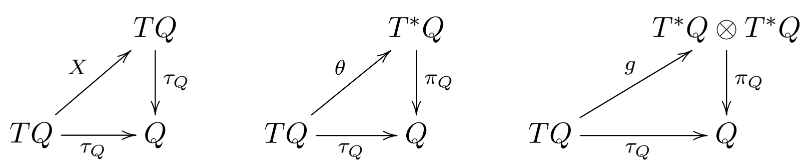


FIGURE 1.1: Tensor fields along τ_Q: $X \in \mathfrak{X}(\tau_Q)$, $\theta \in \Lambda^1(\tau_Q)$ and $g \in T_0^2(\tau_Q)$.

In local cooordinates, tensor fields along τ_Q are written as tensor fields on Q with coefficients on TQ. For instance for the above examples we have

$$\begin{aligned} X &= X^i(q, \dot{q}) \frac{\partial}{\partial q^i} \in \mathfrak{X}(\tau_Q)\,, \\ \theta &= \theta_i(q, \dot{q}) dq^i \in \Lambda^1(\tau_Q)\,, \end{aligned}$$

$$g = g_{ij}(q,\dot q)dq^i \otimes dq^j \in T^2_0(\tau_Q)\,.$$

Non-linear connection from a SODE. A SODE Γ on TQ provides a nonlinear connection on TQ, defined by the horizontal lift

$$\begin{array}{ccc} \mathfrak{X}(Q) & \longrightarrow & \mathfrak{X}(TQ) \\ X & \longmapsto & X^h = \frac{1}{2}(X^c + [X^v, \Gamma])\,, \end{array}$$

where X^c denotes the complete lift of X and X^v its vertical lift, see [53]. A local basis for the horizontal space is

$$H_i = \left(\frac{\partial}{\partial q^i}\right)^h = \frac{\partial}{\partial q^i} - \Gamma^j_i \frac{\partial}{\partial \dot q^j}\,, \quad \text{where} \quad \Gamma^j_i = -\frac{1}{2}\frac{\partial \Gamma^j}{\partial \dot q^i}\,.$$

The Jacobi endomorphism Φ and the dynamical covariant derivative ∇. Every vector field $Z \in \mathfrak{X}(TQ)$ has a unique decomposition as $Z = X^h + Y^v$, where $X, Y \in \mathfrak{X}(\tau_Q)$. In particular we can split the bracket $[\Gamma, X^h]$ as

$$[\Gamma, X^h] = (\nabla X)^h + (\Phi(X))^v\,,$$

from where the Jacobi endomorphism Φ and the dynamical covariant derivative ∇ are defined. Φ is a $(1,1)$-tensor field along τ_Q and ∇ is a degree 0 derivation which can be extended by duality to act on arbitrary tensor fields. In particular ∇ can be extended to act on $\Lambda^1(\tau_Q)$ and $T^2_0(\tau_Q)$ as

$$\begin{aligned} \nabla\theta(X) &= \nabla(\theta(X)) - \theta(\nabla X)\,, \\ \nabla g(X,Y) &= \Gamma(g(X,Y)) - g(\nabla X, Y) - g(X, \nabla Y)\,. \end{aligned}$$

Notice that on the standard local basis of vector fields on Q (a particular case of vector field along τ_Q) we get

$$\begin{aligned} \Phi\left(\frac{\partial}{\partial q^j}\right) &= \Phi^i_j(q,\dot q)\frac{\partial}{\partial q^i} = \left(-\frac{\partial \Gamma^i}{\partial q^j} - \Gamma^k_j\Gamma^i_k - \Gamma(\Gamma^i_j)\right)\frac{\partial}{\partial q^i}\,, \\ \nabla\left(\frac{\partial}{\partial q^j}\right) &= \Gamma^i_j(q,\dot q)\frac{\partial}{\partial q^i} = -\frac{1}{2}\frac{\partial \Gamma^i}{\partial \dot q^j}\frac{\partial}{\partial q^i}\,, \end{aligned}$$

which are the same expressions that appear in the Helmholtz conditions.

Vertical and horizontal derivatives. The vertical and horizontal derivatives D^v_X and D^h_X are defined analogously from the unique decomposition

$$[X^h, Y^v] = (D^h_X Y)^v - (D^v_Y X)^h\,.$$

Both D^v_X and D^h_X are degree 0 derivations. We will also need the operation D^v, whose action on a tensor field U is defined by

$$D^v U(X, ...) = D^v_X U(...)\,.$$

Geometric version of Helmholtz conditions. A regular Lagrangian $L(q,\dot q)$ exists for the system

$$\ddot q^i(t) = \Gamma^i(q^j(t), \dot q^j(t))\,, \quad i,j = 1,\dots,n$$

if and only if there is a nondegenerate symmetric (0,2)-tensor field g along τ_Q (i.e. a multiplier matrix) such that

$$\begin{aligned} g(\Phi(X),Y) &= g(X,\Phi(Y))\,,\\ (D^v_X g)(Y,Z) &= (D^v_Y g)(X,Z)\,,\\ \nabla g &= 0\,. \end{aligned}$$

The Haantjes tensor. The Haantjes tensor is a $(1,2)$-tensor field along τ_Q defined by

$$\begin{aligned} H_\Phi(X,Y) &= D^v\Phi(\Phi X,\Phi Y) - \Phi D^v\Phi(\Phi X,Y)\\ &\quad -\Phi D^v\Phi(X,\Phi Y) + \Phi^2 D^v\Phi(X,Y)\\ &= C^v_\Phi(\Phi(X),Y) - \Phi(C^v_\Phi(X,Y))\,, \end{aligned} \tag{1.19}$$

where $C^v_\Phi(X,Y) = [D^v_X\Phi,\Phi]\,(Y)$, see [47, 49]. This tensor will be useful in the geometric version of Douglas' classification and we will need it in Chapter 2.

All the above ingredients for non-autonomous systems. A vector field along $\pi_1 : \mathbb{R}\times TQ \longrightarrow \mathbb{R}\times Q$ is now a section of the pull-back bundle $\pi_1^*(T(\mathbb{R}\times Q)) \to \mathbb{R}\times TQ$. The $C^\infty(\mathbb{R}\times TQ)$ module of vector fields along π_1 will be denoted by $\mathfrak{X}(\pi_1)$. In local coordinates we have

$$X = X^0(t,q,\dot q)\frac{\partial}{\partial t} + X^i(t,q,\dot q)\frac{\partial}{\partial q^i} \in \mathfrak{X}(\pi_1)\,,$$

One particular example of a vector field along π_1 is the so-called canonical vector field $\mathbf{T} = \partial/\partial t + \dot q^i\partial/\partial q^i$, but also vector fields on Q and on $\mathbb{R}\times TQ$ can be thought of as being vector fields along π_1. In this way, one may see that the set $\{\mathbf{T},\partial/\partial q^i\}$ locally spans $\mathfrak{X}(\pi_1)$. Analogously we can define arbitrary tensor fields along π_1.

A SODE on $\mathbb{R}\times TQ$ also provides a nonlinear connection on $\mathbb{R}\times TQ$. Now the horizontal lift is defined by

$$\begin{array}{ccc} \mathfrak{X}(\mathbb{R}\times Q) & \longrightarrow & \mathfrak{X}(T(\mathbb{R}\times Q))\\ X & \longmapsto & X^h = \frac{1}{2}(X^{(1)} + [X^v,\Gamma] + \langle X, dt\rangle\Gamma)\,, \end{array}$$

where $X^{(1)}$ denotes the prolongation of X, see [146]. Splitting the brackets $[\Gamma, X^h]$ and $[X^h, Y^v]$ again into a horizontal and vertical part defines the operators ∇, Φ, D^v_X and D^h_X. Now we describe them in some more detail. The operator $\nabla : \mathfrak{X}(\pi) \to \mathfrak{X}(\pi)$ is a degree 0 derivation, which means that, for functions $F \in C^\infty(J^1\pi)$ and vector fields $X \in \mathfrak{X}(\pi_1)$ along π_1 it satisfies

$$\nabla(FX) = \Gamma(F)X + F\nabla X\,.$$

In the sequel we only need the action of ∇ on the basis $\{\mathbf{T},\partial/\partial q^i\}$, which is given by

$$\nabla\mathbf{T} = 0, \qquad \nabla\frac{\partial}{\partial q^j} = \Gamma^i_j(t,q,\dot q)\frac{\partial}{\partial q^i}, \quad \text{with} \quad \Gamma^i_j = -\frac{1}{2}\frac{\partial\Gamma^i}{\partial\dot q^j}\,.$$

The second operator Φ defines a (1,1)-tensor field along π_1, meaning that $\Phi(FX) = F\Phi(X)$. We may write it locally as

$$\Phi = \Phi^i_j(t,q,\dot q)\frac{\partial}{\partial q^i}\otimes(dq^j - \dot q^j dt), \quad \text{with} \quad \Phi^i_j = -\frac{\partial\Gamma^i}{\partial q^j} - \Gamma^k_j\Gamma^i_k - \Gamma(\Gamma^i_j)\,.$$

The operation ∇ can be further extended by duality to arbitrary tensor fields along π_1. In particular, $\nabla\Phi$ stands for the (1,1)-tensor field along π_1, given by

$$(\nabla\Phi)(X) = \nabla(\Phi(X)) - \Phi(\nabla X)\,.$$

The coefficients of $\nabla\Phi = (\nabla\Phi)^i_j \partial/\partial q^i \otimes (dq^j - \dot{q}^j dt)$ are then

$$(\nabla\Phi)^i_j = \Gamma(\Phi^i_j) + \Gamma^i_m \Phi^m_j - \Gamma^m_j \Phi^i_m. \tag{1.20}$$

The last operator we need is the vertical derivative D^v_X. For each $X \in \mathfrak{X}(\pi_1)$ it maps vector fields along π to vector fields along π. It can be defined by requiring that it vanishes on both $\mathbf{T}$ and the coordinate vector fields $\partial/\partial q^i$, and that it satisfies $D^v_X(FY) = X^v(F)Y + FD^v_X Y$ for all $F \in C^\infty(J^1\pi)$ and $Y \in \mathfrak{X}(\pi)$.

The aforementioned Helmholtz conditions (1.16)-(1.16) can be written in a form that makes use of the above geometric calculus. In [145] it is shown that a regular Lagrangian exists for the system (1.12) if and only if there is a nondegenerate symmetric (0,2)-tensor field g along π (i.e. a multiplier) such that

$$g(\mathbf{T}, X) = 0, \qquad g(\Phi(X), Y) = g(X, \Phi(Y)), \qquad (D^v_X g)(Y, Z) = (D^v_Y g)(X, Z), \qquad \nabla g = 0\,, \tag{1.21}$$

for arbitrary $X, Y, Z \in \mathfrak{X}(\pi_1)$. We prefer to use this geometric approach to the Helmholtz conditions, over the more analytical style of Douglas' paper, for the reason that it can be conveniently applied in the next chapter to a (noncoordinate) frame of eigenvectors of Φ. More details on this calculus may be found in the review paper [141].

The Φ-condition represents an algebraic relation between the different components of the multiplier g and as such it forms the basis of the classification of the problem in several subcases. The Haantjes tensor can be defined in the time-dependent setting by the same expression as in (1.19).

Geometric Douglas' classification. Now we reproduce the proposal in [49] for a geometric version of Douglas' classification, based on properties of the Jacobi endomorphism Φ and the dynamical covariant derivative ∇ already introduced, and also on the properties of the Haantjes tensor H_Φ for some of the subcases. The cases highlighted in blue are always variational and the cases highlighted in red are never variational, as mentioned already in Section 1.3.1.

- Case I: Φ is a multiple of the identity tensor I.
- Case II: $\nabla\Phi$ is a linear combination of Φ and I.
  - Case IIa: Φ has distinct eigenvalues.
    - Case IIa1: $H_\Phi = 0$.
    - Case IIa2: H_Φ has one independent component.
    - Case IIa3: H_Φ has two independent components.
  - Case IIb: The eigenvalues of Φ coincide.

- $*$ Case IIb1: $H_\Phi = 0$.
  - Case IIb1'.
  - Case IIb1".
- $*$ Case IIb2: $H_\Phi \neq 0$.

- Case III: $\nabla^2\Phi$ is a linear combination of $\nabla\Phi$, Φ and I.
  - Case IIIa: $\det[\Phi, \nabla\Phi] \neq 0$.
  - Case IIIb: $\det[\Phi, \nabla\Phi] = 0$.
- Case IV: $\nabla^2\Phi$, $\nabla\Phi$, Φ and I are linearly independent.

1.4 Alternative characterizations of the inverse problem

Notice that the inverse problem (multiplier version) poses a question involving both the multipliers and the Lagrangian as unknowns but then the Helmholtz conditions provide a way to reformulate the problem just in terms of the existence of multipliers. The same problem can be reformulated in terms of the existence of a Poincaré-Cartan two-form, that is, a two-form on TQ satisfying some properties. This characterization of being variational will be very useful in the sequel and was given in [43] by Crampin.

Let $V(TQ)$ denote the set of all vertical vector fields for $\tau_Q\colon TQ \to Q$, that is, $V(TQ) = \mathrm{Ker}T\tau_Q$.

Theorem 1.4.1 ([43]). *A SODE Γ on TQ is variational if and only if there exists a two-form Ω on TQ of maximal rank such that*

(i) $d\Omega = 0$,

(ii) $\Omega(v_1, v_2) = 0$ *for all* $v_1, v_2 \in V(TQ)$,

(iii) $\mathcal{L}_\Gamma\Omega = 0$.

In [48] an alternative characterization analogous to the one in [43] is given for the time-dependent case:

Theorem 1.4.2 ([48]). *A SODE Γ on $\mathbb{R}\times TQ$ is variational if and only if there exists a two-form Ω on $\mathbb{R}\times TQ$ of maximal rank such that*

(i) $d\Omega = 0$,

(ii) $\Omega(v_1, v_2) = 0$ *for all* $v_1, v_2 \in V(\mathbb{R}\times TQ)$,

(iii) $i_\Gamma\Omega = 0$.

Since 1980, the inverse problem has been considered by many authors [43, 80, 124, 140, 152] giving a geometric interpretation of Douglas' classification and generalizing some of the results to higher dimensions, although no classification is available for dimension greater than two. Some of the cases that have been generalized are Case I and Case IIa1 in Douglas' classification [55], which are always variational for arbitrary dimension, see [143] and [47], respectively. Case I was also proved to be always variational in [3] and [74] using different approaches. In [3] the proof is given using exterior differential systems [26]. The strategy can be roughly summarized as follows. We can define a submodule of two-forms satisfying all conditions in Theorem 1.4.2, but closure. Then we take exterior derivatives and shrink the original submodule until we obtain a differential ideal. In this process all algebraic conditions are imposed, to restrict the possible two-forms. If we still have maximal rank two-forms among them, then we write the closure conditions as a Pfaffian system and apply the Cartan test for involution, see [26]. Exterior differential systems have also been applied more recently to obtain generalizations of other cases in Douglas' classification, see [2, 54].

1.5 New geometric characterization of the inverse problem

In this section we provide a new characterization of the inverse problem of the calculus of variations which will be very useful in Chapters 3, 4 and 5 in order to define, for instance, the variationality of a SODE on a submanifold of TQ or on a Lie algebroid. We regard this result as a previous step towards constrained systems, as well as the other settings considered in this thesis, namely, Lagrangian mechanics on Lie algebroids and discrete mechanics. Some of the results of this section are included in [11].

We will give this new characterization in terms of the existence of a Legendre transformation, using Lagrangian submanifolds both for autonomous and non-autonomous systems. Some references on symplectic and Poisson geometry are [1, 30, 104, 164]. The definition of a Lagrangian submanifold is recalled in Appendix A.

The notion of Legendre transformation is a key element in this section and we quickly review it here. It provides, if the Lagrangian is regular, a transition between the main branches of analytic mechanics, namely Lagrangian mechanics and Hamiltonian mechanics.

Definition 1.5.1. *Let L be a Lagrangian function on TQ. The fiber derivative*

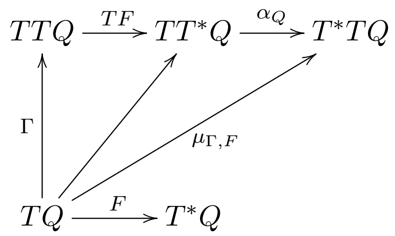
$$\begin{array}{rccc} Leg_L: & TQ & \longrightarrow & T^*Q \\ & v_q & \longmapsto & Leg_L(v_q)\,, \end{array}$$

defined by $\langle Leg_L(v_q), w_q\rangle = \frac{d}{dt}\big|_{t=0} L(v_q + tw_q)$ *is known as the Legendre transformation of L. Locally it is given by* $(q^i, \dot{q}^j) \longmapsto (q^i, p_j = \frac{\partial L}{\partial \dot{q}^j})$.

1.5.1 Autonomous SODEs

We will follow a symplectic approach working with Lagrangian submanifolds of symplectic manifolds [162] associated to the geometry of the tangent bundle, which is the space where a SODE is geometrically defined. We will derive the Helmholtz conditions in terms of the closedness of a suitable one-form, constructed from the given SODE and a transformation between the tangent bundle and its dual, the cotangent bundle.

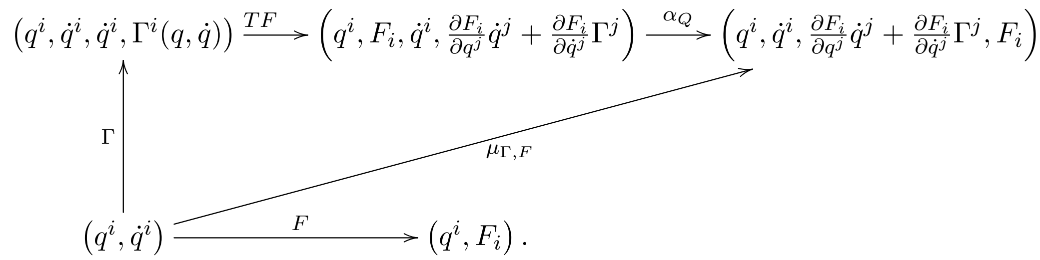
More precisely, for a given SODE $\Gamma : TQ \longrightarrow TTQ$ and a local diffeomorphism $F : TQ \longrightarrow T^*Q$ of fibre bundles over Q, that is, satisfying $\pi_Q \circ F = \tau_Q$, we define a submanifold $\Sigma_{\Gamma,F} := \mathrm{Im}(\mu_{\Gamma,F}) \subset T^*TQ$, where $\mu_{\Gamma,F} = \alpha_Q \circ TF \circ \Gamma$ is a one-form on TQ and α_Q denotes the Tulczyjew isomorphism recalled in Appendix B.



Let $(q^i, \dot{q}^i)$ denote the fibered coordinates on TQ. Then F and Γ are given by

$$F\left(q^i, \dot{q}^i\right) = \left(q^i, F_i(q, \dot{q})\right), \quad \Gamma\left(q^i, \dot{q}^i\right) = \left(q^i, \dot{q}^i, \dot{q}^i, \Gamma^i(q, \dot{q})\right),$$

and the above diagram in coordinates becomes



Note that $\mu_{\Gamma,F}$ is a one-form on TQ locally given by $\left(\frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j\right) dq^i + F_i d\dot{q}^i$. From this last expression it is easy to deduce that

$$\mu_{\Gamma,F} = \mathcal{L}_\Gamma F^* \theta_Q\,, \tag{1.22}$$

where θ_Q denotes the Liouville one-form on T^*Q, see Appendix A.

In this section we will show that the inverse problem of the calculus of variations for a SODE Γ is equivalent to seeing whether or not it is possible to find a local diffeomorphism $F : TQ \longrightarrow T^*Q$ of fibre bundles over Q such that $\Sigma_{\Gamma,F} = \mathrm{Im}(\mu_{\Gamma,F})$ is a Lagrangian submanifold of (T^*TQ, ω_{TQ}). This characterization will be useful for our approach to the inverse problem for constrained systems, as well as Lagrangian mechanics on Lie algebroids and discrete mechanics.

Observe that since $\Sigma_{\Gamma,F}$ is the image of the one-form $\mu_{\Gamma,F}$ on TQ, $\Sigma_{\Gamma,F}$ is a Lagrangian submanifold of (T^*TQ, ω_{TQ}) if and only if $\mu_{\Gamma,F}$ is closed, i.e. $d\mu_{\Gamma,F} = 0$. Therefore, using the Poincaré Lemma we deduce the local existence of a function L on TQ such that $\mu_{\Gamma,F} = dL$.

Theorem 1.5.2. *A SODE Γ on TQ is variational if and only if there exists a local diffeomorphism $F : TQ \longrightarrow T^*Q$ of fibre bundles over Q such that Im$(\mu_{\Gamma,F})$ is a Lagrangian submanifold of (T^*TQ, ω_{TQ}).*

Proof. We use the characterization in Theorem 1.4.1 to prove this result.

$\Leftarrow$ Define $\Omega = -d(F^*\theta_Q)$ and note that if $F(q^i, \dot{q}^i) = (q^i, F_i(q, \dot{q}))$, then

$$\begin{aligned} \mathcal{L}_\Gamma F^*\theta_Q &= \mathcal{L}_\Gamma(F_i dq^i) = \Gamma(F_i)dq^i + F_i d\dot{q}^i \\ &= \left(\frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j\right) dq^i + F_i d\dot{q}^i = \mu_{\Gamma,F}. \end{aligned}$$

Then Ω trivially satisfies all the conditions in Theorem 1.4.1.

$\Rightarrow$ From Theorem 1.4.1 we have that Γ is variational if and only if there exists a nondegenerate two-form Ω on TQ satisfying $\mathcal{L}_\Gamma\Omega = 0$, $\Omega(v, w) = 0$ for all $v, w \in V(TQ)$ and $d\Omega = 0$. From the last condition we deduce that locally $\Omega = d\Theta$ on a neighborhood $U \subseteq TQ$, where Θ is a one-form on U. The restriction of $d\Theta$ to vertical subspaces is zero. Thus the restriction of Θ to each fiber is exact, then there is a function $f : U \to \mathbb{R}$ such that $\Theta(v) = \langle df, v\rangle$ for any $v \in V(TQ)$. Therefore, $\widetilde{\Theta} = \Theta - df$ verifies $\widetilde{\Theta}(v) = 0$ for all $v \in V(TQ)$ and $d\widetilde{\Theta} = \Omega$. Using $\tilde{\Theta}$ we construct the map $F : U \subseteq TQ \to T^*Q$ as follows:

$$\langle F(v_q), w_q\rangle = \langle\widetilde{\Theta}(v_q), W_q\rangle\,,$$

where $v_q \in TQ$, $w_q \in TQ$ and $W_q \in TTQ$ satisfies $T\tau_Q(W_q) = w_q$. This definition does not depend on the choice of W_q since $\widetilde{\Theta}$ vanishes on vertical vector fields. Then, it is easy to show that $\widetilde{\Theta} = F^*\theta_Q$ and from equation (1.22), $\mu_{\Gamma,F} = \mathcal{L}_\Gamma\widetilde{\Theta}$ verifies

$$d\mu_{\Gamma,F} = d\mathcal{L}_\Gamma\widetilde{\Theta} = \mathcal{L}_\Gamma d\widetilde{\Theta} = \mathcal{L}_\Gamma\Omega = 0\,.$$

Hence $\mathrm{Im}(\mu_{\Gamma,F})$ is a Lagrangian submanifold of (T^*TQ, ω_{TQ}). Note that the nondegeneracy of Ω implies that $\det\left(\frac{\partial F_i}{\partial \dot{q}^j}\right) \neq 0$, which is precisely the condition for F to be a local diffeomorphism. ∎

Observe that the submanifold $\Sigma_{\Gamma,F} \subset T^*TQ$, given in local coordinates by $\left(q^i, \dot{q}^i, \frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j, F_i\right)$ is a Lagrangian submanifold of (T^*TQ, ω_{TQ}) if and only if there exists a locally defined function $L\colon TQ \to \mathbb{R}$ such that

$$\frac{\partial L}{\partial \dot{q}^i} = F_i \quad \text{and} \quad \frac{\partial L}{\partial q^i} = \frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j.$$

We have that

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} &= \frac{dF_i}{dt} - \frac{\partial F_i}{\partial q^j}\dot{q}^j - \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j \\ &= \frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\ddot{q}^j - \frac{\partial F_i}{\partial q^j}\dot{q}^j - \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j \\ &= \frac{\partial F_i}{\partial \dot{q}^j}(\ddot{q}^j - \Gamma^j)\,. \end{aligned}$$

Thus the solutions to the Euler-Lagrange equations for L coincide with the solutions to the SODE Γ, since F is a local diffeomorphism, that is, locally the matrix $\left(\frac{\partial F_i}{\partial \dot{q}^j}\right)$ is nondegenerate. Then the multipliers for the Helmholtz conditions are $g_{ij} = \frac{\partial F_i}{\partial \dot{q}^j}$.

Remark 1.5.3. Since $\alpha_Q : TT^*Q \to T^*TQ$ is a symplectomorphism (see Appendix B) then we can alternatively characterize the inverse problem of the calculus of variations for a SODE Γ by checking whether or not the submanifold $S_{\Gamma,F}$ defined by

$$S_{\Gamma,F} = TF(\Gamma(Q)) = \alpha_Q^{-1}(\mu_{\Gamma,F}(Q))$$

is a Lagrangian submanifold of the symplectic manifold $(TT^*Q, d_T\omega_Q)$.

Remark 1.5.4. The submanifold $\Sigma_{F,\Gamma}$ will be Lagrangian if and only if

$$d\left(\left(\frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j\right) dq^i + F_i d\dot{q}^i\right) \;=\; 0\,.$$

Equivalently, we get the following conditions:

$$\frac{\partial F_i}{\partial \dot{q}^k} = \frac{\partial F_k}{\partial \dot{q}^i}\,, \tag{1.23}$$

$$\frac{\partial^2 F_i}{\partial q^k \partial q^j}\dot{q}^j + \frac{\partial^2 F_i}{\partial q^k \partial \dot{q}^j}\Gamma^j + \frac{\partial F_i}{\partial \dot{q}^j}\frac{\partial \Gamma^j}{\partial q^k} = \frac{\partial^2 F_k}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 F_k}{\partial q^i \partial \dot{q}^j}\Gamma^j + \frac{\partial F_k}{\partial \dot{q}^j}\frac{\partial \Gamma^j}{\partial q^i}\,, \tag{1.24}$$

$$\frac{\partial F_k}{\partial q^i} = \frac{\partial^2 F_i}{\partial \dot{q}^k \partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial q^k} + \frac{\partial^2 F_i}{\partial \dot{q}^k \partial \dot{q}^j}\Gamma^j + \frac{\partial F_i}{\partial \dot{q}^j}\frac{\partial \Gamma^j}{\partial \dot{q}^k}\,. \tag{1.25}$$

Direct computations showing the equivalence between the equations (1.23)-(1.25) and the Helmholtz conditions (1.14)-(1.16) for $g_{ij} = \frac{\partial F_i}{\partial \dot{q}^j}$ can be carried out, as shown in Appendix C.

Note that the conditions (1.23)-(1.25) are given in the standard basis. We can also easily recover the usual Helmholtz conditions using the basis $\left\{V_i = \frac{\partial}{\partial \dot{q}^i}, H_i = \frac{\partial}{\partial q^i} + \frac{1}{2}\frac{\partial \Gamma^k}{\partial \dot{q}^i}\frac{\partial}{\partial \dot{q}^k}\right\}$. We just need to impose that $d\mu_{\Gamma,F}$ evaluated on pairs $(H_i, H_j), (H_i, V_j), (V_i, V_j)$ vanishes and also use the condition $d\mu_{\Gamma,F}(H_i, V_j) - d\mu_{\Gamma,F}(H_j, V_i) = 0$, which is the same as $(1.25)_{ik} - (1.25)_{ki} = 0$.

Remark 1.5.5. In Theorem 1.5.2 we are asking for the existence of a Legendre transformation for Γ. In [28, Theorem 5.3] the characterization is given in terms of the existence of a Poincaré-Cartan one-form, so the semi-basic one-form they seek and the local diffeomorphism we seek are simply related by $\theta = F^*\theta_Q$.

Remark 1.5.6. If we admit that the matrix (g_{ij}) could be degenerate, then we get conditions for the existence of a singular Lagrangian L such that

$$g_{ij}(\ddot{q}^j - \Gamma^j) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\,,$$

which implies that the solutions of the SODE are also solutions to the Euler-Lagrange equations for L, but not conversely.

Example 1.5.7. Let $Q = \mathbb{R}^2$ and consider a SODE Γ given by $\ddot{x} = f(x,y)$, $\ddot{y} = f(x,y)$, that is, $\Gamma^1 = \Gamma^2 = f(x,y)$. Then $L = \frac{1}{2}(\dot{x} - \dot{y})^2$ is a singular Lagrangian that gives the dynamics $\ddot{x} = \ddot{y}$, which includes the solutions to Γ, and satisfies

$$g_{ij}(\ddot{q}^j - \Gamma^j) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}$$

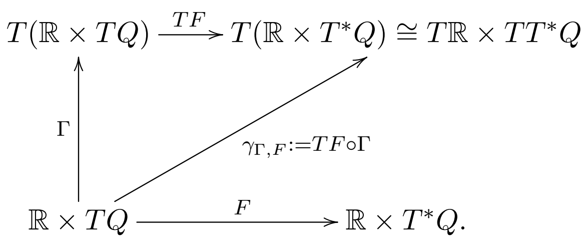
with $g_{11} = g_{22} = 1$ and $g_{12} = g_{21} = -1$. For some choices of $f(x, y)$, the SODE will fall into one of the cases in [55] which do not admit a regular Lagrangian. For instance if we take $f(x, y) = xy$, then, in the notation of [55] (except for the coordinates which we denote as $(t, x, y, \dot{x}, \dot{y})$), we get

$$\begin{array}{rclrclrcl} A &=& -2x, & B &=& (y-x), & C &=& 2y\,, \\ A_1 &=& -2\dot{x}, & B_1 &=& 2(\dot{y}-\dot{x}), & C_1 &=& 2\dot{y}\,, \\ A_2 &=& -2xy, & B_2 &=& 0, & C_2 &=& 2xy\,. \end{array}$$

Then the determinant of $\left(\begin{smallmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{smallmatrix}\right)$ is nonzero and the example falls into the nonvariational Case IV of Douglas [55].

1.5.2 Time-dependent SODEs

Consider now the following diagram, where $F : \mathbb{R} \times TQ \longrightarrow \mathbb{R} \times T^*Q$ is a local diffeomorphism over $\mathbb{R} \times Q$:



In local coordinates, if we write $\Gamma(t, q^i, \dot{q}^i) = (t, q^i, \dot{q}^i, 1, \dot{q}^i, \Gamma^i(t, q^j, \dot{q}^j))$ and $F(t, q^i, \dot{q}^i) = (t, q^i, F_i(t, q, \dot{q}))$, then we get

$$\gamma_{\Gamma,F}(t, q^i, \dot{q}^i) = \left(t, q^i, F_i(t, q, \dot{q}), 1, \dot{q}^i, \frac{\partial F_i}{\partial t} + \dot{q}^j \frac{\partial F_i}{\partial q^j} + \Gamma^j \frac{\partial F_i}{\partial \dot{q}^j}\right).$$

In order to characterize the property of being variational for a time-dependent SODE we will use Lagrangian submanifolds for a Poisson strucutre in $T(\mathbb{R} \times T^*Q)$. We will first recall some definitions that are needed.

Definition 1.5.8. *Let f be a function on a manifold P. We can define the complete and vertical lift of the function f to TP, which will be denoted respectively by f^c and f^v, as*

$$f^c = \frac{d}{dt}(f \circ \tau_P) \quad \textit{and} \quad f^v = f \circ \tau_P\,,$$

where $\tau_P \colon TP \to P$ is the canonical projection, see [167].

Definition 1.5.9. *A Poisson manifold is a pair $(P, \{\cdot,\cdot\})$, where P is a manifold and $\{\cdot,\cdot\}$ is an $\mathbb{R}$-bilinear operation*

$$\{\cdot,\cdot\} : \mathcal{C}^\infty(P) \times \mathcal{C}^\infty(P) \longrightarrow \mathcal{C}^\infty(P)\,,$$

known as Poisson bracket, satisfying

- $\{f, g\} = -\{g, f\}$ *(anticommutativity),*
- $\{fg, h\} = f\{g, h\} + g\{f, h\}$ *(Leibniz's rule),*
- $\{\{f, g\}, h\} + \{\{h, f\}, g\} + \{\{g, h\}, f\} = 0$ *(Jacobi's identity),*

for all $f, g, h \in \mathcal{C}^\infty(P)$. If the Jacobi identity is not satisfied then $(P, \{\cdot,\cdot\})$ is called an almost-Poisson manifold. This case will appear in Chapter 6.

A Poisson bracket defines a skew-symmetric $(2,0)$-tensor field Λ on P by $\{f,g\} = \Lambda(df, dg)$. We will call Λ the Poisson bivector.

Definition 1.5.10 ([41, 73])**.** *Let $(P, \{\cdot,\cdot\})$ be a Poisson manifold. The tangent Poisson bracket on TP is given by*

$$\begin{aligned}\{f^c, g^c\}^T &= \{f,g\}^c,\\ \{f^c, g^v\}^T &= \{f,g\}^v,\\ \{f^v, g^v\}^T &= 0.\end{aligned}$$

If (x^i) denote local coordinates in P and the Poisson bivector is given by

$$\Lambda = \frac{1}{2}\Lambda^{ij}(x)\frac{\partial}{\partial x^i}\wedge\frac{\partial}{\partial x^j},$$

then

$$d_T\Lambda = \Lambda^T := \Lambda^{ij}(x)\frac{\partial}{\partial x^i}\wedge\frac{\partial}{\partial \dot{x}^j} + \frac{1}{2}\frac{\partial\Lambda^{ij}(x)}{\partial x^k}\dot{x}^k\frac{\partial}{\partial \dot{x}^i}\wedge\frac{\partial}{\partial \dot{x}^j}$$

is the Poisson bivector corresponding to the bracket $\{\cdot,\cdot\}^T$.

Definition 1.5.11 ([157])**.** *Let $(P, \{\cdot,\cdot\})$ be a Poisson manifold and N be a submanifold of P. Denote by Λ the Poisson bivector and by $\sharp : T^*P \longrightarrow TP$ the induced morphism of vector bundles. The submanifold N is called Lagrangian if*

$$\sharp(TN^\circ) = TN \cap \mathcal{C},$$

where TN° is the annihilator of TN and $\mathcal{C} := Im(\sharp)$ is the characteristic distribution.

Now we consider the projection $\tilde{\pi} : T^*(\mathbb{R}\times Q) \equiv T^*\mathbb{R}\times T^*Q \to \mathbb{R}\times T^*Q$ given by $\tilde{\pi} = (\pi_{\mathbb{R}}, id_{T^*Q})$, that is, $\tilde{\pi}(\alpha_t, \beta_q) = (t, \beta_q)$, where $\alpha_t \in T_t^*\mathbb{R}$ and $\beta_q \in T_q^*Q$. We induce a Poisson bracket on $\mathbb{R}\times T^*Q$ such that $\tilde{\pi}$ is a Poisson morphism, where we are considering in $T^*(\mathbb{R}\times Q)$ the standard Poisson bracket induced by the symplectic two-form $\omega_{\mathbb{R}\times Q}$. Locally, in coordinates (t, q^i, p_i) for $\mathbb{R}\times T^*Q$, we have that the induced bracket $\{\cdot,\cdot\}$ is defined by

$$\{t, q^i\} = \{t, p_i\} = \{q^i, q^j\} = \{p_i, p_j\} = 0 \quad\text{and}\quad \{q^i, p_j\} = \delta^i_j\,.$$

Then we take its tangent lift to $T(\mathbb{R}\times T^*Q)$, which is defined on the induced coordinate functions $(t, q, p, v_t, \dot{q}, \dot{p})$ by

$$\{q^i, \dot{p}_i\}^T = 1,\ \{\dot{q}^i, p_i\}^T = 1$$

and the remaining Poisson brackets vanish. The variational property of Γ will be characterized in terms of Lagrangian submanifolds for this Poisson structure. To be more precise, $\mathrm{Im}(\gamma_{\Gamma,F})$ must be Lagrangian for some local diffeomorphism $F : \mathbb{R}\times TQ \longrightarrow \mathbb{R}\times T^*Q$.

Now we will write the conditions that arise when forcing $\mathrm{Im}(\mu_{\Gamma,F})$ to be Lagrangian. In local coordinates $(t, q, p, v_t, \dot q, \dot p)$ for $T(\mathbb{R}\times T^*Q)$ we have

$$\begin{aligned}
T(\mathrm{Im}(\gamma_{\Gamma,F})) &= \mathrm{span}\left\{\frac{\partial}{\partial t}+\frac{\partial F_j}{\partial t}\frac{\partial}{\partial p_j}+\frac{\partial\Gamma(F_j)}{\partial t}\frac{\partial}{\partial \dot p_j},\right.\\
&\left.\frac{\partial}{\partial q^i}+\frac{\partial F_j}{\partial q^i}\frac{\partial}{\partial p_j}+\frac{\partial\Gamma(F_j)}{\partial q^i}\frac{\partial}{\partial \dot p_j},\frac{\partial}{\partial \dot q^i}+\frac{\partial F_j}{\partial \dot q^i}\frac{\partial}{\partial p_j}+\frac{\partial\Gamma(F_j)}{\partial \dot q^i}\frac{\partial}{\partial \dot p_j}\right\},\\
T(\mathrm{Im}(\gamma_{\Gamma,F}))^\circ &= \mathrm{span}\left\{dv_t, \frac{\partial F_i}{\partial q^j}dq^j-dp_i+\frac{\partial F_i}{\partial t}dt+\frac{\partial F_i}{\partial \dot q^j}d\dot q^j,\right.\\
&\left.\frac{\partial\Gamma(F_i)}{\partial q^j}dq^j-d\dot p_i+\frac{\partial\Gamma(F_i)}{\partial t}dt+\frac{\partial\Gamma(F_i)}{\partial \dot q^j}d\dot q^j\right\},\\
\sharp(T(\mathrm{Im}(\gamma_{\Gamma,F}))^\circ) &= \mathrm{span}\left\{\frac{\partial F_i}{\partial q^j}\frac{\partial}{\partial \dot p_j}+\frac{\partial}{\partial \dot q^i}+\frac{\partial F_i}{\partial \dot q^j}\frac{\partial}{\partial p_j},\frac{\partial\Gamma(F_i)}{\partial q^j}\frac{\partial}{\partial \dot p_j}+\frac{\partial}{\partial q^i}+\frac{\partial\Gamma(F_i)}{\partial \dot q^j}\frac{\partial}{\partial p_j}\right\}.
\end{aligned}$$

As $\mathcal{C}=\mathrm{span}\{\frac{\partial}{\partial q^i},\frac{\partial}{\partial p_i},\frac{\partial}{\partial \dot q^i},\frac{\partial}{\partial \dot p_i}\}$, the equality $\sharp(T(\mathrm{Im}(\gamma_{\Gamma,F}))^\circ)=T(\mathrm{Im}(\gamma_{\Gamma,F}))\cap\mathcal{C}$ holds if the following conditions are satisfied

$$\frac{\partial F_j}{\partial \dot q^i}=\frac{\partial F_i}{\partial \dot q^j},\quad \frac{\partial F_i}{\partial q^j}=\frac{\partial\Gamma(F_j)}{\partial \dot q^i},\quad \frac{\partial\Gamma(F_j)}{\partial q^i}=\frac{\partial\Gamma(F_i)}{\partial q^j}, \tag{1.26}$$

which in more detail read

$$\frac{\partial F_j}{\partial \dot q^i} = \frac{\partial F_i}{\partial \dot q^j}, \tag{1.27}$$

$$\frac{\partial^2 F_j}{\partial \dot q^i\partial t}+\frac{\partial F_j}{\partial q^i}+\dot q^k\frac{\partial^2 F_j}{\partial \dot q^i\partial q^k}+\frac{\partial\Gamma^k}{\partial \dot q^i}\frac{\partial F_j}{\partial \dot q^k}+\Gamma^k\frac{\partial^2 F_j}{\partial \dot q^i\partial \dot q^k} = \frac{\partial F_i}{\partial q^j}, \tag{1.28}$$

$$\begin{aligned}
\frac{\partial^2 F_j}{\partial q^i\partial t}+\dot q^k\frac{\partial^2 F_j}{\partial q^i\partial q^k}+\frac{\partial\Gamma^k}{\partial q^i}\frac{\partial F_j}{\partial \dot q^k}+\Gamma^k\frac{\partial^2 F_j}{\partial q^i\partial \dot q^k} &= \frac{\partial^2 F_i}{\partial q^j\partial t}+\dot q^k\frac{\partial^2 F_i}{\partial q^j\partial q^k}+\frac{\partial\Gamma^k}{\partial q^j}\frac{\partial F_i}{\partial \dot q^k}\\
&+\Gamma^k\frac{\partial^2 F_i}{\partial q^j\partial \dot q^k}.
\end{aligned} \tag{1.29}$$

Remark 1.5.12. Note that the above conditions are the same that arise if we require that the natural projection of $\mathrm{Im}(\gamma_{\Gamma,F})\subset T(\mathbb{R}\times T^*Q)$ onto TT^*Q be a Lagrangian submanifold for each time coordinate with the symplectic structure $d_T\omega_Q$.

Now we give a characterization of the variational character of a time-dependent SODE in terms of Lagrangian submanifolds of the Poisson manifold $(T(\mathbb{R}\times T^*Q),\{\cdot,\cdot\}^T)$, that is, we provide an analog of Theorem 1.5.2 for the time-dependent setting.

Theorem 1.5.13. *A SODE Γ on $\mathbb{R}\times TQ$ is variational if and only if there is a local diffeomorphism $F:\mathbb{R}\times TQ\longrightarrow\mathbb{R}\times T^*Q$ over $\mathbb{R}\times Q$ such that $Im(\gamma_{\Gamma,F})$ is a Lagrangian submanifold of $(T(\mathbb{R}\times T^*Q),\{\cdot,\cdot\}^T)$.*

Proof. $\Rightarrow$ If Γ is variational then there is a local regular Lagrangian $L:\mathbb{R}\times TQ\longrightarrow\mathbb{R}$ such that

$$\frac{\partial^2 L}{\partial \dot q^i\partial \dot q^j}(\ddot q^j-\Gamma^j)=\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q^i}\right)-\frac{\partial L}{\partial q^i},$$

that is,

$$\frac{\partial^2 L}{\partial \dot q^i\partial \dot q^j}\Gamma^j=\frac{\partial L}{\partial q^i}-\frac{\partial^2 L}{\partial t\partial \dot q^i}-\frac{\partial^2 L}{\partial q^j\partial \dot q^i}\dot q^j.$$

We can define F as the corresponding Legendre transformation, that is, $F_i = \frac{\partial L}{\partial \dot{q}^i}$, and then

$$\gamma_{\Gamma,F}(t, q^i, \dot{q}^i) = \left(t, q^i, \frac{\partial L}{\partial \dot{q}^i}, 1, \dot{q}^i, \frac{\partial L}{\partial q^i}\right),$$

whose image is a Lagrangian submanifold of $(T(\mathbb{R} \times T^*Q), \{\cdot, \cdot\}^T)$.

$\Leftarrow$ Given a local diffeomorphism

$$\begin{array}{rccc} F: & \mathbb{R} \times TQ & \longrightarrow & \mathbb{R} \times T^*Q \\ & (t, q^i, \dot{q}^i) & \longmapsto & (t, q^i, F_i) \end{array}$$

satisfying (1.27), (1.28) and (1.29), we define

$$\Omega = -dF^*\theta_Q - i_\Gamma dF^*\theta_Q \wedge dt = -dF^*\theta_Q + (di_\Gamma F^*\theta_Q - \mathcal{L}_\Gamma F^*\theta_Q) \wedge dt\,,$$

which clearly satisfies $\Omega(v_1, v_2) = 0$ for all $v_1, v_2 \in V(\mathbb{R} \times TQ)$. In local coordinates,

$$\Omega = -\frac{\partial F_i}{\partial q^j} dq^j \wedge dq^i - \frac{\partial F_i}{\partial \dot{q}^j} d\dot{q}^j \wedge dq^i + \left(\frac{\partial F_i}{\partial q^j}\dot{q}^i - \frac{\partial F_j}{\partial q^i}\dot{q}^i - \frac{\partial F_j}{\partial \dot{q}^i}\Gamma^i\right) dq^j \wedge dt + \frac{\partial F_i}{\partial \dot{q}^j}\dot{q}^i d\dot{q}^j \wedge dt\,.$$

Computing the exterior derivative of Ω we get

$$d\Omega = -\frac{\partial \Gamma(F_i)}{\partial q^j} dq^j \wedge dq^i \wedge dt - \frac{\partial \Gamma(F_i)}{\partial \dot{q}^j} d\dot{q}^j \wedge dq^i \wedge dt - \frac{\partial F_i}{\partial q^j} dq^j \wedge d\dot{q}^i \wedge dt - \frac{\partial F_i}{\partial \dot{q}^j} d\dot{q}^j \wedge d\dot{q}^i \wedge dt\,.$$

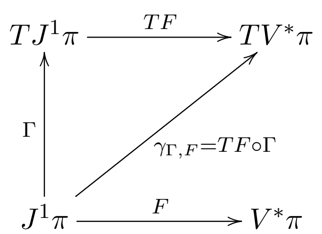
Conditions (1.26) on F yield $d\Omega = 0$. It is also readily checked that $i_\Gamma \Omega = 0$. Since F is a local diffeomorphism, that is, $\operatorname{rank}\left(\frac{\partial F_i}{\partial \dot{q}^j}\right) = n$, the term $\frac{\partial F_i}{\partial \dot{q}^j} d\dot{q}^j \wedge dq^i$ makes Ω have maximal rank. Thus Ω satisfies all the conditions in Theorem 1.4.2 and Γ is variational. ■

Remark 1.5.14. Note that the Poincaré-Cartan two-form Ω_L (see Remark 1.1.4) can be alternatively rewritten as $\Omega_L = \omega + dE_L \wedge dt$, with $\omega = -d\left(dL \circ S - (i_{\frac{\partial}{\partial t}} dL \circ S)dt\right)$ and $E_L = \Delta(L) - L$. If we consider the Legendre transformation $Leg_L : \mathbb{R} \times TQ \to \mathbb{R} \times T^*Q$ locally given by

$$Leg_L(t, q, \dot{q}) = \left(t, q, \frac{\partial L}{\partial \dot{q}^i}\right),$$

then $\omega = -d(Leg_L)^*\theta_Q$, $\Delta(L) = i_{\Gamma_L}(Leg_L)^*\theta_Q$ and $dL \wedge dt = \mathcal{L}_{\Gamma_L}(Leg_L)^*\theta_Q \wedge dt$. These substitutions motivate the definition of Ω in the proof of Theorem 1.5.13 for arbitrary F and Γ instead of Leg_L and Γ_L. For more details on the formulation of time-dependent Lagrangian mechanics see [48, 134] .

Remark 1.5.15. If we replace the trivial bundle $\mathbb{R} \times Q \longrightarrow \mathbb{R}$ by an arbitrary fiber bundle $\pi : E \longrightarrow \mathbb{R}$, then the first jet manifold, denoted by $J^1\pi$, is the generalization of $\mathbb{R} \times TQ$. The generalization of $\mathbb{R} \times T^*Q$ is $V^*\pi$, the dual bundle of the vertical bundle to π. $V^*\pi$ is also equipped with a Poisson structure that can be lifted to $TV^*\pi$, so we could copy the same scheme to study the variationality of a SODE on $J^1\pi$ in terms of $\mathrm{Im}(\gamma_{\Gamma,F})$ being Lagrangian in $(TV^*\pi, \{\cdot, \cdot\}^T)$, see [71].



1.5.3 Implicit second order systems

In this section we go back to autonomous systems, but given in implicit form. Let $T^{(2)}Q$ denote the second order tangent bundle, which is a submanifold of TTQ, given by

$$T^{(2)}Q = \{v \in TTQ : T\pi_Q(v) = \pi_{TQ}(v)\} .$$

Consider now an implicit system of second order differential equations given by a submanifold $M \subset T^{(2)}Q$. Assume M is defined by the vanishing of functions

$$\Phi^i(q, \dot{q}, \ddot{q}) = 0, \quad i = 1, \dots, n\,, \tag{1.30}$$

such that $C := \left(\frac{\partial \Phi}{\partial \ddot{q}}\right)$ is regular, where $C^k_j = \frac{\partial \Phi^k}{\partial \ddot{q}^j}$. We will now derive Helmholtz conditions for the problem of finding a regular Lagrangian L such that the systems

$$\Phi^i(q, \dot{q}, \ddot{q}) = 0 \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0\,, \quad i = 1, \dots, n,$$

have the same solutions. Emulating the explicit case, we aim for a local diffeomorphism over the identity $F : TQ \to T^*Q$ such that $TF(M) \subset TT^*Q$ is a Lagrangian submanifold of $(TT^*Q, d_T\omega_Q)$. If $(q, p, \dot{q}, \dot{p})$ denote fibered coordinates on TT^*Q then locally $d_T\omega_Q = dq \wedge d\dot{p} + d\dot{q} \wedge dp$. The submanifold $TF(M)$ is locally given by

$$\left(q^i, F_i(q, \dot{q}), \dot{q}^i, \frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\ddot{q}^j\right)$$

plus the condition $\Phi^i(q, \dot{q}, \ddot{q}) = 0$ for all $i = 1, \dots, n$. If we write $\omega_{TF} = (TF)^* d_T\omega_Q$ then locally

$$\begin{aligned}
\omega_{TF} &= dq^i \wedge d\left(\frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\ddot{q}^j\right) + d\dot{q}^i \wedge dF_i \\
&= \left(\frac{\partial^2 F_i}{\partial q^k \partial q^j}\dot{q}^j + \frac{\partial^2 F_i}{\partial q^k \partial \dot{q}^j}\ddot{q}^j\right) dq^i \wedge dq^k \\
&\quad + \left(\frac{\partial^2 F_i}{\partial \dot{q}^k \partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial q^k} + \frac{\partial^2 F_i}{\partial \dot{q}^k \partial \dot{q}^j}\ddot{q}^j - \frac{\partial F_k}{\partial q^i}\right) dq^i \wedge d\dot{q}^k \\
&\quad + \frac{\partial F_i}{\partial \dot{q}^k} dq^i \wedge d\ddot{q}^k + \frac{\partial F_i}{\partial \dot{q}^k} d\dot{q}^i \wedge d\dot{q}^k .
\end{aligned}$$

The condition that $TF(M)$ be a Lagrangian submanifold of TT^*Q is equivalent to the condition $(TF \circ i_M)^* d_T\omega_Q = 0$ and can be written as $\omega_{TF}(X, Y) = 0$ for all $X, Y \in \mathfrak{X}(M)$. Therefore we compute a local basis for $\mathfrak{X}(M)$, by imposing that $X \in \mathfrak{X}(T^{(2)}Q)$ satisfies $d\Phi(X) = 0$, and we get

$$A_i = \frac{\partial}{\partial q^i} - \frac{\partial \Phi^j}{\partial q^i}(C^{-1})^k_j \frac{\partial}{\partial \ddot{q}^k}\,, \quad B_i = \frac{\partial}{\partial \dot{q}^i} - \frac{\partial \Phi^j}{\partial \dot{q}^i}(C^{-1})^k_j \frac{\partial}{\partial \ddot{q}^k}\,,$$

because of the regularity of C. Finally the implicit Helmholtz conditions

$$\omega_{TF}(B_i, B_j) = 0, \quad \omega_{TF}(A_i, B_j) = 0 \quad \text{and} \quad \omega_{TF}(A_i, A_j) = 0$$

are respectively given by

$$\frac{\partial F_i}{\partial \dot{q}^j} = \frac{\partial F_j}{\partial \dot{q}^i}\,, \tag{1.31}$$

$$\frac{\partial^2 F_i}{\partial \dot{q}^j \partial q^k}\dot{q}^k + \frac{\partial F_i}{\partial q^j} + \frac{\partial^2 F_i}{\partial \dot{q}^j \partial \dot{q}^k}\ddot{q}^k - \frac{\partial F_j}{\partial q^i} = \frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Phi^r}{\partial \dot{q}^j}(C^{-1})^k_r\,, \tag{1.32}$$

$$\frac{\partial^2 F_i}{\partial q^j \partial q^k}\dot{q}^k + \frac{\partial^2 F_i}{\partial q^j \partial \dot{q}^k}\ddot{q}^k - \frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Phi^r}{\partial q^j}(C^{-1})^k_r = \frac{\partial^2 F_j}{\partial q^i \partial q^k}\dot{q}^k + \frac{\partial^2 F_j}{\partial q^i \partial \dot{q}^k}\ddot{q}^k - \frac{\partial F_j}{\partial \dot{q}^k}\frac{\partial \Phi^r}{\partial q^i}(C^{-1})^k_r\,. \tag{1.33}$$

Remark 1.5.16. Notice that for the system $\Phi^j = \ddot{q}^j - \Gamma^j(q,\dot{q})$, $j = 1,\dots,n$, we have $C = Id$ and therefore we recover conditions (1.23)-(1.25) given earlier in Section 1.5.1.

Using the implicit function theorem to write $\ddot{q}^i = \Gamma^i(q,\dot{q})$ in appropriate neighborhoods, we have that the Helmholtz conditions (1.31)-(1.33) are equivalent to (1.30) being variational. Indeed if (1.30) is variational for a regular Lagrangian L, then $F_i = \frac{\partial L}{\partial \dot{q}^i}$ provides a Lagrangian submanifold $TF(M)$. Conversely, if the there is a local diffeomorphism F satisfying the Helmholtz conditions, then

$$\left\{\Phi^i(q,\dot{q},\ddot{q}) = 0\right\} \equiv \left\{\ddot{q}^i = \Gamma^i(q,\dot{q})\right\} \equiv \left\{\frac{\partial F_i}{\partial \dot{q}^j}\left(\ddot{q}^i - \Gamma^i(q,\dot{q})\right) = 0\right\} \equiv \left\{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0\right\},$$

where in the last equality we use the fact that $TF(M)$ is Lagrangian. Here we have used the notation $\{X(q,\dot{q},\ddot{q}) = 0\} \equiv \{Y(q,\dot{q},\ddot{q}) = 0\}$ to denote that the solutions of the equations $X(q,\dot{q},\ddot{q}) = 0$ and $Y(q,\dot{q},\ddot{q}) = 0$ coincide.

To conclude this chapter we will see a very simple example that clearly shows the difference between the version of the inverse problem that we are discussing now, namely the multiplier version in the implicit description, and the first version of the question raised by Helmholtz [161], introduced in Section 1.2.1.

Example 1.5.17. Consider the system

$$\Phi^1 = e^{\ddot{x}-x} - 1 = 0\,, \quad \Phi^2 = \ddot{y} - y = 0\,, \tag{1.34}$$

which is clearly implicitly variational in the sense that its solutions coincide with the solutions to

$$\Phi^1 = \ddot{x} - x = 0\,, \quad \Phi^2 = \ddot{y} - y = 0\,, \tag{1.35}$$

and for the last system we can find a regular Lagrangian, for instance $L = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + x^2 + y^2\right)$.

Notice that, as it should be, the implicit Helmholtz conditions (1.31)-(1.33) admit solutions, for instance $F_1 = \frac{\partial L}{\partial \dot{x}} = \dot{x}$, $F_2 = \frac{\partial L}{\partial \dot{y}} = \dot{y}$.

On the other hand the original Helmholtz conditions, which can be directly checked on Φ in (1.34), are not satisfied. Indeed, from (1.10) we get

$$2\frac{\partial \Phi^2}{\partial \dot{x}} - 2\frac{d}{dt}\frac{\partial \Phi^1}{\partial \ddot{x}} = -2\frac{d}{dt}e^{\ddot{x}-x} = -2e^{\ddot{x}-x}(\dot{x} - \dddot{x})\,.$$

Chapter 2

Applications to control theory

In this chapter we will see two applications of the inverse problem of the calculus of variations to control theory, more precisely to the problem of stabilization of mechanical systems. In Section 2.1 we will review for completeness the main stability results that will be used. In Section 2.2 we will give an introduction to one fruitful approach to the problem, known as controlled Lagrangian techniques and matching conditions. Then in Section 2.3 we will show an application of Douglas' classification to the problem of stabilization of a class of two-dimensional mechanical systems, based on [57]. More concretely in Section 2.3.2 we provide necessary and sufficient conditions for the variationality of a particular class of SODEs and in Section 2.3.3 we give further conditions for stability. In Section 2.3.4 we apply the results to the inverted pendulum on a cart and the inertia wheel pendulum. In Section 2.4 we will use the Helmholtz conditions to recover some of the matching conditions and we will also derive some new ones for a particular class of mechanical systems, including the inverted pendulum on a cart.

2.1 Stabilization of mechanical systems

With a view on achieving a desired goal, dynamical systems are often modeled in such a way that a controlled quantity may influence its behavior. In this chapter we will consider mechanical systems with a possibly unstable equilibrium. We will be interested in making structural modifications to this system by adding extra controlled external forces or torques to it, in order to arrive at a controlled system where the equilibrium has become stable. In a series of papers by A. M. Bloch, N. Leonard and J. E. Marsden (starting with [21]) it was shown that, subject to a number of assumptions, some of those controlled systems can be seen to be equivalent to the Euler-Lagrange equations of a new Lagrangian, the so-called controlled Lagrangian. This controlled Lagrangian is a modification of the original Lagrangian of the system by means of some control parameters. Sufficient conditions for this situation to occur have been derived in [21] and the technique is often referred to as 'the matching theorems'. Since its first appearance the method of controlled Lagrangians and the matching conditions have been successfully applied in many papers (see [19] for many references). The main advantage of the approach is that, once we know that the controlled system is Lagrangian, we may use energy methods and the available freedom in the choice of controls to analyze the stability of equilibria. An example of a paper that focuses on similar methods for two-dimensional systems is [35]. A recent paper that surveys some aspects of both the method of controlled Lagrangians and the aforementioned inverse problem is [19].

We will now recall some basic definitions and results that will be needed later [96, 127], on the stability of an equilibrium for an autonomous ODE

$$\dot{q} = X(q), \quad X \in \mathfrak{X}(Q)\,. \tag{2.1}$$

Definition 2.1.1 (Stability). *An equilibrium x_e of (2.1) is stable if for every neighborhood U of x_e there is a neighborhood V of x_e such that if $x_0 \in V$ then $x(t) \in U$ for all $t \geq 0$, where $x(t)$ is the solution to (2.1) satisfying $x(0) = x_0$.*

Definition 2.1.2 (Asymptotic stability). *An equilibrium x_e of (2.1) is asymptotically stable if x_e is stable and there is a neighborhood V of x_e such that for all $x_0 \in V$ the solution $x(t)$ to (2.1) satisfying $x(0) = x_0$ converges to x_e as $t \to \infty$.*

Definition 2.1.3 (Lyapunov function). *A smooth function E defined on a neighborhood V of an equilibrium point x_e is a Lyapunov function if it satisfies $E(x_e) = 0$, $E(x) > 0$ for all $x \neq x_e$ and $X(E)(x) \leq 0$ for all $x \in V$.*

Theorem 2.1.4 (Second method of Lyapunov). *If there is a Lyapunov function E defined on a neighborhood V of an equilibrium point x_e, then x_e is stable. If E further satisfies $X(E)(x) < 0$ for all $x \in V - \{x_e\}$ then x_e is asymptotically stable.*

Theorem 2.1.5 (LaSalle's Invariance Principle). *Let E be a Lyapunov function for an equilibrium x_e of (2.1) and write*

$$M = \{x \in V : X(E) = 0\}\,.$$

If the largest invariant set contained in M coincides with $\{x_e\}$ then x_e is asymptotically stable.

In general it can be hard to find a Lyapunov function but for mechanical systems there are some natural candidates, namely conserved quantities such as the total energy of the system. We can also try to use Casimir functions (energy-Casimir method) and momentum conservation in the presence of symmetries (energy-momentum method). For controlled systems we can try to achieve stability, and prove it, by energy shaping methods, in particular kinetic shaping or potential shaping.

2.2 Controlled Lagrangian techniques and matching conditions

Consider a configuration manifold Q and a Lie group G acting freely and properly on Q. Take a Lagrangian function L on TQ of mechanical type, that is, of kinetic minus potential energy type. More precisely

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

for some Riemannian metric g on Q. We will assume that the Lagrangian L is invariant under the action of G on Q. Now we will briefly recall how from L we can define a new Lagrangian function $L_{\tau,\sigma,\rho}$ known as the controlled Lagrangian.

Consider the vertical spaces with respect to the projection $\pi : Q \longrightarrow Q/G$, that is, the tangent spaces to the orbits of the G-action, and the orthogonal complements with respect to the kinetic energy metric, which will be referred to as horizontal spaces. Then for each $v_q \in T_qQ$ we obtain a unique decomposition

$$v_q = \mathrm{Hor} v_q + \mathrm{Ver} v_q \,,$$

where $\mathrm{Ver} v_q \in T_q Orb(q)$ and $\mathrm{Hor} v_q \in T_q Orb(q)^{\perp_g}$.

Let $\mathfrak{g}$ denote the Lie algebra of G and ξ_Q the infinitesimal generator corresponding to $\xi \in \mathfrak{g}$ and let τ be a $\mathfrak{g}$-valued G-equivariant horizontal one-form on Q. We define the τ-horizontal projection and the τ-vertical projection respectively as

$$\mathrm{Hor}_\tau : v_q \longrightarrow \mathrm{Hor} v_q - \tau(v_q)_Q(q) \quad \text{and} \quad \mathrm{Ver}_\tau : v_q \longrightarrow \mathrm{Ver} v_q + \tau(v_q)_Q(q) \quad \text{for} \quad v_q \in T_qQ \,.$$

The freedom in the controlled Lagrangian $L_{\tau,\sigma,\rho}$ comes from the following choices:

- A new choice of horizontal space, corresponding to a choice of τ,
- A change $g \to g_\sigma$ of the metric acting on τ-horizontal vectors,
- A change $g \to g_\rho$ of the metric acting on vertical vectors.

Once these choices have been made, according to [21, Definition 1.2] the controlled Lagrangian can be defined as

$$L_{\tau,\sigma,\rho}(v_q) = \frac{1}{2}\left(g_\sigma(\mathrm{Hor}_\tau v_q, \mathrm{Hor}_\tau v_q) + g_\rho(\mathrm{Ver}_\tau v_q, \mathrm{Ver}_\tau v_q)\right) - V(q) \,.$$

According to [21, Theorem 2.1], if g and g_σ coincide on the horizontal spaces and further it holds that the horizontal and vertical spaces are g_σ-orthogonal, then the controlled Lagrangian can be rewritten as

$$L_{\tau,\sigma,\rho}(v_q) = L(v_q + \tau(v_q)_Q) + \frac{1}{2} g_\sigma(\tau(v_q)_Q, \tau(v_q)_Q) + \frac{1}{2}\varpi(v_q) \,, \tag{2.2}$$

where $v_q \in T_qQ$ and $\varpi(v_q) = (g_\rho - g)(\mathrm{Ver}_\tau(v_q), \mathrm{Ver}_\tau(v_q))$.

Assume now that G is Abelian. If we take local coordinates (x^α, θ^a) on Q such that x^α are coordinates on Q/G and θ^a are coordinates on G, then the given mechanical Lagrangian is written as

$$L(x^\alpha, \dot{x}^\beta, \dot{\theta}^a) = \frac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{\theta}^a + \frac{1}{2} g_{ab}\dot{\theta}^a \dot{\theta}^b - V(x^\alpha) \,.$$

In these coordinates, under the same assumptions that provide (2.2), the controlled Lagrangian becomes

$$\begin{aligned} L_{\tau,\sigma,\rho}\left(x^\alpha, \dot{x}^\beta, \dot{\theta}^a\right) \;=\; & L\left(x^\alpha, \dot{x}^\beta, \dot{\theta}^a + \tau^a_\alpha \dot{x}^\alpha\right) + \frac{1}{2}\sigma_{ab}\tau^a_\alpha \tau^b_\beta \dot{x}^\alpha \dot{x}^\beta \\ & + \frac{1}{2}\varpi_{ab}\left(\dot{\theta}^a + g^{ac} g_{\alpha c}\dot{x}^\alpha + \tau^a_\alpha \dot{x}^\alpha\right)\left(\dot{\theta}^b + g^{bd} g_{\beta d}\dot{x}^\beta + \tau^b_\beta \dot{x}^\beta\right) \,, \end{aligned}$$

where σ_{ab}, ϖ_{ab} are the coefficients of the last two terms in (2.2) and τ^a_α are the coefficients of τ.

As in [21] we will refer to the particular choice $g_\rho = g$ as the **special matching assumption** and write the corresponding controlled Lagrangian as $L_{\tau,\sigma}$. In this case, according to (2.2), the controlled Lagrangian $L_{\tau,\sigma}$ becomes

$$L_{\tau,\sigma,\rho}(v_q) = L(v_q + \tau(v_q)_Q) + \frac{1}{2} g_\sigma(\tau(v_q)_Q, \tau(v_q)_Q)\,.$$

Now we will compare the solutions of the controlled Euler-Lagrange system

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\alpha} - \frac{\partial L}{\partial x^\alpha} = 0\,, \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}^a} = u_a\,, \tag{2.3}$$

with the solutions of the Euler-Lagrange systems

$$\frac{d}{dt}\frac{\partial L_{\tau,\sigma}}{\partial \dot{x}^\alpha} - \frac{\partial L_{\tau,\sigma}}{\partial x^\alpha} = 0\,, \qquad \frac{d}{dt}\frac{\partial L_{\tau,\sigma}}{\partial \dot{\theta}^a} = 0\,, \tag{2.4}$$

$$\frac{d}{dt}\frac{\partial L_{\tau,\sigma,\rho}}{\partial \dot{x}^\alpha} - \frac{\partial L_{\tau,\sigma,\rho}}{\partial x^\alpha} = 0\,, \qquad \frac{d}{dt}\frac{\partial L_{\tau,\sigma,\rho}}{\partial \dot{\theta}^a} = 0\,, \tag{2.5}$$

where u_a is chosen in such a way that the θ^a equation of (2.3) coincides with the θ^a equation of (2.4) or (2.5), depending on whether we are making the special matching assumption or not.

Consider the following two sets of assumptions, known as **matching conditions** and **simplified matching conditions** respectively:

Assumption M1: $\tau^b_\alpha = -\sigma^{ab} g_{\alpha a}$,

Assumption M2: $\sigma^{bd}(\sigma_{ad,\alpha} + g_{ad,\alpha}) = 2g^{bd} g_{ad,\alpha}$,

Assumption M3: $\tau^b_{\alpha,\beta} - \tau^b_{\beta,\alpha} = g^{db} g_{ad,\alpha} \tau^a_\beta$,

Assumption SM1: $\sigma_{ab} = \sigma g_{ab}$ for a constant σ ,

Assumption SM2: g_{ab} is constant ,

Assumption SM3: $\tau^b_\alpha = -(1/\sigma) g^{ab} g_{\alpha a}$,

Assumption SM4: $g_{\alpha a,\delta} = g_{\delta a,\alpha}$,

where $,\alpha$ denotes partial derivative with respect to x^α.

In [21, Theorem 2.2] it is shown that under the **matching conditions** M1-M3 the solutions of (2.3) and (2.4) coincide. In particular, if the **simplified matching conditions** SM1-SM4 hold, then M1-M3 also hold and therefore the solutions of (2.3) and (2.4) coincide.

If the **simplified matching conditions** do not hold, then we may relax the special matching assumption $g_\rho = g$ and consider controlled Lagrangians of the form $L_{\tau,\sigma,\rho}$. In this case we can consider the **generalized matching conditions**, which provide equivalence of (2.3) and (2.5), see [19, Theorem 1.3]:

Assumption GM-1: $\tau^b_\alpha = -\sigma^{ab} g_{\alpha a}$,

Assumption GM-2: $\sigma^{bd}(\sigma_{ad,\alpha} + g_{ad,\alpha}) = 2g^{bd} g_{ad,\alpha}$,

Assumption GM-3: $\varpi_{ab,\alpha} = 0$,

Assumption GM-4: $\tau^b_{\alpha,\delta} - \tau^b_{\delta,\alpha} + \varpi_{ad}\rho^{bd}(\zeta^a_{\alpha,\delta} - \zeta^a_{\delta,\alpha}) - \varpi_{ad}\rho^{dc} g_{ce,\delta}\rho^{eb}\zeta^a_\alpha - \rho^{db} g_{ad,\alpha}\tau^a_\delta = 0$,

where $\zeta^a_\alpha = g^{ac} g_{\alpha c}$.

So far we have assumed that the Lagrangian L is invariant under the action of an Abelian Lie group. If we keep this assumption for the kinetic energy part of the Lagrangian but allow for a symmetry break in the potential energy part then we can add an extra condition to the simplified matching conditions, namely

Assumption SM-5: $V_{,\alpha a}g^{ad}g_{\beta d} = V_{,\beta a}g^{ad}g_{\alpha d}\,,$

which ensures that, if we choose $g_\rho = \rho g_{ab}$ for some constant ρ, then (2.3) is equivalent to

$$\frac{d}{dt}\frac{\partial L_{\tau,\sigma,\rho,\epsilon}}{\partial \dot{x}^\alpha} - \frac{\partial L_{\tau,\sigma,\rho,\epsilon}}{\partial x^\alpha} = 0\,, \qquad \frac{d}{dt}\frac{\partial L_{\tau,\sigma,\rho,\epsilon}}{\partial \dot{\theta}^a} = 0\,, \tag{2.6}$$

where

$$L_{\tau,\sigma,\rho,\epsilon} = L_{\tau,\sigma} + \frac{1}{2}(\rho-1)g_{ab}(\dot{\theta}^a + g^{ac}g_{\alpha c}\dot{x}^\alpha + \tau^a_\alpha \dot{x}^\alpha)(\dot{\theta}^b + g^{ac}g_{\alpha c}\dot{x}^\beta + \tau^b_\beta \dot{x}^\beta) - V_\epsilon(x^\alpha, \theta^a)\,,$$

see [15, Theorem III.1].

Remark 2.2.1. All of the above mentioned matching conditions are sufficient conditions to get equivalence of (2.3) and (2.4) or (2.5) or (2.6), but they are not enough to guarantee stability of the desired equilibrium. Further conditions to get stability are given in [21, 15] using energy shaping methods.

2.3 Application of Douglas' classification to control theory

We want to take a somewhat different approach to the matching theorems and rephrase some aspects of the issue in terms of the inverse problem of the calculus of variations.

The goal of the present section is to give conditions for the stabilization of an unstable equilibrium for a concrete class of two-dimensional underactuated mechanical systems. We will come to our class of interest in two steps of specification. First we will assume that the Lagrangian of the original mechanical system with configuration variables (x, y) is time-independent, that it has x as cyclic variable, and that it is of the form

$$\mathcal{L}(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\left(a_{11}\dot{x}^2 + 2a_{12}(y)\dot{x}\dot{y} + a_{22}(y)\dot{y}^2\right) - \mathcal{V}(y)\,,$$

where a_{11} is a nonzero constant. We also assume that we may add controlled external forces to the system in such a way that the control subbundle is span $\{dx\}$. The equations of motion are then of the type

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) &= u\,, \\ \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}}\right) - \frac{\partial \mathcal{L}}{\partial y} &= 0\,. \end{aligned} \tag{2.7}$$

This class of systems is general enough to include, among others, the main example that has been discussed abundantly throughout the literature, namely the inverted pendulum on a cart, and also the inertia wheel pendulum (see Section 2.3.4). The second order ordinary differential equations (2.7) can be written in normal form as

$$\ddot{x} = a^{12}\left(-\frac{\partial \mathcal{V}}{\partial y} - \frac{1}{2}\frac{\partial a_{22}}{\partial y}\dot{y}^2\right) + a^{11}\left(-\frac{\partial a_{12}}{\partial y}\dot{y}^2 + u\right),$$

$$\ddot{y} \;=\; a^{22}\left(-\frac{\partial \mathcal{V}}{\partial y}-\frac{1}{2}\frac{\partial a_{22}}{\partial y}\dot{y}^2\right)+a^{12}\left(-\frac{\partial a_{12}}{\partial y}\dot{y}^2+u\right),$$

where (a^{ij}) is the inverse matrix of (a_{ij}). When we only consider controls of the form $u(y,\dot{y})$, the above equations are of the type

$$\ddot{x}=f^1(y,\dot{y}), \qquad \ddot{y}=f^2(y,\dot{y}).$$

Our first goal is to understand when such a system is variational. For that purpose, we will rely on Douglas' classification [55] for two-dimensional systems, although we will use the geometric approach to the inverse problem that has been proposed in the papers [47, 49, 143, 144], see also Chapter 1. For most of the cases, Douglas was able to decide whether or not the systems are variational. In our approach, the matching conditions are replaced with sufficient conditions for the system to lie in one of the variational cases of Douglas' classification.

If, in a second step, we only allow controls of the type $u(y,\dot{y})=M(y)\dot{y}^2+N(y)$, the equations (2.7) may even be written in the form

$$\ddot{x}=T(y)\dot{y}^2+U(y), \qquad \ddot{y}=R(y)\dot{y}^2+S(y).$$

Our restriction in the second step is motivated by results in the literature. The condition that a_{11} is constant is in fact, for two-dimensional systems, one of the simplified matching conditions of [21]. More concretely it corresponds to SM2 given in Section 2.2. Under these assumptions, the authors of [21] derive for the system (2.7) a feedback control which may be written as

$$u \;=\; \frac{1}{\sigma}\left(\frac{\partial a_{12}}{\partial y}-\frac{a_{12}}{A_{22}}\left(\frac{1}{2}\frac{\partial a_{22}}{\partial y}-\left(1-\frac{1}{\sigma}\right)\frac{a_{12}}{a_{11}}\frac{\partial a_{12}}{\partial y}\right)\right)\dot{y}^2-\frac{1}{\sigma}\frac{a_{12}}{A_{22}}\frac{\partial \mathcal{V}}{\partial y}, \tag{2.8}$$

where σ is a constant and $A_{22}=a_{22}-\frac{a_{12}^2}{a_{11}}\left(1-\frac{1}{\sigma}\right)$. The feedback control (2.8) clearly fits into the class of controls that we wish to consider.

The strategy in the examples consists of pushing the controlled system into one of the cases of Douglas' classification that is known to be variational. In Section 2.3.2 we will give a necessary and sufficient condition for a system of the above type to be variational. Our approach is, in a sense, more general than that of the matching conditions. In [21], the matching conditions are a consequence of an a priori assumption on the relationship between the original Lagrangian of the original system, and the controlled Lagrangian of the controlled system. In our approach, no such assumption needs to be imposed. Moreover, we will show in Section 2.3.2 that if the system is variational, it admits a Lagrangian function of mechanical type, that is, a Lagrangian whose kinetic energy is related to a positive-definite metric. In that case, the energy function of this Lagrangian is always a first integral of the system. We next show, in Section 2.3.3, that under certain further conditions it can be used as a Lyapunov function. We conclude the section with a sufficient condition, written in terms of the system, that guarantees stability of the equilibrium.

In Section 2.3.4 we discuss some examples. For the example of the inverted pendulum on a cart we give new feedback controls and we also recover the ones given in [21]. For this class of controls, we provide a (slightly) wider class of Lagrangians.

The goal of Section 2.3.5 is to achieve asymptotic stability by allowing dissipative forces into the picture. We first add in extra controls to make the system equivalent to Euler-Lagrange equations with external dissipative forces. We then give sufficient conditions for asymptotic stability, based on LaSalle's invariance principle. We illustrate this method by means of an example.

2.3.1 Discussion of Douglas' classification

We are only interested in SODEs Γ which exhibit very special symmetry properties. In this section we assume that there exists a coordinate change $(t, q^1, q^2) \mapsto (t, x = x(q^1, q^2), y = y(q^1, q^2))$ for which the second order differential equations take the form

$$\ddot{x} = f^1(y, \dot{y}), \qquad \ddot{y} = f^2(y, \dot{y})\,. \tag{2.9}$$

Lemma 2.3.1. *The SODE Γ takes the form (2.9) if and only if $[\Gamma, \partial/\partial t] = 0$ and if there exists a vector field E_1 on Q such that $\Phi(E_1) = \nabla E_1 = 0$.*

Proof. The first condition says that the right-hand sides of the second order differential equations do not depend on t. If the SODE takes the special form (2.9), the vector field $E_1 = \partial/\partial x$ satisfies the conditions. Conversely, if such a vector field $E_1 = X^i(q)\partial/\partial q^i$ on Q exists, we may always straighten it out to become the vector field $\partial/\partial x$. In these coordinates, the condition $\nabla E_1 = 0$ becomes $\Gamma^i_1 = 0$, which means that the functions f^i do not depend on $\dot{x}$. With that, the condition $\Phi(E_1) = 0$ becomes $\Phi^i_1 = \frac{\partial f^i}{\partial x} = 0$, from which it follows that the functions f^i do not depend on x either. Hence, the system takes the form (2.9). ■

The specific form of the SODE (2.9) narrows the number of cases in Douglas' classification to which it may belong. Indeed, since also $\nabla\Phi(E_1) = \nabla^2\Phi(E_1) = 0$, it is easy to see that the system may never belong to Case IV. In coordinates where $E_1 = \partial/\partial x$, the system will lie in Case I if and only if $\Phi^1_2 = \Phi^2_2 = 0$. It will belong to Case II when Φ^1_2 and Φ^2_2 are not both zero, but

$$(\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2 = 0\,. \tag{2.10}$$

The system will belong to Case III whenever $(\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2 \neq 0$. In that case, it is clear that the determinant of the commutator $[\Phi, \nabla\Phi]$ does always vanish, which is the defining property for the system to lie in subcase Case IIIb. Douglas concluded in [55] that this case is never variational.

Case II has been further subdivided in Case IIa (Φ has distinct eigenvalues) and Case IIb (the eigenvalues of Φ coincide). Since $\Phi(E_1) = 0$, Φ has always eigenvalue zero, with eigenvector $E_1 = \partial/\partial x$. The other eigenvalue is given by Φ^2_2. If it is nonzero then the system lies in Case IIa, and a corresponding eigenvector is

$$E_2 = \frac{\partial}{\partial y} + \nu\frac{\partial}{\partial x}, \qquad \text{with} \quad \nu = \frac{\Phi^1_2}{\Phi^2_2}\,.$$

Both Cases IIa and IIb are further subdivided, according to a relation on the Haantjes tensor $H_\Phi(X, Y) = C^v_\Phi(\Phi(X), Y) - \Phi(C^v_\Phi(X, Y))$ of Φ introduced in Section 1.3.2, see also [47, 49]. Case

IIa1 and Case IIb1 correspond to the situation where $H_\Phi = 0$. This tensor field vanishes when all the commutators $C^v_\Phi(X,Y) = [D^v_X\Phi, \Phi](Y)$ vanish. For the SODE (2.9) we get

$$C^v_\Phi = \left(\Phi^2_2 \frac{\partial \Phi^1_2}{\partial \dot y} - \Phi^1_2 \frac{\partial \Phi^2_2}{\partial \dot y}\right) dy \otimes dy \otimes \frac{\partial}{\partial x}\,.$$

From this expression, we may conclude that the last term in the Haantjes tensor always vanishes. The only nonvanishing term in the Haantjes tensor is then $H_\Phi(\partial/\partial y, \partial/\partial y) = \Phi^2_2 C^v_\Phi(\partial/\partial y, \partial/\partial y)$. The necessary and sufficient condition for the Haantjes tensor to vanish is therefore

$$\Phi^2_2 \left(\Phi^2_2 \frac{\partial \Phi^1_2}{\partial \dot y} - \Phi^1_2 \frac{\partial \Phi^2_2}{\partial \dot y}\right) = 0\,. \tag{2.11}$$

If the system belongs to Case IIb (i.e. if $\Phi^2_2 = 0$) the above condition is trivially satisfied. Douglas [55] has one further subdivision of Case IIb1, depending on a further relation of the double eigenvalue of Φ. In the special case when that double eigenvalue happens to be zero, Douglas' Case IIb1' is characterized by the vanishing of the expression

$$\frac{\partial^2}{\partial \dot x^2}\left(\frac{\partial f^1}{\partial \dot x} - \frac{\partial f^2}{\partial \dot y}\right)\,.$$

This is clearly the case for the system (2.9). We may therefore conclude that if the system (2.9) belongs to Case IIb, it can only lie in Case IIb1'. For this case Douglas concluded that it is always variational.

Consider now the situation where the system (2.9) belongs to Case IIa (i.e. $\Phi^2_2 \neq 0$). Since the Haantjes tensor has at most one nonvanishing component what is called Case IIa3 can never occur. The only possibilities are therefore Case IIa1 (with vanishing Haantjes tensor) and Case IIa2 (the one component of the Haantjes tensor does not vanish). Douglas concluded that Case IIa1 is always variational (the same is true in general dimension n, see [47]). The necessary and sufficient condition for this to happen is

$$\Phi^2_2 \frac{\partial \Phi^1_2}{\partial \dot y} - \Phi^1_2 \frac{\partial \Phi^2_2}{\partial \dot y} = 0\,.$$

For a system in Case IIa2 to be variational, further requirements hold.

From all this we may conclude the following.

Proposition 2.3.2. *If the SODE (2.9) is variational then condition (2.10) is satisfied. If the system satisfies the further assumption (2.11), condition (2.10) is both necessary and sufficient for the system to be variational.*

Proof. For systems of the type (2.9) Case IV is excluded. If the system is variational, it cannot belong to Case III, since Case IIIb is never variational. It must therefore lie in either Case I or II, which is characterized by the condition (2.10). If (2.10) and (2.11) are both satisfied, the Haantjes tensor vanishes. If so, we must be either in Case IIa1 or Case IIb1', both of which are variational. ■

Case I is characterized by the fact that both $\Phi^1_2 = \Phi^2_2 = 0$. For Case IIb1, $\Phi^2_2 = (\nabla\Phi)^2_2 = 0$, but $\Phi^1_2 \neq 0$. For Case IIa1 $\Phi^2_2 \neq 0$ and $(\nabla\Phi)^1_2 = \nu(\nabla\Phi)^2_2$.

2.3.2 Conditions for variationality

Our interest in systems of the type (2.9) has been motivated by the fact that control systems of the type (2.7) with controls $u(y,\dot y)$ all fall in this category. In the second step we limit the suitable controls to those of the quadratic type $u(y,\dot y) = M(y)\dot y^2 + N(y)$. As a result, the system (2.7), when written in normal form becomes of the type

$$\ddot x = T(y)\dot y^2 + U(y), \quad \ddot y = R(y)\dot y^2 + S(y)\,. \tag{2.12}$$

For later use, we give a few characterizations for its variationality below. In what follows we will denote a derivative with respect to y simply by a prime $'$.

Proposition 2.3.3. *The SODE (2.12) is variational if and only if*

$$\begin{aligned} 0 &= 2T(S')^2 + S^2\left(TR' - RT'\right) - 2RS'U' + U'S'' - S'U'' \\ &\quad + S\left[S'T' + R^2U' - R'U' - TS'' + R\left(-TS' + U''\right)\right]. \end{aligned} \tag{2.13}$$

On the basis of the value of Φ^2_2 we can further specify the following.

1. *When $\Phi^2_2 = 0$ the SODE (2.12) is always variational.*
2. *When $\Phi^2_2 \neq 0$, the following statements are equivalent:*
   - *the SODE (2.12) is variational,*
   - $(U - \nu S)' = 0$,
   - $\nu' = T - R\nu$, *where* $\nu = \frac{\Phi^1_2}{\Phi^2_2}$.

Proof. One easily verifies that for the system (2.12),

$$\Gamma^1_2 = -T\dot y, \quad \Gamma^2_2 = -R\dot y, \quad \Phi^1_2 = -U' + ST \quad \text{and} \quad \Phi^2_2 = -S' + RS\,, \tag{2.14}$$

from which it follows that the condition (2.11) is always satisfied. The necessary and sufficient condition for variationality is therefore condition (2.10).

Since now

$$(\nabla\Phi)^1_2 = \dot y(2S'T + ST' - U'' - RU') \quad \text{and} \quad (\nabla\Phi)^2_2 = \dot y(SR' + RS' - S'')\,, \tag{2.15}$$

the first statement in the proposition follows.

When we take the value of Φ^2_2 into account, we may further specify the following.

(i) We have already mentioned that the condition $\Phi^2_2 = 0$ is a sufficient condition for (2.12) to be variational since it implies $(\nabla\Phi)^2_2 = 0$. Therefore (2.10) is satisfied.

(ii) In view of the coordinate expression (1.20) for $(\nabla\Phi)^i_j$ and the expressions (2.14) for Γ^i_2, we may write

$$(\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2 = [(\Phi^1_2)'\Phi^2_2 - (\Phi^2_2)'\Phi^1_2 - \Phi^2_2(\Phi^2_2 T - \Phi^1_2 R)]\dot y\,.$$

On the other hand, with $\nu = \Phi^1_2/\Phi^2_2$,

$$
\begin{aligned}
\dot y(U-\nu S)' &= \frac{\dot y}{(\Phi^2_2)^2}\left[U'(\Phi^2_2)^2 - \left((\Phi^1_2)'\Phi^2_2 - (\Phi^2_2)'\Phi^1_2\right)S - \Phi^1_2\Phi^2_2 S'\right] \\
&= -\frac{S}{(\Phi^2_2)^2}\left((\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2\right) + \frac{\dot y}{(\Phi^2_2)^2}\left[-S\Phi^2_2(\Phi^2_2 T - \Phi^1_2 R) + U'(\Phi^2_2)^2 - \Phi^1_2\Phi^2_2 S'\right] \\
&= -\frac{S}{(\Phi^2_2)^2}\left((\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2\right)
\end{aligned}
$$

because, in view of (2.14),

$$-S(\Phi^2_2T-\Phi^1_2R)+U'\Phi^2_2-\Phi^1_2S' = -S(-S'+RS)T+S(-U'+ST)R+U'(-S'+RS)-(-U'+ST)S' = 0\,.$$

We may also write $\dot y\nu' = \Gamma(\nu) = (\Gamma(\Phi^1_2)\Phi^2_2 - \Gamma(\Phi^2_2)\Phi^1_2)/(\Phi^2_2)^2$. With that

$$\dot y\nu' = \frac{(\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2}{(\Phi^2_2)^2} + \frac{\Phi^2_2T - \Phi^1_2R}{\Phi^2_2}\dot y = \frac{(\nabla\Phi)^1_2\Phi^2_2 - (\nabla\Phi)^2_2\Phi^1_2}{(\Phi^2_2)^2} + (T - R\nu)\dot y\,.$$

■

We remark that a sufficient condition for $\Phi^2_2 = 0$ is that $S(y) = 0$.

The above proposition points to some strategies one may follow in the search for controls $u = M(y)\dot y^2 + N(y)$ for which equations (2.7) are variational. For such a control law, Equations (2.7) become of type (2.12) and the conditions given in Proposition 2.3.3 can be interpreted as a PDE in the unknowns $M(y)$ and $N(y)$. In a sense, one may interpret Equation (2.13) as a generalization (to the current setting) of the matching conditions of [21]. We may follow any one of the following paths.

- Find a control u such that the corresponding SODE satisfies condition (2.13).
- Find a control u for which $\Phi^2_2 \neq 0$, but the corresponding SODE satisfies $(U - \nu S)' = 0$ (i.e. lies in Case IIa1).
- Find a control u such that the corresponding SODE satisfies $\Phi^2_2 = 0$ (i.e. lies in Case IIb1').
- Find a control u such that the corresponding SODE is such that $S = 0$.

In this section we will mainly focus on the first and second strategies. The reason for this is that Case IIa1 has been shown to be variational in arbitrary dimensions [47], which leaves the door open to a possible generalization of our results to higher dimensional systems. In the examples we will use an ansatz for $N(y)$ and solve the corresponding PDE for $M(y)$ (mainly because $N(y)$ appears with two derivatives in it and $M(\phi)$ with just one). In the next section we will also show that the last strategy is not the best one to follow, in view of the pursuit for stability.

The multipliers (g_{ij}) of a variational system may in general depend on velocities $\dot q$. As a consequence, a Lagrangian of a variational system is not necessarily of mechanical type. We first prove that, if the system (2.12) is variational, we may always find a Lagrangian L of the form $L(q,\dot q) = g_{ij}(q)\dot q^i\dot q^j - V(q)$, where the multiplier matrix (g_{ij}) is independent of velocities, time-independent and positive-definite.

Proposition 2.3.4. *For a variational SODE of type (2.12) with $\Phi^2_2 \neq 0$ there exists a positive-definite matrix of multipliers (g_{ij}) which only depend on y and for which g_{11} is a constant.*

Proof. Under the assumptions in the statement, the SODE belongs to Case IIa1. This means that the Jacobi endomorphism Φ has two distinct eigenvalues 0 and Φ^2_2, with eigenvectors E_1 and E_2, respectively. We remark that in the case under consideration both E_1 and E_2 may be thought of as vector fields on Q (that is, when considered as vector fields along π_1, they do not depend on t or on $\dot{q}$). One easily verifies that, after taking (2.10) into account, we may write that

$$\nabla E_1 = 0, \qquad \nabla E_2 = \Gamma^1_2 \frac{\partial}{\partial x} + \Gamma^2_2 \frac{\partial}{\partial y} + \Gamma(\nu)\frac{\partial}{\partial x} = \Gamma^2_2 E_2 = -R\dot{y}E_2\,,$$

where we have invoked the third characterization of Proposition 2.3.3.

We will denote the dual basis of one-forms on Q as $\{\theta^1, \theta^2\}$. From the above it follows that

$$\nabla\theta^1 = 0, \qquad \nabla\theta^2 = R\dot{y}\theta^2\,.$$

Since the system is supposed to be variational, we may assume that solutions of the Helmholtz conditions (3.6) exist. We show now that among these solutions there is at least one that satisfies the specifics of the statement. From the Φ-condition we may conclude that the multiplier is of the type $g = \rho_1\theta^1 \otimes \theta^1 + \rho_2\theta^2 \otimes \theta^2$. With this, the condition $\nabla g = 0$ becomes

$$\Gamma(\rho_1) = 0, \qquad \Gamma(\rho_2) = -2R\dot{y}\rho_2\,.$$

We are not interested in the most general solution of these two PDEs in ρ_i. Any positive constant ρ_1 clearly satisfies the first equation, and we may even set it to be simply 1. We now show that the second equation has solutions $\rho_2(y)$ that only depend on y. Indeed, for such functions the equation becomes $\rho_2' = -2R\rho_2$, which has (among others) the solutions $\rho_2(y) = A\exp(-2\int_1^y R(\bar{y})d\bar{y})$. Also the integration constant $A = \rho_2(1)$ can be chosen to be positive. With such functions $\rho_1 = 1$ and $\rho_2(y)$ the D^v-condition of the Helmholtz conditions is automatically satisfied. Clearly, $g = \theta^1 \otimes \theta^1 + \rho_2\theta^2 \otimes \theta^2$ is then a positive-definite metric. ■

In Proposition 2.3.4 we may replace positive-definiteness by negative-definiteness, since in the proof we may choose ρ_1 and A to be both negative.

2.3.3 Lyapunov stability

In this section we assume again that a mechanical system of the type (2.7) is given, with an arbitrary quadratic feedback control $u = M(y)\dot{y}^2 + N(y)$. The relevant equations are then of type (2.12). For such systems x is clearly a cyclic variable, and it generates a symmetry for the system. We may therefore reduce the two second order differential equations in (x, y) by that symmetry to a system of three first order equations in (y, v_y, v_x), by canceling out the variable x:

$$\dot{y} = v_y, \quad \dot{v}_y = R(y)v_y^2 + S(y), \quad \dot{v}_x = T(y)v_y^2 + U(y). \tag{2.16}$$

If we assume that $U(0) = S(0) = 0$, the reduced system has an equilibrium at $(y = 0, v_x = 0, v_y = 0)$ (or, equivalently, the original system (2.12) has a relative equilibrium $(x, y = 0, \dot{x} = 0, \dot{y} = 0)$).

We wish to find sufficient conditions for that equilibrium to be stable. Note first that the Jacobian of the system (2.16) in the equilibrium has a zero eigenvalue, and that, as a consequence, the equilibrium can never be linearly stable. Second, for systems of second order differential equations, one may also consider a second, more geometric, linearization process, where the linearized equations are given by the matrix that corresponds to the Jacobi endomorphism Φ (see for instance [136]). Since, for systems of the type (2.12), Φ has always eigenvalue zero, we can also not conclude that the equilibrium is Jacobi stable.

We are therefore left with trying to find a Lyapunov function for the system (2.12). For that reason, we now assume that we were able to find a feedback control $u = M(y)\dot{y}^2 + N(y)$ for which the SODE (2.12) is variational, and for which Φ^2_2 is not zero. Our method will rely on the use of the energy function of the variational system as a Lyapunov function (see for instance [166] or Section 2.1 for the definition of a Lyapunov function).

The multiplier we found in the proof of Proposition 2.3.4 may also be written in a coordinate basis as

$$\begin{aligned} g &= dx \otimes dx - \nu(dx \otimes dy + dy \otimes dx) + \left(\nu^2 + \rho_2\right) dy \otimes dy \\ &= g_{11} dx \otimes dx + g_{12}(dx \otimes dy + dy \otimes dx) + g_{22} dy \otimes dy\,. \end{aligned} \tag{2.17}$$

Recall that the relation between possible Lagrangians and multipliers is such that the multiplier is the Hessian of the Lagrangian with respect to the velocities. From this it follows that the Lagrangian that corresponds with the multiplier (2.17) must be of the type

$$L = g_{11}\dot{x}^2 + 2g_{12}(y)\dot{x}\dot{y} + g_{22}(y)\dot{y}^2 + A_1(x,y)\dot{x} + A_2(x,y)\dot{y} - V(x,y)\,.$$

The Euler-Lagrange equations of L provide further conditions on the functions A_i and V. We obtain

$$\begin{aligned} \frac{\partial V}{\partial x} &= -g_{11}U - g_{12}S, \qquad \frac{\partial A_1}{\partial y} - \frac{\partial A_2}{\partial x} = 0\,, \\ \frac{\partial V}{\partial y} &= -g_{12}U - g_{22}S\,. \end{aligned}$$

The equation which involves A_i simply says that we may take any total time derivative $(\partial f/\partial q^i)\dot{q}^i$ for the linear part $A_i\dot{q}^i$, for example simply $f = 0$. The validity of the Helmholtz conditions (3.6) with the multiplier g_{ij} ensures that a function $V(x, y)$ exists for the equations in the first column. This is clear from Proposition 2.3.3, which shows that the integrability condition of this system of PDEs in V, namely $(g_{11}U + g_{12}S)' = (U - \nu S)' = 0$, is guaranteed by the variationality of the system.

If we assume as before that the system is such that $S(0) = 0$ and $U(0) = 0$, then

$$g_{11}U + g_{12}S = g_{11}(0)U(0) + g_{12}(0)S(0) = 0.$$

The potentials which further satisfy $V(x, 0) = 0$ are then

$$V(x,y) = -\int_0^y (g_{12}U + g_{22}S)d\bar{y} = \int_0^y (\nu U - \nu^2 S - \rho_2 S)d\bar{y} = -\int_0^y \rho_2 S d\bar{y}.$$

We will denote this potential simply by $V(y)$. At $y = 0$, it has the properties that

$$\frac{\partial V}{\partial y}(0) = 0, \qquad \frac{\partial^2 V}{\partial y^2}(0) = -\rho_2'(0)S(0) - \rho_2(0)S'(0) = -\rho_2(0)S'(0).$$

From the above we may conclude that if we assume $S'(0) < 0$, then $y = 0$ is a local minimum for V.

Proposition 2.3.5. *Suppose given a variational system (2.12) with $\Phi_2^2 \neq 0$, $U(0) = S(0) = 0$ and $S'(0) < 0$. Then $(y = 0, \dot{x} = 0, \dot{y} = 0)$ represents a stable relative equilibrium.*

Proof. Consider the energy function of the Lagrangian L we found above, that is,

$$E_L(y, \dot{x}, \dot{y}) = \frac{1}{2}(g_{11}\dot{x}^2 + 2g_{12}(y)\dot{x}\dot{y} + g_{22}(y)\dot{y}^2) + V(y). \tag{2.18}$$

Since the Lagrangian is autonomous, this function is always a first integral of the system. It can now be used as a Lyapunov function. Indeed, since $V(0) = 0$, we have $E_L(0, 0, 0) = 0$. Since $y = 0$ is always a stationary point for V, so will also be $(0, 0, 0)$ for E_L. Moreover since g is positive-definite, and since $y = 0$ is a minimum for V, we know that in a neighborhood of $(0, 0, 0)$, $E_L(y, \dot{x}, \dot{y}) > 0$. We conclude therefore that $(y = 0, \dot{x} = 0, \dot{y} = 0)$ is Lyapunov stable in the reduced space. ■

Notice that, although the reasoning in the proof relies on the fact that we have chosen the multiplier matrix (g_{ij}) to be positive-definite, the condition $S'(0) < 0$ does not. If we had chosen to work with a negative-definite multiplier, then ρ_2 would be negative, and with $S'(0) < 0$ we would get that $(\partial^2 V/\partial y^2)(0) < 0$, but then $E_L(y, \dot{x}, \dot{y}) < 0$ in a neighborhood of $(0, 0, 0)$, which gives the same result.

2.3.4 Examples

In this section we will derive stabilizing controls for two systems, the inverted pendulum on a cart and the inertia wheel pendulum. For the first example we can both derive new controls and recover the ones from [21].

2.3.4.1 The inverted pendulum on a cart

Definition of the system. The system consists of a pendulum of length l and a bob of mass m. The pendulum is attached to the top of a cart of mass M. The configuration manifold of the system is $Q = S^1 \times \mathbb{R}$ with coordinates $(x = s, y = \phi)$. The upright position of the pendulum corresponds to $\phi = 0$ (see Figure 2.1). The Lagrangian is given by kinetic minus potential energy, that is,

$$\mathcal{L}(s, \phi, \dot{s}, \dot{\phi}) = \frac{1}{2}(\gamma\dot{s}^2 + 2\beta\cos(\phi)\dot{s}\dot{\phi} + \alpha\dot{\phi}^2) + \delta\cos(\phi),$$

where $\alpha = ml^2$, $\beta = ml$, $\gamma = M + m$ and $\delta = -mgl$ are constants related to the dimensions of the system, and g denotes the standard acceleration due to gravity.

The control subbundle is span $\{ds\}$ and we have here

$$a^{-1}(uds) = u\left(-\frac{\beta\cos(\phi)}{\alpha\gamma - \beta^2\cos^2(\phi)}\frac{\partial}{\partial\phi} + \frac{\alpha}{\alpha\gamma - \beta^2\cos^2(\phi)}\frac{\partial}{\partial s}\right),$$

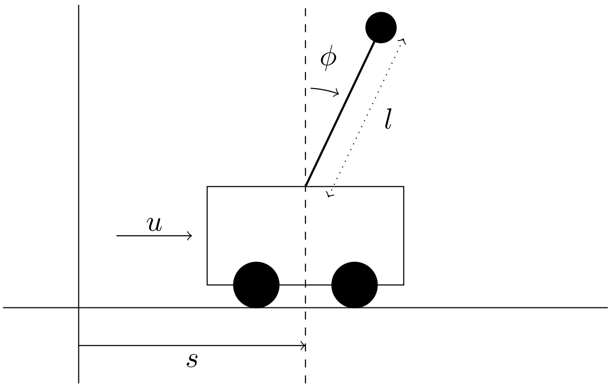


FIGURE 2.1: The inverted pendulum on a cart.

where a has components $a_{11} = \gamma, a_{12} = a_{21} = \beta\cos(\phi)$ and $a_{22} = \alpha$. If we consider controls of the type $u(\phi, \dot\phi) = M(\phi)\dot\phi^2 + N(\phi)$, the controlled Euler-Lagrange equations (2.7), written in normal form, are

$$\begin{aligned}\ddot s &= \frac{\beta\delta\sin(\phi)\cos(\phi) + \alpha\beta\sin(\phi)\dot\phi^2 + \alpha u}{\alpha\gamma - \beta^2\cos^2(\phi)},\\ \ddot\phi &= \frac{-\gamma\delta\sin(\phi) - \beta^2\sin(\phi)\cos(\phi)\dot\phi^2 - \beta\cos(\phi)u}{\alpha\gamma - \beta^2\cos^2(\phi)}.\end{aligned}$$

A new stabilizing control. We will give a new class of feedback controls which turn the upright position of the pendulum into a stable equilibrium, modulo the translational symmetry. For this purpose we look for solutions of the equation (2.13). We will require that $\Phi_2^2 \neq 0$, which means that we aim for a controlled SODE that lies in Case IIa1.

If we take $N(\phi) = d\cos(\phi)\sin(\phi)$, where d is a constant, one may verify that the pair (L, M) with

$$M(\phi) = -\frac{d\left(2\beta^2\delta - 2\alpha\gamma\delta + \alpha\beta d + \beta(2\beta\delta + \alpha d)\cos(2\phi)\right)\sin(\phi)}{\delta\left(2\gamma\delta + \beta d + \beta d\cos(2\phi)\right)}$$

solves the PDE (2.13). The controlled SODE is then given by

$$\begin{aligned}\ddot s &= \left(\frac{\alpha\beta}{\alpha\gamma - \beta^2\cos^2(\phi)} - \frac{\alpha d\left(2\beta^2\delta - 2\alpha\gamma\delta + \alpha\beta d + \beta(2\beta\delta + \alpha d)\cos(2\phi)\right)}{\delta(2\gamma\delta + \beta d + \beta d\cos(2\phi))(\alpha\gamma - \beta^2\cos^2(\phi))}\right)\sin(\phi)\dot\phi^2\\ &\quad + \frac{(\beta\delta + \alpha d)\cos(\phi)\sin(\phi)}{\alpha\gamma - \beta^2\cos^2(\phi)}\\ &= T(\phi)\dot\phi^2 + U(\phi),\\ \ddot\phi &= \left(\frac{\beta(\beta\delta + \alpha d)(-2\gamma\delta + \beta d + \beta d\cos(2\phi))\cos(\phi)\sin(\phi)}{\delta\left(\alpha\gamma - \beta^2\cos(\phi)^2\right)\left(2\gamma\delta + \beta d + \beta d\cos(2\phi)\right)}\right)\dot\phi^2\\ &\quad - \frac{(2\gamma\delta + \beta d + \beta d\cos(2\phi))\sin(\phi)}{2\left(\alpha\gamma - \beta^2\cos^2(\phi)\right)}\\ &= R(\phi)\dot\phi^2 + S(\phi).\end{aligned}$$

We clearly have $U(0) = S(0) = 0$. For the first denominator, we have that

$$\alpha\gamma - \beta^2\cos^2(\phi) = m^2l^2(1 - \cos^2(\phi)) + mMl^2 > 0 \text{ for all } \phi.$$

If we fix some $\phi_{max} \in (-\pi/2, \pi/2)$, we may choose d in such a way that $d > \frac{2(M+m)g}{1+\cos(2\phi_{max})}$. If so, we get that $2\gamma\delta + \beta d + \beta d \cos(2\phi) > 0$ in the range $(-\phi_{max}, \phi_{max})$.

The components of the Jacobi endomorphism for this SODE are given by

$$\begin{aligned}\Phi_2^2 &= \frac{\cos(\phi)\left(2\gamma\delta + \beta d + \beta d\cos(2\phi)\right)\left(-2\beta^2\delta + 2\alpha\gamma\delta - \alpha\beta d + \alpha\beta d\cos(2\phi)\right)}{2\delta\left(\alpha\gamma - \beta^2\cos^2(\phi)\right)^2}, \\ \Phi_2^1 &= -\frac{(\beta\delta + \alpha d)\cos^2(\phi)\left(-2\beta^2\delta + 2\alpha\gamma\delta - \alpha\beta d + \alpha\beta d\cos(2\phi)\right)}{\delta\left(\alpha\gamma - \beta^2\cos^2(\phi)\right)^2}.\end{aligned}$$

Since we also have $\alpha\beta d(\cos(2\phi) - 1) - 2\delta(\beta^2 - \alpha\gamma) < 0$ we get $\Phi_2^2 \neq 0$ for all $\phi \in (-\pi/2, \pi/2)$. This means that the SODE belongs to Case IIa1. By Proposition 2.3.4 we can find a positive-definite matrix of multipliers. Since $S'(0) = -\frac{\gamma\delta+\beta d}{\alpha\gamma-\beta^2}$ and $\alpha\gamma - \beta^2 = mMl^2 > 0$, we know from Proposition 2.3.5 that the equilibrium $\phi = 0, \dot{\phi} = 0, \dot{s} = 0$ will be stable in the reduced space when $d > \frac{-\gamma\delta}{\beta} = (m+M)g$.

If we fix the values of the parameters to be $M = 2$, $m = 1$, $l = 1$ the control discussed above will stabilize the upright position of the pendulum for $d > 3g$. If we also choose $\phi_{max} = \frac{\pi}{4}$ then we need to require $d > 6g$ for the SODE to be defined. For these parameters and with $d = 7g$ the SODE is

$$\begin{aligned}\ddot{s} &= -\frac{6\sin(\phi)\left(g\cos(\phi)(1 + 7\cos(2\phi)) + (13 + 7\cos(2\phi))\dot{\phi}^2\right)}{\left(-3 + \cos(\phi)^2\right)\left(1 + 7\cos(2\phi)\right)}, \\ \ddot{\phi} &= \frac{\sin(\phi)\left(g(1 + 7\cos(2\phi))^2 + 6(33\cos(\phi) + 7\cos(3\phi))\dot{\phi}^2\right)}{(-5 + \cos(2\phi))(1 + 7\cos(2\phi))}.\end{aligned}$$

Figure 2.2 shows a simulation of this example with MATLAB, with initial conditions $\phi(0) = 0.4$, $\dot{\phi}(0) = 0.1$, $s(0) = 0$, $\dot{s}(0) = -1.5$, and $g = 9.81$. The position s of the cart is not stabilized, since it represents a cyclic variable.

The controls of [21]. In this paragraph we recover the control given in [21] for the inverted pendulum on a cart and we give additional multipliers for the Lagrangian. We consider again a control of the type $u(\phi, \dot{\phi}) = M(\phi)\dot{\phi}^2 + N(\phi)$, but now we take

$$N(\phi) = \frac{\kappa\beta\delta\cos(\phi)\sin(\phi)}{\alpha - \frac{\beta^2}{\gamma}(1+\kappa)\cos^2(\phi)},$$

where κ is a constant. With this $N(\phi)$,

$$M(\phi) = \frac{\kappa\beta\alpha\sin(\phi)}{\alpha - \frac{\beta^2}{\gamma}(1+\kappa)\cos^2(\phi)}$$

is a solution of the PDE (2.13). This control coincides with the one given in [21]. The controlled SODE is then

$$\begin{aligned}\ddot{\phi} &= \frac{\gamma\delta\sin(\phi)}{-\alpha\gamma + \beta^2(1+\kappa)\cos^2(\phi)} + \frac{\beta^2(1+\kappa)\cos(\phi)\sin(\phi)}{-\alpha\gamma + \beta^2(1+\kappa)\cos^2(\phi)}\dot{\phi}^2, \\ \ddot{s} &= -\frac{\beta\delta(1+\kappa)\cos(\phi)\sin(\phi)}{-\alpha\gamma + \beta^2(1+\kappa)\cos^2(\phi)} - \frac{\alpha\beta(1+\kappa)\sin(\phi)}{-\alpha\gamma + \beta^2(1+\kappa)\cos^2(\phi)}\dot{\phi}^2.\end{aligned}$$

The value of $S'(0) = \gamma\delta/(\beta^2(1+\kappa) - \alpha\gamma)$ will be negative when $\kappa > (\alpha\gamma - \beta^2)/\beta^2 = M/m$. For such values of κ we will get a stable equilibrium.

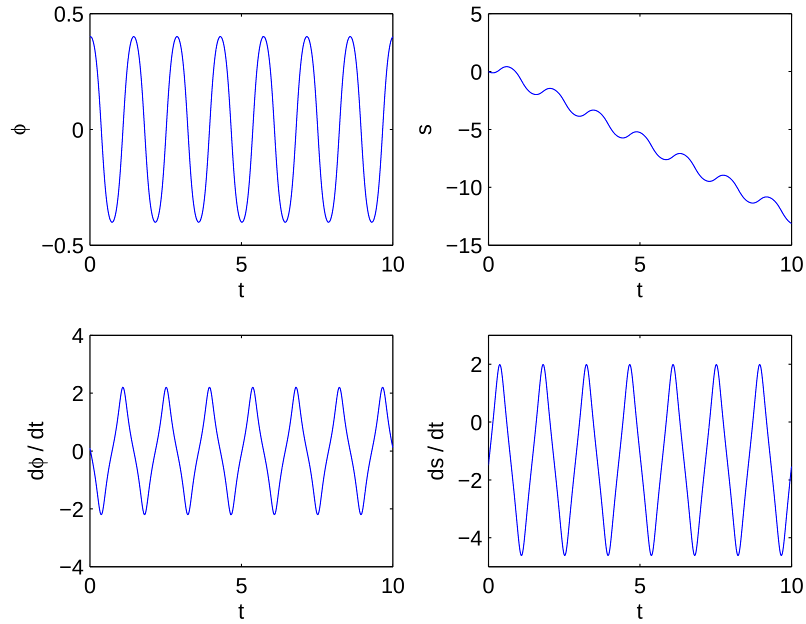


FIGURE 2.2: Simulations for the inverted pendulum on a cart.

Our criterion for stability does not involve the multipliers, nor the potential energy. If we compute the multipliers, we may compare them with the Hessian of the Lagrangian given in [21]. In the current setting we have

$$\nu = -\frac{\beta(1+\kappa)\cos(\phi)}{\gamma}\,.$$

The equation for ρ_2 was $\rho_2' = -2R\rho_2$. One may easily verify that

$$\rho_2 = A(\beta^2(\kappa+1)\cos^2(\phi) - \alpha\gamma)$$

(with constant A) is a solution of it which is, however, not always positive. Since we are only interested in proving stability in a small region around the equilibrium, we may restrict our analysis to $\phi \in (-\phi_{max}, \phi_{max})$, with

$$\sin^2(\phi_{max}) = \frac{\kappa - \frac{M}{m}}{1+\kappa}.$$

(Under the current assumption on κ the constant on the right-hand side is indeed positive.) The above region for ϕ coincides with the one that is also adopted in [21]. In this region, the denominator of the control never vanishes and, for every positive choice of A, the function ρ_2 remains positive and the multiplier matrix positive-definite. The corresponding multiplier is in fact

$$g_{11} = 1, \qquad g_{12} = \frac{\beta(1+\kappa)\cos(\phi)}{\gamma}, \qquad g_{22} = \left(\frac{\beta^2(1+\kappa)^2}{\gamma^2} + 2A\beta^2(1+\kappa)\right)\cos^2(\phi) - 2A\alpha\gamma\,, \quad (2.19)$$

from which one may derive, up to a constant, the Lagrangian

$$L = \frac{1}{2}\left(\dot{s}^2 + 2g_{12}\dot{s}\dot{\phi} + g_{22}\dot{\phi}^2\right) - 2A\gamma\delta\cos(\phi)\,.$$

The Lagrangian that has been proposed in [21] is

$$L = \frac{1}{2}\left(\alpha\dot{\phi}^2 + 2\beta\cos(\phi)\dot{s}\dot{\phi} + 2\beta\cos(\phi)K\dot{\phi}^2 + \gamma\dot{s}^2 + 2\gamma K\dot{s}\dot{\phi} + \gamma K^2\dot{\phi}^2\right) + \frac{\sigma}{2}\gamma K^2\dot{\phi}^2 + \delta\cos(\phi)$$

with $K = \kappa\frac{\beta}{\gamma}\cos(\phi)$, $\sigma = -1/\kappa$ and κ a constant (satisfying $\kappa > \frac{\alpha\gamma-\beta^2}{\beta^2}$). Its multipliers are

$$g_{11} = \gamma, \qquad g_{12} = \beta(1+\kappa)\cos(\phi), \qquad g_{22} = \frac{\beta^2\kappa}{\gamma}(1+\kappa)\cos^2(\phi) + \alpha\,. \tag{2.20}$$

For better comparison, we may rescale both this multiplier and its Lagrangian with a constant factor $1/\gamma$, to get also $g_{11} = 1$. The multiplier matrix (2.20) then agrees with (2.19), if we set the integration constant A to be $-1/(2\gamma^2)$. The negative choice for A is not in disagreement with what we said before, since the multiplier matrix (2.20) of [21] is, surprisingly, non-definite in the region $(-\phi_{max}, \phi_{max})$.

2.3.4.2 The inertia wheel pendulum

Definition of the system. The system consists of an inverted pendulum with an actuated wheel at the end. The configuration space is $S^1 \times S^1$. We will denote the coordinates of the system by $(x = \varphi, y = \theta)$, where φ and θ are the angles of the wheel and the pendulum, respectively (see Figure 2.3). The upright position of the pendulum corresponds to $\theta = 0$. The Lagrangian is given by

$$\mathcal{L}(\varphi,\theta,\dot{\varphi},\dot{\theta}) = \frac{1}{2}(b\dot{\varphi}^2 + 2b\dot{\theta}\dot{\varphi} + a\dot{\theta}^2) - m(1+\cos(\theta))\,,$$

where m, a and b are positive constants with $a > b$. These constants are defined from the physical parameters of the system as

$$a = m_1 l_1^2 + m_2 l_2^2 + I_1 + I_2\,, \quad b = I_2\,, \quad \text{and} \quad m = m_1 l_1 + m_2 l_2\,,$$

where m_1, I_1, m_2, I_2 denote respectively the masses and moments of inertia of the pendulum and the wheel, and l_1, l_2 denote, respectively, the distances from the origin to the center of mass of the pendulum and the wheel, as shown in Figure 2.3. See [150] for more details.

The controlled Euler-Lagrange equations are $a\ddot{\theta} + b\ddot{\varphi} = m\sin(\theta)$ and $b\ddot{\theta} + b\ddot{\varphi} = u$, which in normal form become

$$\ddot{\theta} = \frac{bm\sin(\theta) - bu}{b(a-b)}, \quad \ddot{\varphi} = \frac{-bm\sin(\theta) + au}{b(a-b)}\,. \tag{2.21}$$

Stabilizing control. In view of the lack of quadratic terms in $\dot{\theta}$ in the above equations, we try to find a control $u = N(\theta)$, i.e. with $M(\theta) = 0$, such that SODE (2.21) lies in Case IIa1. Equation (2.13) is then

$$\frac{4m\dot{\theta}(\sin(\theta)N' + \cos(\theta)N'')}{(a-b)b} = 0\,.$$

It admits a solution $N(\theta) = d_2 + d_1\sin(\theta)$, where d_1 and d_2 are integration constants. Since we want the state $(\theta = 0, \dot{\theta} = 0, \dot{\varphi} = 0)$ to be an equilibrium we must take $N(\theta) = d_1\sin(\theta)$. In that case

$$\Phi^2_2 = \frac{2(d_1 - m)\cos(\theta)}{a-b} \quad \text{and} \quad \Phi^1_2 = \frac{2(ad_1 - bm)\cos(\theta)}{b(b-a)}\,, \tag{2.22}$$

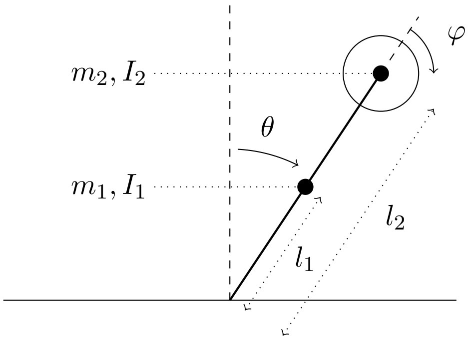


FIGURE 2.3: The inertia wheel pendulum.

so we will have $\Phi_2^2 \neq 0$ around the equilibrium as long as we require $d_1 \neq m$.

The controlled SODE (in Case IIa1) is then given by

$$\ddot{\theta} = \frac{(m - d_1)\sin(\theta)}{a - b} = S(\theta)\,,$$
$$\ddot{\varphi} = \frac{(-ad_1 + bm)\sin(\theta)}{b(b - a)} = U(\theta)\,.$$

By Proposition 2.3.5 it is enough to choose $d_1 > m$ to get stability for the equilibrium $\theta = 0$, $\dot{\theta} = 0$, $\dot{\varphi} = 0$.

We choose the parameters of the system to be $a = 0.4846$, $b = 0.0032$ and $m = 37.98$ (as in a simulation of [75]). If we set the constant in the control to be $d_1 = 60$ and take the initial conditions to be $\theta_0 = 0.0001$, $\varphi_0 = 0.1$, $\dot{\theta}_0 = 0.0001$ and $\dot{\varphi}_0 = 0.1$, we get the MATLAB simulations of Figure 2.4.

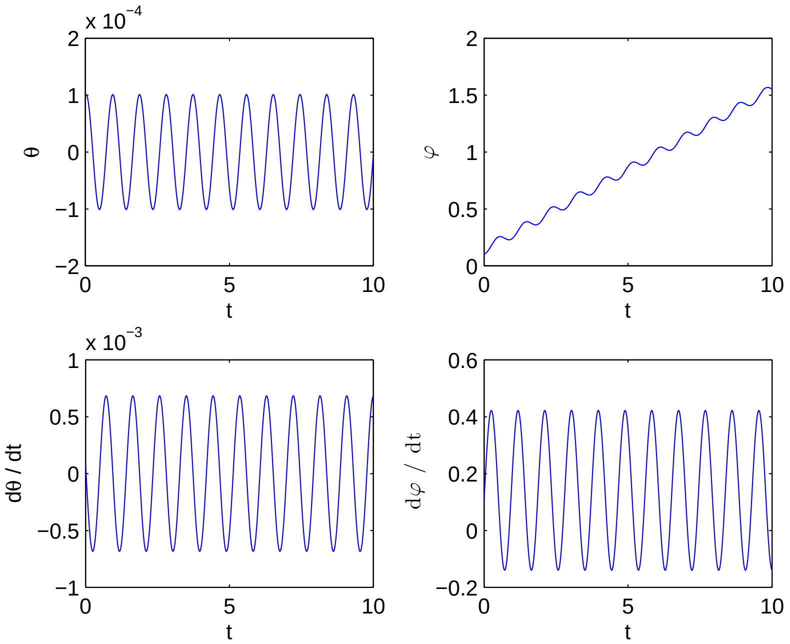


FIGURE 2.4: Simulations for the inertia wheel pendulum.

In the introduction to Section 2.3 we mentioned the feedback control (2.8) of [21]. Although the example is not explicitly treated in [21], one may still calculate the corresponding control. Since the multipliers of the given Lagrangian are constant, it reduces to

$$u = -\frac{1}{\sigma}\frac{a_{12}}{A_{22}}\frac{\partial \mathcal{V}}{\partial \theta} = \frac{-bm}{\sigma(a-b)+b}\sin(\theta).$$

This coincides with our control $u = d_1 \sin(\theta)$ if we take $\sigma = \frac{b(m+d_1)}{d_1(b-a)}$. Our stability condition $d_1 > m$ is then equivalent to $\frac{2b}{b-a} < \sigma < \frac{b}{b-a}$.

2.3.5 Asymptotic stability

In Section 2.3.3 we only gave a criterion for stability of Lyapunov type. Along the lines of [21], we now modify the control that gives Lyapunov stability in such a way that the system becomes dissipative, and the equilibrium asymptotically stable.

We use, as before, the notation q^i for the variables (x, y). Assume that we had found a control $u(y, \dot{y}) = M(y)\dot{y}^2 + N(y)$ for which the system (2.7) is variational, that is, assume that we know of multipliers (g_{ij}) and a regular Lagrangian L such that

$$g_{ij}\left(\ddot{q}^j - f^j\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\,. \tag{2.23}$$

Now, we further add control forces to this system with the goal of modifying it into a set of Euler-Lagrange equations with external dissipative forces. More precisely, we put $u = M(y)\dot{y}^2 + N(y) + u_2$ in (2.7). Then, in normal form, we are considering systems of the type

$$\ddot{q}^j = f^j + a^{-1}(u_2 dx)^j, \tag{2.24}$$

where a is the metric of the original Lagrangian and the second control u_2 is chosen in such a way that

$$g_{ij}\left(\ddot{q}^j - f^j - a^{-1}(u_2 dx)^j\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} - \frac{\partial D}{\partial \dot{q}^i}\,. \tag{2.25}$$

The term in D is called a dissipative force if it has the effect that, along trajectories, the energy E_L has the property $\dot{E}_L < 0$. In view of (2.23), condition (2.25) will hold when

$$g_{ij}(a^{-1}u_2 dx)^j = u_2(g_{i1}a^{11} + g_{i2}a^{12}) = \frac{\partial D}{\partial \dot{q}^i}\,. \tag{2.26}$$

We may think of the above as a PDE in D. If we introduce the simplified notation

$$\Box = g_{11}a^{11} + g_{12}a^{12} = a^{11} - \nu a^{12}\,, \quad \diamond = g_{12}a^{11} + g_{22}a^{12} = -\nu a^{11} + (\nu^2 + \rho_2)a^{12}\,,$$

the integrability condition is $\frac{\partial u_2}{\partial \dot{y}}\Box = \frac{\partial u_2}{\partial \dot{x}}\diamond$ (mixed derivatives of D coincide). The functions u_2 that satisfy this condition are of the type

$$u_2 = f(x, y)\left(\Box\dot{x} + \diamond\dot{y}\right) + g(x, y)\,.$$

With this,

$$D = f(x,y)\left(\frac{\Box^2}{2}\dot{x}^2 + \Box\diamond\,\dot{x}\dot{y} + \frac{\diamond^2}{2}\dot{y}^2\right) + g(x,y)\,(\Box\dot{x} + \diamond\dot{y}) + h(x,y)$$

satisfies the condition (2.26).

We had already established in Section 2.3.3 that, when $\Phi_2^2 \neq 0$, $S(0) = 0, U(0) = 0$ and $S'(0) < 0$, there exists a positive-definite multiplier and a potential V such that, in a neighborhood around $(x, y = 0, \dot{x} = 0, \dot{y} = 0)$, $E_L > 0$. It is easy to see that, along trajectories of the system, $\dot{E}_L = \dot{q}^i(\partial D/\partial \dot{q}^i)$. If we choose $g = h = 0$ and f to be a strictly negative function, then

$$D = \frac{f}{2}\left(\Box\dot{x} + \diamond\dot{y}\right)^2, \qquad u_2 = f\left(\Box\dot{x} + \diamond\dot{y}\right)$$

and $\dot{E}_L \leq 0$. It is also clear that the equilibrium does not change under the extra control law, since $u_2(x, y = 0, \dot{x} = 0, \dot{y} = 0) = 0$. From LaSalle's invariance principle (see for instance [166] or Section 2.1) it follows that if the only trajectory of (2.24) contained in the set

$$M = \{(x, y, \dot{x}, \dot{y}) : \dot{E}_L = 0\} = \{(x, y, \dot{x}, \dot{y}) : D = 0\} = \{(x, y, \dot{x}, \dot{y}) : \Box\dot{x} + \diamond\dot{y} = 0\}$$

is $(x, 0, 0, 0)$, then the relative equilibrium is asymptotically stable.

Proposition 2.3.6. *Assume that the system (2.12) is variational with* $\Phi_2^2 \neq 0$*, and that* $S(0) = U(0) = 0$ *and* $S'(0) < 0$*. If there exists no solution* $(x(t), y(t))$ *of (2.24), other than the equilibrium* $(x, 0)$*, that satisfies*

$$(\Box T + \diamond R)\dot{y}^2 + \dot{\Box}\dot{x} + \dot{\diamond}\dot{y} + \Box U + \diamond S = 0, \tag{2.27}$$

then the relative equilibrium is asymptotically stable.

Proof. Suppose there is a solution that satisfies $\Box\dot{x} + \diamond\dot{y} = 0$. Then it has $u_2 = f\left(\Box\dot{x} + \diamond\dot{y}\right) = 0$ and thus also

$$0 = \dot{\Box}\dot{x} + \Box\ddot{x} + \dot{\diamond}\dot{y} + \diamond\ddot{y} = (\Box T + \diamond R)\dot{y}^2 + \dot{\Box}\dot{x} + \dot{\diamond}\dot{y} + \Box U + \diamond S.$$

■

The condition (2.27) will be useful in the example below.

Example 1: Asymptotic stabilization of the inertia wheel pendulum. We will modify the control that we found in Section 2.3.4, in accordance with the considerations above. We first compute the multipliers of the new Lagrangian. Notice that

$$\nu = -\frac{ad_1 - bm}{bd_1 - bm}.$$

Since $R = 0$, we may take the function ρ_2 of the multiplier to be any positive constant. The multiplier is then the constant matrix

$$g_{11} = 1, \qquad g_{12} = -\nu, \qquad g_{22} = \nu^2 + \rho_2.$$

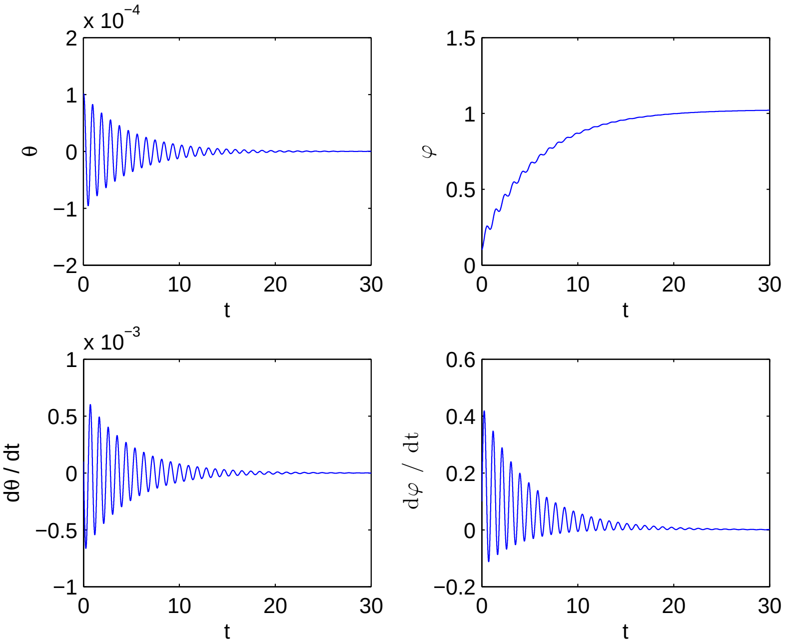


FIGURE 2.5: Simulations of asymptotic stability for the inertia wheel pendulum.

The controlled SODE, with extra control u_2, is

$$\begin{aligned}\ddot{\theta} &= \frac{(m-d_1)\sin(\theta)}{a-b} - \frac{bu_2}{ab-b^2},\\ \ddot{\varphi} &= \frac{(-ad_1+bm)\sin(\theta)}{b(b-a)} + \frac{au_2}{ab-b^2},\end{aligned}$$

where $u_2 = f(\Box\dot{\varphi} + \diamond\dot{\theta})$. Since also the matrix (a_{ij}) is constant, both $\Box$ and $\diamond$ are constants. With all this, condition (2.27) takes a very simple form. If a solution $(\theta(t), \varphi(t))$, other than the equilibrium, exists in the set where $\Box\dot{\varphi} + \diamond\dot{\theta} = 0$, then this solution also satisfies

$$0 = \Box U + \diamond S = \left(\Box\frac{-ad_1+bm}{b(b-a)} + \diamond\frac{m-d_1}{a-b}\right)\sin(\theta(t)) = \frac{(d_1-m)\rho_2}{(a-b)^2}\sin(\theta(t))\,.$$

Since we had chosen $d_1 \neq m$ and $\rho_2 > 0$, we get that the first factor never vanishes. The only possible solution with the above property is therefore given by $\sin(\theta(t)) = 0$, and thus $\theta(t) = 0$. We may conclude that the equilibrium is asymptotically stable.

If we take $u_2 = \frac{-0.1}{\nu^2}(\diamond\dot{\theta} + \Box\dot{\varphi})$, the same parameters and initial conditions as in Section 2.3.4, and $\rho_2 = \nu^2$ then we get the MATLAB simulation from Figure 2.5.

Example 2: The inverted pendulum on a cart. Consider again the new stabilizing control found in Section 2.3.4. In this case we get

$$\nu = \frac{-2(\beta\delta + \alpha d)\cos(\phi)}{(2\gamma\delta + \beta d + \beta d\cos(2\phi))}, \quad \rho_2 = A\frac{(\beta^2 - 2\alpha\gamma + \beta^2\cos(2\phi))^{1-\frac{\alpha d}{\beta\delta}}}{(2\gamma\delta + \beta d + \beta d\cos(2\phi))^2}\,.$$

If we choose $A = 0.04$ and take the control $u_2 = -0.03s^2(\Box\dot{s} + \diamond\dot{\phi})$, then with the same parameters and initial conditions as in Section 2.3.4 we get the MATLAB simulation from Figure 2.6.

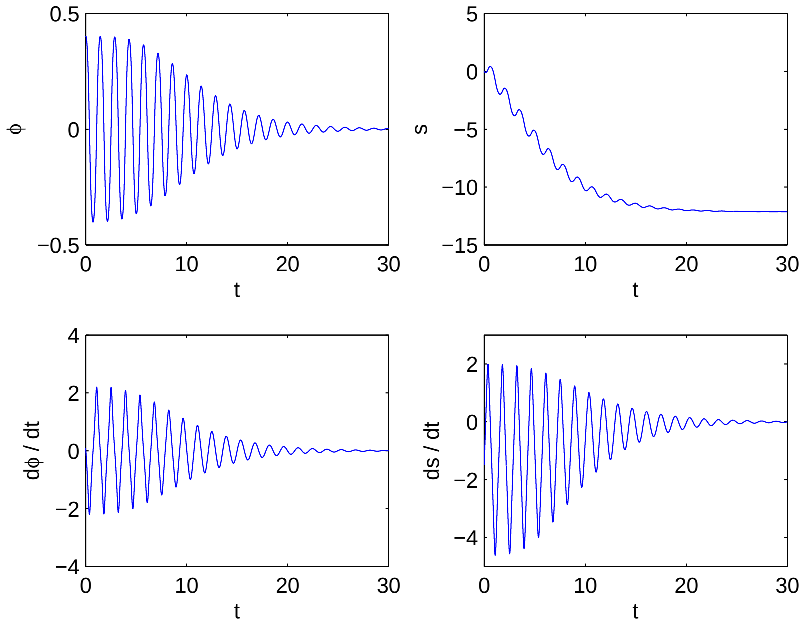


FIGURE 2.6: Simulations of asymptotic stability for the inverted pendulum on a cart.

2.4 Application of Helmholtz conditions to control theory

All of the matching conditions mentioned in Section 2.2 give particular solutions of the Helmholtz conditions (1.14)-(1.16) or equivalently of the implicit Helmholtz conditions introduced in Section 1.5.3, if we consider the Legendre transformation corresponding to $L_{\tau,\sigma}$, $L_{\tau,\sigma,\rho}$ or $L_{\tau,\sigma,\rho,\epsilon}$.

We rewrite the implicit Helmholtz conditions here as

$$\frac{\partial F_i}{\partial \dot q^j} = \frac{\partial F_j}{\partial \dot q^i}\,, \tag{2.28}$$

$$\frac{\partial^2 F_i}{\partial \dot q^j \partial q^k}\dot q^k + \frac{\partial F_i}{\partial q^j} + \frac{\partial^2 F_i}{\partial \dot q^j \partial \dot q^k}\ddot q^k - \frac{\partial F_j}{\partial q^i} = \frac{\partial F_i}{\partial \dot q^k}\frac{\partial \Phi_r}{\partial \dot q^j}(C^{-1})^{kr}\,, \tag{2.29}$$

$$\frac{\partial^2 F_i}{\partial q^j \partial q^k}\dot q^k + \frac{\partial^2 F_i}{\partial q^j \partial \dot q^k}\ddot q^k - \frac{\partial F_i}{\partial \dot q^k}\frac{\partial \Phi_r}{\partial q^j}(C^{-1})^{kr} = \frac{\partial^2 F_j}{\partial q^i \partial q^k}\dot q^k + \frac{\partial^2 F_j}{\partial q^i \partial \dot q^k}\ddot q^k - \frac{\partial F_j}{\partial \dot q^k}\frac{\partial \Phi_r}{\partial q^i}(C^{-1})^{kr}\,, \tag{2.30}$$

where we are now considering the constraints as $\Phi_i = \delta_{ij}\Phi^j$.

Under the matching conditions M1-M3 we obtain a SODE for which $L_{\tau,\sigma}$ solves the problem (1.13). Therefore the multipliers $g_{ij} = \frac{\partial^2 L_{\tau,\sigma}}{\partial \dot q^i \partial \dot q^j}$ must satisfy the Helmholtz conditions (1.14)-(1.16), and the Legendre transformation components $F_i = \frac{\partial L_{\tau,\sigma}}{\partial \dot q^i}$ must satisfy the implicit Helmholz conditions (2.28)-(2.30), and the same holds for $L_{\tau,\sigma,\rho}$ and $L_{\tau,\sigma,\rho,\epsilon}$.

In this section we will slightly modify the expression of the controlled Lagrangian. We will consider Lagrangians of the form

$$\tilde L_{\tau,\sigma} = K_{\tau,\sigma} - \tilde V_{\tau,\sigma}(x^\alpha, \theta^a)\,,$$

where $K_{\tau,\sigma}$ denotes the kinetic energy part of $L_{\tau,\sigma}$, but the potential energy part $\tilde V_{\tau,\sigma}$ does not necessarily coincide with the one in the original Lagrangian.

We will take the corresponding Legendre transformations and impose them as solutions of the implicit Helmholz conditions (2.28)-(2.30). By doing so, we should recover the matching conditions as solutions, but may find new ones due to the freedom in the potential energy part. Now the unknowns are the free parameters that appear both in the controlled Lagrangian and the controlled SODE.

We will follow this approach in Section 2.4.1, where we will see explicitly how the matching conditions M1-M3 arise from the Helmholtz conditions (2.28)-(2.30) if we choose $\tilde{L}_{\tau,\sigma}$ as the new Lagrangian.

In Section 2.4.2, for the case of one degree of underactuation and (g_{ab}) constant we will give an additional solution to the one provided by SM3, using the Helmholtz conditions (2.28)-(2.30). In this case we obtain a feedback control which is independent of velocities. Finally we will derive this new solution for the example of the inverted pendulum on a cart and check stability.

2.4.1 Arbitrary dimension with special matching assumption

In this section we will show how the matching conditions M1-M3 arise from the implicit Helmholtz conditions (2.28)-(2.30) using the special matching assumption, that is, choosing a controlled Lagrangian with $g_\rho = g$. Since we will use the Legendre transformation of the controlled Lagrangian $L_{\tau,\sigma}$, the potential energy of the new Lagrangian will not play any role in satisfying (2.28)-(2.30).

As a starting point consider a given mechanical Lagrangian of the form

$$L(x^\alpha, \dot{x}^\alpha, \dot{\theta}^a) = \frac{1}{2}\left(g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta + 2g_{\alpha a}\dot{x}^\alpha\dot{\theta}^a + g_{ab}\dot{\theta}^a\dot{\theta}^b\right) - V(x^\alpha)\,, \tag{2.31}$$

with corresponding Euler-Lagrange equations given by

$$\begin{aligned}
\Phi_\alpha &= (g_{\alpha\beta,\gamma} - \frac{1}{2}g_{\gamma\beta,\alpha})\dot{x}^\gamma\dot{x}^\beta + (g_{\alpha a,\gamma} - g_{\gamma a,\alpha})\dot{x}^\gamma\dot{\theta}^a \\
&\quad + g_{\alpha\beta}\ddot{x}^\beta + g_{\alpha a}\ddot{\theta}^a - \frac{1}{2}g_{ab,\alpha}\dot{\theta}^a\dot{\theta}^b + \frac{\partial V}{\partial x^\alpha} = 0\,, \\
\Phi_a &= g_{\alpha a,\gamma}\dot{x}^\gamma\dot{x}^\alpha + g_{ab,\gamma}\dot{x}^\gamma\dot{\theta}^b + g_{\alpha a}\ddot{x}^\alpha + g_{ab}\ddot{\theta}^b = 0\,.
\end{aligned}$$

Now consider a controlled Lagrangian with the special matching assumption $g_\rho = g$, that is,

$$L_{\tau,\sigma} = L(x^\alpha, \dot{x}^\alpha, \dot{\theta}^a + \tau^a_\alpha\dot{x}^\alpha) + \frac{1}{2}\sigma_{ab}\tau^a_\alpha\tau^b_\beta\dot{x}^\alpha\dot{x}^\beta\,,$$

and choose feedback controls u_a for $a = 2, \ldots, n$ such that the θ^a-equations of both L and $L_{\tau,\sigma}$ coincide, that is,

$$\begin{aligned}
u_a &= \left(\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}^a}\right) - \frac{\partial L}{\partial\theta^a}\right) - \left(\frac{d}{dt}\left(\frac{\partial L_{\tau,\sigma}}{\partial\dot{\theta}^a}\right) - \frac{\partial L_{\tau,\sigma}}{\partial\theta^a}\right) \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}^a} - \frac{\partial L_{\tau,\sigma}}{\partial\dot{\theta}^a}\right) = \frac{d}{dt}\left(-g_{ab}\tau^b_\beta\dot{x}^\beta\right) = -(g_{ab}\tau^b_\beta)_{,\gamma}\dot{x}^\beta\dot{x}^\gamma - g_{ab}\tau^b_\beta\ddot{x}^\beta\,.
\end{aligned} \tag{2.32}$$

Then the controlled Euler-Lagrange equations (2.3) are

$$\tilde{\Phi}_\alpha = \Phi_\alpha = 0\,, \tag{2.33}$$

$$\tilde{\Phi}_a = \Phi_a + (g_{ab}\tau^b_\beta)_{,\gamma}\dot{x}^\beta\dot{x}^\gamma + g_{ab}\tau^b_\beta\ddot{x}^\beta = 0\,. \tag{2.34}$$

From $\tilde{\Phi}$ we can compute

$$\tilde{C} = \left(\frac{\partial\tilde{\Phi}}{\partial\ddot{q}}\right) = \begin{pmatrix} g_{\alpha\beta} & g_{\alpha b} \\ g_{a\beta} + g_{ad}\tau^d_\beta & g_{ab} \end{pmatrix},$$

which is assumed to be regular. If we introduce the notation $A_{\alpha\beta} = g_{\alpha\beta} - g_{\alpha b}g^{ab}(g_{a\beta} + g_{ad}\tau^d_\beta)$, where (g^{ab}) denotes the inverse matrix of (g_{ab}), and denote the inverse of $(A_{\alpha\beta})$ by $(A^{\alpha\beta})$ then

$$\tilde{C}^{-1} = \tilde{W} = \begin{pmatrix} \tilde{W}^{\alpha\beta} & \tilde{W}^{\alpha b} \\ \tilde{W}^{a\beta} & \tilde{W}^{ab} \end{pmatrix} = \begin{pmatrix} A^{\alpha\beta} & -A^{\alpha\gamma}g_{\gamma d}g^{db} \\ -(g^{ab}g_{b\gamma} + \tau^a_\gamma)A^{\gamma\beta} & g^{ab} + (g^{ad}g_{d\gamma} + \tau^a_\gamma)A^{\gamma\nu}g_{\nu e}g^{eb} \end{pmatrix}.$$

Now we will impose that the system (2.33)-(2.34) be variational using equations (2.28)-(2.30), and with F_j given by the Legendre transformation of the controlled Lagrangian $L_{\tau,\sigma}$, that is,

$$\begin{aligned} \tilde{F}_\alpha &= (g_{\alpha\beta} + g_{\alpha a}\tau^a_\beta + g_{\beta a}\tau^a_\alpha + g_{ab}\tau^a_\alpha\tau^b_\beta + \sigma_{ab}\tau^a_\alpha\tau^b_\beta)\dot{x}^\beta + (g_{\alpha b} + g_{db}\tau^d_\alpha)\dot{\theta}^b, \\ \tilde{F}_a &= (g_{\alpha a} + g_{da}\tau^d_\alpha)\dot{x}^\alpha + g_{ab}\dot{\theta}^b. \end{aligned}$$

If we plug $\tilde{\Phi}_\alpha$, $\tilde{\Phi}_a$ and the proposed solutions $\tilde{F}_\alpha$, $\tilde{F}_a$ into equations (2.28)-(2.30) we get the following:

- Equation (2.28): vanishes identically for all indices
- Equation (2.29):
  - * $(2.29)_{ab}$ vanishes identically
  - * $(2.29)_{a\beta}$ vanishes identically
  - * $(2.29)_{\alpha b}$ vanishes using M1 and M3 . Alternatively it vanishes if (g_{ab}) is constant and the system has one degree of underactuation
  - * $(2.29)_{\alpha\beta}$ vanishes using M1, M2 and M3.
- Equation (2.30):
  - * $(2.30)_{ab}$ vanishes identically
  - * $(2.30)_{\alpha b}$ vanishes identically
  - * $(2.30)_{\alpha\beta}$ vanishes using M1, M2 and M3. Alternatively it vanishes for systems with one degree of underactuation.

Now detailed computations follow, until Section 2.4.2, proving the above statement. We will need

$$\frac{\partial\tilde{\Phi}_c}{\partial\dot{x}^\beta} = \left(g_{\beta c,\gamma} + g_{\gamma c,\beta} + (g_{cd}\tau^d_\beta)_{,\gamma} + (g_{cd}\tau^d_\gamma)_{,\beta}\right)\dot{x}^\gamma + g_{cd,\beta}\dot{\theta}^d, \tag{2.35}$$

$$\begin{aligned} \frac{\partial\tilde{\Phi}_a}{\partial x^\alpha} &= (g_{\nu a,\gamma\alpha} + (g_{ad}\tau^d_\nu)_{\alpha\gamma})\dot{x}^\gamma\dot{x}^\nu + g_{ab,\gamma\alpha}\dot{x}^\gamma\dot{\theta}^b + g_{ab,\alpha}\ddot{\theta}^b \\ &\quad + (g_{\gamma a,\alpha} + g_{ad,\alpha}\tau^d_\gamma + g_{ad}\tau^d_{\gamma,\alpha})\ddot{x}^\gamma, \end{aligned} \tag{2.36}$$

$$\frac{\partial\tilde{\Phi}_\nu}{\partial\dot{x}^\beta} = (g_{\nu\beta,\gamma} + g_{\nu\gamma,\beta} - g_{\gamma\beta,\nu})\dot{x}^\gamma + (g_{\nu a,\beta} - g_{\beta a,\nu})\dot{\theta}^a, \tag{2.37}$$

$$\frac{\partial\tilde{\Phi}_\gamma}{\partial x^\alpha} = g_{\gamma\nu,\eta\alpha}\dot{x}^\eta\dot{x}^\nu + g_{\gamma a,\nu\alpha}\dot{x}^\nu\dot{\theta}^a + g_{\gamma\nu,\alpha}\ddot{x}^\nu + g_{\gamma a,\alpha}\ddot{\theta}^a$$

$$-\frac{1}{2}g_{\nu\eta,\gamma\alpha}\dot{x}^\nu\dot{x}^\eta - g_{\nu a,\gamma\alpha}\dot{x}^\nu\dot{\theta}^a - \frac{1}{2}g_{ab,\gamma\alpha}\dot{\theta}^a\dot{\theta}^b + \frac{\partial^2 V}{\partial x^\alpha \partial x^\gamma}\,, \tag{2.38}$$

and also the following expressions:

$$\begin{aligned}
\left(\frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\tilde{W}^{\gamma\nu} + \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\tilde{W}^{a\nu}\right) &= (g_{\alpha\gamma} + g_{\alpha d}\tau^d_\gamma + g_{\gamma d}\tau^d_\alpha + g_{fd}\tau^f_\alpha\tau^d_\gamma + \sigma_{fd}\tau^f_\alpha\tau^d_\gamma)A^{\gamma\nu} \\
&\quad -(g_{\alpha a} + g_{da}\tau^d_\alpha)\left(g^{ae}g_{e\gamma} + \tau^a_\gamma\right)A^{\gamma\nu} \\
&= (g_{\alpha\gamma} - g_{\alpha a}g^{ae}g_{e\gamma} + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu} \\
&= \delta^\nu_\alpha + (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\,, \qquad (2.39) \\
\left(\frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\tilde{W}^{\gamma c} + \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\tilde{W}^{ac}\right) &= (g_{\alpha\gamma} + g_{\alpha d}\tau^d_\gamma + g_{\gamma d}\tau^d_\alpha + g_{fd}\tau^f_\alpha\tau^d_\gamma + \sigma_{fd}\tau^f_\alpha\tau^d_\gamma)\left(-A^{\gamma\eta}g_{\eta e}g^{ec}\right) \\
&\quad +(g_{\alpha a} + g_{da}\tau^d_\alpha)\left(g^{ac} + (g^{af}g_{f\eta} + \tau^a_\eta)A^{\eta\nu}g_{\nu e}g^{ec}\right) \\
&= (-g_{\alpha\gamma} - \sigma_{ad}\tau^a_\alpha\tau^d_\gamma + g_{\alpha a}g^{af}g_{f\gamma})A^{\gamma\eta}g_{\eta e}g^{ec} + g_{\alpha a}g^{ac} + \tau^c_\alpha \\
&= \tau^c_\alpha - (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\,. \qquad (2.40)
\end{aligned}$$

Equation (2.28) vanishes identically for all indices:

$$\begin{aligned}
(2.28)_{ab} &= \frac{\partial\tilde{F}_a}{\partial\dot{\theta}^b} - \frac{\partial\tilde{F}_b}{\partial\dot{\theta}^a} = g_{ab} - g_{ba} = 0\,, \\
(2.28)_{a\beta} &= \frac{\partial\tilde{F}_a}{\partial\dot{x}^\beta} - \frac{\partial\tilde{F}_\beta}{\partial\dot{\theta}^a} = g_{\beta a} + g_{ad}\tau^d_\beta - (g_{\beta a} + g_{da}\tau^d_\beta) = 0\,, \\
(2.28)_{\alpha\beta} &= \frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\beta} - \frac{\partial\tilde{F}_\beta}{\partial\dot{\theta}^\alpha} = (g_{\alpha\beta} + g_{\alpha a}\tau^a_\beta + g_{\beta a}\tau^a_\alpha + g_{ab}\tau^a_\alpha\tau^b_\beta + \sigma_{ab}\tau^a_\alpha\tau^b_\beta) \\
&\quad -(g_{\beta\alpha} + g_{\beta a}\tau^a_\alpha + g_{\alpha a}\tau^a_\beta + g_{ab}\tau^a_\beta\tau^b_\alpha + \sigma_{ab}\tau^a_\beta\tau^b_\alpha) = 0\,.
\end{aligned}$$

Equation (2.29) vanishes identically for indices ab and $a\beta$:

$$\begin{aligned}
(2.29)_{ab} &= \frac{\partial^2\tilde{F}_a}{\partial\dot{\theta}^b\partial q^k}\dot{q}^k + \frac{\partial\tilde{F}_a}{\partial\theta^b} + \frac{\partial^2\tilde{F}_a}{\partial\dot{\theta}^b\partial\dot{q}^k}\ddot{q}^k - \frac{\partial\tilde{F}_b}{\partial\theta^a} - \frac{\partial\tilde{F}_a}{\partial\dot{q}^k}\frac{\partial\tilde{\Phi}_r}{\partial\dot{\theta}^b}\tilde{W}^{kr} \\
&= \frac{\partial^2\tilde{F}_a}{\partial\dot{\theta}^b\partial x^\gamma}\dot{x}^\gamma - \frac{\partial\tilde{F}_a}{\partial\dot{x}^\gamma}\left(\frac{\partial\Phi_c}{\partial\dot{\theta}^b}\tilde{W}^{\gamma c} + \frac{\partial\Phi_\beta}{\partial\dot{\theta}^b}\tilde{W}^{\gamma\beta}\right) - \frac{\partial\tilde{F}_a}{\partial\dot{\theta}^d}\left(\frac{\partial\Phi_c}{\partial\dot{\theta}^b}\tilde{W}^{dc} + \frac{\partial\Phi_\beta}{\partial\dot{\theta}^b}\tilde{W}^{d\beta}\right) \\
&= g_{ab,\gamma}\dot{x}^\gamma - (\tilde{C}_{a\gamma}\tilde{W}^{\gamma c} + \tilde{C}_{ad}\tilde{W}^{dc})\frac{\partial\Phi_c}{\partial\dot{\theta}^b} - (\tilde{C}_{a\gamma}\tilde{W}^{\gamma\beta} + \tilde{C}_{ad}\tilde{W}^{d\beta})\frac{\partial\Phi_\beta}{\partial\dot{\theta}^b} \\
&= g_{ab,\gamma}\dot{x}^\gamma - \frac{\partial\Phi_a}{\partial\dot{\theta}^b} = 0 \\
(2.29)_{a\beta} &= \frac{\partial^2\tilde{F}_a}{\partial\dot{x}^\beta\partial q^k}\dot{q}^k + \frac{\partial\tilde{F}_a}{\partial x^\beta} + \frac{\partial^2\tilde{F}_a}{\partial\dot{x}^\beta\partial\dot{q}^k}\ddot{q}^k - \frac{\partial\tilde{F}_\beta}{\partial\theta^a} - \frac{\partial\tilde{F}_a}{\partial\dot{q}^k}\frac{\partial\tilde{\Phi}_r}{\partial\dot{x}^\beta}\tilde{W}^{kr} \\
&= (g_{\beta a} + g_{ab}\tau^b_\beta)_{,\gamma}\dot{x}^\gamma + (g_{\gamma a} + g_{ab}\tau^b_\gamma)_{,\beta}\dot{x}^\gamma + g_{ab,\beta}\dot{\theta}^b \\
&\quad -\frac{\partial\tilde{F}_a}{\partial\dot{x}^\gamma}\left(\frac{\partial\tilde{\Phi}_c}{\partial\dot{x}^\beta}\tilde{W}^{\gamma c} + \frac{\partial\Phi_\nu}{\partial\dot{x}^\beta}\tilde{W}^{\gamma\nu}\right) - \frac{\partial\tilde{F}_a}{\partial\dot{\theta}^d}\left(\frac{\partial\tilde{\Phi}_c}{\partial\dot{x}^\beta}\tilde{W}^{dc} + \frac{\partial\Phi_\nu}{\partial\dot{x}^\beta}\tilde{W}^{d\nu}\right) \\
&= (g_{\beta a} + g_{ab}\tau^b_\beta)_{,\gamma}\dot{x}^\gamma + (g_{\gamma a} + g_{ab}\tau^b_\gamma)_{,\beta}\dot{x}^\gamma + g_{ab,\beta}\dot{\theta}^b \\
&\quad -(\tilde{C}_{a\gamma}\tilde{W}^{\gamma c} + \tilde{C}_{ad}\tilde{W}^{dc})\frac{\partial\tilde{\Phi}_c}{\partial\dot{x}^\beta} - (\tilde{C}_{a\gamma}\tilde{W}^{\gamma\nu} + \tilde{C}_{ad}\tilde{W}^{d\nu})\frac{\partial\Phi_\nu}{\partial\dot{x}^\beta}
\end{aligned}$$

$$
\begin{aligned}
&= (g_{\beta a} + g_{ab}\tau^b_\beta)_{,\gamma}\dot{x}^\gamma + (g_{\gamma a} + g_{ab}\tau^b_\gamma)_{,\beta}\dot{x}^\gamma + g_{ab,\beta}\dot{\theta}^b - \frac{\partial\tilde{\Phi}_a}{\partial\dot{x}^\beta} \\
&\overset{(2.35)}{=} 0\,.
\end{aligned}
$$

For indices αb Equation (2.29) vanishes using M1 and M3 . Alternatively it vanishes if (g_{ab}) is constant and the system has one degree of underactuation:

$$
\begin{aligned}
(2.29)_{\alpha b} \quad &= \frac{\partial^2\tilde{F}_\alpha}{\partial\dot{\theta}^b\partial q^k}\dot{q}^k + \frac{\partial\tilde{F}_\alpha}{\partial\theta^b} + \frac{\partial^2\tilde{F}_\alpha}{\partial\dot{\theta}^b\partial\dot{q}^k}\ddot{q}^k - \frac{\partial\tilde{F}_b}{\partial x^\alpha} - \frac{\partial\tilde{F}_\alpha}{\partial\dot{q}^k}\frac{\partial\tilde{\Phi}_r}{\partial\dot{\theta}^b}\tilde{W}^{kr} \\
&= \frac{\partial^2\tilde{F}_\alpha}{\partial\dot{\theta}^b\partial x^\gamma}\dot{x}^\gamma - \frac{\partial\tilde{F}_b}{\partial x^\alpha} - \frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\left(\frac{\partial\tilde{\Phi}_c}{\partial\dot{\theta}^b}\tilde{W}^{\gamma c} + \frac{\partial\tilde{\Phi}_\nu}{\partial\dot{\theta}^b}\tilde{W}^{\gamma\nu}\right) - \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\left(\frac{\partial\tilde{\Phi}_c}{\partial\dot{\theta}^b}\tilde{W}^{ac} + \frac{\partial\tilde{\Phi}_\nu}{\partial\dot{\theta}^b}\tilde{W}^{a\nu}\right) \\
&= \frac{\partial^2\tilde{F}_\alpha}{\partial\dot{\theta}^b\partial x^\gamma}\dot{x}^\gamma - \frac{\partial\tilde{F}_b}{\partial x^\alpha} - \left(\frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\tilde{W}^{\gamma c} + \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\tilde{W}^{ac}\right)\frac{\partial\tilde{\Phi}_c}{\partial\dot{\theta}^b} - \left(\frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\tilde{W}^{\gamma\nu} + \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\tilde{W}^{a\nu}\right)\frac{\partial\tilde{\Phi}_\nu}{\partial\dot{\theta}^b} \\
&\overset{(2.39),(2.40)}{=} (g_{\alpha b} + g_{db}\tau^d_\alpha)_{,\gamma}\dot{x}^\gamma - (g_{\gamma b} + g_{bd}\tau^d_\gamma)_{,\alpha}\dot{x}^\gamma - g_{bd,\alpha}\dot{\theta}^d \\
&\quad - \left(\tau^c_\alpha - (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\right)g_{cb,\gamma}\dot{x}^\gamma \\
&\quad - \left(\delta^\nu_\alpha + (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\right)\left((g_{\nu b,\gamma} - g_{\gamma b,\nu})\,\dot{x}^\gamma - g_{db,\nu}\dot{\theta}^d\right)\,.
\end{aligned}
$$

The $\dot{\theta}$ component becomes

$$
\left(-g_{bd,\alpha} + \left(\delta^\nu_\alpha + (g_{\alpha e}\tau^e_\gamma + \sigma_{ae}\tau^a_\alpha\tau^e_\gamma)A^{\gamma\nu}\right)g_{db,\nu}\right)\dot{\theta}^d = (g_{\alpha e} + \sigma_{ae}\tau^a_\alpha)\tau^e_\gamma A^{\gamma\nu}g_{db,\nu}\dot{\theta}^d\,,
$$

from where we can clearly see the solution M1 (but there are more). The $\dot{x}$ component becomes

$$
\begin{aligned}
&(g_{\alpha b} + g_{db}\tau^d_\alpha)_{,\gamma}\dot{x}^\gamma - (g_{\gamma b} + g_{bd}\tau^d_\gamma)_{,\alpha}\dot{x}^\gamma - \left(\tau^c_\alpha - (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\right)g_{cb,\gamma}\dot{x}^\gamma \\
&- \left(\delta^\nu_\alpha + (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\right)(g_{\nu b,\gamma} - g_{\gamma b,\nu})\,\dot{x}^\gamma \\
= \;&\left(g_{db}\left(\tau^d_{\alpha,\gamma} - \tau^d_{\gamma,\alpha}\right) - g_{bd,\alpha}\tau^d_\gamma + (g_{\alpha d} + \sigma_{ad}\tau^a_\alpha)\tau^d_\gamma A^{\gamma\eta}\left(g_{\eta e}g^{ec}g_{cb,\gamma} - g_{\eta b,\gamma} + g_{\gamma b,\eta}\right)\right)\dot{x}^\gamma\,, \quad (2.41)
\end{aligned}
$$

from where assuming M1 we obtain M3 as a solution.

On the other hand notice that if (g_{ab}) is constant, that is, SM2 holds, and the system has one degree of underacuation then the equation vanishes identically without imposing M1 nor M3 (conditions which involve τ).

It is also enough to assume the simplified matching conditions SM2, SM4 and the matching condition M3 (without M1).

For indices $\alpha\beta$ Equation (2.29) vanishes using M1, M2 and M3:

$$
\begin{aligned}
(2.29)_{\alpha\beta} \quad &= \frac{\partial^2\tilde{F}_\alpha}{\partial\dot{x}^\beta\partial q^k}\dot{q}^k + \frac{\partial\tilde{F}_\alpha}{\partial x^\beta} + \frac{\partial^2\tilde{F}_\alpha}{\partial\dot{x}^\beta\partial\dot{q}^k}\ddot{q}^k - \frac{\partial\tilde{F}_\beta}{\partial x^\alpha} - \frac{\partial\tilde{F}_\alpha}{\partial\dot{q}^k}\frac{\partial\tilde{\Phi}_r}{\partial\dot{x}^\beta}\tilde{W}^{kr} \\
&= \left(-\frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\tilde{W}^{\gamma c} - \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\tilde{W}^{ac}\right)\frac{\partial\tilde{\Phi}_c}{\partial\dot{x}^\beta} + \left(-\frac{\partial\tilde{F}_\alpha}{\partial\dot{x}^\gamma}\tilde{W}^{\gamma\nu} - \frac{\partial\tilde{F}_\alpha}{\partial\dot{\theta}^a}\tilde{W}^{a\nu}\right)\frac{\partial\tilde{\Phi}_\nu}{\partial\dot{x}^\beta} \\
&\quad + (g_{\alpha\beta} + g_{\alpha d}\tau^d_\beta + g_{\beta d}\tau^d_\alpha + g_{de}\tau^d_\beta\tau^e_\alpha + \sigma_{ed}\tau^e_\alpha\tau^d_\beta)_{,\gamma}\dot{x}^\gamma + \frac{\partial\tilde{F}_\alpha}{\partial x^\beta} - \frac{\partial\tilde{F}_\beta}{\partial x^\alpha} \\
&\overset{(2.39),(2.40)}{=} -\left(\tau^c_\alpha - (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\right)\frac{\partial\tilde{\Phi}_c}{\partial\dot{x}^\beta} - \left(\delta^\nu_\alpha + (g_{\alpha d}\tau^d_\gamma + \sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\right)\frac{\partial\tilde{\Phi}_\nu}{\partial\dot{x}^\beta}
\end{aligned}
$$

$$
\begin{aligned}
&+(g_{\alpha\beta}+g_{\alpha d}\tau^d_\beta+g_{\beta d}\tau^d_\alpha+g_{de}\tau^d_\beta\tau^e_\alpha+\sigma_{ed}\tau^e_\alpha\tau^d_\beta)_{,\gamma}\dot{x}^\gamma+\frac{\partial\tilde{F}_\alpha}{\partial x^\beta}-\frac{\partial\tilde{F}_\beta}{\partial x^\alpha}\\
= \quad &-\left(\tau^c_\alpha-(g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\right)\left(g_{\beta c,\gamma}+g_{\gamma c,\beta}+(g_{cd}\tau^d_\beta)_{,\gamma}+(g_{cd}\tau^d_\gamma)_{,\beta}\right)\dot{x}^\gamma\\
&-\left(\tau^c_\alpha-(g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\right)g_{cf,\beta}\dot{\theta}^f\\
&-\left(\delta^\nu_\alpha+(g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\right)(g_{\nu\beta,\gamma}+g_{\nu\gamma,\beta}-g_{\gamma\beta,\nu})\dot{x}^\gamma\\
&-\left(\delta^\nu_\alpha+(g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\right)(g_{\nu f,\beta}-g_{\beta f,\nu})\dot{\theta}^f\\
&+(g_{\alpha\beta}+g_{\alpha d}\tau^d_\beta+g_{\beta d}\tau^d_\alpha+g_{de}\tau^d_\beta\tau^e_\alpha+\sigma_{ed}\tau^e_\alpha\tau^d_\beta)_{,\gamma}\dot{x}^\gamma\\
&+(g_{\alpha\gamma}+g_{\alpha d}\tau^d_\gamma+g_{\gamma d}\tau^d_\alpha+g_{ab}\tau^a_\alpha\tau^b_\gamma+\sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_{,\beta}\dot{x}^\gamma+(g_{\alpha f}+g_{ef}\tau^e_\alpha)_{,\beta}\dot{\theta}^f\\
&-(g_{\beta\gamma}+g_{\beta d}\tau^d_\gamma+g_{\gamma d}\tau^d_\beta+g_{ab}\tau^a_\beta\tau^b_\gamma+\sigma_{ab}\tau^a_\beta\tau^b_\gamma)_{,\alpha}\dot{x}^\gamma-(g_{\beta f}+g_{ef}\tau^e_\beta)_{,\alpha}\dot{\theta}^f\,.
\end{aligned}
$$

The $\dot{\theta}$ component becomes

$$
\left((g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}\left(g_{\eta e}g^{ec}g_{cf,\beta}-g_{\eta f,\beta}+g_{\beta f,\eta}\right)+g_{ef}\left(\tau^e_{\alpha,\beta}-\tau^e_{\beta,\alpha}\right)-g_{ef,\alpha}\tau^e_\beta\right)\dot{\theta}^f\,, \tag{2.42}
$$

which vanishes with the same assumptions as (2.41).

The $\dot{x}$ component becomes

$$
\begin{aligned}
&-\left(\tau^c_\alpha-(g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\eta}g_{\eta e}g^{ec}\right)\left(g_{\beta c,\gamma}+g_{\gamma c,\beta}+(g_{cd}\tau^d_\beta)_{,\gamma}+(g_{cd}\tau^d_\gamma)_{,\beta}\right)\dot{x}^\gamma\\
&-\left(\delta^\nu_\alpha+(g_{\alpha d}\tau^d_\gamma+\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)A^{\gamma\nu}\right)(g_{\nu\beta,\gamma}+g_{\nu\gamma,\beta}-g_{\gamma\beta,\nu})\dot{x}^\gamma\\
&+(g_{\alpha\beta}+g_{\alpha d}\tau^d_\beta+g_{\beta d}\tau^d_\alpha+g_{de}\tau^d_\beta\tau^e_\alpha+\sigma_{ed}\tau^e_\alpha\tau^d_\beta)_{,\gamma}\dot{x}^\gamma\\
&+(g_{\alpha\gamma}+g_{\alpha d}\tau^d_\gamma+g_{\gamma d}\tau^d_\alpha+g_{ab}\tau^a_\alpha\tau^b_\gamma+\sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_{,\beta}\dot{x}^\gamma\\
&-(g_{\beta\gamma}+g_{\beta d}\tau^d_\gamma+g_{\gamma d}\tau^d_\beta+g_{ab}\tau^a_\beta\tau^b_\gamma+\sigma_{ab}\tau^a_\beta\tau^b_\gamma)_{,\alpha}\dot{x}^\gamma\\
\overset{\text{M1}}{=}\quad &(g_{\alpha a}\tau^a_\beta)_{,\gamma}+g_{\beta a}\tau^a_{\alpha,\gamma}+g_{ad}\tau^d_\beta\tau^a_{\alpha,\gamma}+(\sigma_{ad}\tau^a_\alpha\tau^d_\beta)_{,\gamma}+(g_{\alpha d}\tau^d_\gamma)_{,\beta}+g_{\gamma d}\tau^d_{\alpha,\beta}+g_{ab}\tau^b_\gamma\tau^a_{\alpha,\beta}\\
&+(\sigma_{ad}\tau^a_\alpha\tau^d_\gamma)_{,\beta}-(g_{\beta d}\tau^d_\gamma)_{,\alpha}-(g_{\gamma d}\tau^d_\beta)_{,\alpha}-(g_{ab}\tau^a_\beta\tau^b_\gamma)_{,\alpha}-(\sigma_{ad}\tau^a_\beta\tau^d_\gamma)_{,\alpha}\\
=\quad &g_{ab}\tau^b_\gamma(\tau^a_{\alpha,\beta}-\tau^a_{\beta,\alpha})+g_{\beta a}(\tau^a_{\alpha,\gamma}-\tau^a_{\gamma,\alpha})+g_{ad}\tau^d_\beta(\tau^a_{\alpha,\gamma}-\tau^a_{\gamma,\alpha})+g_{\gamma d}\tau^d_{\alpha,\beta}-g_{\beta d,\alpha}\tau^d_\gamma-g_{ab,\alpha}\tau^a_\beta\tau^b_\gamma
\end{aligned}
$$

where we have used again M1 and the symmetry of σ_{ad}. Now using M3 to cancel the first and last terms we have

$$
\begin{aligned}
&(g_{\beta a}+g_{ad}\tau^d_\beta)(\tau^a_{\alpha,\gamma}-\tau^a_{\gamma,\alpha})+g_{\gamma d}\tau^d_{\alpha,\beta}-g_{\beta d,\alpha}\tau^d_\gamma\\
\overset{\text{M1,M3}}{=}\quad &(g_{\beta a}+g_{ad}\tau^d_\beta)(g^{ea}g_{fe,\alpha}\tau^f_\gamma)+g_{\gamma d}\tau^d_{\alpha,\beta}+(\sigma_{de}\tau^e_\beta)_{,\alpha}\tau^d_\gamma\\
\overset{\text{M1}}{=}\quad &g_{\beta a}g^{ea}g_{fe,\alpha}\tau^f_\gamma+\tau^d_\beta g_{fd,\alpha}\tau^f_\gamma-\sigma_{de}\tau^e_\gamma\tau^d_{\alpha,\beta}+\sigma_{de,\alpha}\tau^e_\beta\tau^d_\gamma+\sigma_{de}\tau^e_{\beta,\alpha}\tau^d_\gamma\\
\overset{\text{M1}}{=}\quad &-\sigma_{ah}\tau^h_\beta g^{ea}g_{fe,\alpha}\tau^f_\gamma+\tau^d_\beta g_{fd,\alpha}\tau^f_\gamma+\sigma_{de,\alpha}\tau^e_\beta\tau^d_\gamma+\sigma_{de}\tau^d_\gamma(\tau^e_{\beta,\alpha}-\tau^e_{\alpha,\beta})\\
\overset{\text{M3}}{=}\quad &-\sigma_{ah}\tau^h_\beta g^{ea}g_{fe,\alpha}\tau^f_\gamma-\sigma_{de}\tau^d_\gamma g^{fe}g_{hf,\alpha}\tau^h_\beta+(g_{de,\alpha}+\sigma_{de,\alpha})\tau^d_\gamma\tau^e_\beta\,.
\end{aligned}
$$

Summing up, for the $\dot{x}$ component, using M1 and M3 we get

$$
(g_{de,\alpha}+\sigma_{de,\alpha})\tau^d_\gamma\tau^e_\beta-\sigma_{ah}\tau^h_\beta g^{ea}g_{fe,\alpha}\tau^f_\gamma-\sigma_{de}\tau^d_\gamma g^{fe}g_{hf,\alpha}\tau^h_\beta\,. \tag{2.43}
$$

Now using M2 we have

$$
\begin{aligned}
& (g_{de,\alpha}+\sigma_{de,\alpha})\tau^d_\gamma\tau^e_\beta - \sigma_{ah}\tau^h_\beta g^{ea} g_{fe,\alpha}\tau^f_\gamma - \sigma_{de}\tau^d_\gamma g^{fe} g_{hf,\alpha}\tau^h_\beta \\
\overset{\text{M2}}{=} \; & (g_{de,\alpha}+\sigma_{de,\alpha})\tau^d_\gamma\tau^e_\beta - \sigma_{ah}\tau^h_\beta \frac{1}{2}\sigma^{ea}(\sigma_{de,\alpha}+g_{de,\alpha})\tau^d_\gamma - \sigma_{de}\tau^d_\gamma\frac{1}{2}\sigma^{fe}(\sigma_{hf,\alpha}+g_{hf,\alpha})\tau^h_\beta \\
= \; & (g_{de,\alpha}+\sigma_{de,\alpha})\tau^d_\gamma\tau^e_\beta - \frac{1}{2}(\sigma_{dh,\alpha}+g_{dh,\alpha})\tau^d_\gamma\tau^h_\beta - \frac{1}{2}(\sigma_{hd,\alpha}+g_{hd,\alpha})\tau^h_\beta\tau^d_\gamma = 0\,.
\end{aligned}
$$

That is, using all M1, M2 and M3 we get that the Helmholtz condition $(2.29)_{\alpha\beta}$ vanishes.

Equation (2.30) vanishes identically for indices ab and αb:

$$
\begin{aligned}
(2.30)_{ab} \;=\; & \frac{\partial^2\tilde F_a}{\partial\theta^b\partial q^k}\dot q^k + \frac{\partial^2\tilde F_a}{\partial\theta^b\partial\dot q^k}\ddot q^k - \frac{\partial\tilde F_a}{\partial\dot q^k}\frac{\partial\tilde\Phi_r}{\partial\theta^b}\tilde W^{kr} - \frac{\partial^2\tilde F_b}{\partial\theta^a\partial q^k}\dot q^k - \frac{\partial^2\tilde F_b}{\partial\theta^a\partial\dot q^k}\ddot q^k + \frac{\partial\tilde F_b}{\partial\dot q^k}\frac{\partial\tilde\Phi_r}{\partial\theta^a}\tilde W^{kr} = 0\,,\\
(2.30)_{\alpha b} \;=\; & \frac{\partial^2\tilde F_\alpha}{\partial\theta^b\partial q^k}\dot q^k + \frac{\partial^2\tilde F_\alpha}{\partial\theta^b\partial\dot q^k}\ddot q^k - \frac{\partial\tilde F_\alpha}{\partial\dot q^k}\frac{\partial\tilde\Phi_r}{\partial\theta^b}\tilde W^{kr} - \frac{\partial^2\tilde F_b}{\partial x^\alpha\partial q^k}\dot q^k - \frac{\partial^2\tilde F_b}{\partial x^\alpha\partial\dot q^k}\ddot q^k + \frac{\partial\tilde F_b}{\partial\dot q^k}\frac{\partial\tilde\Phi_r}{\partial x^\alpha}\tilde W^{kr}\\
=\; & -\frac{\partial^2\tilde F_b}{\partial x^\alpha\partial x^\gamma}\dot x^\gamma - \frac{\partial^2\tilde F_b}{\partial x^\alpha\partial\dot q^k}\ddot q^k + \frac{\partial\tilde F_b}{\partial\dot q^k}\left(\frac{\partial\tilde\Phi_\nu}{\partial x^\alpha}\tilde W^{k\nu} + \frac{\partial\tilde\Phi_d}{\partial x^\alpha}\tilde W^{kd}\right)\\
=\; & -\frac{\partial^2\tilde F_b}{\partial x^\alpha\partial x^\gamma}\dot x^\gamma - \frac{\partial^2\tilde F_b}{\partial x^\alpha\partial\dot q^k}\ddot q^k + \left(\tilde C_{b\gamma}\tilde W^{\gamma\nu} + \tilde C_{bc}\tilde W^{c\nu}\right)\frac{\partial\tilde\Phi_\nu}{\partial x^\alpha} + \left(\tilde C_{b\gamma}\tilde W^{\gamma d} + \tilde C_{bc}\tilde W^{cd}\right)\frac{\partial\tilde\Phi_d}{\partial x^\alpha}\\
=\; & -\frac{\partial^2\tilde F_b}{\partial x^\alpha\partial x^\gamma}\dot x^\gamma - \frac{\partial^2\tilde F_b}{\partial x^\alpha\partial\dot x^\gamma}\ddot x^\gamma - \frac{\partial^2\tilde F_b}{\partial x^\alpha\partial\dot\theta^c}\ddot\theta^c + \frac{\partial\tilde\Phi_b}{\partial x^\alpha}\\
=\; & -(g_{\nu b}+g_{bc}\tau^c_\nu)_{,\alpha\gamma}\dot x^\nu\dot x^\gamma - g_{bc,\alpha\gamma}\dot x^\gamma\dot\theta^c - (g_{\gamma b}+g_{bc}\tau^c_\gamma)\ddot x^\gamma - g_{bc,\alpha}\ddot\theta^c + \frac{\partial\tilde\Phi_b}{\partial x^\alpha}\\
\overset{(2.36)}{=}\; & 0\,.
\end{aligned}
$$

For indices $\alpha\beta$ Equation (2.30) vanishes using M1, M2 and M3 or alternatively for one degree of underactuation, since it is symmetric in α and β :

$$
\begin{aligned}
(2.30)_{\alpha\beta} \;=\; & \frac{\partial^2\tilde F_\alpha}{\partial x^\beta\partial x^\gamma}\dot x^\gamma + \frac{\partial^2\tilde F_\alpha}{\partial x^\beta\partial\dot q^k}\dot v^k - \frac{\partial^2\tilde F_\beta}{\partial x^\alpha\partial x^\gamma}\dot x^\gamma - \frac{\partial^2\tilde F_\beta}{\partial x^\alpha\partial\dot q^k}\dot v^k\\
& -\frac{\partial\tilde F_\alpha}{\partial\dot q^k}\frac{\partial\tilde\Phi_r}{\partial x^\beta}\tilde W^{kr} + \frac{\partial\tilde F_\beta}{\partial\dot q^k}\frac{\partial\tilde\Phi_r}{\partial x^\alpha}\tilde W^{kr}\\
\overset{\text{M1}}{=}\; & (g_{\alpha b}+g_{ab}\tau^a_\alpha)_{,\beta\gamma}\dot\theta^b\dot x^\gamma + (g_{\alpha\nu}+g_{\alpha a}\tau^a_\nu+g_{\nu a}\tau^a_\alpha+g_{ab}\tau^a_\alpha\tau^b_\nu+\sigma_{ab}\tau^a_\alpha\tau^b_\nu)_{,\beta\gamma}\dot x^\nu\dot x^\gamma\\
& +(g_{\alpha b}+g_{ab}\tau^a_\alpha)_{,\beta}\ddot\theta^b + (g_{\alpha\gamma}+g_{\alpha a}\tau^a_\gamma+g_{\gamma a}\tau^a_\alpha+g_{ab}\tau^a_\alpha\tau^b_\gamma+\sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_{,\beta}\ddot x^\gamma\\
& -(g_{\beta b}+g_{ab}\tau^a_\beta)_{,\alpha\gamma}\dot\theta^b\dot x^\gamma - (g_{\beta\nu}+g_{\beta a}\tau^a_\nu+g_{\nu a}\tau^a_\beta+g_{ab}\tau^a_\beta\tau^b_\nu+\sigma_{ab}\tau^a_\beta\tau^b_\nu)_{,\alpha\gamma}\dot x^\nu\dot x^\gamma\\
& -(g_{\beta b}+g_{ab}\tau^a_\beta)_{,\alpha}\ddot\theta^b - (g_{\beta\gamma}+g_{\beta a}\tau^a_\gamma+g_{\gamma a}\tau^a_\beta+g_{ab}\tau^a_\beta\tau^b_\gamma+\sigma_{ab}\tau^a_\beta\tau^b_\gamma)_{,\alpha}\ddot x^\gamma\\
& -\frac{\partial\Phi_\alpha}{\partial x^\beta} - \tau^c_\alpha\frac{\partial\tilde\Phi_c}{\partial x^\beta} + \frac{\partial\Phi_\beta}{\partial x^\alpha} + \tau^c_\beta\frac{\partial\tilde\Phi_c}{\partial x^\alpha}\\
\overset{(2.36),(2.38)}{=}\; & (g_{\alpha b}+g_{ab}\tau^a_\alpha)_{,\beta\gamma}\dot\theta^b\dot x^\gamma + (g_{\alpha\nu}+g_{\alpha a}\tau^a_\nu+g_{\nu a}\tau^a_\alpha+g_{ab}\tau^a_\alpha\tau^b_\nu+\sigma_{ab}\tau^a_\alpha\tau^b_\nu)_{,\beta\gamma}\dot x^\nu\dot x^\gamma\\
& +(g_{\alpha b}+g_{ab}\tau^a_\alpha)_{,\beta}\ddot\theta^b + (g_{\alpha\gamma}+g_{\alpha a}\tau^a_\gamma+g_{\gamma a}\tau^a_\alpha+g_{ab}\tau^a_\alpha\tau^b_\gamma+\sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_{,\beta}\ddot x^\gamma\\
& -(g_{\beta b}+g_{ab}\tau^a_\beta)_{,\alpha\gamma}\dot\theta^b\dot x^\gamma - (g_{\beta\nu}+g_{\beta a}\tau^a_\nu+g_{\nu a}\tau^a_\beta+g_{ab}\tau^a_\beta\tau^b_\nu+\sigma_{ab}\tau^a_\beta\tau^b_\nu)_{,\alpha\gamma}\dot x^\nu\dot x^\gamma\\
& -(g_{\beta b}+g_{ab}\tau^a_\beta)_{,\alpha}\ddot\theta^b - (g_{\beta\gamma}+g_{\beta a}\tau^a_\gamma+g_{\gamma a}\tau^a_\beta+g_{ab}\tau^a_\beta\tau^b_\gamma+\sigma_{ab}\tau^a_\beta\tau^b_\gamma)_{,\alpha}\ddot x^\gamma\\
& -(g_{\alpha\nu,\eta\beta}\dot x^\eta\dot x^\nu + g_{\alpha a,\nu\beta}\dot x^\nu\dot\theta^a + g_{\alpha\nu,\beta}\ddot x^\nu + g_{\alpha a,\beta}\ddot\theta^a)
\end{aligned}
$$

$$
\begin{aligned}
&+(g_{\beta\nu,\eta\alpha}\dot{x}^\eta\dot{x}^\nu + g_{\beta a,\nu\alpha}\dot{x}^\nu\dot{\theta}^a + g_{\beta\nu,\alpha}\ddot{x}^\nu + g_{\beta a,\alpha}\ddot{\theta}^a)\\
&-\tau^c_\alpha(g_{\nu c,\gamma\beta} + (g_{cb}\tau^b_\nu)_{,\beta\gamma})\dot{x}^\nu\dot{x}^\gamma - \tau^c_\alpha g_{cb,\gamma\beta}\dot{x}^\gamma\dot{\theta}^b\\
&-\tau^c_\alpha(g_{\gamma c,\beta} + g_{cb,\beta}\tau^b_\gamma + g_{cb}\tau^b_{\gamma,\beta})\ddot{x}^\gamma - \tau^c_\alpha g_{cb,\beta}\ddot{\theta}^b\\
&+\tau^c_\beta(g_{\nu c,\gamma\alpha} + (g_{cb}\tau^b_\nu)_{,\alpha\gamma})\dot{x}^\nu\dot{x}^\gamma + \tau^c_\beta g_{cb,\gamma\alpha}\dot{x}^\gamma\dot{\theta}^b\\
&+\tau^c_\beta(g_{\gamma c,\alpha} + g_{cb,\alpha}\tau^b_\gamma + g_{cb}\tau^b_{\gamma,\alpha})\ddot{x}^\gamma + \tau^c_\beta g_{cb,\alpha}\ddot{\theta}^b\,.
\end{aligned}
$$

The $\ddot{\theta}$ component becomes

$$
(g_{\alpha b} + g_{cb}\tau^c_\alpha)_{,\beta}\ddot{\theta}^b - (g_{\beta b} + g_{cb}\tau^c_\beta)_{,\alpha}\ddot{\theta}^b - g_{\alpha b,\beta}\ddot{\theta}^b + g_{\beta b,\alpha}\ddot{\theta}^b - \tau^c_\alpha g_{cb,\beta}\ddot{\theta}^b + \tau^c_\beta g_{cb,\alpha}\ddot{\theta}^b = g_{cb}(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha})\ddot{\theta}^b\,.
$$

The $\ddot{x}$ component becomes

$$
\begin{aligned}
&(g_{\alpha\gamma} + g_{\alpha a}\tau^a_\gamma + g_{\gamma a}\tau^a_\alpha + g_{ab}\tau^a_\alpha\tau^b_\gamma + \sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_{,\beta}\ddot{x}^\gamma - (g_{\beta\gamma} + g_{\beta a}\tau^a_\gamma + g_{\gamma a}\tau^a_\beta + g_{ab}\tau^a_\beta\tau^b_\gamma + \sigma_{ab}\tau^a_\beta\tau^b_\gamma)_{,\alpha}\ddot{x}^\gamma\\
&-g_{\alpha\gamma,\beta}\ddot{x}^\gamma + g_{\beta\gamma,\alpha}\ddot{x}^\gamma - \tau^c_\alpha(g_{\gamma c,\beta} + g_{cb,\beta}\tau^b_\gamma + g_{cb}\tau^b_{\gamma,\beta})\ddot{x}^\gamma + \tau^c_\beta(g_{\gamma c,\alpha} + g_{cb,\alpha}\tau^b_\gamma + g_{cb}\tau^b_{\gamma,\alpha})\ddot{x}^\gamma\\
=\;& g_{\alpha a,\beta}\tau^a_\gamma + g_{\alpha a}\tau^a_{\gamma,\beta} + g_{\gamma a}\tau^a_{\alpha,\beta} + g_{ab}\tau^b_\gamma\tau^a_{\alpha,\beta} + (\sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_\beta\\
&-g_{\beta a,\alpha}\tau^a_\gamma - g_{\beta a}\tau^a_{\gamma,\alpha} - g_{\gamma a}\tau^a_{\beta,\alpha} - g_{ab}\tau^b_\gamma\tau^a_{\beta,\alpha} - (\sigma_{ab}\tau^a_\beta\tau^b_\gamma)_\alpha\\
=\;& (g_{\alpha a}\tau^a_\gamma)_\beta + g_{\gamma a}\tau^a_{\alpha,\beta} + g_{ab}\tau^b_\gamma\tau^a_{\alpha,\beta} + (\sigma_{ab}\tau^a_\alpha\tau^b_\gamma)_\beta - (g_{\beta a}\tau^a_\gamma)_\alpha - g_{\gamma a}\tau^a_{\beta,\alpha} - g_{ab}\tau^b_\gamma\tau^a_{\beta,\alpha} - (\sigma_{ab}\tau^a_\beta\tau^b_\gamma)_\alpha\\
\overset{\text{M1}}{=}\;& (g_{\gamma c} + g_{cb}\tau^b_\gamma)(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha})\ddot{x}^\gamma\,.
\end{aligned}
$$

The $\dot{\theta}\dot{x}$ component becomes

$$
\begin{aligned}
&(g_{\alpha b} + g_{ab}\tau^a_\alpha)_{,\beta\gamma}\dot{\theta}^b\dot{x}^\gamma - (g_{\beta b} + g_{ab}\tau^a_\beta)_{,\alpha\gamma}\dot{\theta}^b\dot{x}^\gamma - g_{\alpha b,\gamma\beta}\dot{x}^\gamma\dot{\theta}^b\\
&+g_{\beta b,\gamma\alpha}\dot{x}^\gamma\dot{\theta}^b - \tau^c_\alpha g_{cb,\gamma\beta}\dot{x}^\gamma\dot{\theta}^b + \tau^c_\beta g_{cb,\gamma\alpha}\dot{x}^\gamma\dot{\theta}^b\\
=\;& (g_{ab}\tau^a_\alpha)_{,\beta\gamma}\dot{\theta}^b\dot{x}^\gamma - (g_{ab}\tau^a_\beta)_{,\alpha\gamma}\dot{\theta}^b\dot{x}^\gamma - \tau^c_\alpha g_{cb,\gamma\beta}\dot{x}^\gamma\dot{\theta}^b + \tau^c_\beta g_{cb,\gamma\alpha}\dot{x}^\gamma\dot{\theta}^b\\
=\;& \big(g_{ca,\gamma}(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha}) + (g_{ca}\tau^c_{\alpha,\gamma})_{,\beta} - (g_{ca}\tau^c_{\beta,\gamma})_{,\alpha}\big)\,\dot{\theta}^a\dot{x}^\gamma\\
\overset{\text{M3}}{=}\;& g_{ca,\gamma}(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha})\dot{\theta}^a\dot{x}^\gamma\,,
\end{aligned}
$$

where in the last equality we have used M3 in the following way:

$$
\begin{aligned}
-(g_{ca}\tau^c_{\beta,\gamma})_{,\alpha} + (g_{ca}\tau^c_{\alpha,\gamma})_{,\beta} \overset{\text{M3}}{=}\;& -(g_{ca}\tau^c_{\gamma,\beta} + g_{ca}g^{cd}g_{ed,\beta}\tau^e_\gamma)_{,\alpha} + (g_{ca}\tau^c_{\gamma,\alpha} + g_{ca}g^{cd}g_{ed,\alpha}\tau^e_\gamma)_{,\beta}\\
=\;& -(g_{ca,\alpha}\tau^c_{\gamma,\beta} + g_{ca}\tau^c_{\gamma,\beta\alpha} + g_{ea,\beta\alpha}\tau^e_\gamma + g_{ea,\beta}\tau^e_{\gamma,\alpha})\\
&+g_{ca,\beta}\tau^c_{\gamma,\alpha} + g_{ca}\tau^c_{\gamma,\alpha\beta} + g_{ea,\alpha\beta}\tau^e_\gamma + g_{ea,\alpha}\tau^e_{\gamma,\beta} = 0\,.
\end{aligned}
$$

Finally the $\dot{x}\dot{x}$ component becomes

$$
\begin{aligned}
&(g_{\alpha\nu} + g_{\alpha a}\tau^a_\nu + g_{\nu a}\tau^a_\alpha + g_{ab}\tau^a_\alpha\tau^b_\nu + \sigma_{ab}\tau^a_\alpha\tau^b_\nu)_{,\beta\gamma}\dot{x}^\nu\dot{x}^\gamma\\
&-(g_{\beta\nu} + g_{\beta a}\tau^a_\nu + g_{\nu a}\tau^a_\beta + g_{ab}\tau^a_\beta\tau^b_\nu + \sigma_{ab}\tau^a_\beta\tau^b_\nu)_{,\alpha\gamma}\dot{x}^\nu\dot{x}^\gamma\\
&-g_{\alpha\nu,\gamma\beta}\dot{x}^\gamma\dot{x}^\nu + g_{\beta\nu,\gamma\alpha}\dot{x}^\gamma\dot{x}^\nu - \tau^c_\alpha(g_{\nu c,\gamma\beta} + (g_{cb}\tau^b_\nu)_{,\beta\gamma})\dot{x}^\nu\dot{x}^\gamma + \tau^c_\beta(g_{\nu c,\gamma\alpha} + (g_{cb}\tau^b_\nu)_{,\alpha\gamma})\dot{x}^\nu\dot{x}^\gamma\\
=\;& (g_{ab}\tau^b_\nu)_{,\beta}\tau^a_{\alpha,\gamma} + (g_{ab}\tau^b_\nu)_{,\gamma}\tau^a_{\alpha,\beta} + g_{ab}\tau^b_\nu\tau^a_{\alpha,\beta\gamma} + (g_{\alpha a}\tau^a_\nu)_{,\beta\gamma} + g_{\nu a,\beta}\tau^a_{\alpha,\gamma} + g_{\nu a,\gamma}\tau^a_{\alpha,\beta} + g_{\nu a}\tau^a_{\alpha,\beta\gamma}\\
&+(\sigma_{ab}\tau^a_\alpha\tau^b_\nu)_{,\beta\gamma} - (g_{ab}\tau^b_\nu)_{,\alpha}\tau^a_{\beta,\gamma} - (g_{ab}\tau^b_\nu)_{,\gamma}\tau^a_{\beta,\alpha} - g_{ab}\tau^b_\nu\tau^a_{\beta,\alpha\gamma} - (g_{\beta a}\tau^a_\nu)_{,\alpha\gamma} - g_{\nu a,\alpha}\tau^a_{\beta,\gamma} - g_{\nu a,\gamma}\tau^a_{\beta,\alpha}
\end{aligned}
$$

$$
\begin{aligned}
& -g_{\nu a}\tau^a_{\beta,\alpha\gamma} - (\sigma_{ab}\tau^a_\beta\tau^b_\nu)_{,\alpha\gamma} \\
= \;& g_{ab}\tau^b_\nu\tau^a_{\alpha,\beta\gamma} - g_{ab}\tau^b_\nu\tau^a_{\beta,\alpha\gamma} + g_{\nu a,\beta}\tau^a_{\alpha,\gamma} + g_{\nu a,\gamma}\tau^a_{\alpha,\beta} + g_{\nu a}\tau^a_{\alpha,\beta\gamma} - g_{\nu a,\alpha}\tau^a_{\beta,\gamma} - g_{\nu a,\gamma}\tau^a_{\beta,\alpha} - g_{\nu a}\tau^a_{\beta,\alpha\gamma} \\
& +(g_{ab}\tau^b_\nu)_{,\beta}\tau^a_{\alpha,\gamma} + (g_{ab}\tau^b_\nu)_{,\gamma}\tau^a_{\alpha,\beta} - (g_{ab}\tau^b_\nu)_{,\alpha}\tau^a_{\beta,\gamma} - (g_{ab}\tau^b_\nu)_{,\gamma}\tau^a_{\beta,\alpha} \\
= \;& (g_{\nu a,\gamma} + (g_{ab}\tau^b_\nu)_{,\gamma})(\tau^a_{\alpha,\beta} - \tau^a_{\beta,\alpha}) \\
& +g_{ab}\tau^b_\nu\tau^a_{\alpha,\beta\gamma} - g_{ab}\tau^b_\nu\tau^a_{\beta,\alpha\gamma} + g_{\nu a,\beta}\tau^a_{\alpha,\gamma} + g_{\nu a}\tau^a_{\alpha,\beta\gamma} - g_{\nu a,\alpha}\tau^a_{\beta,\gamma} - g_{\nu a}\tau^a_{\beta,\alpha\gamma} \\
& +(g_{ab}\tau^b_\nu)_{,\beta}\tau^a_{\alpha,\gamma} - (g_{ab}\tau^b_\nu)_{,\alpha}\tau^a_{\beta,\gamma}\,,
\end{aligned}
$$

and adding all of the components we get

$$
\begin{aligned}
(2.30)_{\alpha\beta} \;=\; & g_{cb}(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha})\ddot\theta^b + (g_{\gamma c} + g_{cb}\tau^b_\gamma)(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha})\ddot x^\gamma + g_{ca,\gamma}(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha})\dot\theta^a\dot x^\gamma \\
& +(g_{\nu a,\gamma} + (g_{ab}\tau^b_\nu)_{,\gamma})(\tau^a_{\alpha,\beta} - \tau^a_{\beta,\alpha})\dot x^\nu\dot x^\gamma \\
& +(g_{ab}\tau^b_\nu\tau^a_{\alpha,\beta\gamma} - g_{ab}\tau^b_\nu\tau^a_{\beta,\alpha\gamma} + g_{\nu a,\beta}\tau^a_{\alpha,\gamma} + g_{\nu a}\tau^a_{\alpha,\beta\gamma} - g_{\nu a,\alpha}\tau^a_{\beta,\gamma} - g_{\nu a}\tau^a_{\beta,\alpha\gamma} \\
& +(g_{ab}\tau^b_\nu)_{,\beta}\tau^a_{\alpha,\gamma} - (g_{ab}\tau^b_\nu)_{,\alpha}\tau^a_{\beta,\gamma})\dot x^\nu\dot x^\gamma \\
=\; & \tilde\Phi_c(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha}) + (g_{ab}\tau^b_\nu\tau^a_{\alpha,\beta\gamma} - g_{ab}\tau^b_\nu\tau^a_{\beta,\alpha\gamma} + g_{\nu a,\beta}\tau^a_{\alpha,\gamma} + g_{\nu a}\tau^a_{\alpha,\beta\gamma} - g_{\nu a,\alpha}\tau^a_{\beta,\gamma} - g_{\nu a}\tau^a_{\beta,\alpha\gamma} \\
& +(g_{ab}\tau^b_\nu)_{,\beta}\tau^a_{\alpha,\gamma} - (g_{ab}\tau^b_\nu)_{,\alpha}\tau^a_{\beta,\gamma})\dot x^\nu\dot x^\gamma =: \tilde\Phi_c(\tau^c_{\alpha,\beta} - \tau^c_{\beta,\alpha}) + R\dot x^\nu\dot x^\gamma\,.
\end{aligned}
$$

Next we will show that the term $R\dot x^\nu\dot x^\gamma$ vanishes using M1, M2 and M3. First we compute some expressions that we will need. We will consistently omit writing the term $\dot x^\nu\dot x^\gamma$ but will take it into account and cancel any symmetric terms in ν and γ. From M2 we get $\sigma_{ab,\beta} = -g_{ab,\beta} + 2\sigma_{be}g^{ed}g_{ad,\beta}$ and therefore

$$(g_{ab} - \sigma_{ab})_{,\beta} = 2g_{ab,\beta} - 2\sigma_{be}g^{ed}g_{ad,\beta}\,. \tag{2.44}$$

Using $g^{da}_{,\alpha} = -g^{ea}g^{dh}g_{eh,\alpha}$ we get

$$
\begin{aligned}
(g^{da}g_{ed,\alpha}\tau^e_\gamma)_{,\beta} - (g^{da}g_{ed,\beta}\tau^e_\gamma)_{,\alpha} \;=\; & g^{da}(g_{ed,\alpha\beta}\tau^e_\gamma + g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta\alpha}\tau^e_\gamma - g_{ed,\beta}\tau^e_{\gamma,\alpha}) \\
& +g^{da}_{,\beta}g_{ed,\alpha}\tau^e_\gamma - g^{da}_{,\alpha}g_{ed,\beta}\tau^e_\gamma \\
=\; & g^{da}(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - g^{ea}g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma \\
& +g^{ea}g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma\,,
\end{aligned}
$$

and therefore we have

$$
\begin{aligned}
& (g_{ab} - \sigma_{ab})\tau^b_\nu\left((g^{da}g_{ed,\alpha}\tau^e_\gamma)_{,\beta} - (g^{da}g_{ed,\beta}\tau^e_\gamma)_{,\alpha}\right) \\
= \;& (g_{ab} - \sigma_{ab})\tau^b_\nu\left(g^{da}(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - g^{ea}g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma + g^{ea}g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma\right) \\
= \;& g_{ab}\tau^b_\nu g^{da}(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - g_{ab}\tau^b_\nu g^{ea}g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma + g_{ab}\tau^b_\nu g^{ea}g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma \\
& -\sigma_{ab}\tau^b_\nu(g^{da}(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - g^{ea}g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma + g^{ea}g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma) \\
= \;& \tau^d_\nu(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - \tau^e_\nu g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma + \tau^e_\nu g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma \\
& -\sigma_{ab}\tau^b_\nu(g^{da}(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - g^{ea}g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma + g^{ea}g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma) \\
= \;& \tau^d_\nu(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - \sigma_{ab}\tau^b_\nu g^{da}g_{ed,\alpha}\tau^e_{\gamma,\beta} + \sigma_{ab}\tau^b_\nu g^{da}g_{ed,\beta}\tau^e_{\gamma,\alpha}
\end{aligned}
$$

$$
\begin{aligned}
&+\sigma_{ab}\tau^b_\nu g^{ea}g^{dh}g_{eh,\beta}g_{kd,\alpha}\tau^k_\gamma - \sigma_{ab}\tau^b_\nu g^{ea}g^{dh}g_{eh,\alpha}g_{kd,\beta}\tau^k_\gamma \\
\overset{M2}{=}\;& \tau^d_\nu(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - \frac{1}{2}\sigma_{ab}\tau^b_\nu\sigma^{da}(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta} + \frac{1}{2}\sigma_{ab}\tau^b_\nu\sigma^{da}(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha} \\
&+\frac{1}{4}\sigma_{ab}\tau^b_\nu\sigma^{ea}(\sigma_{eh,\beta}+g_{eh,\beta})\sigma^{dh}(\sigma_{kd,\alpha}+g_{kd,\alpha})\tau^k_\gamma - \frac{1}{4}\sigma_{ab}\tau^b_\nu\sigma^{ea}(\sigma_{eh,\alpha}+g_{eh,\alpha})\sigma^{dh}(\sigma_{kd,\beta}+g_{kd,\beta})\tau^k_\gamma \\
=\;& \tau^d_\nu(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - \frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta} + \frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha} \\
&+\frac{1}{4}\tau^e_\nu(\sigma_{eh,\beta}+g_{eh,\beta})\sigma^{dh}(\sigma_{kd,\alpha}+g_{kd,\alpha})\tau^k_\gamma - \frac{1}{4}\tau^e_\nu(\sigma_{eh,\alpha}+g_{eh,\alpha})\sigma^{dh}(\sigma_{kd,\beta}+g_{kd,\beta})\tau^k_\gamma \\
=\;& \tau^d_\nu(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - \frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta} + \frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha}\,. \qquad (2.45)
\end{aligned}
$$

We will also use

$$
\begin{aligned}
&-2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu g^{ha}g_{kh,\alpha}\tau^k_\gamma + 2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu g^{ha}g_{kh,\beta}\tau^k_\gamma \\
\overset{M2}{=}\;& -\frac{1}{2}\sigma_{be}\sigma^{ed}(\sigma_{ad,\beta}+g_{ad,\beta})\tau^b_\nu\sigma^{ha}(\sigma_{kh,\alpha}+g_{kh,\alpha})\tau^k_\gamma + \frac{1}{2}\sigma_{be}\sigma^{ed}(\sigma_{ad,\alpha}+g_{ad,\alpha})\tau^b_\nu\sigma^{ha}(\sigma_{kh,\beta}+g_{kh,\beta})\tau^k_\gamma \\
=\;& -\frac{1}{2}(\sigma_{ab,\beta}+g_{ab,\beta})\tau^b_\nu\sigma^{ha}(\sigma_{kh,\alpha}+g_{kh,\alpha})\tau^k_\gamma + \frac{1}{2}(\sigma_{ab,\alpha}+g_{ab,\alpha})\tau^b_\nu\sigma^{ha}(\sigma_{kh,\beta}+g_{kh,\beta})\tau^k_\gamma = 0\,. \qquad (2.46)
\end{aligned}
$$

Now we will finally check that R vanishes using M1, M2 and M3. Recall that in the computation below we omit the term $\dot{x}^\nu\dot{x}^\gamma$:

$$
\begin{aligned}
R \;=\;& ((g_{ab}\tau^b_\nu + g_{\nu a})\tau^a_{\alpha,\gamma})_{,\beta} - ((g_{ab}\tau^b_\nu + g_{\nu a})\tau^a_{\beta,\gamma})_{,\alpha} \\
\overset{M1}{=}\;& ((g_{ab}-\sigma_{ab})\tau^b_\nu\tau^a_{\alpha,\gamma})_{,\beta} - ((g_{ab}-\sigma_{ab})\tau^b_\nu\tau^a_{\beta,\gamma})_{,\alpha} \\
\overset{M3}{=}\;& ((g_{ab}-\sigma_{ab})\tau^b_\nu(\tau^a_{\gamma,\alpha} + g^{da}g_{ed,\alpha}\tau^e_\gamma))_{,\beta} - ((g_{ab}-\sigma_{ab})\tau^b_\nu(\tau^a_{\gamma,\beta} + g^{da}g_{ed,\beta}\tau^e_\gamma))_{,\alpha} \\
=\;& (g_{ab}-\sigma_{ab})_{,\beta}\tau^b_\nu(\tau^a_{\gamma,\alpha} + g^{da}g_{ed,\alpha}\tau^e_\gamma) + (g_{ab}-\sigma_{ab})\tau^b_{\nu,\beta}(\tau^a_{\gamma,\alpha} + g^{da}g_{ed,\alpha}\tau^e_\gamma) \\
&+(g_{ab}-\sigma_{ab})\tau^b_\nu(\tau^a_{\gamma,\alpha} + g^{da}g_{ed,\alpha}\tau^e_\gamma)_{,\beta} - (g_{ab}-\sigma_{ab})_{,\alpha}\tau^b_\nu(\tau^a_{\gamma,\beta} + g^{da}g_{ed,\beta}\tau^e_\gamma) \\
&-(g_{ab}-\sigma_{ab})\tau^b_{\nu,\alpha}(\tau^a_{\gamma,\beta} + g^{da}g_{ed,\beta}\tau^e_\gamma) - (g_{ab}-\sigma_{ab})\tau^b_\nu(\tau^a_{\gamma,\beta} + g^{da}g_{ed,\beta}\tau^e_\gamma)_{,\alpha} \\
=\;& (g_{ab}-\sigma_{ab})_{,\beta}\tau^b_\nu(\tau^a_{\gamma,\alpha} + g^{da}g_{ed,\alpha}\tau^e_\gamma) + (g_{ab}-\sigma_{ab})\tau^b_{\nu,\beta}(g^{da}g_{ed,\alpha}\tau^e_\gamma) \\
&+(g_{ab}-\sigma_{ab})\tau^b_\nu(g^{da}g_{ed,\alpha}\tau^e_\gamma)_{,\beta} - (g_{ab}-\sigma_{ab})_{,\alpha}\tau^b_\nu(\tau^a_{\gamma,\beta} + g^{da}g_{ed,\beta}\tau^e_\gamma) \\
&-(g_{ab}-\sigma_{ab})\tau^b_{\nu,\alpha}(g^{da}g_{ed,\beta}\tau^e_\gamma) - (g_{ab}-\sigma_{ab})\tau^b_\nu(g^{da}g_{ed,\beta}\tau^e_\gamma)_{,\alpha} \\
\overset{(2.44),(2.45)}{=}\;& 2g_{ab,\beta}\tau^b_\nu\tau^a_{\gamma,\alpha} - 2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu\tau^a_{\gamma,\alpha} + 2g_{ab,\beta}\tau^b_\nu g^{da}g_{ed,\alpha}\tau^e_\gamma - 2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu g^{da}g_{ed,\alpha}\tau^e_\gamma \\
&-2g_{ab,\alpha}\tau^b_\nu\tau^a_{\gamma,\beta} + 2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu\tau^a_{\gamma,\beta} - 2g_{ab,\alpha}\tau^b_\nu g^{da}g_{ed,\beta}\tau^e_\gamma + 2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu g^{da}g_{ed,\beta}\tau^e_\gamma \\
&+\tau^d_{\nu,\beta}g_{ed,\alpha}\tau^e_\gamma - \sigma_{ab}\tau^b_{\nu,\beta}g^{da}g_{ed,\alpha}\tau^e_\gamma - \tau^d_{\nu,\alpha}g_{ed,\beta}\tau^e_\gamma + \sigma_{ab}\tau^b_{\nu,\alpha}g^{da}g_{ed,\beta}\tau^e_\gamma \\
&+\tau^d_\nu(g_{ed,\alpha}\tau^e_{\gamma,\beta} - g_{ed,\beta}\tau^e_{\gamma,\alpha}) - \frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta} + \frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha} \\
=\;& -2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu\tau^a_{\gamma,\alpha} + 2g_{ab,\beta}\tau^b_\nu g^{da}g_{ed,\alpha}\tau^e_\gamma - 2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu g^{da}g_{ed,\alpha}\tau^e_\gamma \\
&+2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu\tau^a_{\gamma,\beta} - 2g_{ab,\alpha}\tau^b_\nu g^{da}g_{ed,\beta}\tau^e_\gamma + 2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu g^{da}g_{ed,\beta}\tau^e_\gamma \\
&-\sigma_{ab}\tau^b_{\nu,\beta}g^{da}g_{ed,\alpha}\tau^e_\gamma + \sigma_{ab}\tau^b_{\nu,\alpha}g^{da}g_{ed,\beta}\tau^e_\gamma - \frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta} + \frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha} \\
=\;& -2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu\tau^a_{\gamma,\alpha} - 2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu g^{da}g_{ed,\alpha}\tau^e_\gamma + 2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu\tau^a_{\gamma,\beta} \\
&+2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu g^{da}g_{ed,\beta}\tau^e_\gamma - \sigma_{ab}\tau^b_{\nu,\beta}g^{da}g_{ed,\alpha}\tau^e_\gamma + \sigma_{ab}\tau^b_{\nu,\alpha}g^{da}g_{ed,\beta}\tau^e_\gamma
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta}+\frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha}\\
\overset{(2.46)}{=}\quad &-2\sigma_{be}g^{ed}g_{ad,\beta}\tau^b_\nu\tau^a_{\gamma,\alpha}+2\sigma_{be}g^{ed}g_{ad,\alpha}\tau^b_\nu\tau^a_{\gamma,\beta}-\sigma_{ab}\tau^b_{\nu,\beta}(g^{da}g_{ed,\alpha}\tau^e_\gamma)+\sigma_{ab}\tau^b_{\nu,\alpha}(g^{da}g_{ed,\beta}\tau^e_\gamma)\\
&-\frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta}+\frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha}\\
\overset{\mathrm{M2}}{=}\quad &-\sigma_{be}\sigma^{ed}(\sigma_{ad,\beta}+g_{ad,\beta})\tau^b_\nu\tau^a_{\gamma,\alpha}+\sigma_{be}\sigma^{ed}(\sigma_{ad,\alpha}+g_{ad,\alpha})\tau^b_\nu\tau^a_{\gamma,\beta}\\
&-\frac{1}{2}\sigma_{ab}\tau^b_{\nu,\beta}\sigma^{da}(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_\gamma+\frac{1}{2}\sigma_{ab}\tau^b_{\nu,\alpha}\sigma^{da}(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_\gamma\\
&-\frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta}+\frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha}\\
=\quad &-(\sigma_{ab,\beta}+g_{ab,\beta})\tau^b_\nu\tau^a_{\gamma,\alpha}+(\sigma_{ab,\alpha}+g_{ab,\alpha})\tau^b_\nu\tau^a_{\gamma,\beta}-\frac{1}{2}\tau^b_{\nu,\beta}(\sigma_{eb,\alpha}+g_{eb,\alpha})\tau^e_\gamma\\
&+\frac{1}{2}\tau^b_{\nu,\alpha}(\sigma_{eb,\beta}+g_{eb,\beta})\tau^e_\gamma-\frac{1}{2}\tau^d_\nu(\sigma_{ed,\alpha}+g_{ed,\alpha})\tau^e_{\gamma,\beta}+\frac{1}{2}\tau^d_\nu(\sigma_{ed,\beta}+g_{ed,\beta})\tau^e_{\gamma,\alpha}=0\,.
\end{aligned}
$$

2.4.2 Arbitrary dimension under assumptions SM1, SM2 and one degree of underactuation

Recall that as mentioned at the beginning of Section 2.4, when solving the Helmholtz conditions we have used the Legendre transformation of the controlled Lagragian $L_{\tau,\sigma}$. The Helmholtz conditions guarantee the existence of a Lagrangian with the same Legendre transformation as the controlled Lagrangian which we have used but the potential energy terms need not coincide. Therefore we consider controlled Lagrangians of the form $\tilde{L}_{\tau,\sigma} = K_{\tau,\sigma} - \tilde{V}(x^\alpha, \theta^a)$ with arbitrary $\tilde{V}$.

Since in this section we deal with systems with one degree of underactuation, we will now use the notation $\tau^c_\alpha = \tau^c_1 =: \tau^c$ and also use a $'$ intead of $_{,1}$ to denote derivative with respect to $x^1 =: x$.

Theorem 2.4.1. *Under assumptions SM1, SM2 and one degree of underactuation, there is a controlled Lagrangian $\tilde{L}_{\tau,\sigma}$ such that the Euler-Lagrange equations for $\tilde{L}_{\tau,\sigma}$ are equivalent to the controlled Euler-Lagrange equations for L if τ^a satisfies the ODE system*

$$
2\tau^a g_{1e}g^{ec}g'_{1c}+2\tau^a g_{1e}(\tau^e)'-\tau^a g'_{11}+2g_{11}(\tau^a)'-2g_{1c}g^{dc}g_{d1}(\tau^a)'-2g_{1c}\tau^c(\tau^a)'=0\,, \tag{2.47}
$$

for all $a = 2, \ldots, n$. In the particular case when $dim(Q) = 2$ we obtain the new solution

$$
\tau(x) = k\sqrt{g_{11}(x)g_{22}-g_{12}(x)^2}\,, \tag{2.48}
$$

where k is an arbitrary constant. Notice that one degree of underactuation implies that SM4 holds and therefore we are providing an alternative to the solution given by SM3.

Proof. From the previous computations we can see that Equation (2.30) vanishes identically since the assumption of one degree of underactuation implies that Equation $(2.30)_{\alpha\beta}$ is void and also that Equation (2.29) vanishes for indices ab, $a\beta$, and αb. Now under assumption SM1, that is $\sigma_{ab} = \sigma g_{ab}$ for some constant σ, we compute Equation (2.29) for indices $\alpha\beta$, which are just $\alpha = \beta = 1$. This gives an ODE system as an extra solution, alternative to M1.

Indeed, the $\dot\theta$ components (2.42) vanish identically using that (g_{ab}) is constant and $\alpha = \beta = \eta$. Imposing that the $\dot x$ component vanishes we get

$$\begin{aligned}
0 &= -\left(\tau^c - (g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}g_{1e}g^{ec}\right)\left(g'_{1c} + g'_{1c} + (g_{cd}\tau^d)' + (g_{cd}\tau^d)'\right)\dot x \\
&\quad -\left(\delta^1_1 + (g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}\right)(g'_{11} + g'_{11} - g'_{11})\dot x \\
&\quad +(g_{11} + g_{1d}\tau^d + g_{1d}\tau^d + g_{de}\tau^d\tau^e + \sigma_{ed}\tau^e\tau^d)'\dot x \\
&= -\left(\tau^c - (g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}g_{1e}g^{ec}\right)\left(2g'_{1c} + 2(g_{cd}\tau^d)'\right)\dot x \\
&\quad -g'_{11}\dot x - \left((g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}\right)g'_{11}\dot x + g'_{11}\dot x + (2g_{1d}\tau^d + g_{de}\tau^d\tau^e + \sigma_{ed}\tau^e\tau^d)'\dot x \\
&= -\tau^c\left(2g_{1c,1} + 2(g_{cd}\tau^d)'\right)\dot x + (g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}g_{1e}g^{ec}\left(2g'_{1c} + 2(g_{cd}\tau^d)'\right)\dot x \\
&\quad -\left((g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}\right)g'_{11}\dot x + (2g_{1d}\tau^d + g_{de}\tau^d\tau^e + \sigma_{ed}\tau^e\tau^d)'\dot x \\
&= -\tau^c\left(2g'_{1c} + 2(g_{cd}\tau^d)'\right)\dot x + (g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}g_{1e}g^{ec}\left(2g'_{1c} + 2(g_{cd}\tau^d)'\right)\dot x \\
&\quad -\left((g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}\right)g'_{11}\dot x + (2g'_{1d}\tau^d + 2g_{1d}(\tau^d)' + g'_{de}\tau^d\tau^e + g_{de}(\tau^d)'\tau^e + g_{de}\tau^d(\tau^e)')\dot x \\
&\quad +(\sigma_{ed}\tau^e\tau^d)'\dot x \\
&= (g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}g_{1e}g^{ec}\left(2g'_{1c} + 2(g_{cd}\tau^d)'\right)\dot x \\
&\quad -\left((g_{1d}\tau^d + \sigma_{ad}\tau^a\tau^d)A^{11}\right)g'_{11}\dot x + (2g_{1d}(\tau^d)' + \sigma_{ed}(\tau^e)'\tau^d + \sigma_{ed}\tau^e(\tau^d)')\dot x \\
&= A^{11}\left((g_{1d} + \sigma_{ad}\tau^a)\tau^d\left(g_{1e}g^{ec}\left(2g'_{1c} + 2g_{cd}(\tau^d)'\right) - g'_{11}\right) + 2A_{11}(g_{1d}(\tau^d)' + \sigma_{ed}(\tau^e)'\tau^d)\right) \\
&= A^{11}\left((g_{1d} + \sigma_{ad}\tau^a)\tau^d\left(g_{1e}g^{ec}\left(2g'_{1c} + 2g_{cd}(\tau^d)'\right) - g'_{11}\right) + 2\left(g_{1d} + \sigma_{da}\tau^a\right)(\tau^d)'A_{11}\right) \\
&= A^{11}(g_{1d} + \sigma_{ad}\tau^a)\left(\tau^d\left(g_{1e}g^{ec}\left(2g'_{1c} + 2g_{cd}(\tau^d)'\right) - g'_{11}\right) + 2(\tau^d)'A_{11}\right) \\
&= A^{11}(g_{1d} + \sigma_{ad}\tau^a)\left(2\tau^d g_{1e}g^{ec}g'_{1c} + 2\tau^d g_{1e}(\tau^e)' - \tau^d g'_{11} + 2g_{11}(\tau^d)' - 2g_{1c}g^{ec}g_{e1}(\tau^d)' - 2g_{1c}\tau^c(\tau^d)'\right).
\end{aligned}$$

Therefore we have the two solutions M1 and

$$2\tau^a g_{1e}g^{ec}g'_{1c} + 2\tau^a g_{1e}(\tau^e)' - \tau^a g'_{11} + 2g_{11}(\tau^a)' - 2g_{1c}g^{dc}g_{d1}(\tau^a)' - 2g_{1c}\tau^c(\tau^a)' = 0\,,$$

for each $a = 2, \ldots, n$. Notice that in the case when $\dim(Q) = 2$ the system (2.47) becomes

$$g^{22}(2g_{12}g'_{12}\tau - g'_{11}g_{22}\tau + 2g_{11}g_{22}\tau' - 2g_{12}^2\tau') = 0\,,$$

and the solution is given by

$$\tau(x) = k\sqrt{g_{11}(x)g_{22} - g_{12}(x)^2}\,,$$

where k is an arbitrary constant. ■

Proposition 2.4.2. *Under the assumptions of Theorem 2.4.1 and using the new solution given by (2.47) we have that the control (2.32) is independent of velocities.*

Proof. Indeed we have that the equations $\tilde\Phi_1 = 0$ and $\tilde\Phi_a = 0$ are given by

$$g_{11}\ddot x + g_{1a}\ddot\theta^a = -\frac{1}{2}g'_{11}\dot x^2 - V'\,,$$

$$(g_{1a} + g_{ab}\tau^b)\ddot{x} + g_{ab}\ddot{\theta}^b \quad = \quad -(g'_{1a} + g_{ab}(\tau^b)')\dot{x}^2\,.$$

Therefore, since $\tilde{C}$ is regular we have

$$\ddot{x} = A^{11}\left(-\frac{1}{2}g'_{11}\dot{x}^2 + g_{1d}g^{de}g'_{1e}\dot{x}^2 + g_{1d}(\tau^d)'\dot{x}^2 - V'\right) \tag{2.49}$$

and the control (2.32) is given by

$$u_a = -g_{ab}(\tau^b)'\dot{x}^2 - g_{ab}\tau^b A^{11}\left(-\frac{1}{2}g'_{11}\dot{x}^2 + g_{1d}g^{de}g'_{1e}\dot{x}^2 + g_{1d}(\tau^d)'\dot{x}^2 - V'\right) = g_{ab}\tau^b A^{11}V'\,,$$

where in the last equality we have used that $A_{11} = g_{11} - g_{1f}g^{ef}(g_{e1} + g_{ed}\tau^d)$ is nonvanishing in order to get

$$\begin{aligned}
& -A_{11}g_{ab}(\tau^b)' - g_{ab}\tau^b\left(-\frac{1}{2}g'_{11} + g_{1d}g^{de}g'_{1e} + g_{1d}(\tau^d)'\right) \\
= \quad & -g_{11}g_{ab}(\tau^b)' + g_{1f}g^{ef}(g_{e1} + g_{ed}\tau^d)g_{ab}(\tau^b)' + \frac{1}{2}g'_{11}g_{ab}\tau^b - g_{1d}g^{de}g'_{1e}g_{ab}\tau^b - g_{1d}(\tau^d)'g_{ab}\tau^b \\
= \quad & g_{ab}\left(-g_{11}(\tau^b)' + g_{1f}g^{ef}g_{e1}(\tau^b)' + g_{1d}\tau^d(\tau^b)' + \frac{1}{2}g'_{11}\tau^b - \tau^b g_{1d}g^{de}g'_{1e} - g_{1d}\tau^b(\tau^d)'\right) \\
\overset{(2.47)}{=} \quad & 0\,.
\end{aligned}$$

■

Example 2.4.3 (Inverted pendulum on a cart)**.** Consider again the inverted pendulum on a cart introduced in Example 2.3.4.1. We will now provide a new stabilizing control using the solution provided by Theorem 2.4.1. We will now denote the coordinates of the system by (x, s) instead of (ϕ, s). The upright position of the pendulum corresponds to $x = 0$.

Recall that the Lagrangian is given by

$$L = \frac{1}{2}\left(\alpha\dot{x}^2 + 2\beta\dot{s}\dot{x}\cos(x) + \gamma\dot{s}^2\right) + d\cos(x)\,,$$

where $\alpha = ml^2, \beta = ml, \gamma = m + M$ and $d = -mgl$ are constants.

If we choose the solution provided by (2.48), that is,

$$\tau(x) = k\sqrt{\alpha\gamma - \beta^2\cos^2(x)}\,,$$

then we obtain the control

$$\begin{aligned}
u \quad &= \quad -g_{22}(x)\left(\dot{x}^2\tau'(x) + \tau(x)\ddot{x}\right) \\
&= \quad -\frac{\gamma k\left(\alpha\gamma\ddot{x} + \beta^2\cos(x)\left(\dot{x}^2\sin(x) - \ddot{x}\cos(x)\right)\right)}{\sqrt{\alpha\gamma - \beta^2\cos^2(x)}}\,.
\end{aligned}$$

The controlled Euler -Lagrange equations are

$$\begin{aligned}
\tilde{\Phi}_1 \quad &= \quad \alpha\ddot{x} + \beta\ddot{s}\cos(x) + d\sin(x) = 0\,, \\
\tilde{\Phi}_2 \quad &= \quad \frac{\gamma k\left(\alpha\gamma\ddot{x} + \beta^2\cos(x)\left(\dot{x}^2\sin(x) - \ddot{x}\cos(x)\right)\right)}{\sqrt{\alpha\gamma - \beta^2\cos^2(x)}}
\end{aligned}$$

$$+\beta\ddot{x}\cos(x) - \beta\dot{x}^2\sin(x) + \gamma\ddot{s} = 0\,.$$

Using the first equation to eliminate $\ddot{s}$ from the second one we obtain

$$\ddot{x}\left(\gamma k\sqrt{\alpha\gamma - \beta^2\cos^2(x)} - \frac{\alpha\gamma\sec(x)}{\beta} + \beta\cos(x)\right)$$
$$+\beta\dot{x}^2\sin(x)\left(\frac{\beta\gamma k\cos(x)}{\sqrt{\alpha\gamma - \beta^2\cos^2(x)}} - 1\right) - \frac{d\gamma\tan(x)}{\beta} = 0\,.$$

Using this last expression we can eliminate the accelerations from the control which becomes

$$u = -\frac{d\gamma^2 k\sin(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)}}{\beta\gamma k\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)} - \alpha\gamma + \beta^2\cos^2(x)}\,. \tag{2.50}$$

We will now check the stability of the upright position of the pendulum with this control. To this end we will use the energy function corresponding to the new Lagrangian $\tilde{L}_{\tau,\sigma}$ (with the same Legendre transformation as $L_{\tau,\sigma}$ but a possibly different potential energy term, as remarked above).

When written in explicit form, the controlled Euler-Lagrange equations become

$$\ddot{x} = \frac{\sin(x)\left(\frac{d\gamma\left(\beta^2\cos^2(x)-\alpha\gamma\right)}{-\beta\gamma k\cos(x)\sqrt{\alpha\gamma-\beta^2\cos^2(x)}+\alpha\gamma-\beta^2\cos^2(x)} - \beta^2\dot{x}^2\cos(x)\right)}{\alpha\gamma - \beta^2\cos^2(x)} =: F\,, \tag{2.51}$$

$$\ddot{s} = \sin(x)\left(-\frac{\alpha d\gamma^2 k}{\sqrt{\alpha\gamma - \beta^2\cos^2(x)}\left(\beta\gamma k\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)} - \alpha\gamma + \beta^2\cos^2(x)\right)}\right.$$
$$\left.+\frac{\beta d\cos(x)}{\alpha\gamma - \beta^2\cos^2(x)} + \frac{\alpha\beta\dot{x}^2}{\alpha\gamma - \beta^2\cos^2(x)}\right) =: G\,. \tag{2.52}$$

We can write the new Lagrangian as

$$\tilde{L}_{\tau,\sigma} = \frac{1}{2}\left(\tilde{g}_{11}(x)\dot{x}^2 + 2\tilde{g}_{12}(x)\dot{x}\dot{s} + \tilde{g}_{22}\dot{s}^2\right) - \tilde{V}(x,s)\,,$$

where

$$\begin{aligned}
\tilde{g}_{11}(x) &= \gamma k^2(\sigma+1)\left(\alpha\gamma - \beta^2\cos^2(x)\right) + 2\beta k\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)} + \alpha\,,\\
\tilde{g}_{12}(x) &= \gamma k\sqrt{\alpha\gamma - \beta^2\cos^2(x)} + \beta\cos(x)\,,\\
\tilde{g}_{22} &= \gamma\,.
\end{aligned}$$

Then the equivalence conditions (1.13) are

$$\begin{aligned}
-\tilde{g}_{11}F - \tilde{g}_{12}G &= \frac{\partial\tilde{g}_{11}}{\partial x}\dot{x}^2 + \frac{\partial\tilde{g}_{12}}{\partial x}\dot{s}\dot{x} - \left(\frac{1}{2}\frac{\partial\tilde{g}_{11}}{\partial x}\dot{x}^2 + \frac{\partial\tilde{g}_{12}}{\partial x}\dot{s}\dot{x} + \frac{\partial\tilde{V}}{\partial x}\right),\\
-\tilde{g}_{21}F - \tilde{g}_{22}G &= \frac{\partial\tilde{g}_{21}}{\partial x}\dot{x}^2 - \frac{\partial\tilde{V}}{\partial s}\,,
\end{aligned}$$

from where we get

$$\frac{\partial\tilde{V}}{\partial x} = -\frac{d\left(\gamma^2k^2\sigma + 1\right)\sin(x)\left(\alpha\gamma - \beta^2\cos^2(x)\right)}{\beta\gamma k\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)} - \alpha\gamma + \beta^2\cos^2(x)}\,,$$

$$\frac{\partial \tilde{V}}{\partial s} \quad = \quad 0\,.$$

Now we impose conditions so that the new multiplier matrix $(\tilde{g}_{ij})$ will be positive-definite. If we introduce the notation

$$D = g_{11}g_{22} - g_{12}^2 \qquad \text{and} \qquad \tilde{D} = \tilde{g}_{11}\tilde{g}_{22} - \tilde{g}_{12}^2\,,$$

then we have $\tilde{D} = D + \sigma(g_{22}\tau)^2$, and therefore we need to choose $\sigma > \frac{-D}{(g_{22}\tau)^2} = \frac{-1}{\gamma^2 k^2}$. We also need $\tilde{g}_{11} > 0$, for which it is enough to take $\tau > \frac{-2g_{12}}{g_{22}(1+\sigma)}$. Therefore it is enough to choose $\sigma > 0$ and $k > 0$.

On the other hand, looking at $\partial\tilde{V}/\partial x$, notice that we have

$$d < 0, \quad \alpha\gamma - \beta^2\cos^2(x) > 0 \quad \text{and} \quad \gamma^2k^2\sigma + 1 > 0$$

from the previous choice. Then, in order to get a positive-definite potential energy, we need to impose

$$\beta\gamma k\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)} - \alpha\gamma + \beta^2\cos^2(x) > 0\,,$$

which, taking $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, reduces to

$$k > \frac{\alpha\gamma - \beta^2\cos^2(x)}{\beta\gamma\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)}}\,, \tag{2.53}$$

(but stability of $x = 0$ is guaranteed with $k > \frac{\alpha\gamma - \beta^2}{\beta\gamma\sqrt{\alpha\gamma-\beta^2}}$).

Summing up, we can choose σ to guarantee that the new kinetic energy is positive-definite and we can choose the constant k in the control to guarantee that the potential energy is positive-definite. Then the energy is a Lyapunov function for (2.51)-(2.52). Notice that the requirement (2.53) corresponds to $A_{11} < 0$.

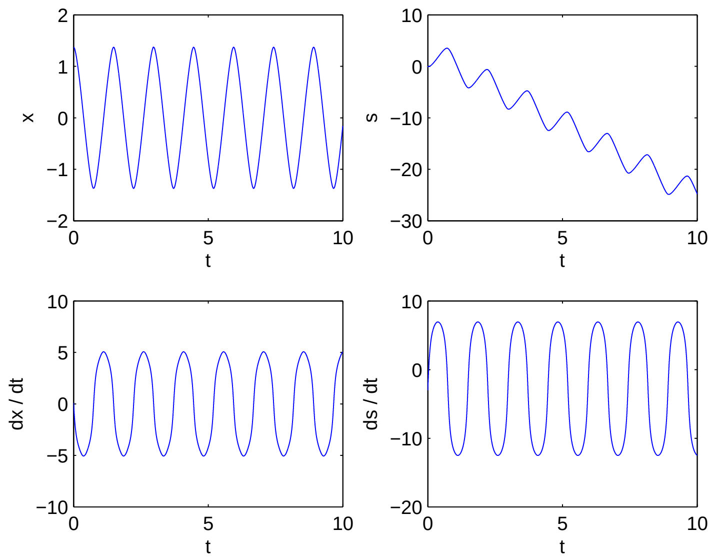


FIGURE 2.7: Simulations for the inverted pendulum on a cart.

We fix the parameters of the system to be $m = 0.14$ kg, $M = 0.44$ kg and $l = 0.215$ m as in [21] and take the initial conditions to be $\phi(0) = \pi/2 - 0.2$ rad, $\dot{\phi}(0) = 0.1$ rad/s, $s(0) = 0$ m, and $\dot{s}(0) = -3$ m/s, also as in [21]. In Figure 2.7 there is a MATLAB simulation of this situation with $k = 35$.

Remark 2.4.4. The system (2.51)-(2.52) fits into the class of systems dealt with in Section 2.3 and belongs to Case IIa1 from Douglas' classification since

$$\Phi_2^2 = d\gamma\left(2(\beta^2 - \alpha\gamma)\cos(x)\sqrt{\alpha\gamma - \beta^2\cos^2(x)} - 2\beta\gamma k(\alpha\gamma + \beta^2\cos^2(x))\right) \neq 0\,.$$

This is the same case as the controlled systems that appear in [21] and Section 2.3 for the example of the inverted pendulum on a cart.

Chapter 3

Inverse problem for SODEs on Lie algebroids

In the nineties of the last century two important contributions showed how Lie algebroids and Lie groupoids [106] are very useful in order to describe Lagrangian mechanics [103, 165]. From then on, the benefits of Lie algebroids to describe Lagrangian and Hamiltonian mechanics have become very clear in the literature, see [51, 113] and references therein. For instance, using the Atiyah algebroid framework, Lagrange-Poincaré and Hamilton-Poincaré equations are naturally obtained [90].

In this chapter we extend the use of Lagrangian submanifolds to geometrically characterize the inverse problem of the calculus of variations on regular Lie algebroids, giving a different approach from [131] and extending the results for Lie algebras described in [45]. On Lie algebroids the role of the SODE (second order differential equation) is played by a SODE section [51, 113]. Locally, the system $\ddot{x}^i = \Gamma^i(x, \dot{x})$ is replaced by

$$\dot{x}^i = \rho^i_\alpha(x) y^\alpha \quad \text{and} \quad \dot{y}^\alpha = \Gamma^\alpha(x, y) \,,$$

where (x^i, y^α) are local coordinates on a Lie algebroid E. The inverse problem on Lie algebroids poses the same question as the classical inverse problem [84, 148, 161]. When is the above system equivalent to the Euler-Lagrange equations for some regular Lagrangian? More precisely, when is it possible to find a nondegenerate matrix of multipliers $(g_{\alpha\beta}(x, y))$ such that

$$g_{\alpha\beta}(\dot{y}^\alpha - \Gamma^\alpha) = \frac{d}{dt}\left(\frac{\partial L}{\partial y^\beta}\right) - \rho^i_\beta \frac{\partial L}{\partial x^i} + C^\gamma_{\beta\nu} y^\nu \frac{\partial L}{\partial y^\gamma}$$

has a regular solution L?

Lie algebroids as a concept unifying tangent bundles and Lie algebras have deserved a lot of interest in the last years [103, 113, 163]. For instance, many mechanical systems are not defined on tangent bundles but on quotients by a symmetry Lie group, and then the equations of motion are not the standard Euler-Lagrange equations but the so-called Lagrange-Poincaré equations. In many interesting cases it is simpler to analyze the reduced equations instead of the original equations. This is why studying the possible variational origin of the reduced equations can be a problem of great interest. If the reduced equations admit a Lagrangian formulation, then automatically the unreduced ones are also variational in the classical sense, but not conversely. Moreover, the equations of motion often appear in a reduced version and, using our methods, it is possible to analyze the existence of a possible Lagrangian formulation on the Lie algebroid setting. We expect that these results will be useful for developing methods for the stabilization of controlled mechanical systems with symmetry based on the techniques of controlled Lagrangians introduced in Chapter 2 and then to use energy methods to find control gains that yield closed-loop stability (see [21, 57]).

This chapter is organized as follows. In Section 3.1 we give the necessary background on the theory of Lie algebroids, including prolongations of Lie algebroids, the Tulczyjew isomorphism and symplectic Lie algebroids. In Section 3.2 we discuss the lack of the Poincaré Lemma for the differential associated to general Lie algebroids and give a characterization of locally exact sections of the dual of a regular Lie algebroid. This is a key lemma in Section 3.3.2. In Section 3.3.1 we review the derivation of the Euler-Lagrange equations on a Lie algebroid in the way given in [113]. In Section 3.3.2 we identify the insufficiency of the Helmholtz conditions as the lack of the Poincaré Lemma and give a characterization of the variationality of a SODE on a regular Lie algebroid using Lemma 3.2.3. We also give a generalization to SODEs on regular Lie algebroids of Crampin's characterization for SODEs on tangent bundles [43] weakening the notion of variationality and we include an example of a SODE on a Lie algebroid that is not variational but satisfies the Helmholtz conditions. In Section 3.4 we study how morphisms of Lie algebroids treat the variational condition for SODE sections. This generalizes with an intrinsic proof results in [45] about the inverse problem on a Lie group and the corresponding reduced inverse problem on the Lie algebra. An interesting application appears in Section 3.5 where the inverse problem on Atiyah algebroids is considered, including some illustrative examples. In Section 3.6 we give the equivalence between the Helmholtz conditions derived in this chapter and the Helmholtz conditions given in [45] for Lie algebras and in [131] for Lie algebroids. Note that in the last paper the insufficiency of the Helmholtz conditions is not discussed.

3.1 Background on Lie algebroids

In this section we give the background on the theory of Lie algebroids that will be needed later on. This includes prolongations of Lie algebroids, the Tulczyjew isomorphism for Lie algebroids, symplectic Lie algebroids and Lagrangian submanifolds. For further details we refer the reader to [51, 106] and references therein.

Definition 3.1.1. *A Lie algebroid is a vector bundle $\tau : E \longrightarrow M$ together with a morphism of vector bundles $\rho : E \longrightarrow TM$, called the anchor map, and a Lie bracket $[\cdot, \cdot]$ in $\Gamma(E)$, the $\mathcal{C}^\infty(M)$-module of sections of E, satisfying the Leibniz rule*

$$[X, fY] = \rho(X)(f)Y + f[X, Y] \quad \textit{for all} \quad X, Y \in \Gamma(E) \textit{ and } f \in C^\infty(M)\,.$$

Note that the notion of Lie algebroid is, in particular, a generalization of tangent bundles and Lie algebras.

Let (x^i) denote local coordinates on M and $\{e_1, \ldots, e_n\}$ be a basis of local sections of E. With respect to this basis, the structure functions ρ^i_α and $C^\gamma_{\alpha\beta}$ of the Lie algebroid are functions on M defined by

$$\rho(e_\alpha) = \rho^i_\alpha \frac{\partial}{\partial x^i} \quad \text{and} \quad [e_\alpha, e_\beta] = C^\gamma_{\alpha\beta} e_\gamma\,. \tag{3.1}$$

Since the anchor map ρ is a Lie algebra morphism, that is $[\rho(e_\alpha), \rho(e_\beta)] = \rho[e_\alpha, e_\beta]$, and the Jacobi identity $\sum_{(\alpha,\beta,\gamma)}[e_\gamma, [e_\alpha, e_\beta]] = 0$ holds, the structure functions must satisfy the structure equations

$$\rho^j_\alpha \frac{\partial \rho^i_\beta}{\partial x^j} - \rho^j_\beta \frac{\partial \rho^i_\alpha}{\partial x^j} = \rho^i_\gamma C^\gamma_{\alpha\beta} \quad \text{and} \quad \sum_{(\alpha,\beta,\gamma)} \left[\rho^i_\alpha \frac{\partial C^\nu_{\beta\gamma}}{\partial x^i} + C^\nu_{\alpha\mu} C^\mu_{\beta\gamma} \right] = 0\,.$$

A Lie algebroid structure in a vector bundle $\tau : E \longrightarrow M$ is equivalent to an exterior differential d^E in the dual vector bundle $\tau^* : E^* \longrightarrow M$, that is, an operator $d^E : \Gamma(\Lambda^* E^*) \longrightarrow \Gamma(\Lambda^{*+1} E^*)$ satisfying

$$d^E \circ d^E = 0\,,$$
$$d^E(\alpha \wedge \beta) = d^E\alpha \wedge \beta + (-1)^{\deg(\alpha)} \alpha \wedge d^E\beta\,,$$

where $\alpha, \beta \in \Gamma(\Lambda^* E^*)$ and $\deg(\alpha)$ denotes the degree of α.

If $\alpha \in \Gamma(\Lambda^n E^*)$, the exterior differential $d^E\alpha$ is defined from the bracket and the anchor map by

$$\begin{aligned} d^E\alpha(e_0, \ldots, e_n) &= \sum_{i=0}^{n} (-1)^i \rho(e_i)\alpha(e_0, \ldots, \widehat{e_i}, \ldots, e_n) \\ &\quad + \sum_{i<j} (-1)^{i+j} \alpha([e_i, e_j], e_0, \ldots, \widehat{e_i}, \ldots, \widehat{e_j}, \ldots, e_n)\,, \end{aligned}$$

where $e_0 \ldots, e_n \in \Gamma(E)$. On the other hand, given an exterior differential d^E, the equations

$$\rho(e)(f) = \langle d^E f, e\rangle \quad \text{and} \quad \langle \alpha, [e_1, e_2]\rangle = \rho(e_1)\langle \alpha, e_2\rangle - \rho(e_2)\langle \alpha, e_1\rangle - d^E\alpha(e_1, e_2)\,,$$

where $\alpha \in \Gamma(E^*)$, $e, e_1, e_2 \in \Gamma(E)$ and $f \in C^\infty(M)$, define an anchor map ρ and a Lie bracket of sections $[\cdot, \cdot]$ for E.

If $\{e^\alpha\}$ denotes the dual basis to $\{e_\alpha\}$, then for each $f \in \mathcal{C}^\infty(M)$ and $\theta = \theta_\alpha e^\alpha \in \Gamma(E^*)$ the local expressions of the corresponding exterior differentials are

$$d^E f = \frac{\partial f}{\partial x^i} \rho^i_\alpha e^\alpha \quad \text{and} \quad d^E\theta = \left(\frac{\partial \theta_\gamma}{\partial x^i} \rho^i_\beta - \frac{1}{2} \theta_\alpha C^\alpha_{\beta\gamma} \right) e^\beta \wedge e^\gamma\,.$$

The following definitions will be used in order to introduce Lagrangian submanifolds on Lie algebroids later in Section 3.1.2.

Definition 3.1.2. *A Lie algebroid morphism is a morphism of vector bundles $F : E \longrightarrow E'$ over $f : M \longrightarrow M'$ such that $d^E((F, f)^*\phi') = (F, f)^*(d^{E'}\phi')$, for all $\phi' \in \Gamma(\Lambda^k(E')^*)$. A Lie algebroid epimorphism is a Lie algebroid morphism (F, f) such that f is a surjective submersion and $F|_{E_x} : E_x \to E'_{f(x)}$ is a linear epimorphism for all $x \in M$.*

Definition 3.1.3. *A Lie subalgebroid is a morphism of Lie algebroids $j : F \longrightarrow E$, $i : N \longrightarrow M$ such that the pair (j, i) is a monomorphism of vector bundles and i is an injective immersion.*

3.1.1 Prolongations of Lie algebroids

Now we will introduce the prolongation of a Lie algebroid over a smooth map $f : M' \longrightarrow M$. This notion will allow a derivation of the Euler-Lagrange equations without using the Poisson bracket on the dual of the Lie algebroid (as was done in [165]). This is the analog of the Klein formalism for tangent bundles [97] and it was given by E. Martínez in [113] for Lie algebroids. We will recall it in the next section.

In order to guarantee that the following construction is a vector bundle, a constant c is needed such that

$$\dim\left(\rho(E_{f(x')}) + (T_{x'}f)(T_{x'}M')\right) = c \quad \text{for all} \quad x' \in M'. \tag{3.2}$$

This condition implies that the dimension of the fibers must be constant.

Definition 3.1.4 ([51, 83])**.** *Let $(E, [\cdot,\cdot], \rho)$ be a Lie algebroid over a manifold M with projection denoted by τ, and $f : M' \longrightarrow M$ a smooth map satisfying (3.2). The prolongation of E over f is the Lie algebroid $(\mathcal{L}^f E, [\cdot,\cdot]^f, \rho^f)$ over M' with total space*

$$\mathcal{L}^f E = \left\{(b, v') \in E \times TM' : \rho(b) = (Tf)(v')\right\}$$

and projection $\tau^f : \mathcal{L}^f E \longrightarrow M'$ given by $\tau^f(b, v') = \tau_{M'}(v')$. The sections of $\mathcal{L}^f E$ are of the form $(h^i(X_i \circ f), X')$, where $X' \in \mathfrak{X}(M')$, $X_i \in \Gamma(E)$ and $h^i \in \mathcal{C}^\infty(M')$. Then the Lie bracket is defined by

$$[(h^i(X_i \circ f), X'), (s^j(Y_j \circ f), Y')]^f = (h^i s^j([X_i, Y_j] \circ f) + X'(s^j)(Y_j \circ f) - Y'(h^i)(X_i \circ f), [X', Y'])$$

where $X', Y' \in \mathfrak{X}(M')$, $X_i, Y_i \in \Gamma(E)$ and $h^i, s^i \in \mathcal{C}^\infty(M')$. Note that the bracket in the second factor denotes the usual bracket of vector fields. Finally the anchor map is given by the projection onto the second factor:

$$\begin{array}{rccc} \rho^f : & \mathcal{L}^f E & \longrightarrow & TM' \\ & (b, v') & \longmapsto & v'. \end{array}$$

In particular if we take f to be the projections $\tau : E \longrightarrow M$ and $\tau^* : E^* \longrightarrow M$ respectively then the prolongations $\mathcal{L}^\tau E$ and $\mathcal{L}^{\tau^*} E$ play the roles of TTQ and TT^*Q respectively, which are recovered when $E = TQ$. These are the prolongations that we will use in this chapter, so we will now introduce local coordinates for them.

Let $\{e_\alpha\}$ denote a basis of local sections of $\tau : E \longrightarrow M$ and (x^i, y^α) the corresponding coordinates on E. Having in mind the structure functions defined in (3.1), we consider the basis of local sections of $\mathcal{L}^\tau E \longrightarrow E$ given by

$$\tilde{T}_\alpha(a) = \left(e_\alpha(\tau(a)), \rho^i_\alpha \left.\frac{\partial}{\partial x^i}\right|_a\right), \quad \tilde{V}_\alpha(a) = \left(0, \left.\frac{\partial}{\partial y^\alpha}\right|_a\right), \quad a \in E \tag{3.3}$$

following the notation in [51].

With respect to this basis the structure functions are given by

$$[\tilde{T}_\alpha, \tilde{T}_\beta]^\tau = C^\gamma_{\alpha\beta} \tilde{T}_\gamma, \quad [\tilde{T}_\alpha, \tilde{V}_\beta]^\tau = 0, \quad [\tilde{V}_\alpha, \tilde{V}_\beta]^\tau = 0\,,$$

$$\rho^\tau(\tilde{T}_\alpha) = \rho^i_\alpha \frac{\partial}{\partial x^i}, \quad \rho^\tau(\tilde{V}_\alpha) = \frac{\partial}{\partial y^\alpha}\,,$$

and the local coordinates induced on $\mathcal{L}^\tau E$ will be denoted by $(x^i, y^\alpha, z^\alpha, v^\alpha)$.

Let $\{e^\alpha\}$ be the dual basis to $\{e_\alpha\}$ and (x^i, y_α) the corresponding coordinates on E^*. We consider the basis of local sections of $\mathcal{L}^{\tau^*}E \longrightarrow E^*$

$$\tilde{T}_\alpha(a^*) = \left(e_\alpha(\tau^*(a^*)), \rho^i_\alpha \left.\frac{\partial}{\partial x^i}\right|_{a^*}\right), \quad \tilde{V}^\alpha(a^*) = \left(0, \left.\frac{\partial}{\partial y_\alpha}\right|_{a^*}\right), \quad a^* \in E^*$$

with structure functions given by

$$[\tilde{T}_\alpha, \tilde{T}_\beta]^{\tau^*} = C^\gamma_{\alpha\beta}\tilde{T}_\gamma, \quad [\tilde{T}_\alpha, \tilde{V}^\beta]^{\tau^*} = 0, \quad [\tilde{V}^\alpha, \tilde{V}^\beta]^{\tau^*} = 0\,,$$

$$\rho^{\tau^*}(\tilde{T}_\alpha) = \rho^i_\alpha \frac{\partial}{\partial x^i}, \quad \rho^{\tau^*}(\tilde{V}_\alpha) = \frac{\partial}{\partial y_\alpha}\,.$$

The local coordinates induced on $\mathcal{L}^{\tau^*}E \longrightarrow E^*$ will be denoted by $(x^i, y_\alpha, z^\alpha, v_\alpha)$.

Remark 3.1.5. A map $F : E \longrightarrow E^*$ over M induces a map $\mathcal{L}F : \mathcal{L}^\tau E \longrightarrow \mathcal{L}^{\tau^*}E$ defined by

$$\mathcal{L}F(b, X_a) := (b, T_aF(X_a))\,.$$

If locally $F(x^i, y^\alpha) = (x^i, F_\alpha(x, y))$, then the local expression for $\mathcal{L}F$ is

$$\mathcal{L}F(x^i, y^\alpha, z^\alpha, v^\alpha) = \left(x^i, F_\alpha, z^\alpha, \rho^i_\beta z^\beta \frac{\partial F_\alpha}{\partial x^i} + v^\beta \frac{\partial F_\alpha}{\partial y^\beta}\right)\,.$$

3.1.2 Lagrangian submanifolds of symplectic Lie algebroids

According to the philosophy in Section 1.5, we must define Lagrangian submanifolds of symplectic Lie algebroids, see [51] for more details.

Definition 3.1.6. *A symplectic section Ω on a Lie algebroid $(E, [\cdot,\cdot], \rho)$ is a closed section of the vector bundle $E^* \wedge E^* \longrightarrow M$ satisfying that $\Omega_x : E_x \wedge E_x \longrightarrow \mathbb{R}$ is nondegenerate, that is, each fiber is a symplectic vector space. A Lie algebroid with a symplectic section will be called a symplectic Lie algebroid.*

Example 3.1.7. The Lie algebroid $\mathcal{L}^{\tau^*}E$ has a canonical symplectic section defined as $\Omega_E = -d^{\mathcal{L}^{\tau^*}E}\lambda_E$, where λ_E is the canonical section of $(\mathcal{L}^{\tau^*}E)^* \longrightarrow E^*$ given by

$$\lambda_E(a^*)(b, v) = a^*(b) \quad \text{for all } a^* \in E^*$$

and is called the Liouville section.

Once a symplectic section has been defined, we can introduce Lagrangian submanifolds of the Lie algebroid. As mentioned before, we are interested in Lagrangian submanifolds because we want to extend the geometric characterization of the inverse problem presented in Section 1.5 to the Lie algebroid setting.

Definition 3.1.8. *Let Ω be a symplectic section on E. The Lie subalgebroid $j : F \longrightarrow E$, $i : N \longrightarrow M$ is called Lagrangian if $j(F_x)$ is a Lagrangian subspace of $(E_{i(x)}, \Omega_{i(x)})$ for each $x \in N$.*

The Tulczyjew isomorphism in classical mechanics can be extended to the Lie algebroid setting. In this context the canonical isomorphism is between $\rho^*(TE^*)$ and $(\mathcal{L}^\tau E)^*$

$$\begin{array}{ccccc} & \mathcal{L}^{\tau^*}E \equiv \rho^*(TE^*) & \xrightarrow{A_E} & (\mathcal{L}^\tau E)^* & \\ \swarrow & & \searrow & \swarrow & \\ E^* & & E & & \end{array}$$

and is locally given by $A_E(x^i, y_\alpha, z^\alpha, v_\alpha) = (x^i, z^\alpha, v_\alpha + C^\gamma_{\alpha\beta} y_\gamma z^\beta, y_\alpha)$. For an intrinsic definition and more details we refer the reader to [51].

Remark 3.1.9. The vector bundles $\mathcal{L}^{\tau^*}E \longrightarrow E^*$ and $\rho^*(TE^*) \longrightarrow E$ have the same total spaces but different projections.

Now we will recall Proposition 7.8 in [51] which will be used in the sequel. Let τ^N denote the projection $\tau : E \longrightarrow M$ restricted to a submanifold $i : N \hookrightarrow E$, that is, $\tau^N = \tau \circ i$.

Proposition 3.1.10. *Given a section $\tilde{X}$ of the pull-back vector bundle $\rho^*(TE^*) \longrightarrow E$ define $\alpha_{\tilde{X}} = A_E \circ \tilde{X}$, which is a section of $(\mathcal{L}^\tau E)^* \longrightarrow E$, and put $N = \tilde{X}(E)$. Then the Lie subalgebroid (Id, Ti) : $\mathcal{L}^{(\tau^{\tau^*})^N}(\mathcal{L}^{\tau^*}E) \longrightarrow \mathcal{L}^{\tau^{\tau^*}}(\mathcal{L}^{\tau^*}E)$, $i : N \longrightarrow \mathcal{L}^{\tau^*}E$ is Lagrangian if and only if $d^{\mathcal{L}^\tau E}\alpha_{\tilde{X}} = 0$, where $\mathcal{L}^{(\tau^{\tau^*})^N}(\mathcal{L}^{\tau^*}E)$ is the prolongation of $\mathcal{L}^{\tau^*}E$ over the map $(\tau^{\tau^*})^N : N \longrightarrow E^*$.*

Remark 3.1.11. According to Definition 8.1 in [51], $N = \tilde{X}(E)$ is a Lagrangian submanifold of $\mathcal{L}^{\tau^*}E$.

3.2 Closed sections versus exact sections

On Lie algebroids the Poincaré Lemma does not hold in general for the differential d^E, that is, the closedness of a section does not guarantee its local exactness.

Example 3.2.1. Consider Example 3.3.6 in [106], that is, the Lie algebroid with total space $E = T\mathbb{R}$, base space $M = \mathbb{R}$, Lie bracket defined by

$$\left[\xi\frac{d}{dt}, \eta\frac{d}{dt}\right]' = t\left(\frac{d\eta}{dt}\xi - \frac{d\xi}{dt}\eta\right)\frac{d}{dt}$$

for functions $\xi, \eta : \mathbb{R} \longrightarrow \mathbb{R}$, where t denotes the coordinate on $\mathbb{R}$, and anchor map given by

$$\begin{array}{rccc} \rho : & T\mathbb{R} & \longrightarrow & T\mathbb{R} \\ & \xi\frac{d}{dt} & \longmapsto & t\xi\frac{d}{dt}. \end{array}$$

Thus, the structure functions are $\rho^1_1 = t$ and $C^1_{11} = 0$. Note that this algebroid is not regular since $\rho(E_0) = 0$ while $\text{rank}(\rho(E_t)) = 1$ for $t \neq 0$, where E_t denotes the fiber of E over t in M.

We want to detect a section of $T^*\mathbb{R} \longrightarrow \mathbb{R}$ which is closed but not locally exact. Note first that, by dimension, $d^{T\mathbb{R}}\theta = 0$ for all $\theta = \alpha(t)dt \in \Gamma(T^*\mathbb{R})$. Since $d^{T\mathbb{R}}f = \frac{df}{dt}tdt$ for $f : \mathbb{R} \longrightarrow \mathbb{R}$, it suffices to take $\alpha(t)$ equal to a nonzero constant c so that the equation $\alpha(t) = t\frac{df}{dt}$ is not satisfied around 0.

We will give a characterization of the local exactness of a section of the dual of a regular Lie algebroid. For that we use some suitable coordinates given by the local splitting theorem in [60]. If E is regular, that is, ρ has constant rank q, then the theorem reduces to the following one:

Theorem 3.2.2 ([60])**.** *Let $(E, [\cdot,\cdot], \rho)$ be a regular Lie algebroid over M and let $x_0 \in M$. There exist coordinates (x^i), $i = 1, \dots, m = \dim(M)$ in a neighborhood U of x_0 and a basis of sections $\{e_1, \dots, e_n\}$ of $\tau^{-1}(U) \longrightarrow U$ such that*

$$\rho(e_i) = \frac{\partial}{\partial x^i}, \quad i = 1, \dots, q\,,$$
$$\rho(e_s) = 0, \quad s = q+1, \dots, n\,.$$

Moreover $C^{\alpha}_{\beta\gamma} = 0$ for all $\alpha \leq q$.

As a consequence of the above theorem, we have the following lemma that characterizes locally exact sections. To our best knowledge, this lemma does not appear in the literature.

Lemma 3.2.3. *Let $(E, [\cdot,\cdot], \rho)$ be a regular Lie algebroid over M. A section θ of $\tau^* : E^* \longrightarrow M$ is locally exact if and only if it is closed and it satisfies $\theta(Z) = 0$ for all $Z \in \Gamma(Ker(\rho))$.*

Proof. $\Rightarrow$ Let $\{e^1, \dots, e^n\}$ denote a local basis of $\tau^* : E^* \longrightarrow M$ and write $\theta = \theta_\gamma(x) e^\gamma$. If $\theta = d^E f$ locally, then $d^E\theta = 0$. The second condition also holds since $\theta_\gamma = \frac{\partial f}{\partial x^i}\rho^i_\gamma$ and then for each $X = X^\gamma e_\gamma \in \Gamma(\mathrm{Ker}(\rho))$ we have

$$\theta(X) = \theta_\gamma X^\gamma = \frac{\partial f}{\partial x^i} \underbrace{\rho^i_\gamma X^\gamma}_{=0} = 0\,.$$

$\Leftarrow$ To prove the converse result take the coordinates (x^i) on M and the basis $\{e_1, \dots, e_n\}$ of sections of $E \longrightarrow M$ given in the splitting Theorem 3.2.2, so that $\{e_{q+1}, \dots, e_n\}$ is a basis of $\Gamma(\mathrm{Ker}(\rho))$. Let $\{e^1, \dots, e^n\}$ denote the dual basis. If θ annihilates the sections $\Gamma(\mathrm{Ker}(\rho))$, then it is written as $\theta = \theta_\gamma(x^i)e^\gamma$ for $\gamma = 1, \dots, q$.

Locally, the condition $d^E\theta = 0$ reads

$$\frac{\partial \theta_\gamma}{\partial x^i}\rho^i_\beta - \frac{\partial \theta_\beta}{\partial x^i}\rho^i_\gamma - \theta_\alpha C^\alpha_{\beta\gamma} = 0 \quad \text{for all} \quad \beta, \gamma = 1, \dots n, \quad i = 1 \dots m\,.$$

Using that $\rho^i_\beta = 0$ for $\beta > q$, $\rho^i_\beta = \delta^i_\beta$ for $\beta \leq q$ and $C^\alpha_{\beta\gamma} = 0$ for $\alpha \leq q$ in the chosen coordinates and also that $\theta_\gamma = 0$ for $\gamma > q$ the above condition reduces to

$$\frac{\partial \theta_\gamma}{\partial x^\beta} - \frac{\partial \theta_\beta}{\partial x^\gamma} = 0, \quad \beta, \gamma = 1, \dots, q\,,$$

which is precisely the integrability condition that provides locally a function $f(x)$ such that $\theta_\gamma = \frac{\partial f}{\partial x^\gamma}$, $\gamma = 1 \dots q$. ■

Next we give an example of a regular Lie algebroid for which the Poincaré Lemma is not satisfied.

Example 3.2.4. Consider the Lie algebra $E = \mathfrak{se}(2)$ with generators

$$e_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

Lie bracket given by

$$[e_1, e_2] = 0, \quad [e_1, e_3] = -e_2 \quad \text{and} \quad [e_2, e_3] = e_1$$

and anchor map $\rho \equiv 0$. Let $\{e^1, e^2, e^3\}$ denote the dual basis. Observe that $d^E(e^3)$=0, since $C^3_{\alpha\beta} = 0$. Note also that $\mathrm{Ker}(\rho) = \{e_1, e_2, e_3\}$ and $e^3(e_3) = 1 \neq 0$, that is, the second condition in Lemma 3.2.3 is not satisfied and therefore e^3 is not locally exact.

3.3 The inverse problem of the calculus of variations on Lie algebroids

We first need to introduce briefly Lagrangian mechanics on Lie algebroids so that the geometric framework of the inverse problem on Lie algebroids can be described.

3.3.1 Lagrangian mechanics on Lie algebroids

In this section we will derive the Euler-Lagrange equations for a Lagrangian function on a Lie algebroid following [113]. These equations were previously derived in [165] using the Poisson structure in the dual bundle.

The vertical endomorphism and the Liouville vector field on tangent bundles can be generalized to Lie algebroids. Note that these are the two ingredients needed to define the concept of a SODE. First we give the definitions of the vertical and complete lifts of a section of $E \longrightarrow M$ to a section of $\mathcal{L}^\tau E \longrightarrow E$.

Definition 3.3.1. *Let $X \in \Gamma(E)$.*

- *The vertical lift of X is the section $X^v \in \Gamma(\mathcal{L}^\tau E)$ defined by $X^v(a) = (0, X(\tau(a))^v_a), a \in E$, where for each pair $a, b \in E$, b^v_a acts on a function $F \in \mathcal{C}^\infty(E)$ as*

$$b^v_a(F) = \left.\frac{d}{dt}\right|_{t=0} F\,(a + tb)\ .$$

- *The complete lift of X is the unique section $X^c \in \Gamma(\mathcal{L}^\tau E)$ that projects over X and satisfies*

$$\rho^\tau(X^c)\widehat{\theta} = \widehat{\mathcal{L}^E_X \theta} \quad \textit{for all} \quad \theta \in \Gamma(E^*)\,,$$

where $\widehat{\theta} : E \longrightarrow \mathbb{R}$ is the linear function defined by the pairing $\widehat{\theta}(e) = \langle \theta(\tau^*(e)), e\rangle$ and $\mathcal{L}^E_X := i_X \circ d^E + d^E \circ i_X$ is the Lie derivative.

Definition 3.3.2. *Given a Lie algebroid $E \longrightarrow M$,*

- *the vertical endomorphism S is the unique section of $\mathcal{L}^\tau E \otimes \mathcal{L}^\tau E \longrightarrow E$ satisfying*

$$S(X^v) = 0 \quad \text{and} \quad S(X^c) = X^v \quad \text{for all} \quad X \in \Gamma(E)\,,$$

- *the Euler section Δ is the section of $\mathcal{L}^\tau E \longrightarrow E$ defined by*

$$\Delta(a) = (0, a_a^v) \quad \text{for all} \quad a \in E\,.$$

Definition 3.3.3. *A section Γ of $\mathcal{L}^\tau E \longrightarrow E$ is a SODE (second order differential equation) if it satisfies $S(\Gamma) = \Delta$. We will use the expressions SODE section and SODE field to distinguish the section Γ from the vector field $\rho^\tau(\Gamma)$.*

With respect to the basis $\{\tilde{T}_\alpha, \tilde{V}_\alpha\}$ defined in (3.3), the local expression of a SODE section is

$$\Gamma = y^\alpha \tilde{T}_\alpha + \Gamma^\alpha \tilde{V}_\alpha\,.$$

As $\rho^\tau(\tilde{T}_\alpha) = \rho^i_\alpha \frac{\partial}{\partial x^i}$ and $\rho^\tau(\tilde{V}_\alpha) = \frac{\partial}{\partial y^\alpha}$, the local expression for the SODE field is

$$\rho^\tau(\Gamma) = y^\alpha \rho^i_\alpha \frac{\partial}{\partial x^i} + \Gamma^\alpha \frac{\partial}{\partial y^\alpha}\,,$$

so the integral curves of $\rho^\tau(\Gamma)$ are the solutions to $\dot{x}^i = \rho^i_\alpha y^\alpha$ and $\dot{y}^\alpha = \Gamma^\alpha(x, y)$.

If we also have a Lagrangian function $L : E \longrightarrow \mathbb{R}$ on the Lie algebroid, then we can define the Poincaré-Cartan one-section θ_L, the Poincaré-Cartan two-section ω_L and the energy function E_L as follows:

$$\theta_L = S(d^E L), \quad \omega_L = -d^E \theta_L, \quad E_L = \rho^\tau(\Delta)(L) - L\,.$$

If L is regular, that is, if the matrix $\left(\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right)$ is regular, then ω_L is a symplectic section and the Hamiltonian equation

$$i_\Gamma \omega_L = d^E E_L$$

has a unique solution Γ_L. The integral curves of Γ_L are the integral curves of $\rho^\tau(\Gamma_L)$, which are those locally satisfying the Euler-Lagrange equations for a Lie algebroid:

$$\begin{aligned} \frac{dx^i}{dt} &= \rho^i_\alpha y^\alpha\,, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) &= \rho^i_\alpha \frac{\partial L}{\partial x^i} - C^\gamma_{\alpha\beta} y^\beta \frac{\partial L}{\partial y^\gamma}\,, \end{aligned}$$

where (x^i) are the coordinates on M and (x^i, y^α) the coordinates on E.

Note that for the special cases ($E = TQ, [\cdot,\cdot], \rho = Id$) and ($\mathfrak{g}, [\cdot,\cdot], \rho = 0$) we recover the Euler-Lagrange equations and the Euler-Poincaré equations respectively.

3.3.2 The inverse problem on Lie algebroids

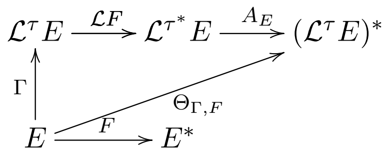
In this section we recover the Helmholtz conditions for a SODE on a Lie algebroid and give a characterization of the inverse problem for regular Lie algebroids.

Let Γ be a SODE on E, locally written as $\Gamma = y^\alpha \tilde{T}_\alpha + \Gamma^\alpha \tilde{V}_\alpha$. The inverse problem poses the following question: When is it possible to find a nondegenerate matrix of multipliers $(g_{\alpha\beta}(x, y))$ such that

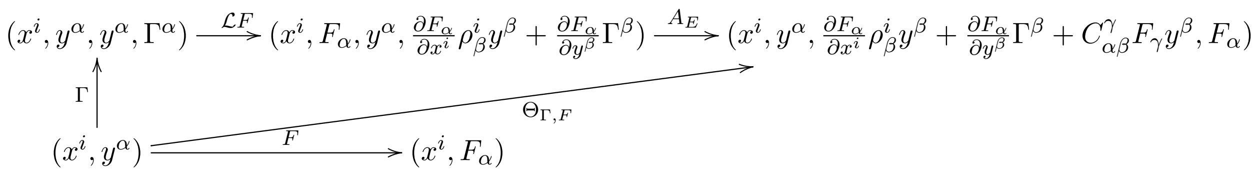
$$g_{\alpha\beta}(\dot{y}^\alpha - \Gamma^\alpha) = \frac{d}{dt}\left(\frac{\partial L}{\partial y^\beta}\right) - \rho^i_\beta \frac{\partial L}{\partial x^i} + C^\gamma_{\beta\nu} y^\nu \frac{\partial L}{\partial y^\gamma} \tag{3.4}$$

has a regular solution L? If it is possible then Γ is called variational.

Given a SODE Γ on E and a local diffeomorphism $F : E \longrightarrow E^*$, we define a section of $(\mathcal{L}^\tau E)^* \longrightarrow E$ by $\Theta_{\Gamma,F} := A_E \circ \mathcal{L}F \circ \Gamma$, as shown in the following diagram:



In local coordinates the above diagram becomes the following:



Let $\{\tilde{T}^\gamma, \tilde{V}^\gamma\}$ denote the dual basis of $\{\tilde{T}_\gamma, \tilde{V}_\gamma\}$. Then locally we can write $\Theta_{\Gamma,F} = \theta_\alpha \tilde{T}^\alpha + F_\alpha \tilde{V}^\alpha$, where

$$\theta_\alpha = \frac{\partial F_\alpha}{\partial x^i}\rho^i_\beta y^\beta + \frac{\partial F_\alpha}{\partial y^\beta}\Gamma^\beta + C^\gamma_{\alpha\beta}F_\gamma y^\beta. \tag{3.5}$$

The exterior differential of $\Theta_{\Gamma,F}$ is

$$d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = \left(\frac{\partial\theta_\gamma}{\partial x^i}\rho^i_\beta - \frac{1}{2}\theta_\alpha C^\alpha_{\beta\gamma}\right)\tilde{T}^\beta \wedge \tilde{T}^\gamma + \frac{\partial\theta_\gamma}{\partial y^\beta}\tilde{V}^\beta \wedge \tilde{T}^\gamma + \frac{\partial F_\gamma}{\partial x^i}\rho^i_\beta \tilde{T}^\beta \wedge \tilde{V}^\gamma + \frac{\partial F_\beta}{\partial y^\gamma}\tilde{V}^\beta \wedge \tilde{V}^\gamma\,.$$

Imposing $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = 0$ we obtain the Helmholtz conditions

$$\frac{\partial F_\beta}{\partial y^\gamma} = \frac{\partial F_\gamma}{\partial y^\beta}, \quad \frac{\partial\theta_\gamma}{\partial y^\beta} = \frac{\partial F_\beta}{\partial x^i}\rho^i_\gamma, \quad \frac{\partial\theta_\gamma}{\partial x^i}\rho^i_\beta - \frac{1}{2}\theta_\alpha C^\alpha_{\beta\gamma} = \frac{\partial\theta_\beta}{\partial x^i}\rho^i_\gamma - \frac{1}{2}\theta_\alpha C^\alpha_{\gamma\beta}\,. \tag{3.6}$$

As mentioned earlier, these conditions are not enough to guarantee the existence of a Lagrangian function on E, since the Poincaré Lemma does not hold for an arbitrary Lie algebroid. We need to ask for the additional condition

$$\Theta_{\Gamma,F}(Z) = 0 \quad \text{for all} \quad Z \in \Gamma(\mathrm{Ker}(\rho^\tau))\,.$$

Let $\{e_I\}$ denote a local basis of $\Gamma(\mathrm{Ker}(\rho))$. Then $\{\tilde{T}_I\}$ is a local basis of $\Gamma(\mathrm{Ker}(\rho^\tau))$ and the condition on $\Gamma(\mathrm{Ker}(\rho^\tau))$ is

$$\theta_I = \frac{\partial F_I}{\partial x^i}\rho^i_\beta y^\beta + \frac{\partial F_I}{\partial y^\beta}\Gamma^\beta + C^\gamma_{I\beta}F_\gamma y^\beta = 0, \quad I = 1, \ldots, d = \dim(\mathrm{Ker}(\rho)) \le n\,. \tag{3.7}$$

Using the local basis $\{e_I, e_a\}$ adapted to $\mathrm{Ker}(\rho)$, the anchor map has the local expression $\rho^i_I = 0$. Then the Helmholtz conditions in (3.6) become

$$\frac{\partial F_\beta}{\partial y^\gamma} = \frac{\partial F_\gamma}{\partial y^\beta}\,,$$
$$\frac{\partial \theta_a}{\partial y^\beta} = \frac{\partial F_\beta}{\partial x^i}\rho^i_a, \quad \frac{\partial \theta_I}{\partial y^\beta} = 0\,, \tag{3.8}$$
$$\frac{\partial \theta_a}{\partial x^i}\rho^i_b - \theta_\alpha C^\alpha_{ba} = \frac{\partial \theta_b}{\partial x^i}\rho^i_a, \quad \frac{\partial \theta_I}{\partial x^i}\rho^i_a - \theta_\alpha C^\alpha_{aI} = 0, \quad \theta_\alpha C^\alpha_{JI} = 0\,. \tag{3.9}$$

From the second equation in (3.8) we deduce that $\theta_I(x,y) = \theta_I(x)$. Then the additional condition in (3.7) will be satisfied if $\theta_I(x) = 0$.

Theorem 3.3.4. *A SODE section Γ on a regular Lie algebroid E is variational if and only if there is a local diffeomorphism $F : E \longrightarrow E^*$ such that $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = 0$ and $\Theta_{\Gamma,F}(Z) = 0$ for all $Z \in \Gamma(Ker(\rho^\tau))$.*

Proof. $\Leftarrow$ If there is a local diffeomorphism F such that $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = 0$ and $\Theta_{\Gamma,F}(Z) = 0$ for all $Z \in \Gamma(Ker(\rho^\tau))$ then by Lemma 3.2.3 we have $\Theta_{\Gamma,F} = d^{\mathcal{L}^\tau E}L$ for a locally defined function $L : E \longrightarrow \mathbb{R}$. In local coordinates we get

$$F_\beta = \frac{\partial L}{\partial y^\beta} \quad \text{and} \quad \frac{\partial F_\gamma}{\partial y^\beta}\Gamma^\beta + y^\beta \frac{\partial F_\gamma}{\partial x^i}\rho^i_\beta + y^\beta F_\alpha C^\alpha_{\gamma\beta} = \frac{\partial L}{\partial x^i}\rho^i_\gamma\,.$$

Therefore $\frac{d}{dt}\left(\frac{\partial L}{\partial y^\beta}\right) - \rho^i_\beta \frac{\partial L}{\partial x^i} + C^\gamma_{\beta\nu} y^\nu \frac{\partial L}{\partial y^\gamma} = \frac{\partial F_\beta}{\partial y^\gamma}(\dot y^\gamma - \Gamma^\gamma)$, $g_{\beta\gamma} = \frac{\partial F_\beta}{\partial y^\gamma}$ are the multipliers for the problem and L is regular since $(g_{\beta\gamma})$ is nondegenerate.

$\Rightarrow$ If Γ is variational then there is a regular Lagrangian L such that Equation (3.4) is satisfied with $g_{\alpha\beta} = \frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}$. Taking F to be the Legendre transformation, which is a local diffeomorphism, it is straightforward to check that Equations (3.6) are satisfied, using $\theta_\gamma = \frac{\partial L}{\partial x^i}\rho^i_\gamma$ and the structure equation $\rho^j_\alpha \frac{\partial \rho^i_\beta}{\partial x^j} - \rho^j_\beta \frac{\partial \rho^i_\alpha}{\partial x^j} = \rho^i_\gamma C^\gamma_{\alpha\beta}$ for the last set. If $Z \in \Gamma(Ker(\rho^\tau))$, $Z = z^\alpha \tilde{T}_\alpha$, then we also get

$$\Theta_{\Gamma,F}(Z) = \left(\frac{\partial^2 L}{\partial y^\beta \partial y^\gamma}\Gamma^\beta + y^\beta \frac{\partial^2 L}{\partial x^i \partial y^\gamma}\rho^i_\beta + y^\beta \frac{\partial L}{\partial y^\alpha} C^\alpha_{\gamma\beta}\right) z^\gamma = \frac{\partial L}{\partial x^i}\rho^i_\gamma z^\gamma = 0\,.$$

■

Example 3.3.5. Note that for the Lie algebroid $(E = TQ, [\cdot,\cdot], \rho = Id)$, where $[\cdot,\cdot]$ is the Lie bracket of vector fields, we recover the equations

$$\frac{\partial F_\beta}{\partial y^\gamma} = \frac{\partial F_\gamma}{\partial y^\beta}, \quad \frac{\partial \Gamma(F_\gamma)}{\partial y^\beta} = \frac{\partial F_\beta}{\partial x^\gamma}, \quad \frac{\partial \Gamma(F_\gamma)}{\partial x^\beta} = \frac{\partial \Gamma(F_\beta)}{\partial x^\gamma}$$

given in Section 1.5 and the condition involving $\mathrm{Ker}(\rho^\tau)$ is void.

Example 3.3.6. For a Lie algebra $(\mathfrak{g}, [\cdot,\cdot], \rho = 0)$ we get the Helmholtz conditions

$$\frac{\partial F_\beta}{\partial y^\gamma} = \frac{\partial F_\gamma}{\partial y^\beta}, \quad \frac{\partial\left(\frac{\partial F_\gamma}{\partial y^\tau}\Gamma^\tau + C^r_{\gamma\tau}F_r y^\tau\right)}{\partial y^\beta} = 0, \quad \left(\frac{\partial F_\alpha}{\partial y^\tau}\Gamma^\tau + C^r_{\alpha\tau}F_r y^\tau\right)C^\alpha_{\beta\gamma} = 0\,.$$

In this case the condition on $\Gamma(\mathrm{Ker}(\rho^\tau))$ is $\frac{\partial F_\alpha}{\partial y^\tau}\Gamma^\tau + C^r_{\alpha\tau}F_r y^\tau = 0$, which makes the last two conditions always true. Note that the symmetry gives a function L such that $F_\gamma = \frac{\partial L}{\partial y^\gamma}$ and then the remaining conditions are just the Euler-Poincaré equations for L.

Now we will introduce the notion of weak variational SODE in order to avoid confusion with the meaning of variationality and its relationship with the Helmholtz conditions.

Definition 3.3.7. *A SODE Γ on E will be called weak variational if there is a local diffeomorphism $F : E \longrightarrow E^*$ such that $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = 0$.*

Hence, a SODE Γ on E is variational if it is weak variational and satisfies $\Theta_{\Gamma,F}(Z) = 0$ for all $Z \in \Gamma(\mathrm{Ker}(\rho^\tau))$. This definition, for the case of a Lie algebra, is equivalent to satisfying the reduced Helmholtz conditions given in [45].

Due to the lack of a Poincaré Lemma we give a generalization of Theorem 1.4.1 for weak variational SODEs substituting the closedness condition by local exactness of a section of the bundle $(\mathcal{L}^\tau E)^* \wedge (\mathcal{L}^\tau E)^* \longrightarrow E$, which plays the role of the Poincaré-Cartan two-section generalizing the Poincaré-Cartan two-form.

Theorem 3.3.8. *A SODE Γ on a regular Lie algebroid E is weak variational if and only if there is a nondegenerate section Ω of $(\mathcal{L}^\tau E)^* \wedge (\mathcal{L}^\tau E)^* \longrightarrow E$ such that*

(i) $\mathcal{L}_\Gamma \Omega = 0$,

(ii) $\Omega = d^{\mathcal{L}^\tau E}\Theta$ *for some locally defined section* Θ *of* $(\mathcal{L}^\tau E)^* \longrightarrow E$,

(iii) $\Omega(\tilde{V}_\alpha, \tilde{V}_\beta) = 0$ *for all* $\alpha, \beta = 1, \ldots, n$.

Proof. $\Rightarrow$ If Γ is weak variational then there is a local diffeomorphism $F : E \longrightarrow E^*$ over M such that $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = 0$. Then we can define $\Omega = d^{\mathcal{L}^\tau E}(F^*\lambda_E)$, which clearly satisfies the second condition and $\Omega(\tilde{V}_\alpha, \tilde{V}_\beta) = 0$. Note that $\mathcal{L}_\Gamma F^*\lambda_E = \Theta_{\Gamma,F}$ and hence it also satisfies $\mathcal{L}_\Gamma \Omega = d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F}=0$. Finally the nondegeneracy of $\left(\frac{\partial F_\gamma}{\partial y^\beta}\right)$ implies the nondegeneracy of Ω.

$\Leftarrow$ If we write $\Theta = \mu_\alpha \tilde{T}^\alpha + \nu_\alpha \tilde{V}^\alpha$ then the condition $d^{\mathcal{L}^\tau E}\Theta(\tilde{V}_\alpha, \tilde{V}_\beta) = \frac{\partial \nu_\alpha}{\partial y^\beta} - \frac{\partial \nu_\beta}{\partial y^\alpha} = 0$ gives a locally defined function $f : E \longrightarrow \mathbb{R}$ such that $\nu_\alpha = \frac{\partial f}{\partial y^\alpha}$ and then $d^{\mathcal{L}^\tau E} f(\tilde{V}_\alpha) = \Theta(\tilde{V}_\alpha) = \nu_\alpha$. Define $\tilde{\Theta} = \Theta - d^{\mathcal{L}^\tau E} f$, which satisfies $\tilde{\Theta}(\tilde{V}_\alpha) = 0$ and $d^{\mathcal{L}^\tau E}\tilde{\Theta} = \Omega$. We seek to have $\tilde{\Theta} = F^*\lambda_E$ for some local diffeomorphism F, so we define $F : E \longrightarrow E^*$ by $\langle F(v_x), w_x\rangle = \langle \tilde{\Theta}(v_x), W_x\rangle$, where $x \in M$, $v_x, w_x \in E$ and $W_x \in \mathcal{L}^\tau E$ is such that $\tau^\tau(W_x) = w_x$. This definition does not depend on the choice of W_x since $\tilde{\Theta}$ vanishes on vertical sections. Locally if we write $\tilde{\Theta} = A_\alpha \tilde{T}^\alpha + B_\alpha \tilde{V}^\alpha$, then $\tilde{\Theta} = \left(\frac{\partial f}{\partial x^i}\rho^i_\alpha - A_\alpha\right)\tilde{T}^\alpha =: F_\alpha \tilde{T}^\alpha$ and the nondegeneracy of $\left(\frac{\partial F_\alpha}{\partial y^\beta}\right)$ follows from the nondegeneracy of Ω. Finally we have $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F} = d^{\mathcal{L}^\tau E}\mathcal{L}_\Gamma F^*\lambda_E = d^{\mathcal{L}^\tau E}\mathcal{L}_\Gamma\tilde{\Theta} = \mathcal{L}_\Gamma\Omega = 0$, that is, Γ is weak variational.■

In [45] some variational examples are found by requiring only that the Helmholtz conditions are satisfied, but this is not generally the case. As we have seen, in order to guarantee the existence of a Lagrangian for a SODE on a Lie algebroid we need to ask for an extra condition. That extra requirement for the *invariant inverse problem* is written in terms of cohomology classes in [45], while here the integrability condition in Lemma 3.2.3 is used. Next we give an example of a SODE on a Lie algebra which is weak variational but not variational.

Example 3.3.9. Let (y^1, y^2, y^3) denote the coordinates for $\mathfrak{g} = \mathfrak{se}(2)$ corresponding to the basis given in Example 3.2.4 and define the following SODE:

$$\begin{array}{rccc} \Gamma : & \mathfrak{se}(2) & \longrightarrow & \mathcal{L}^\tau\mathfrak{se}(2) \cong 2\mathfrak{g} \\ & (y^1, y^2, y^3) & \longmapsto & (y^1, y^2, y^3, \Gamma^1 = y^2y^3, \Gamma^2 = -y^1y^3, \Gamma^3 = 1)\,. \end{array}$$

Consider the local diffeomorphism from $\mathfrak{se}(2)$ to $\mathfrak{se}(2)^*$ given by $F_1 = y^1, F_2 = y^2, F_3 = y^3$ and compute $\theta_1 = \theta_2 = 0$ and $\theta_3 = 1$ to get $\Theta_{\Gamma,F} = \tilde{T}^3 + F_\alpha \tilde{V}^\alpha$. Then $d^{\mathcal{L}^\tau\mathfrak{se}(2)}\Theta_{\Gamma,F} = -\frac{1}{2}\theta_3 C^3_{\beta\gamma}\tilde{T}^\beta \wedge \tilde{T}^\gamma = 0$ since $C^3_{\beta\gamma} = 0$, that is, the Helmholtz conditions are satisfied, so Γ is weak variational, but since $\Theta_{\Gamma,F}(e^c_3) = 1 \neq 0$, Γ is not variational using the ansatz $F_1 = y^1, F_2 = y^2, F_3 = y^3$.

The corresponding left-invariant SODE on TG is given by

$$\ddot{x} = 0, \quad \ddot{y} = 0, \quad \ddot{\theta} = 1,$$

where (x, y, θ) are coordinates on $SE(2)$. According to [45, Theorem 3] this SODE is variational, that is, we can find a Lagrangian on TG, but not an invariant one. It is actually straightforward to obtain the Lagrangian $L = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2\right) + \theta$.

We will now see that the SODE Γ on $\mathfrak{se}(2)$ is not variational. According to Example 3.3.6 we have to check that the system of PDEs

$$\frac{\partial F_1}{\partial y^1}y^2y^3 - \frac{\partial F_1}{\partial y^2}y^1y^3 + \frac{\partial F_1}{\partial y^3} = F_2y^3\,, \tag{3.10}$$

$$\frac{\partial F_2}{\partial y^1}y^2y^3 - \frac{\partial F_2}{\partial y^2}y^1y^3 + \frac{\partial F_2}{\partial y^3} = -F_1y^3\,, \tag{3.11}$$

$$\frac{\partial F_2}{\partial y^1}y^2y^3 - \frac{\partial F_2}{\partial y^2}y^1y^3 + \frac{\partial F_2}{\partial y^3} = F_1y^2 - F_2y^1\,, \tag{3.12}$$

$$\frac{\partial F_\alpha}{\partial y^\beta} = \frac{\partial F_\beta}{\partial y^\alpha}\,, \tag{3.13}$$

does not admit a solution such that $\det\left(\frac{\partial F_\alpha}{\partial y^\beta}\right) \neq 0$. So far we have only seen that the proposed map $F_1 = y^1, F_2 = y^2, F_3 = y^3$ is not a solution. If we first use the method of characteristics to solve the system (3.10)-(3.12), we obtain solutions F_1, F_2, F_3 that depend on three arbitrary functions of two variables, $f(\alpha_1, \alpha_2), g(\alpha_1, \alpha_2)$ and $h(\alpha_1, \alpha_2)$. Then (3.13) implies that h is a constant and gives a system of PDEs for f and g,

$$-\alpha_2\frac{\partial g}{\partial \alpha_1}(\alpha_1, \alpha_2) + \alpha_1\frac{\partial g}{\partial \alpha_2}(\alpha_1, \alpha_2) = f(\alpha_1, \alpha_2)\,, \tag{3.14}$$

$$-\alpha_2\frac{\partial f}{\partial \alpha_1}(\alpha_1, \alpha_2) + \alpha_1\frac{\partial f}{\partial \alpha_2}(\alpha_1, \alpha_2) = -g(\alpha_1, \alpha_2)\,, \tag{3.15}$$

which can also be solved using the method of characteristics. Finally, using (3.13), we compute

$$\det\left(\frac{\partial F_\alpha}{\partial y^\beta}\right) = f^2\frac{\partial f}{\partial \alpha_2} - g^2\frac{\partial f}{\partial \alpha_2} + fg\left(\frac{\partial g}{\partial \alpha_2} - \frac{\partial f}{\partial \alpha_1}\right)\,,$$

which vanishes on the possible solutions to (3.14)-(3.15).

Remark 3.3.10. It is also possible to give an example of a SODE on a Lie algebra $\mathfrak{g}$ which is not variational on $\mathfrak{g}$ and also not weak variational but variational on TG. See Example 8.3 Case 2C in [45].

3.4 Morphisms and the variational problem

The geometric description of the inverse problem on Lie algebroids given in the previous section leads to a generalization of some results in [45], where the relationship between the inverse problem on the tangent bundle to a Lie group and the corresponding reduced inverse problem on the Lie algebra is studied. By means of morphisms of Lie algebroids the same relationship can be studied for the inverse problem on a general Lie algebroid. The proof of the following result is intrinsic, no coordinates are used.

Theorem 3.4.1. *Let $\Psi : E \to E'$ be a morphism of Lie algebroids, and consider its prolongation $\mathcal{L}\Psi : \mathcal{L}^\tau E \to \mathcal{L}^{\tau'} E'$. Let Γ and Γ' be SODE sections on E and E' respectively such that*

$$\mathcal{L}\Psi \circ \Gamma = \Gamma' \circ \Psi \,.$$

If Γ' is weak variational (variational) then Γ is weak variational (variational).

Proof. Since $\mathcal{L}\Psi$ is a morphism of Lie algebroids we have that $(\mathcal{L}\Psi)^* d^{\mathcal{L}^{\tau'} E'} = d^{\mathcal{L}^\tau E}(\mathcal{L}\Psi)^*$. From Theorem 3.3.8 there exists an exact section $\Theta' \in \Gamma((\mathcal{L}^{\tau'} E')^*)$ such that $\Omega' = d^{\mathcal{L}^{\tau'} E'}\Theta'$ satisfies $\mathcal{L}_{\Gamma'}\Omega' = 0$ and the restriction of Ω' to vertical sections vanishes, that is, $\Omega'(U', V') = 0$, where U', V' are vertical sections ($S(U') = S(V') = 0$).

As Ψ is a morphism of Lie algebroids, we have that $\Theta = (\mathcal{L}\Psi)^*\Theta'$ also satisfies the conditions of Theorem 3.3.8. In fact, for every $Z \in \mathcal{L}^\tau E$ we have

$$\begin{aligned}
\langle \mathcal{L}_\Gamma \Theta, Z\rangle &= \rho^\tau(\Gamma)(\langle (\mathcal{L}\Psi)^*\Theta', Z\rangle) - \langle \Theta', \mathcal{L}\Psi([\Gamma, Z]^\tau)\rangle \\
&= \rho^{\tau'}(\Gamma')(\langle \Theta', \mathcal{L}\Psi(Z)\rangle) - \langle \Theta', [\Gamma', \mathcal{L}\Psi(Z)]^{\tau'}\rangle \\
&= \langle (\mathcal{L}\Psi)^*(\mathcal{L}_{\Gamma'}\Theta'), Z\rangle \,.
\end{aligned}$$

Therefore, $\mathcal{L}_\Gamma\Theta = (\mathcal{L}\Psi)^*(\mathcal{L}_{\Gamma'}\Theta')$ and also

$$\mathcal{L}_\Gamma \Omega = (\mathcal{L}\Psi)^*(\mathcal{L}_{\Gamma'}\Omega') = 0 \,.$$

Moreover, for all $Z_1, Z_2 \in \mathcal{L}^\tau E$ we get

$$\begin{aligned}
\Omega(S(Z_1), S(Z_2)) &= (\mathcal{L}\Psi)^*\Omega'(S(Z_1), S(Z_2)) \\
&= \Omega'(\mathcal{L}\Psi(S(Z_1)), \mathcal{L}\Psi(S(Z_2))) \\
&= \Omega'(S'(\mathcal{L}\Psi(Z_1)), S'(\mathcal{L}\Psi(Z_2))) \\
&= 0 \,,
\end{aligned}$$

using that $\mathcal{L}\Psi \circ S = S' \circ \mathcal{L}\Psi$ (see [37]). This proves that if Γ' is weak variational then Γ is also weak variational. Obviously if Θ' is exact, then Θ is also exact. Therefore, if Γ' is variational then Γ is also variational. ■

Now we write the converse to the previous result for the case of a fiberwise surjective morphism satisfying an extra assumption.

Theorem 3.4.2. *Let $\Psi : E \to E'$ be a fiberwise surjective morphism of Lie algebroids. Let Γ and Γ' be SODE sections on E and E' respectively such that $\mathcal{L}\Psi \circ \Gamma = \Gamma' \circ \Psi$. If Γ is weak variational (variational) and it admits a solution Θ of Theorem 3.3.8 such that $\Theta = (\mathcal{L}\Psi)^*\Theta'$ for some $\Theta' \in \Gamma((\mathcal{L}^{\tau'}E')^*)$, then Γ' is weak variational (variational).*

Proof. The proof follows the same lines that Theorem 3.4.1 using that Ψ is a fiberwise surjective morphism. ∎

Remark 3.4.3. See Example 3.3.9 for a case in which there is no section Θ' satisfying the property $\Theta = (\mathcal{L}\Psi)^*\Theta'$.

3.5 The inverse problem for Atiyah algebroids

The theory developed in Section 3.3 has a very interesting application when Atiyah algebroids are considered. We first review the main notions of Atiyah algebroids, see [51] and references therein for more details, and then we geometrically characterize the inverse problem on Atiyah algebroids. As shown in [51], the Euler-Lagrange equations of a G-invariant Lagrangian can be reduced to Lagrange-Poincaré equations by using the morphism of Lie algebroids between TQ and TQ/G (see [34]). Thus the results in Section 3.4 can be applied to establish some relationship between the inverse problem and its reduced version.

3.5.1 Atiyah algebroid associated to a principal bundle

Let $\pi : Q \to Q/G = M$ be a principal G-bundle and $\Phi : G \times Q \to Q$, $\Phi_g(q) = \Phi(g, q)$, the corresponding G-action. Denote by $\Phi^T : G \times TQ \to TQ$ the tangent lift of Φ, that is, $\Phi^T_g = T\Phi_g$ for all $g \in G$. Now consider the quotient vector bundle $\tau_{Q/G} : TQ/G \to M$ whose space of sections $\Gamma(TQ/G)$ is identified with the G-invariant vector fields on Q.

Let $\mathfrak{g}$ be the Lie algebra of G and take the action of G on $Q \times \mathfrak{g}$ given by

$$\begin{array}{rcl} G \times (Q \times \mathfrak{g}) & \longrightarrow & Q \times \mathfrak{g} \\ (g, (q, \xi)) & \longmapsto & (\Phi_g(q), \mathrm{Ad}_g(\xi)) \,, \end{array}$$

where $\mathrm{Ad}\colon G \times \mathfrak{g} \to \mathfrak{g}$ is the adjoint representation of G on $\mathfrak{g}$. The quotient vector bundle $\tilde{\mathfrak{g}} = (Q \times \mathfrak{g})/G$ is called the adjoint bundle associated with the principal bundle $\pi\colon Q \to M$. If ξ_Q is the infinitesimal generator of the action Φ associated with $\xi \in \mathfrak{g}$, that is,

$$\xi_Q(q) = \left.\frac{d}{dt}\right|_{t=0} \Phi(\exp(t\xi), q) \,,$$

then we have the following monomorphism of vector bundles:

$$\begin{array}{rrcl} j : & \tilde{\mathfrak{g}} & \longrightarrow & TQ/G \\ & [(q, \xi)] & \longmapsto & [\xi_Q(q)] \,. \end{array}$$

Moreover, we have the following exact sequence called the Atiyah sequence [106]:

$$0 \longrightarrow \tilde{\mathfrak{g}} \stackrel{j}{\longrightarrow} TQ/G \stackrel{[T\pi]}{\longrightarrow} TM \longrightarrow 0 \,.$$

Assume that we have a principal connection A on Q, that is, $A : TQ \to \mathfrak{g}$ satisfying $A(\xi_Q(q)) = \xi$ and A is equivariant with respect to the actions $\phi^T : G \times TQ \to TQ$ and $Ad : G \times \mathfrak{g} \to \mathfrak{g}$. Every principal connection A induces the following vector bundle isomorphism over the identity:

$$\begin{array}{rcl} TQ/G & \longrightarrow & T(Q/G) \oplus \tilde{\mathfrak{g}} \\ [X_q] & \longmapsto & T_q\pi(X_q) \oplus [(q, A(X_q))] \,, \end{array}$$

where $X_q \in T_qQ$. Therefore we have an identification $\Gamma(TQ/G) \cong \mathfrak{X}(M) \oplus \Gamma(\tilde{\mathfrak{g}})$, where $\Gamma(\tilde{\mathfrak{g}})$ is identified with the set of vector fields on Q which are π-vertical and G-invariant. Let $B : TQ \oplus TQ \to \mathfrak{g}$ be the curvature of the connection A in the principal bundle π. The Lie bracket $[\![\cdot, \cdot]\!]$ on $\Gamma(TQ/G) \cong \Gamma(TM \oplus \tilde{\mathfrak{g}}) \cong \mathfrak{X}(M) \oplus \Gamma(\tilde{\mathfrak{g}})$ is defined as

$$[\![X \oplus \tilde{\xi}, Y \oplus \tilde{\eta}]\!] = [X, Y] \oplus ([\tilde{\xi}, \tilde{\eta}] + [X^h, \tilde{\eta}] - [Y^h, \tilde{\xi}] - B(X^h, Y^h)) \,,$$

for $X, Y \in \mathfrak{X}(M)$ and $\tilde{\xi}, \tilde{\eta} \in \Gamma(\tilde{\mathfrak{g}})$, where $X^h, Y^h \in \mathfrak{X}(Q)$ are the horizontal lift of X, Y, respectively, via the principal connection A. The anchor map $\rho : \Gamma(TQ/G) \cong \mathfrak{X}(M) \oplus \Gamma(\tilde{\mathfrak{g}}) \to \mathfrak{X}(M)$ is given by

$$\rho(X \oplus \tilde{\xi}) = X \,.$$

Now we will give a local description. Let $U \times G$ be a local trivialization of the principal bundle $\pi : Q \to M$, where U is an open subset of M with local coordinates (x^i). Then we consider the trivial principal bundle $\pi : U \times G \to U$, where the action of G on $U \times G$ is given by left multiplication on the second factor, that is, $\Phi_g(m, h) = (m, gh)$, where $m \in U$ and $g, h \in G$. For a basis $\{\xi_a\}$ of $\mathfrak{g}$, $1 \le a \le n$, we denote by $\{\overleftarrow{\xi_a}\}$ the corresponding left-invariant vector fields on G. Then the principal connection is specified by coefficients $A_i^a(x)$ satisfying

$$A\left(\left.\frac{\partial}{\partial x^i}\right|_{(x,e)}\right) = A_i^a(x)\xi_a, \quad 1 \le i \le m \,,$$

where $x \in U$ and e is the identity element of G. The horizontal lift of a coordinate vector field $\frac{\partial}{\partial x^i}$ on U is the vector field $\left(\frac{\partial}{\partial x^i}\right)^h$ on $U \times G$ given by $\left(\frac{\partial}{\partial x^i}\right)^h = \frac{\partial}{\partial x^i} - A_i^a(x)\overleftarrow{\xi_a}$. Thus, the vector fields

$$\left\{ e_i = \frac{\partial}{\partial x^i} - A_i^a \overleftarrow{\xi_a}, e_b = \overleftarrow{\xi_b} \right\} \tag{3.16}$$

on $U \times G$ are left G-invariant and define a local basis $\{e'_i, e'_b\}$ of $\Gamma(TQ/G) \cong \mathfrak{X}(M) \oplus \Gamma(\tilde{\mathfrak{g}})$. We will denote by (x^i, y^i, y^b) the corresponding fibered coordinates on TQ/G.

The curvature of the principal connection is given by

$$B\left(\left.\frac{\partial}{\partial x^i}\right|_{(x,e)}, \left.\frac{\partial}{\partial x^j}\right|_{(x,e)}\right) = B_{ij}^a(x)\xi_a \,,$$

for $i, j \in \{1, \ldots, m\}$ and $x \in U$. If c_{ab}^c are the structure constants of $\mathfrak{g}$ with respect to the basis $\{\xi_a\}$, then

$$B_{ij}^c = \frac{\partial A_i^c}{\partial x^j} - \frac{\partial A_j^c}{\partial x^i} - c_{ab}^c A_i^a A_j^b \,.$$

Then for the previous local basis $\{e'_i, e'_b\}$ of $\Gamma(TQ/G)$ we deduce that

$$[\![e'_i, e'_j]\!] = -B^c_{ij}e'_c, \quad [\![e'_i, e'_a]\!] = c^c_{ab}A^b_i e'_c, \quad [\![e'_a, e'_b]\!] = c^c_{ab}e'_c\,,$$

$$\rho(e'_i) = \frac{\partial}{\partial x^i}, \quad \rho(e'_a) = 0\,,$$

for $i, j \in \{1, \ldots, m\}$ and $a, b \in \{1, \ldots, n\}$. Thus, the local structure functions of the Atiyah algebroid $\tau_{Q/G} : TQ/G \to M = Q/G$ with respect to the local coordinates (x^i) and to the local basis $\{e'_i, e'_a\}$ of $\Gamma(TQ/G)$ are

$$\begin{array}{c} C^k_{ij} = C^j_{ia} = -C^j_{ai} = C^i_{ab} = 0, \quad C^a_{ij} = -B^a_{ij}, \quad C^c_{ia} = -C^c_{ai} = c^c_{ab}A^b_i, \quad C^c_{ab} = c^c_{ab}\,, \\ \rho^j_i = \delta_{ij}, \quad \rho^a_i = \rho^i_a = \rho^b_a = 0\,. \end{array} \tag{3.17}$$

3.5.2 The inverse problem for Atiyah algebroids

In the case of an Atiyah algebroid the section $\Theta_{\Gamma,F} = \theta_\alpha \tilde{T}^\alpha + F_\alpha \tilde{V}^\alpha = \theta_i \tilde{T}^i + \theta_a \tilde{T}^a + F_i \tilde{V}^i + F_a \tilde{V}^a$ of $(\mathcal{L}^\tau E)^* \to E$ defined at the beginning of Section 3.3.2 has the following local components:

$$\begin{array}{rcl} \theta_i & = & \dfrac{\partial F_i}{\partial x^j}y^j + \dfrac{\partial F_i}{\partial y^j}\Gamma^j + \dfrac{\partial F_i}{\partial y^a}\Gamma^a - B^a_{ij}F_a y^j + c^c_{ab}A^b_i F_c y^a\,, \\ \theta_a & = & \dfrac{\partial F_a}{\partial x^j}y^j + \dfrac{\partial F_a}{\partial y^j}\Gamma^j + \dfrac{\partial F_a}{\partial y^b}\Gamma^b - c^c_{ab}A^b_i F_c y^i + c^c_{ab}F_c y^b\,. \end{array}$$

In this case the Helmholtz conditions, given by $d^{\mathcal{L}^{\tau_{Q/G}}}\Theta_{\Gamma,F} = 0$, are

$$\begin{array}{l} \dfrac{\partial F_\beta}{\partial y^\gamma} = \dfrac{\partial F_\gamma}{\partial y^\beta}\,, \\ \dfrac{\partial \theta_j}{\partial y^\beta} = \dfrac{\partial F_\beta}{\partial y^j}\,, \quad \dfrac{\partial \theta_b}{\partial y^\beta} = 0\,, \\ \dfrac{\partial \theta_i}{\partial x^j} + \theta_a B^a_{ji} = \dfrac{\partial \theta_j}{\partial x^i}\,, \quad \dfrac{\partial \theta_b}{\partial x^i} = \theta_a c^a_{bd}A^d_i\,, \quad \theta_c c^c_{ab} = 0\,, \end{array}$$

compared with (3.8) and (3.9). From the last two equations we conclude that $\partial\theta_b/\partial x^i = 0$. Thus θ_b is a constant function.

The extra condition for exactness of $\Theta_{\Gamma,F}$ is given by

$$\theta_a = \frac{\partial F_a}{\partial x^i}y^i + \frac{\partial F_a}{\partial y^\beta}\Gamma^\beta - c^d_{ab}A^b_i F_d y^i + c^d_{ab}F_d y^b = 0\,,$$

since $\mathrm{Ker}\rho = \mathrm{span}\{e'_a\}$. Therefore, $\Theta_{\Gamma,F}$ is exact if $\theta_a(x, y) = \theta_a = 0$. Recall that in the general Lie algebroid case from Section 3.3 the condition of exactness was $\theta_a(x, y) = \theta_a(x) = 0$.

Remark 3.5.1. Note that among the set of Helmholtz conditions, the equations

$$\frac{\partial \theta_b}{\partial y^\beta} = 0, \quad \frac{\partial \theta_b}{\partial x^i} = \theta_a c^a_{bd}A^d_i \quad \text{and} \quad \theta_c c^c_{ab} = 0$$

are immediately satisfied because of the extra condition $\theta_a = 0$. Hence, a set of necessary and sufficient conditions for variationality is

$$\frac{\partial F_\beta}{\partial y^\gamma} = \frac{\partial F_\gamma}{\partial y^\beta}, \quad \frac{\partial \theta_j}{\partial y^\beta} = \frac{\partial F_\beta}{\partial y^j}, \quad \frac{\partial \theta_i}{\partial x^j} = \frac{\partial \theta_j}{\partial x^i}, \quad \theta_a = 0\,,$$

in contrast to Example 3.3.5, where the Helmholtz conditions are necessary and sufficient and Example 3.3.6, where most of the Helmholtz conditions are implied by the extra condition on the kernel. In this example the kernel of ρ is neither trivial nor the whole domain, so we get less overlap between the two sets of conditions.

Remark 3.5.2. If we are given a SODE on TQ and a SODE on TQ/G related as in Theorem 3.4.1 then it is enough to solve the Helmholtz conditions on the Atiyah algebroid in order to get a Lagrangian on TQ since on tangent bundles the notions of variational and weak variational coincide.

Example 3.5.3. Let ξ be a SODE on $T\mathbb{R}^3$ given by

$$\xi = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{z}\frac{\partial}{\partial z} + (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\frac{\partial}{\partial \dot{x}}\,,$$

whose solutions are specified by the following system of second order differential equations:

$$\ddot{x} = \dot{x}^2 + \dot{y}^2 + \dot{z}^2\,, \quad \ddot{y} = 0\,, \quad \ddot{z} = 0\,. \tag{3.18}$$

It is easy to check that the previous equations are the Euler-Lagrange equations for the Lagrangian [13]

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{e^{-2x}\left(\dot{x}^2 - \dot{y}^2 - \dot{z}^2\right)}{\sqrt{\dot{y}^2 + \dot{z}^2}} + \dot{y}\dot{z}\,.$$

Now, Equations (3.18) are invariant by translations in $\mathbb{R}^3$, given by $(x, y, z) \longmapsto (x + a, y + b, z + c)$ with $(a, b, c) \in \mathbb{R}^3$, and the system (3.18) reduces to the following system of differential equations in $T\mathbb{R}^3/\mathbb{R}^3 \equiv \mathbb{R}^3$:

$$\dot{y}^1 = (y^1)^2 + (y^2)^2 + (y^3)^2\,, \quad \dot{y}^2 = 0\,, \quad \dot{y}^3 = 0\,. \tag{3.19}$$

It is easy to see that these reduced equations are not variational since conditions (3.7) are

$$\frac{\partial F_1}{\partial y^1}\left((y^1)^2 + (y^2)^2 + (y^3)^2\right) = \frac{\partial F_2}{\partial y^1}\left((y^1)^2 + (y^2)^2 + (y^3)^2\right) = \frac{\partial F_3}{\partial y^1}\left((y^1)^2 + (y^2)^2 + (y^3)^2\right) = 0\,.$$

Therefore, we obtain $\frac{\partial F_1}{\partial y^1} = \frac{\partial F_2}{\partial y^1} = \frac{\partial F_3}{\partial y^1} = 0$, which is not compatible with the nondegeneracy condition.

Another option is to reduce the system using the action by translations in the y and z direction, that is, $(x, y, z) \longmapsto (x, y + a, z + b)$ with $(a, b) \in \mathbb{R}^2$. The reduced space is the Atiyah algebroid $T\mathbb{R}^3/\mathbb{R}^2 \equiv \mathbb{R}^4 \to \mathbb{R}$. Given the basis of sections $e_1(x) = (x; (1, 0, 0))$, $e_2(x) = (x; (0, 1, 0))$ and $e_3(x) = (x; (0, 0, 1))$ with Lie bracket of sections $[\![e_i, e_j]\!] = 0$, $1 \le i, j \le 3$, we have induced coordinates (x, y^1, y^2, y^3) on $\mathbb{R}^4$ and the anchor map $\rho : \mathbb{R}^4 \to T\mathbb{R}$ is given by $\rho(x, y^1, y^2, y^3) = (x, y^1)$.

The reduced equations in this vector bundle $\mathbb{R}^4 \to \mathbb{R}$ are

$$\dot{y}^1 = (y^1)^2 + (y^2)^2 + (y^3)^2\,, \quad \dot{y}^2 = 0\,, \quad \dot{y}^3 = 0 \;\text{ and }\; \dot{x} = y^1\,,$$

or, equivalently,

$$\ddot{x} = \dot{x}^2 + (y^2)^2 + (y^3)^2\,, \quad \dot{y}^2 = 0\,, \quad \dot{y}^3 = 0\,.$$

As mentioned before, these equations are variational with the Lagrangian

$$l(x, y^1, y^2, y^3) = \frac{e^{-2x}\left((y^1)^2 - (y^2)^2 - (y^3)^2\right)}{\sqrt{(y^2)^2 + (y^3)^2}} + y^2y^3\,.$$

Example 3.5.4. Consider the SODE given by

$$\ddot{x} = -x, \quad \ddot{y} = -y\,, \tag{3.20}$$

which describes the dynamics of a two-dimensional harmonic oscillator [108]. These equations are the Euler-Lagrange equations corresponding to the Lagrangian $L : T(\mathbb{R}^2\backslash\{(0,0)\}) \to \mathbb{R}$, $L(x,y,\dot{x},\dot{y}) = \frac{1}{2}\left(\dot{x}^2 - \dot{y}^2\right) - \frac{1}{2}\left(x^2 - y^2\right)$.

It is easy to check that the solutions of the SODE (3.20) are invariant under the flow of the vector field $X = -x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. Observe that the integral curves of X are $\sigma(t) = (C_1 e^{-t}, C_2 e^{t})$ and the corresponding flow is $\Psi_a : \mathbb{R}^2\backslash\{(0,0)\} \to \mathbb{R}^2\backslash\{(0,0)\}$, $\Psi_a(x,y) = (xe^{-a}, ye^{a})$, $a \in \mathbb{R}$. Then, if $(x(t), y(t))$ is a solution of the SODE (3.20) then so is $\Psi_a(x(t), y(t))$. The Lagrangian L is not invariant since $X^c(L) = (y^2 + x^2) - (\dot{y}^2 + \dot{x}^2)$, where $X^c = -x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y} - \dot{x}\dfrac{\partial}{\partial \dot{x}} + \dot{y}\dfrac{\partial}{\partial \dot{y}}$.

In order to simplify the reduction procedure, consider the change of coordinates $x = ue^{v}$ and $y = ue^{-v}$. In this new set of coordinates the vector field X is rewritten as

$$X = -\frac{\partial}{\partial v}\,.$$

Now the reduction is trivial. The reduced space is an Atiyah algebroid $\tau : (\mathbb{R}\backslash\{0\}) \times \mathbb{R}^2 \to \mathbb{R}\backslash\{0\}$ with coordinates $(u; y^1, y^2)$ induced by a basis of sections $e_1(u) = (u; (1,0))$ and $e_2(u) = (u; (0,1))$ with Lie bracket of sections $[\![e_i, e_j]\!] = 0$, $1 \leq i, j \leq 2$. The projection is given by $\tau(u; y^1, y^2) = u$ and the anchor map $\rho : (\mathbb{R}\backslash\{0\}) \times \mathbb{R}^2 \to T(\mathbb{R}\backslash\{0\})$ is $\rho(u, y^1, y^2) = (u, y^1)$. Therefore, the reduced equations are

$$\begin{aligned} \frac{dy^1}{dt} + 2y^1y^2 + u\frac{dy^2}{dt} + u(y^2)^2 &= -u\,,\\ \frac{dy^1}{dt} - 2y^1y^2 - u\frac{dy^2}{dt} + u(y^2)^2 &= -u\,,\\ y^1 &= \frac{du}{dt}\,. \end{aligned}$$

Even though the Lagrangian L does not reduce to this particular Atiyah algebroid, the equations have a variational nature since they are the Euler-Lagrange equations for the alternative Lagrangian

$$\tilde{l}(u; y^1, y^2) = (y^1)^2 - u^2(y^2)^2 - u^2\,.$$

Observe that $\tilde{l}$ is the reduction of the alternative Lagrangian $\tilde{L}(x,y,\dot{x},\dot{y}) = \dot{x}\dot{y} - xy$ for Equations (3.20) which is now invariant by X (see [108]).

3.6 Relation to other approaches

In Section 3.3.2 we recover the Helmholtz conditions given in [131] as the vanishing of $d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F}$ on a certain basis of sections of $\mathcal{L}^\tau E \longrightarrow E$.

In the previous section we worked in the basis $\{\tilde{T}_\alpha, \tilde{V}_\alpha\}$ of local sections of $\mathcal{L}^\tau E$, constructed from a basis $\{e_\alpha\}$ of local sections of E. Another common basis of sections of $\mathcal{L}^\tau E$ is $\{e_\alpha^C, e_\alpha^V\}$, the set of complete and vertical lifts of $\{e_\alpha\}$. The relationship between both is

$$\tilde{T}_\alpha = e_\alpha^C + C_{\alpha\beta}^\gamma y^\beta \tilde{V}_\gamma \quad \text{and} \quad \tilde{V}_\gamma = e_\gamma^V\,.$$

As in the tangent bundle case, a SODE on a Lie algebroid defines a connection (see [131]). Then the horizontal lift of a section $X \in \Gamma(E)$ can be defined from its complete and vertical lift and the SODE as

$$X^H = \frac{1}{2}\left(X^C - [\Gamma, X^V]\right),$$

and we get another basis $\{H_\alpha := e_\alpha^H, e_\alpha^V\}$ of sections of $\mathcal{L}^\tau E$. The relationship with the above is given by

$$H_\alpha = e_\alpha^H = \tilde{T}_\alpha + \frac{1}{2}\left(\frac{\partial \Gamma^\gamma}{\partial y^\alpha} - C_{\alpha\beta}^\gamma y^\beta\right)\tilde{V}_\gamma = \tilde{T}_\alpha + \Lambda_\alpha^\gamma \tilde{V}_\gamma .$$

Note that if Γ is variational we have $\mathcal{L}_\Gamma F^*\lambda_E = \Theta_{\Gamma,F}$ for some local diffeomorphism F and hence $\mathcal{L}_\Gamma d^{\mathcal{L}^\tau E} F^*\lambda_E = d^{\mathcal{L}^\tau E}\Theta_{\Gamma,F}$. Then the equations

$$\mathcal{L}_\Gamma d^{\mathcal{L}^\tau E} F^*\lambda_E(H_\eta, H_\beta) = 0, \quad \mathcal{L}_\Gamma d^{\mathcal{L}^\tau E} F^*\lambda_E(H_\eta, \tilde{V}_\beta) = 0 \text{ and } \mathcal{L}_\Gamma d^{\mathcal{L}^\tau E} F^*\lambda_E(\tilde{V}_\eta, \tilde{V}_\beta) = 0\,,$$

together with

$$\mathcal{L}_\Gamma d^{\mathcal{L}^\tau E} F^*\lambda_E(H_\eta, \tilde{V}_\beta) - \mathcal{L}_\Gamma d^{\mathcal{L}^\tau E} F^*\lambda_E(H_\beta, \tilde{V}_\eta) = 0$$

yield the Helmholtz conditions given in [131].

In order to check this we first compute $[\Gamma, \tilde{V}_\eta]$ and $[\Gamma, H_\eta]$ in terms of the basis $\{H_\alpha, \tilde{V}_\alpha\}$:

$$[\Gamma, \tilde{V}_\eta] = -\tilde{T}_\eta - \frac{\partial \Gamma^\alpha}{\partial y^\eta}\tilde{V}_\alpha = -(H_\eta - \Lambda_\eta^\beta \tilde{V}_\beta) - \frac{\partial \Gamma^\alpha}{\partial y^\eta}\tilde{V}_\alpha = -H_\eta + \frac{1}{2}\left(C_{\beta\eta}^\gamma y^\beta - \frac{\partial \Gamma^\gamma}{\partial y^\eta}\right)\tilde{V}_\gamma\,,$$

$$\begin{aligned}
[\Gamma, H_\eta] &= [\Gamma, \tilde{T}_\eta] + [\Gamma, \Lambda_\eta^\gamma \tilde{V}_\gamma] = [\Gamma, \tilde{T}_\eta] + \rho(\Gamma)(\Lambda_\eta^\gamma)\tilde{V}_\gamma + \Lambda_\eta^\gamma[\Gamma, \tilde{V}_\gamma] \\
&= -\left(\rho_\eta^i \frac{\partial \Gamma^\alpha}{\partial x^i}\tilde{V}_\alpha + y^\alpha C_{\eta\alpha}^\gamma (H_\gamma - \Lambda_\gamma^\nu \tilde{V}_\nu)\right) + \left(y^\alpha \rho_\alpha^i \frac{\partial \Lambda_\eta^\gamma}{\partial x^i} + \Gamma^\alpha \frac{\partial \Lambda_\eta^\gamma}{\partial y^\alpha}\right)\tilde{V}_\gamma \\
&\quad + \Lambda_\eta^\gamma\left(-H_\gamma + \frac{1}{2}\left(C_{\beta\gamma}^\nu y^\beta - \frac{\partial \Gamma^\nu}{\partial y^\gamma}\right)\tilde{V}_\nu\right) = \frac{1}{2}\left(y^\beta C_{\beta\alpha}^\gamma - \frac{\partial \Gamma^\gamma}{\partial y^\alpha}\right)H_\gamma \\
&\quad + \left(y^\alpha \rho_\alpha^i \frac{\partial \Lambda_\eta^\gamma}{\partial x^i} + \Gamma^\alpha \frac{\partial \Lambda_\eta^\gamma}{\partial y^\alpha} + \Lambda_\eta^\nu \Lambda_\nu^\gamma - \Lambda_\eta^\nu \frac{\partial \Gamma^\gamma}{\partial y^\nu} - \rho_\eta^i \frac{\partial \Gamma^\gamma}{\partial x^i} + y^\alpha C_{\eta\alpha}^\nu \Lambda_\nu^\gamma\right)\tilde{V}_\gamma\,.
\end{aligned}$$

We introduce the notation

$$D_\eta^\gamma := \frac{1}{2}\left(y^\beta C_{\beta\alpha}^\gamma - \frac{\partial \Gamma^\gamma}{\partial y^\alpha}\right) \text{ and } \Phi_\eta^\gamma := \left(y^\alpha \rho_\alpha^i \frac{\partial \Lambda_\eta^\gamma}{\partial x^i} + \Gamma^\alpha \frac{\partial \Lambda_\eta^\gamma}{\partial y^\alpha} + \Lambda_\eta^\nu \Lambda_\nu^\gamma - \Lambda_\eta^\nu \frac{\partial \Gamma^\gamma}{\partial y^\nu} - \rho_\eta^i \frac{\partial \Gamma^\gamma}{\partial x^i} + y^\alpha C_{\eta\alpha}^\nu \Lambda_\nu^\gamma\right)$$

so that $[\Gamma, \tilde{V}_\eta] = -H_\eta + D_\eta^\gamma \tilde{V}_\gamma$ and $[\Gamma, H_\eta] = D_\eta^\gamma H_\gamma + \Phi_\eta^\gamma \tilde{V}_\gamma$.

We will also need the expression of $d^{\mathcal{L}^\tau E} F^*\lambda_E$ in terms of $\{\theta^\alpha := \tilde{V}^\alpha - \Lambda_\beta^\alpha \tilde{T}^\beta, \tilde{T}^\alpha\}$, the dual basis of $\{H_\alpha, \tilde{V}_\alpha\}$:

$$d^{\mathcal{L}^\tau E} F^*\lambda_E = \left(\rho(H_\gamma)(F_\alpha) - \frac{1}{2}F_\nu C_{\gamma\alpha}^\nu\right)\tilde{T}^\gamma \wedge \tilde{T}^\alpha + \frac{\partial F_\alpha}{\partial y^\gamma}\theta^\gamma \wedge \tilde{T}^\alpha = A_{\gamma\alpha}\tilde{T}^\gamma \wedge \tilde{T}^\alpha + \frac{\partial F_\alpha}{\partial y^\gamma}\theta^\gamma \wedge \tilde{T}^\alpha$$

where $A_{\gamma\alpha} = \rho(H_\gamma)(F_\alpha) - \frac{1}{2}F_\nu C_{\gamma\alpha}^\nu$.

Now we introduce the notation $T_F := d^{\mathcal{L}^\tau E} F^*\lambda_E$ and write the Helmholtz conditions in local coordinates as follows:

$$\mathcal{L}_\Gamma T_F(H_\eta, H_\beta) = \Gamma(T_F(H_\eta, H_\beta)) - T_F([\Gamma, H_\eta], H_\beta) - T_F(H_\eta, [\Gamma, H_\beta])$$

$$\begin{aligned} &= \Gamma(A_{\eta\beta}) - \Gamma(A_{\beta\eta}) - \left[A_{\gamma\beta}D^\gamma_\eta - A_{\beta\gamma}D^\gamma_\eta + \frac{\partial F_\beta}{\partial y^\gamma}\Phi^\gamma_\eta\right] \\ &\quad - \left[A_{\eta\gamma}D^\gamma_\beta - A_{\gamma\eta}D^\gamma_\beta - \frac{\partial F_\eta}{\partial y^\gamma}\Phi^\gamma_\beta\right] = 0\,, \end{aligned} \tag{3.21}$$

$$\begin{aligned} \mathcal{L}_\Gamma T_F(H_\eta, \tilde V_\beta) &= \Gamma(T_F(H_\eta, \tilde V_\beta)) - T_F([\Gamma, H_\eta], \tilde V_\beta) - T_F(H_\eta, [\Gamma, \tilde V_\beta]) \\ &= \Gamma\left(-\frac{\partial F_\eta}{\partial y^\beta}\right) + \frac{\partial F_\gamma}{\partial y^\beta}D^\gamma_\eta - \left[-A_{\eta\beta} + A_{\beta\eta} - \frac{\partial F_\eta}{\partial y^\gamma}D^\gamma_\beta\right] = 0\,, \end{aligned} \tag{3.22}$$

$$\begin{aligned} \mathcal{L}_\Gamma T_F(\tilde V_\eta, \tilde V_\beta) &= -T_F([\Gamma, \tilde V_\eta], \tilde V_\beta) - T_F(\tilde V_\eta, [\Gamma, \tilde V_\beta]) \\ &= T_F(H_\eta, \tilde V_\beta) + T_F(\tilde V_\eta, H_\beta) = -\frac{\partial F_\eta}{\partial y^\beta} + \frac{\partial F_\beta}{\partial y^\eta} = 0\,. \end{aligned} \tag{3.23}$$

Now we compute

$$\begin{aligned} \mathcal{L}_\Gamma T_F(H_\eta, \tilde V_\beta) - \mathcal{L}_\Gamma T_F(H_\beta, \tilde V_\eta) &= \Gamma\left(-\frac{\partial F_\eta}{\partial y^\beta}\right) + \frac{\partial F_\gamma}{\partial y^\beta}D^\gamma_\eta - \left[-A_{\eta\beta} + A_{\beta\eta} - \frac{\partial F_\eta}{\partial y^\gamma}D^\gamma_\beta\right] \\ &\quad -\Gamma\left(-\frac{\partial F_\beta}{\partial y^\eta}\right) - \frac{\partial F_\gamma}{\partial y^\eta}D^\gamma_\beta + \left[-A_{\beta\eta} + A_{\eta\beta} - \frac{\partial F_\beta}{\partial y^\gamma}D^\gamma_\eta\right] \end{aligned}$$

and use (3.23) to obtain

$$A_{\eta\beta} = A_{\beta\eta}\,. \tag{3.24}$$

Substituting (3.24) into (3.21) and (3.22) these equations become

$$\frac{\partial F_\beta}{\partial y^\gamma}\Phi^\gamma_\eta = \frac{\partial F_\eta}{\partial y^\gamma}\Phi^\gamma_\beta \quad \text{and} \quad \Gamma\left(\frac{\partial F_\eta}{\partial y^\beta}\right) - \frac{\partial F_\gamma}{\partial y^\beta}D^\gamma_\eta - \frac{\partial F_\eta}{\partial y^\gamma}D^\gamma_\beta = 0\,,$$

which are the equations given in [131]. This can be checked directly by making the substitution

$$\Lambda^\nu_\eta\frac{\partial \Gamma^\gamma}{\partial y^\nu} = 2\Lambda^\nu_\eta\Lambda^\gamma_\nu + \Lambda^\nu_\eta C^\gamma_{\nu\tau}y^\tau\,.$$

Beware that the notation in [131] is $N^\gamma_\eta = -\Lambda^\gamma_\eta$.

Note that the Helmholtz conditions for invariant Lagrangians on the tangent bundle of a Lie group G given in [45] are also recovered. Indeed, by dropping the terms where derivatives with respect to x^i appear and substituting $\Lambda^\gamma_\eta = \frac{\partial\Gamma^\gamma}{\partial y^\eta} - C^\gamma_{\eta\beta}y^\beta$ we get

$$\Phi^\gamma_\eta = \frac{1}{2}\Gamma^\alpha\frac{\partial^2\Gamma^\gamma}{\partial y^\alpha\partial y^\eta} - \frac{1}{2}\Gamma^\alpha C^\gamma_{\eta\alpha} - \frac{1}{4}\frac{\partial\Gamma^\nu}{\partial y^\eta}\frac{\partial\Gamma^\gamma}{\partial y^\nu} - \frac{1}{4}C^\nu_{\eta\beta}y^\beta C^\gamma_{\nu\tau}y^\tau - \frac{1}{4}\frac{\partial\Gamma^\nu}{\partial y^\eta}C^\gamma_{\nu\tau}y^\tau + \frac{3}{4}C^\nu_{\eta\beta}y^\beta\frac{\partial\Gamma^\gamma}{\partial y^\nu}\,.$$

Chapter 4

Inverse problem for second order difference equations

In this chapter we will explore the inverse problem for discrete systems, that is, second order difference equations. We are particularly interested in the case when such systems are numerical integrators for a continuous system, and it will be common to find them in implicit form. In the discrete case an implicit system of second order difference equations will be given by a submanifold $M \subset Q\times Q\times Q$. We will assume that the submanifold M can be described as the vanishing of functions $\Phi^i(q_{k-1}, q_k, q_{k+1})$, $i = 1, \ldots, n$, such that the matrix $\left(\frac{\partial \Phi}{\partial q_{k+1}}\right)$ is regular. Then a natural discrete formulation of the classical inverse problem would be to ask whether or not it is possible to find a regular discrete Lagrangian $L_d : Q \times Q \longrightarrow \mathbb{R}$ such that both systems

$$\Phi^i(q_{k-1}, q_k, q_{k+1}) = 0 \quad \text{and} \quad D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k) = 0$$

admit the same solutions. A different and more restrictive version of the problem, which is concerned with the equality

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

has been addressed in [23, 42, 88].

In Section 4.1 we give a brief introduction to discrete mechanics and variational integrators. In Section 4.2 we define variational second order difference equations and give a characterization in terms of Lagrangian submanifolds of $T^*Q \times T^*Q$. We also give an analog of Crampin's Theorem and discuss the one-dimensional case. In Section 4.3 we show how to go from the continuous to the discrete setting. Finally in Section 4.4 we show how two alternative discrete Lagrangian formulations can lead to constants of motion. This may be applied to obtain integrable discretizations of integrable systems.

4.1 Variational integrators and discrete mechanics

We will now recall, following mostly [112], how to derive the discrete Euler-Lagrange equations and obtain discretizations of the continuous equations of motion by discretizing Hamilton's variational principle, instead of directly discretizing the equations.

We will consider $Q \times Q$ as a discrete version of TQ and therefore $Q \times Q \times Q \times Q$ as a discrete analogue of TTQ, see [112]. Instead of curves on Q, the solutions are replaced by sequences of points on Q. If we fix some $N \in \mathbb{N}$ then we use the notation

$$\mathcal{C}_d(Q) = \Big\{ q_d : \{k\}_{k=0}^N \longrightarrow Q \Big\}$$

for the set of possible solutions, which can be identified with the manifold $Q \times \overset{(N+1)}{\cdots} \times Q$. Define a functional, the discrete action map, on the space of sequences $\mathcal{C}_d(Q)$ by

$$S_d(q_d) = \sum_{k=0}^{N-1} L_d(q_k, q_{k+1}), \quad q_d \in \mathcal{C}_d(Q)\,.$$

If we consider variations of q_d with fixed end points q_0 and q_N and extremize S_d over $q_1, \ldots, q_{N-1}$, we obtain the discrete Euler-Lagrange equations (DEL for short)

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

where $D_1 L_d(q_{k-1}, q_k) \in T^*_{q_{k-1}}Q$ and $D_2 L_d(q_{k-1}, q_k) \in T^*_{q_k}Q$ correspond to $dL_d(q_{k-1}, q_k)$ under the identification $T^*_{(q_{k-1},q_k)}(Q \times Q) \cong T^*_{q_{k-1}}Q \times T^*_{q_k}Q$.

If L_d is regular, that is, $D_{12}L_d$ is regular, then we obtain a well defined discrete Lagrangian map

$$\begin{array}{rccc} F_{L_d}: & Q \times Q & \longrightarrow & Q \times Q \\ & (q_{k-1}, q_k) & \longmapsto & (q_k, q_{k+1}(q_{k-1}, q_k))\,. \end{array}$$

If L_d is regular then we can further assure that the discrete Lagrangian map is invertible, so we can also write $q_{k-1} = q_{k-1}(q_k, q_{k+1})$, see [112, Theorem 1.5.1].

In this setting we can define two discrete Legendre transformations

$$\mathbb{F}^+ L_d, \mathbb{F}^- L_d : Q \times Q \longrightarrow T^*Q\,,$$

since each projection is equally eligible for the base point. They can be defined intrinsically and are given in coordinates as

$$\begin{aligned} \mathbb{F}^+ L_d &: (q_{k-1}, q_k) \longmapsto (q_k, D_2 L_d(q_{k-1}, q_k))\,, \\ \mathbb{F}^- L_d &: (q_{k-1}, q_k) \longmapsto (q_{k-1}, -D_1 L_d(q_{k-1}, q_k))\,. \end{aligned}$$

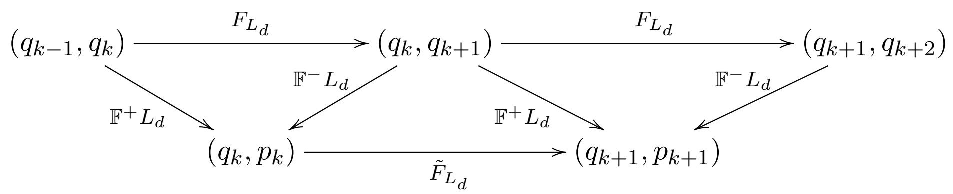
It holds that $(\mathbb{F}^+ L_d)^* \omega_Q = (\mathbb{F}^- L_d)^* \omega_Q =: \Omega_{L_d}$, with local expression

$$\Omega_{L_d}(q_{k-1}, q_k) = \frac{\partial^2 L_d}{\partial q^i_{k-1} \partial q^j_k} dq^i_{k-1} \wedge dq^j_k\,.$$

We can also define the evolution of the discrete system on the Hamiltonian side, $\tilde{F}_{L_d} : T^*Q \longrightarrow T^*Q$, by any of the formulas

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

because of the commutativity of the following diagram:



The discrete Hamiltonian map $\tilde{F}_{L_d} : (T^*Q, \omega_Q) \longrightarrow (T^*Q, \omega_Q)$ is symplectic. Therefore the submanifold

$$\begin{aligned} (q_k, p_k, q_{k+1}, p_{k+1}) &= \big(q_k, \mathbb{F}^- L_d(q_k, q_{k+1}), q_{k+1}, \mathbb{F}^- L_d(q_{k+1}, q_{k+2})\big) \\ &= \big(q_k, \mathbb{F}^+ L_d(q_{k-1}, q_k), q_{k+1}, \mathbb{F}^+ L_d(q_k, q_{k+1})\big) \end{aligned}$$

is Lagrangian in $(T^*Q \times T^*Q, \Omega_Q)$, where $\Omega_Q := pr_2^*\omega_Q - pr_1^*\omega_Q$ is a symplectic form.

So far we have taken as the starting point a discrete Lagrangian $L_d : Q \times Q \longrightarrow \mathbb{R}$. However, if we start with a continuous Lagrangian and take an appropriate discrete Lagrangian then the DEL become a geometric integrator for the continuous Euler-Lagrange system, known as a variational integrator. Hence, given a regular Lagrangian function $L : TQ \longrightarrow \mathbb{R}$, we define a discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$ as an approximation to the action of the continuous Lagrangian. More precisely, for a regular Lagrangian L, and appropriate h, q_0, q_1, we can define the exact discrete Lagrangian as

$$L_d^E(q_0, q_1, h) = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))dt\,,$$

where $q_{0,1}(t)$ is the unique solution of the Euler-Lagrange equations for L satisfying $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$, see [79, 110]. Then for a sufficiently small h, the solutions of the DEL for L_d^E lie on the solutions of the Euler-Lagrange equations for L, see [112, Theorem 1.6.4].

In practice, $L_d^E(q_0, q_1, h)$ will not be available. Therefore we will take

$$L_d(q_0, q_1, h) \approx L_d^E(q_0, q_1, h)\,,$$

using some quadrature rule. We obtain symplectic integrators in this way, see [129]. A discrete version of Noether's theorem also holds, [112, Theorem 1.3.3]. Therefore, if the discrete Lagrangian is invariant under the same group action as the continuous Lagrangian then we obtain symplectic-momentum integrators.

Some properties of the symplectic integrator we obtain, the discrete Hamiltonian map $\tilde{F}_{L_d}$, such as the order or self-adjointness, can be deduced from properties of the discrete Lagrangian, see [112, 129].

Example 4.1.1. If Q is a vector space then the discrete Lagrangians

$$\begin{aligned} L_d^{\frac{1}{2}}(q_0, q_1, h) &= hL\left(\frac{q_0 + q_1}{2}, \frac{q_1 - q_0}{h}\right), \\ L_d^{sym,\alpha}(q_0, q_1, h) &= \frac{h}{2}L\left((1-\alpha)q_0 + \alpha q_1, \frac{q_1 - q_0}{h}\right) + \frac{h}{2}L\left(\alpha q_0 + (1-\alpha)q_1, \frac{q_1 - q_0}{h}\right), \end{aligned}$$

provide second order methods, for $\alpha \in [0, 1]$. Actually for a regular Lagrangian L, the first discrete Lagrangian $L_d^{\frac{1}{2}}$ has as discrete Hamiltonian map the midpoint rule on T^*Q for the corresponding Hamiltonian system, which is a well-known symplectic integrator, given by

$$\begin{aligned} \frac{q_1 - q_0}{h} &= \frac{\partial H}{\partial p}\left(\frac{q_1 + q_0}{2}, \frac{p_1 + p_0}{2}\right), \\ \frac{p_1 - p_0}{h} &= -\frac{\partial H}{\partial q}\left(\frac{q_1 + q_0}{2}, \frac{p_1 + p_0}{2}\right), \end{aligned}$$

where H is the Hamiltonian function corresponding to L. According to [112, Theorem 2.1.1], any symplectic integrator for the Hamiltonian equations derived from a regular Lagrangian can be locally seen as a discrete Hamiltonian map associated to some discrete Lagrangian L_d. See [112, Section 2.6] for further examples, including the Störmer-Verlet method and symplectic partitioned Runge-Kutta methods. The corresponding discrete Lagrangians are provided.

4.2 Inverse problem in the discrete setting

In this section we will provide definitions of variationality for explicit and implicit second order difference equations. We will also derive Helmholtz type conditions and prove an analog of Theorem 1.4.1 in this setting.

We will consider separately the explicit case, in which the second order difference equation is given as a map

$$\begin{array}{rccc} \Gamma : & Q\times Q & \longrightarrow & Q\times Q\times Q\times Q \\ & (q_{k-1},q_k) & \longmapsto & (q_{k-1},q_k,q_k,\tilde{\Gamma}(q_{k-1},q_k))\,, \end{array}$$

and the implicit case, in which the second order difference equation is given as a submanifold M of $Q\times Q\times Q$ satisfying some regularity condition.

4.2.1 Explicit second order difference equations

Recall that we consider $Q\times Q$ as a discretization of TQ and $Q\times Q\times Q\times Q$ as a discretization of TTQ. The discrete second order submanifold is given by

$$\ddot{Q}_d = \{\omega_d \in (Q\times Q)\times(Q\times Q) : pr_1\circ pr_2(\omega_d) = pr_2\circ pr_1(\omega_d)\} \cong Q\times Q\times Q\,.$$

A map $\Gamma : Q\times Q \longrightarrow \ddot{Q}_d \subset (Q\times Q)\times(Q\times Q)$ satisfying $pr_1\circ\Gamma = Id$ will be referred to as an explicit second order difference equation (SOdE for short).

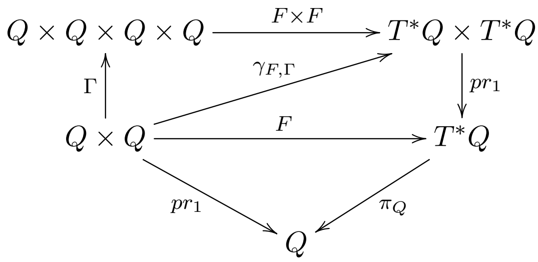
Definition 4.2.1. *The explicit second order difference equation $q_{k+1} = \tilde{\Gamma}(q_{k-1},q_k)$ is variational if and only if there is a locally defined regular discrete Lagrangian $L_d : Q\times Q \longrightarrow \mathbb{R}$ such that both systems*

$$q_{k+1} = \tilde{\Gamma}(q_{k-1},q_k) \quad \textit{and} \quad D_1L_d(q_k,q_{k+1}) + D_2L_d(q_{k-1},q_k) = 0$$

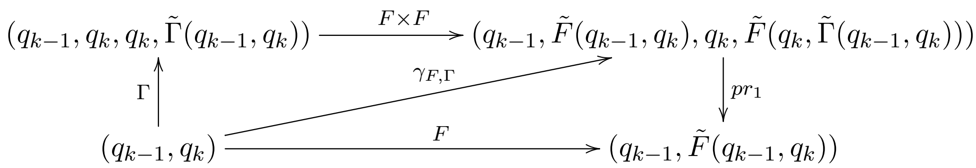
admit the same solutions.

To avoid technical difficulties we are assuming that (q_{k-1},q_k) and $(q_k,\tilde{\Gamma}(q_{k-1},q_k))$ belong to the same neighborhood where L_d is defined.

Consider first $pr_1 : Q\times Q \longrightarrow Q$ substituting the projection $\tau_Q : TQ \longrightarrow Q$ (later in Proposition 4.2.6 we will see that we could also have chosen $pr_2 : Q\times Q \longrightarrow Q$ as a discretization). For a given explicit second order difference equation $q_{k+1} = \tilde{\Gamma}(q_{k-1},q_k)$ and a local diffeomorphism $F : Q\times Q \longrightarrow T^*Q$ over the identity, we define $\gamma_{F,\Gamma} := (F\times F)\circ\Gamma$ as shown in the following commutative diagram:



For (q_{k-1}, q_k) on $Q \times Q$ the diagram is the following:



Proposition 4.2.2. *The second order difference equation $q_{k+1} = \tilde{\Gamma}(q_{k-1}, q_k)$ is variational if and only if there is a local diffeomorphism $F : Q \times Q \longrightarrow T^*Q$ satisfying $pr_1 = \pi_Q \circ F$ and such that $Im(\gamma_{F,\Gamma})$ is a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$.*

Proof. Assume there is an F as in the statement. Then $\mathrm{Im}(\gamma_{F,\Gamma})$ is a submanifold of half the dimension of $T^*Q \times T^*Q$ and $\mathrm{Im}(\gamma_{F,\Gamma})$ is a Lagrangian submanifold if the isotropy condition $\gamma^*_{F,\Gamma}\Omega_Q = 0$ is satisfied. Since

$$\gamma^*_{F,\Gamma}\Omega_Q = d((pr_2 \circ \gamma_{F,\Gamma})^*\theta_Q - (pr_1 \circ \gamma_{F,\Gamma})^*\theta_Q)$$

is an exact two-form on $Q \times Q$, by the Poincaré Lemma the condition $\gamma^*_{F,\Gamma}\Omega_Q = 0$ implies

$$(pr_2 \circ \gamma_{F,\Gamma})^*\theta_Q - (pr_1 \circ \gamma_{F,\Gamma})^*\theta_Q = dL_d$$

for a locally defined map $L_d : Q \times Q \longrightarrow \mathbb{R}$ called the discrete Lagrangian.

In local coordinates we get

$$-F_i(q_{k-1}, q_k)dq^i_{k-1} + F_i(q_k, \tilde{\Gamma}(q_{k-1}, q_k))dq^i_k = \frac{\partial L_d}{\partial q^i_{k-1}}(q_{k-1}, q_k)dq^i_{k-1} + \frac{\partial L_d}{\partial q^i_k}(q_{k-1}, q_k)dq^i_k \ ,$$

that is, $D_1L_d(q_{k-1}, q_k) = -F(q_{k-1}, q_k)$ and $D_2L_d(q_{k-1}, q_k) = F(q_k, \tilde{\Gamma}(q_{k-1}, q_k))$. In particular $F = \mathbb{F}^-L_d$ and the admissibility condition $\tilde{F}(q_k, q_{k+1}) = \tilde{F}(q_k, \tilde{\Gamma}(q_{k-1}, q_k))$ gives the discrete Euler-Lagrange equations $-D_1L_d(q_k, q_{k+1}) = D_2L_d(q_{k-1}, q_k)$, see [91, Section 3.2].

Assume now that $q_{k+1} = \tilde{\Gamma}(q_{k-1}, q_k)$ is variational. Then we can define $F(q_{k-1}, q_k) = -D_1L_d(q_{k-1}, q_k)$ to get

$$\begin{aligned}
\{D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k) = 0\} &\equiv \left\{q_{k+1} = \tilde{\Gamma}(q_{k-1}, q_k)\right\} \\
&\equiv \left\{-D_1L_d(q_k, q_{k+1}) = -D_1L_d(q_k, \tilde{\Gamma}(q_{k-1}, q_k))\right\} ,
\end{aligned}$$

which implies

$$-D_2L_d(q_{k-1}, q_k) = D_1L_d(q_k, \tilde{\Gamma}(q_{k-1}, q_k))\,.$$

Then $\mathrm{Im}(\gamma_{F,\Gamma})$ is given by

$$\left(q_{k-1}, -D_1L_d(q_{k-1},q_k), q_k, -D_1L_d(q_k,\tilde{\Gamma}(q_{k-1},q_k))\right) = (q_{k-1}, -D_1L_d(q_{k-1},q_k), q_k, D_2L_d(q_{k-1},q_k))$$

which is clearly a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$. ■

Remark 4.2.3. Notice that we can equivalently work with Lagrangian submanifolds of $T^*(Q \times Q)$. First consider the symplectomorphism

$$\begin{array}{rccc} \Psi: & (T^*(Q\times Q), \omega_{Q\times Q}) & \longrightarrow & (T^*Q\times T^*Q, \Omega_Q) \\ & (\alpha_{q_0}, \alpha_{q_1}) & \longmapsto & (-\alpha_{q_0}, \alpha_{q_1}) \end{array}$$

and define the one-form $\Psi^{-1}\circ\gamma_{F,\Gamma}$ on $Q\times Q$. Then ask for it to be closed, that is, for $\mathrm{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma})$ to be a Lagrangian submanifold of $(T^*(Q\times Q), \omega_{Q\times Q})$.

If we impose $\mathrm{Im}(\gamma_{F,\Gamma})$ to be a Lagrangian submanifold of $(T^*Q\times T^*Q, \Omega_Q)$ for a given SOdE Γ then we get the following conditions on F:

$$\frac{\partial F_i}{\partial q^j_{k-1}}(q_{k-1},q_k) = \frac{\partial F_j}{\partial q^i_{k-1}}(q_{k-1},q_k)\,, \tag{4.1}$$

$$\frac{\partial F_i}{\partial Q^j_2}(q_{k-1},q_k) + \frac{\partial F_j}{\partial Q^l_2}(q_k,\Gamma_{k-1,k})\frac{\partial \Gamma^l}{\partial q^i_{k-1}} = 0\,, \tag{4.2}$$

$$\frac{\partial F_i}{\partial Q^j_1}(q_k,\Gamma_{k-1,k}) + \frac{\partial F_i}{\partial Q^l_2}(q_k,\Gamma_{k-1,k})\frac{\partial \Gamma^l}{\partial q^j_k} = \frac{\partial F_j}{\partial Q^i_1}(q_k,\Gamma_{k-1,k}) + \frac{\partial F_j}{\partial Q^l_2}(q_k,\Gamma_{k-1,k})\frac{\partial \Gamma^l}{\partial q^i_k}\,, \tag{4.3}$$

where $\Gamma_{k-1,k}$ is short notation for $\tilde{\Gamma}(q_{k-1},q_k)$ and $\partial/\partial Q_1$, $\partial/\partial Q_2$ denote partial derivatives with respect to the first and second slot respectively. We will refer to these equations as **discrete Helmholtz conditions**.

Since we are assuming $\Gamma: U \longrightarrow U\times U$, for some open subset $U\subset Q\times Q$, the last condition can be reduced to

$$\frac{\partial F_i}{\partial Q^l_2}(q_k,\Gamma_{k-1,k})\frac{\partial \Gamma^l}{\partial q^j_k} = \frac{\partial F_j}{\partial Q^l_2}(q_k,\Gamma_{k-1,k})\frac{\partial \Gamma^l}{\partial q^i_k}\,. \tag{4.4}$$

Equivalently, following Remark 4.2.3, the Helmholtz conditions can be written as the closure condition $d(\Psi^{-1}\circ\gamma_{F,\Gamma}) = 0$.

Example 4.2.4 (Toy example)**.** Consider the second order difference equation

$$x_{k+1} = 2x_k - x_{k-1}, \quad y_{k+1} = 2y_k - y_{k-1}\,, \tag{4.5}$$

which is a discretization of the variational SODE $\ddot{x} = 0$, $\ddot{y} = 0$. In this case we already know a Lagrangian function for the continuous system, for instance $L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2)$, so we define a discrete Lagrangian by

$$L_d(q_{k-1},q_k) := \frac{1}{2}\left(\left(\frac{x_k - x_{k-1}}{h}\right)^2 + \left(\frac{y_k - y_{k-1}}{h}\right)^2\right)\,, \tag{4.6}$$

where $q_k = (x_k, y_k)$. Then we can take F to be $\mathbb{F}^-L_d$, that is

$$F: (x_{k-1}, y_{k-1}, x_k, y_k) \longmapsto \left(x_{k-1}, y_{k-1}, \frac{x_k - x_{k-1}}{h^2}, \frac{y_k - y_{k-1}}{h^2}\right)$$

and $\mathrm{Im}(\gamma_{\Gamma,F})$, given by

$$\left(x_{k-1}, y_{k-1}, \frac{x_k - x_{k-1}}{h^2}, \frac{y_k - y_{k-1}}{h^2}, x_k, y_k, \frac{x_k - x_{k-1}}{h^2}, \frac{y_k - y_{k-1}}{h^2}\right),$$

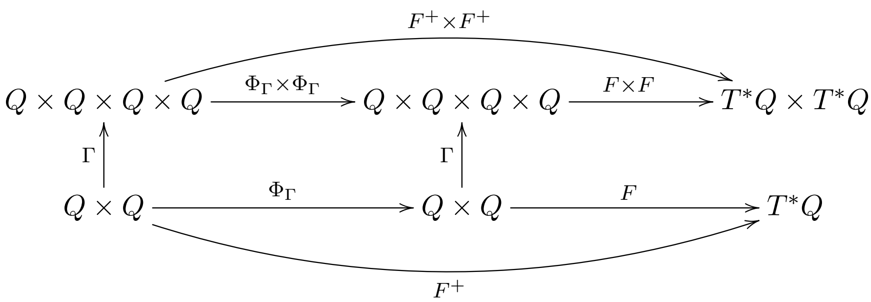
is a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$. Therefore (4.5) is variational, according to Proposition 4.2.2. Indeed the Lagrangian (4.6) has (4.5) as DEL.

Remark 4.2.5. There is no preferred role between q_{k-1} and q_k, therefore Definition 4.2.1 could be rewritten in terms of the existence of a local diffeomorphism $F^+ : Q \times Q \longrightarrow T^*Q$ satisfying $pr_2 = \pi_Q \circ F^+$, that is, $F^+ : (q_{k-1}, q_k) \longmapsto (q_k, \tilde{F}^+(q_{k-1}, q_k))$. Then we would get $F^+ = \mathbb{F}^+ L_d$. More precisely we have the following equivalence result.

Let $\Phi_\Gamma : Q \times Q \longrightarrow Q \times Q$ denote the flow of Γ, that is, $\Phi_\Gamma(q_{k-1}, q_k) = (q_k, \tilde{\Gamma}(q_{k-1}, q_k))$.

Proposition 4.2.6. *There is a local diffeomorphism $F : (Q \times Q, pr_1) \longrightarrow (T^*Q, \pi_Q)$ over the identity such that $Im(\gamma_{F,\Gamma})$ is a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$ if and only if there is a local diffeomorphism $F^+ : (Q \times Q, pr_2) \longrightarrow (T^*Q, \pi_Q)$ over the identity such that $Im(\gamma_{F^+,\Gamma})$ is a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$.*

Proof. If F as in the statement exists then we can define $F^+ = F \circ \Phi_\Gamma$. From condition (4.2) we see that Φ_Γ is a local diffeomorphism and therefore so is F^+.



On the other hand, if we impose that $\mathrm{Im}(F^+ \times F^+) \circ \Gamma$ be Lagrangian, then the condition we obtain corresponding to the vanishing of the $dq_k \wedge dq_{k-1}$ factor is

$$\frac{\partial F_r}{\partial q^j_{k-1}}(q_{k-1}, q_k) = -\frac{\partial F_i}{\partial Q^r_1}(q_k, \Gamma_{k-1,k})\frac{\partial \Gamma^i}{\partial q^j_{k-1}} - \frac{\partial F_i}{\partial Q^s_2}(q_k, \Gamma_{k-1,k})\frac{\partial \Gamma^s}{\partial q^r_k}\frac{\partial \Gamma^i}{\partial q^j_{k-1}} + \frac{\partial F_s}{\partial Q^i_2}(q_k, \Gamma_{k-1,k})\frac{\partial \Gamma^s}{\partial q^r_k}\frac{\partial \Gamma^i}{\partial q^j_{k-1}},$$

which implies that Φ_Γ is a local diffeomorphism since F^+ is a local diffeomorphism. Therefore we can locally define $F = F^+ \circ \Phi_\Gamma^{-1}$, which is also a local diffeomorphism.

Finally, from the commutativity of the above diagram, we have that $\mathrm{Im}(F^+ \times F^+) \circ \Gamma$ is Lagrangian if and only if $\mathrm{Im}(F \times F) \circ \Gamma$ is Lagrangian. ■

The following result is a discrete analogue of Theorem 1.4.1.

Proposition 4.2.7. *An explicit second order difference equation $\Gamma : Q \times Q \longrightarrow Q \times Q \times Q \times Q$ is variational if and only if there is a nondegenerate two-form Ω_d on $Q \times Q$ such that*

(i) $\mathcal{L}^d_\Gamma \Omega_d = 0$,

(ii) $\Omega_d(V_1, V_2) = 0$ *for all* $V_1, V_2 \in Ker(Tpr_1)$,

(iii) $d\Omega_d = 0$,

where $\mathcal{L}^d_\Gamma \Omega_d := (\Phi_\Gamma)^*\Omega_d - \Omega_d$ *is regarded as a discrete analogue of the Lie derivative.*

Remark 4.2.8. The second condition can be replaced by $\Omega_d(V_1, V_2) = 0$ for all $V_1, V_2 \in Ker(Tpr_2)$ (no $dq_k \wedge dq_k$ versus no $dq_{k-1} \wedge dq_{k-1}$ term in Ω_d).

Proof. If Γ is variational define the two-form $\Omega_d := F^*\omega_Q = (F^+)^*\omega_Q$, which clearly satisfies (iii) and the nondegeneracy requirement. From its coordinate expression,

$$\begin{aligned} dq^i_{k-1} \wedge dF_i(q_{k-1}, q_k) &= \left(\frac{\partial F_i}{\partial q^j_{k-1}}\right) dq^i_{k-1} \wedge dq^j_{k-1} + \left(\frac{\partial F_i}{\partial q^j_k}\right) dq^i_{k-1} \wedge dq^j_k \\ &\overset{(4.1)}{=} \left(\frac{\partial F_i}{\partial q^j_k}\right) dq^i_{k-1} \wedge dq^j_k \,, \end{aligned}$$

(ii) is also clear. Finally

$$\begin{aligned} \mathcal{L}^d_\Gamma \Omega_d = (\Phi_\Gamma)^*\Omega_d - \Omega_d &= \left(\frac{\partial F_i(q_k, \Gamma_{k-1,k})}{\partial Q^j_2}\right) dq^i_k \wedge d\Gamma^j - \left(\frac{\partial F_i}{\partial q^j_k}\right) dq^i_{k-1} \wedge dq^j_k \\ &= \frac{\partial F_i}{\partial Q^l_2}(q_k, \Gamma_{k-1,k}) \frac{\partial \Gamma^l}{\partial q^j_k} dq^i_k \wedge dq^j_k \\ &\quad + \left(\frac{\partial F_i}{\partial Q^l_2}(q_k, \Gamma_{k-1,k}) \frac{\partial \Gamma^l}{\partial q^j_{k-1}} + \frac{\partial F_j}{\partial Q^i_2}(q_{k-1}, q_k)\right) dq^i_k \wedge dq^j_{k-1} = 0 \end{aligned}$$

since the discrete Helmholtz conditions (4.2) and (4.3) are satisfied by F.

Conversely let Ω_d be a nondegenerate two-form on $Q \times Q$ satisfying (i)-(iii). From (iii), locally $\Omega_d = d\Theta$ and from (ii) Θ has the local expression

$$\Theta = \alpha_i dq^i_{k-1} + \frac{\partial h}{\partial q^i_k}(q_{k-1}, q_k) dq^i_k$$

for a locally defined map $h : Q \times Q \longrightarrow \mathbb{R}$. Then define $\bar{\Theta} = \Theta - dh$, which satisfies $\bar{\Theta}(V) = 0$ for all $V \in Ker(Tpr_1)$ and $d\bar{\Theta} = \Omega_d$. Then $F : Q \times Q \longrightarrow T^*Q$ can be defined by

$$\langle F(q_{k-1}, q_k), v_{q_{k-1}} \rangle = \langle \bar{\Theta}(q_{k-1}, q_k), V_{v_{q_{k-1}}} \rangle \quad \text{for all} \quad v_{q_{k-1}} \in TQ \,,$$

where $V_{v_{q_{k-1}}} \in T(Q \times Q)$ is any vector satisfying $Tpr_1(V_{v_{q_{k-1}}}) = v_{q_{k-1}}$. The first condition,

$$\mathcal{L}^d_\Gamma \Omega_d = (\Phi_\Gamma)^* d\bar{\Theta} - d\bar{\Theta} = d((\Phi_\Gamma)^*\bar{\Theta} - \bar{\Theta}) = d\mathcal{L}^d_\Gamma \bar{\Theta} = 0 \,,$$

implies the local existence of a discrete Lagrangian $L_d : Q \times Q \longrightarrow \mathbb{R}$ such that $\mathcal{L}^d_\Gamma \bar{\Theta} = dL_d$. From the local expression it is clear that $\gamma_{F,\Gamma} = \Psi \circ dL_d$. As $\mathrm{Im}(dL_d)$ is a Lagrangian submanifold of $T^*(Q \times Q)$, $\mathrm{Im}(\gamma_{F,\Gamma})$ is also a Lagrangian submanifold of $T^*Q \times T^*Q$, that is, Γ is variational. Finally, for Ω_d to be nondegenerate it is necessary to have $\left(\frac{\partial F}{\partial q_k}\right)$ nondegenerate, that is, F is a local diffeomorphism. ■

Remark 4.2.9 (The one-dimensional case)**.** As recalled in Chapter 1, an old result by Sonin [148] shows that the continuous one-dimensional case is always variational. This can be proved by showing that the only Helmholtz condition that remains admits always a nonzero solution, see Remark 1.2.1.

In the discrete setting the only Helmholtz condition to be studied is

$$\frac{\partial F}{\partial Q_2}(q_{k-1}, q_k) + \frac{\partial F}{\partial Q_2}(q_k, \Gamma(q_{k-1}, q_k))\frac{\partial \Gamma}{\partial q_{k-1}}(q_{k-1}, q_k) = 0\,, \tag{4.7}$$

that is, the problem reduces to determining whether or not the functional equation

$$g(y, f(x,y))\frac{\partial f}{\partial x}(x,y) = -g(x,y) \tag{4.8}$$

has a nonzero solution g for a given map $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$.

Assume Γ is linear, that is, $q_{k+1} = aq_{k-1} + bq_k$ for some constants a and b, with $a \neq 0$. Does (4.8) admit a solution $g \not\equiv 0$? We don't have a classification even for this linear case, but some positive examples follow.

- If $a = -1$ then any constant g is a solution.
- If $a = 1$, $b = 0$ then $g(y,x) = -g(x,y)$ admits a solution, for instance $g(x,y) = x - y$.
- If $a < 0$, $b = 0$ then $g(y, ax)a = -g(x,y)$ admits a solution $g(x,y) = \frac{1}{|xy|}$ away from $xy = 0$.
- If $a > 0$, $b = 0$ then $g(x,y) = \frac{1}{x|y|} - \frac{1}{y|x|}$ is a solution away from $xy \geq 0$.
- If $a \neq 0$, $b = \frac{a^3-1}{a}$ then $g(x,y) = -a^2Bx + By$ is a solution for all $B \neq 0$, away from $(0,0)$.

4.2.2 Implicit second order difference equations

Now we go back to the implicit case, where a system of second order difference equations is given by a submanifold $M \subset Q \times Q \times Q$. We assume that M is given by the vanishing of functions $\Phi^i(q_{k-1}, q_k, q_{k+1})$ for $i = 1, \ldots, n$, such that $C := \left(\frac{\partial \Phi}{\partial q_{k+1}}\right)$ is invertible. The problem then consists in deciding whether the original system is equivalent to a discrete Lagrangian system.

Definition 4.2.10. *The implicit system of second order difference equations $\Phi^i(q_{k-1}, q_k, q_{k+1}) = 0$, $i = 1, \ldots, n$, is variational if and only if there is a regular discrete Lagrangian $L_d : Q \times Q \longrightarrow \mathbb{R}$ such that both systems*

$$\Phi^i(q_{k-1}, q_k, q_{k+1}) = 0 \quad \textit{and} \quad D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k) = 0$$

admit the same solutions.

Proposition 4.2.11. *An implicit SOdE M locally given by the vanishing of constraints*

$$\Phi^i(q_{k-1}, q_k, q_{k+1}) = 0, \quad i = 1, \ldots, n\,,$$

is variational if and only if there is a local diffeomorphism $F : Q \times Q \to T^*Q$ satisfying $pr_1 = \pi_Q \circ F$ and such that Im$(F \times F|_M)$ is a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$, where

$$M = \{(q_{k-1}, q_k, q_{k+1}) \in Q \times Q \times Q : \Phi(q_{k-1}, q_k, q_{k+1}) = 0\}\ .$$

Proof. Assume first that a local diffeomorphism F with the stated properties exists. Since we have assumed that C is regular, we can use the implicit function theorem to get for each $(q_{k-1}, q_k, q_{k+1}) \in M$ neighborhoods U of (q_{k-1}, q_k), V of q_{k+1} and $\tilde{\Gamma} : U \longrightarrow V$ such that

$$\begin{aligned}
& \{(q_{k-1}, q_k, q_{k+1}) \in U \times V : \Phi(q_{k-1}, q_k, q_{k+1}) = 0\} \\
= & \left\{(q_{k-1}, q_k, q_{k+1}) \in U \times V : q_{k+1} = \tilde{\Gamma}(q_{k-1}, q_k)\right\} \\
= & \left\{(q_{k-1}, q_k, q_{k+1}) \in U \times V : \tilde{F}(q_k, q_{k+1}) - \tilde{F}(q_k, \tilde{\Gamma}(q_{k-1}, q_k)) = 0\right\} \\
= & \{(q_{k-1}, q_k, q_{k+1}) \in U \times V : D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0\}
\end{aligned}$$

for some locally defined Lagrangian L_d, as explained in Section 4.2.1.

Now assume $\Phi(q_{k-1}, q_k, q_{k+1}) = 0$ is variational, that is, the two sets of equations

$$\Phi(q_{k-1}, q_k, q_{k+1}) = 0 \quad \text{and} \quad D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0$$

have the same solutions for some locally defined Lagrangian L_d. If we choose

$$F(q_{k-1}, q_k) = (q_{k-1}, \tilde{F}(q_{k-1}, q_k)) = (q_{k-1}, -D_1 L_d(q_{k-1}, q_k))$$

then we have

$$\begin{aligned}
\{D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0\} &\equiv \{\Phi(q_{k-1}, q_k, q_{k+1}) = 0\} \\
\equiv \left\{q_{k+1} = \tilde{\Gamma}(q_{k-1}, q_k)\right\} &\equiv \left\{-D_1 L_d(q_k, q_{k+1}) = -D_1 L_d(q_k, \tilde{\Gamma}(q_{k-1}, q_k))\right\}
\end{aligned}$$

which implies

$$-D_2 L_d(q_{k-1}, q_k) = D_1 L_d(q_k, \tilde{\Gamma}(q_{k-1}, q_k)) \,.$$

Then $(F \times F)(S)$ is locally given by

$$\left(q_{k-1}, -D_1 L_d(q_{k-1}, q_k), q_k, -D_1 L_d(q_k, \tilde{\Gamma}(q_{k-1}, q_k))\right) = (q_{k-1}, -D_1 L_d(q_{k-1}, q_k), q_k, D_2 L_d(q_{k-1}, q_k))$$

which is clearly Lagrangian in $(T^*Q \times T^*Q, \Omega_Q)$. ■

In [11] the Helmholtz conditions for explicit SODEs are derived using Lagrangian submanifolds, as explained in Chapter 1. For an implicit SODE we can also derive Helmholtz conditions using Lagrangian submanifolds, as described in Section 1.5.3 of Chapter 1. Analogously now we can obtain the implicit discrete Helmholtz conditions.

The submanifold $(F \times F)(M)$ is locally given by

$$\left(q_{k-1}, \tilde{F}(q_{k-1}, q_k), q_k, \tilde{F}(q_k, q_{k+1})\right)$$

plus the condition $\Phi^i(q_{k-1}, q_k, q_{k+1}) = 0$ for all $i = 1, \ldots, n$. If we write $\tilde{\omega} = (F \times F)^* \Omega_Q$ then locally

$$\begin{aligned}
\tilde{\omega} &= dF_i(q_k, q_{k+1}) \wedge dq_k^i - dF_i(q_{k-1}, q_k) \wedge dq_{k-1}^i \\
&= \frac{\partial F_i}{\partial Q_1^j}(q_k, q_{k+1}) dq_k^j \wedge dq_k^i + \frac{\partial F_i}{\partial Q_2^j}(q_k, q_{k+1}) dq_{k+1}^j \wedge dq_k^i
\end{aligned}$$

$$-\frac{\partial F_i}{\partial Q_1^j}(q_{k-1},q_k)dq_{k-1}^j\wedge dq_{k-1}^i-\frac{\partial F_i}{\partial Q_2^j}(q_{k-1},q_k)dq_k^j\wedge dq_{k-1}^i\,.$$

Then the condition that $(F\times F)(M)$ be a Lagrangian submanifold of $(T^*Q\times T^*Q,\Omega_Q)$ is equivalent to the condition $((F\times F)\circ i_M)^*\Omega_Q=0$ and can be written as $\tilde{\omega}(X,Y)=0$ for all $X,Y\in\mathfrak{X}(M)$. Therefore we will compute a local basis for $\mathfrak{X}(M)$. By imposing that $X\in\mathfrak{X}(Q\times Q\times Q)$ satisfies $d\Phi(X)=0$ we get

$$A_i=\frac{\partial}{\partial q_{k-1}^i}-\frac{\partial\Phi^j}{\partial q_{k-1}^i}W_j^r\frac{\partial}{\partial q_{k+1}^r}\,,\quad B_i=\frac{\partial}{\partial q_k^i}-\frac{\partial\Phi^j}{\partial q_k^i}W_j^r\frac{\partial}{\partial q_{k+1}^r}\,,$$

where W denotes the inverse matrix of C. Finally the **implicit discrete Helmholtz conditions**

$$\tilde{\omega}(A_i,A_j)=0,\quad \tilde{\omega}(A_i,B_j)=0\quad\text{and}\quad\tilde{\omega}(B_i,B_j)=0$$

are respectively given by

$$\frac{\partial F_i}{\partial Q_1^j}(q_{k-1},q_k) = \frac{\partial F_j}{\partial Q_1^i}(q_{k-1},q_k)\,, \tag{4.9}$$

$$\frac{\partial F_j}{\partial Q_1^i}(q_k,q_{k+1})+\frac{\partial F_j}{\partial Q_2^k}(q_k,q_{k+1})\frac{\partial\Phi^r}{\partial q_k^i}W_r^k = \frac{\partial F_i}{\partial Q_1^j}(q_k,q_{k+1})+\frac{\partial F_i}{\partial Q_2^k}(q_k,q_{k+1})\frac{\partial\Phi^r}{\partial q_k^j}W_r^k\,, \tag{4.10}$$

$$\frac{\partial F_i}{\partial Q_2^j}(q_{k-1},q_k)-\frac{\partial F_j}{\partial Q_2^k}(q_k,q_{k+1})\frac{\partial\Phi^r}{\partial q_{k-1}^i}W_r^k = 0\,. \tag{4.11}$$

Remark 4.2.12. If an implicit system is variational then it is possible to find functions $g_{ij}(q_{k-1},q_k,q_{k+1})$ such that

$$g_{ij}(q_{k-1},q_k,q_{k+1})\Phi^j(q_{k-1},q_k,q_{k+1})=[D_1L_d(q_k,q_{k+1})+D_2L_d(q_{k-1},q_k)]_i=:G_i(q_{k-1},q_k,q_{k+1})\,,$$

as shown for instance in [82]. Indeed, since $\left(\frac{\partial\Phi}{\partial q_{k+1}}\right)$ is regular, we can consider a coordinate change

$$(q_{k-1},q_k,q_{k+1})\longrightarrow(q_{k-1}^i,q_k^i,y^i:=\Phi^i(q_{k-1},q_k,q_{k+1}))\,,$$

for which now the submanifold M defines the second order difference equation $y^i=0$. If we let $G_j(q_{k-1}^i,q_k^i,y^i)$ denote the function G_j expressed in the new coordinates, then we have

$$G_j(q_{k-1}^i,q_k^i,y^i)=\int_0^1\frac{d}{dt}G_j(q_{k-1},q_k,ty)dt=y^i\overbrace{\int_0^1\frac{\partial G_j}{\partial y^i}(q_{k-1},q_k,ty)dt}^{g_{ij}(q_{k-1},q_k,q_{k+1})}=\Phi^ig_{ij}\,.$$

4.3 From the continuous to the discrete setting

In this section we will see how to associate a discrete variational SOdE with a continuous variational SODE Γ on TQ. Denote by Φ^Γ the flow of Γ,

$$\Phi^\Gamma:U\subseteq\mathbb{R}\times TQ\to TQ\,,$$

where U is an open subset of $\mathbb{R}\times TQ$. For expository simplicity we will assume that Γ is complete and $U=\mathbb{R}\times TQ$. We will use the notation $\Phi_t^\Gamma(v_q)=\Phi^\Gamma(t,v_q)$.

Proposition 4.3.1. *A complete SODE* Γ *on* TQ *is variational if and only if for all* $t \in \mathbb{R}$, $Im(F \times F) \circ (id \times \Phi_t^\Gamma)$ *is a Lagrangian submanifold of* $(T^*Q \times T^*Q, \Omega_Q = pr_2^*\omega_Q - pr_1^*\omega_Q)$.

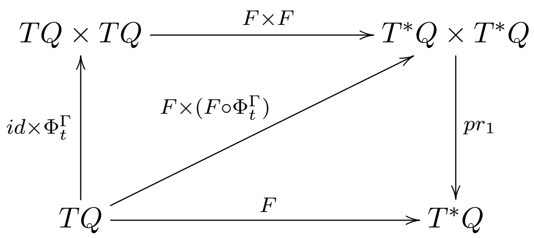
Proof. Assume first that Γ is variational. Then according to the characterization given in [11], there exists a bundle isomorphism $F : TQ \to T^*Q$ over Q such that $S_{\Gamma,F} = \mathrm{Im}(TF \circ \Gamma)$ is a Lagrangian submanifold of $(T^*TQ, d_T\omega_Q)$. This is equivalent to the condition

$$L_\Gamma F^*\omega_Q = 0$$

and therefore

$$(\Phi_t^\Gamma)^*(F^*\omega_Q) = F^*\omega_Q\,.$$

The last equality is equivalent to the statement that $\mathrm{Im}(F \times F) \circ (id \times \Phi_t^\Gamma)$ is a Lagrangian submanifold of $(T^*Q \times T^*Q, \Omega_Q)$.



■

Now in order to define a variational discrete second order system, we need to introduce the exponential map associated with a second order differential equation Γ. Given a point $q_0 \in Q$ and a positive real number h_0, we define

$$exp^\Gamma_{(h_0,q_0)}(v) = \tau_Q((\Phi^\Gamma_{h_0})(v)), \quad \text{for } v \in T_{q_0}Q,$$

(assuming that Γ is complete). For h_0 small enough it is possible to find an open subset $\mathcal{U} \subseteq TQ$ and an open subset U of Q, with $q_0 \in U$ such that the map

$$\begin{array}{rccl} exp^\Gamma_{h_0}: & \mathcal{U} & \longrightarrow & U \times U \subseteq Q \times Q \\ & v & \longmapsto & (\tau_Q(v), exp^\Gamma_{h_0}(v)) \end{array}$$

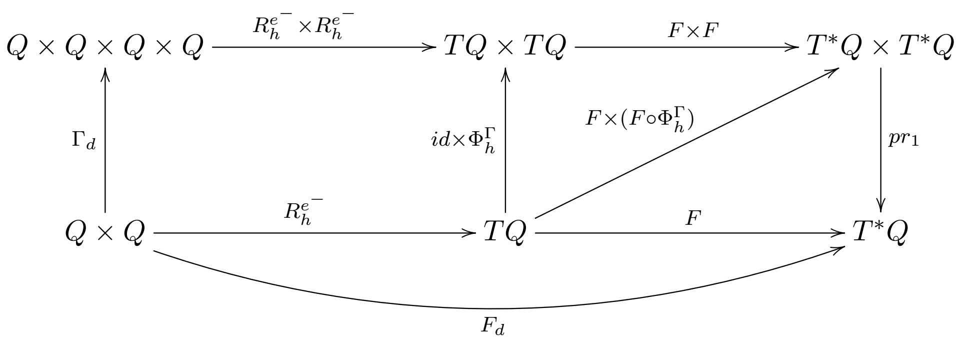
is a diffeomorphism (see [110] for details). Denote by $R^{e^-}_{h_0} : U \times U \to \mathcal{U}$ the inverse map of the diffeomorphism $exp^\Gamma_{h_0} : \mathcal{U} \to U \times U$.

Theorem 4.3.2. *Given a SODE* Γ *on* TQ, *define the discrete second order difference equation*

$$\begin{array}{rccl} \Gamma_d: & U \times U & \longrightarrow & U \times U \times U \times U \\ & (q_{k-1}, q_k) & \longmapsto & (q_{k-1}, q_k, q_k, (\tau_Q \circ \Phi^\Gamma_{2h} \circ R^{e^-}_h)(q_{k-1}, q_k))\,. \end{array}$$

If Γ *is variational then so is* Γ_d *for* h *small enough.*

Proof. Define $F_d : U \times U \subseteq Q \times Q \longrightarrow T^*Q$ by $F_d = F \circ R_h^{e^-}$. Then the proof is a consequence of the commutativity of the following diagram



and Proposition 4.3.1, taking into account that

$$\begin{aligned} (q_k, (\tau_Q \circ \Phi_{2h}^\Gamma \circ R_h^{e^-})(q_{k-1}, q_k)) &= (exp_h^\Gamma \circ \Phi_h^\Gamma \circ R_h^{e^-})(q_{k-1}, q_k) \\ &= ((R_h^{e^-})^{-1} \circ \Phi_h^\Gamma \circ R_h^{e^-})(q_{k-1}, q_k)\,. \end{aligned}$$

■

4.4 Alternative Lagrangian formulations

In this section we will first recall how a class of constants of motion arises from alternative Lagrangian formulations of a SODE, with the two alternative Lagrangians being genuinely different in the sense that they should not differ by a constant and/or addition of a total time derivative [44, 107]. Then we show that the same phenomenon occurs in the discrete setting.

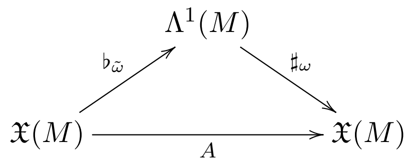
4.4.1 Continuous SODEs

As noted in [44, 107], given a vector field Γ on a manifold M, if we can find a (1,1)-tensor field A on M such that $\mathcal{L}_\Gamma A = 0$ then also $\mathcal{L}_\Gamma A^k = 0$ and therefore $\mathrm{Tr}(A^k)$ is a constant of motion for Γ, for all k.

It is possible to construct such a (1,1)-tensor field if we have the following ingredients. Assume (M, ω) is a symplectic manifold, Γ is a Hamiltonian vector field on M with respect to ω and $\tilde{\omega}$ is a two-form on M such that $\mathcal{L}_\Gamma \tilde{\omega} = 0$. Then we can define A from the condition

$$i_X \tilde{\omega} = i_{A(X)} \omega \quad \text{for all } X \in \mathfrak{X}(M)\,, \tag{4.12}$$

that is, $A(X) = (\sharp_\omega \circ \flat_{\tilde{\omega}})(X)$, where $\flat_\omega : \mathfrak{X}(M) \longrightarrow \Lambda^1(M)$ denotes the contraction map $\flat_\omega(X) = i_X \omega$ and in case $\flat_\omega$ has an inverse then we denote it by $\sharp_\omega$.



The conditions $\mathcal{L}_\Gamma \omega = \mathcal{L}_\Gamma \tilde{\omega} = 0$ imply $\mathcal{L}_\Gamma A = 0$. Indeed, taking Lie derivatives with respect to Γ on both sides of (4.12) we obtain

$$i_{[\Gamma,X]}\tilde{\omega} + i_X \mathcal{L}_\Gamma \tilde{\omega} = i_{[\Gamma,A(X)]}\omega + i_{A(X)}\mathcal{L}_\Gamma \omega\,,$$

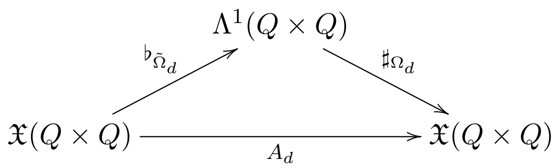
that is, $i_{[\Gamma,X]}\tilde{\omega} = i_{[\Gamma,A(X)]}\omega$ and therefore, again from (4.12), we get $A\,[\Gamma, X] = [\Gamma, A(X)]$, that is, $\mathcal{L}_\Gamma A = 0$.

The above situation arises for instance if we have two alternative Lagrangian formulations for Γ, with Lagrangian functions L and $\tilde{L}$ (see [107]). Since we don't need to make any assumptions on the rank of the two-form $\tilde{\omega}$, it is enough to require that one of the Lagrangians, say L, is regular. Then the corresponding Poincaré-Cartan two-forms Ω_L and $\Omega_{\tilde{L}}$ can be used to construct the (1,1)-tensor field $A = \sharp_{\Omega_L} \circ \flat_{\Omega_{\tilde{L}}}$, which satisfies $\mathcal{L}_\Gamma A = 0$ since Γ is Hamiltonian with respect to both Ω_L and $\Omega_{\tilde{L}}$ and therefore $\mathcal{L}_\Gamma \Omega_L = \mathcal{L}_\Gamma \Omega_{\tilde{L}} = 0$.

4.4.2 Discrete SOdEs

Assume there are two alternative regular discrete Lagrangians L_d and $\tilde{L}_d$ for a discrete second order difference equation Γ on $Q \times Q$. Then we get two discrete Lagrangian symplectic forms Ω_{L_d} and $\Omega_{\tilde{L}_d}$ [112] (equivalently, if we can find F and $\tilde{F}$ then from Propositon 4.2.7 we obtain Ω_d and $\tilde{\Omega}_d$). We can define a (1,1)-tensor field A_d on $Q \times Q$ as before, from the condition

$$i_X \tilde{\Omega}_d = i_{A_d(X)}\Omega_d \quad \text{for all} \quad X \in \mathfrak{X}(Q \times Q)\,. \tag{4.13}$$



Notice again that only the regularity of L_d is actually needed. We define the discrete Lie derivative of A_d along Γ by

$$\mathcal{L}^d_\Gamma A_d = \Phi^*_\Gamma \circ A_d - A_d \circ \Phi^*_\Gamma\,,$$

where $\Phi^*_\Gamma = (\Phi^{-1}_\Gamma)_*$. The conditions $\mathcal{L}^d_\Gamma \Omega_d = \mathcal{L}^d_\Gamma \tilde{\Omega}_d = 0$ imply, as in the continuous case, that $\mathcal{L}^d_\Gamma A_d = 0$. Indeed, if we take discrete Lie derivatives with respect to Γ on both sides of (4.13), we obtain

$$i_{\Phi^*_\Gamma X}\Phi^*_\Gamma \tilde{\Omega}_d = \Phi^*_\Gamma(i_X \tilde{\Omega}_d) = \Phi^*_\Gamma(i_{A_d(X)}\Omega_d) = i_{\Phi^*_\Gamma A_d(X)}\Phi^*_\Gamma \Omega_d\,,$$

which using $\mathcal{L}^d_\Gamma \Omega_d = \mathcal{L}^d_\Gamma \tilde{\Omega}_d = 0$ becomes $i_{\Phi^*_\Gamma X}\tilde{\Omega}_d = i_{\Phi^*_\Gamma A_d X}\Omega_d$. Then by definition of A_d we obtain that $A_d \Phi^*_\Gamma X = \Phi^*_\Gamma A_d X$. Observe that the condition $A_d \Phi^*_\Gamma X = \Phi^*_\Gamma A_d X$ is equivalent to $A_d(\Phi_\Gamma)_* X = (\Phi_\Gamma)_* A_d X$.

Choose a basis $\{X_1, \ldots, X_{2n}\}$ of $\mathfrak{X}(Q \times Q)$ and write $A_d(X_a) = \mathcal{A}^b_a X_b$, $(\Phi_\Gamma)_*(X_a) = \phi^b_a X_b$. Then the above condition takes the form

$$0 \;=\; (\Phi_\Gamma)_* \circ A_d(X_a) - A_d \circ (\Phi_\Gamma)_*(X_a)$$

$$\begin{aligned} &= (\Phi_\Gamma)_*(\mathcal{A}^b_a(x)X_b(x)) - A_d(\phi^b_a(x)X_b(\Phi_\Gamma(x))) \\ &= \mathcal{A}^b_a(x)\phi^c_b(x)X_c(\Phi_\Gamma(x)) - \phi^b_a(x)\mathcal{A}^c_b(\Phi_\Gamma(x))X_c(\Phi_\Gamma(x))\,, \end{aligned}$$

from which we get $\mathcal{A}^b_a(x)\phi^c_b(x) = \phi^b_a(x)\mathcal{A}^c_b(\Phi_\Gamma(x))$, that is, $\mathcal{A}^c_d(\Phi_\Gamma(x)) = (\phi^{-1})^a_d(x)\mathcal{A}^b_a(x)\phi^c_b(x)$. Therefore the eigenvalues of $A_d(x)$ and $A_d(\Phi_\Gamma(x))$ coincide and $\mathrm{Tr}A^k_d(x) = \mathrm{Tr}A^k_d(\Phi_\Gamma(x))$, that is, $\mathrm{Tr}A^k_d$ is a constant of motion for Γ.

Example 4.4.1. Consider the second order differential equation $\ddot{x} + x = 0$ on $\mathbb{R}$, that is, the SODE $\Gamma = \dot{x}\frac{\partial}{\partial x} - x\frac{\partial}{\partial \dot{x}} \in \mathfrak{X}(\mathbb{R}^2)$. We will find a discretization of the system, which admits two alternative discrete Lagrangians L_{d1} and L_{d2}, and for which $A_d = \Omega^{-1}_{L_{d1}} \circ \Omega_{L_{d2}}$ provides constants of motion. The solutions to the continuous system are given by $x(t) = a\cos(t) + b\sin(t)$, where a and b are constants. Therefore the exponential map associated with Γ is given by

$$\begin{array}{rccc} exp^\Gamma_{(x,h)}: & TQ & \longrightarrow & Q\times Q \\ & (x,v) & \longmapsto & (x, x\cos(h) + v\sin(h)) \end{array}$$

and the flow at time h is

$$\begin{array}{rccc} \Phi^\Gamma_h: & TQ & \longrightarrow & TQ \\ & (x,v) & \longmapsto & (x\cos(h) + v\sin(h), v\cos(h) - x\sin(h))\,. \end{array}$$

Notice that for the continuous system we have the two alternative Lagrangians

$$L = \frac{1}{2}(\dot{x}^2 - x^2) \quad \text{and} \quad \tilde{L} = \frac{1}{3}\dot{x}^4 + 2x^2\dot{x}^2 - x^4\,,$$

with corresponding Legendre transformations $F_1(x,\dot{x}) = (x,\dot{x})$ and $F_2(x,\dot{x}) = (x, \frac{4}{3}\dot{x}^3 + 4x^2\dot{x})$. Therefore $\mathrm{Im}(F_1 \times F_1) \circ (id \times \Phi^\Gamma_t)$ and $\mathrm{Im}(F_2 \times F_2) \circ (id \times \Phi^\Gamma_t)$ are both Lagrangian submanifolds of $(T^*Q \times T^*Q, \Omega_Q)$. Then we can define the discrete SOdE

$$\begin{array}{rccc} \Gamma_d = (exp^\Gamma_h \times exp^\Gamma_h) \circ (id \times \Phi^\Gamma_h) \circ R^{e^-}_h: & Q\times Q & \longrightarrow & Q\times Q\times Q\times Q \\ & (x_0,x_1) & \longmapsto & (x_0, x_1, x_1, 2x_1\cos(h) - x_0) \end{array}$$

and the discrete Lengendre transformations $F_{d1} = F_1 \circ R^{e^-}_h$ and $F_{d2} = F_2 \circ R^{e^-}_h$ that provide Lagrangian submanifolds $\mathrm{Im}(F_{d1} \times F_{d1}) \circ \Gamma_d$ and $\mathrm{Im}(F_{d2} \times F_{d2}) \circ \Gamma_d$ according to Theorem 4.3.2.

The Lagrangian submanifolds $\mathrm{Im}(F_{d1} \times F_{d1}) \circ \Gamma_d$ and $\mathrm{Im}(F_{d2} \times F_{d2}) \circ \Gamma_d$ are given respectively by

$$\left(x_0, \frac{x_1 - x_0\cos(h)}{\sin(h)}, x_1, \frac{x_1\cos(h) - x_0}{\sin(h)}\right) \quad \text{and}$$

$$\left(x_0, \frac{4}{3}\left(\frac{x_1 - x_0\cos(h)}{\sin(h)}\right)^3 + 4x_0^2\left(\frac{x_1 - x_0\cos(h)}{\sin(h)}\right), x_1, \frac{4}{3}\left(\frac{x_1\cos(h) - x_0}{\sin(h)}\right)^3 + 4x_1^2\left(\frac{x_1\cos(h) - x_0}{\sin(h)}\right)\right),$$

from where we get

$$\Omega_{L_{d1}} = \frac{-1}{\sin(h)}dx_0 \wedge dx_1 \quad \text{and} \quad \Omega_{L_{d2}} = -4\left(\frac{x_1^2 - 2x_0x_1\cos(h) + x_0^2}{\sin^3(h)}\right)dx_0 \wedge dx_1\,.$$

Therefore we have

$$A_d = \frac{4}{\sin^2(h)}(x_1^2 - 2x_0x_1\cos(h) + x_0^2)dx_0 \otimes \frac{\partial}{\partial x_0} + \frac{4}{\sin^2(h)}(x_1^2 - 2x_0x_1\cos(h) + x_0^2)dx_1 \otimes \frac{\partial}{\partial x_1}$$

and we obtain the conserved quantity $x_1^2 - 2x_0x_1\cos(h) + x_0^2$ for the SOdE Γ_d, which is a discretization of the conserved quantity $\dot{x}^2 + x^2$ for Γ.

Although they are not needed in order to get A_d, the two discrete Lagrangians that we obtain are

$$\begin{aligned}
L_{d1}(x_0, x_1) &= \frac{\cos(h)}{2\sin(h)}\left(x_0^2 + x_1^2\right) - \frac{x_0x_1}{\sin(h)}\,, \\
L_{d2}(x_0, x_1) &= x_1^4\cot(h) - \frac{4}{3}x_0x_1^3\csc(h) + \frac{1}{3}x_0^4\cos(2h)\csc(h)\sec(h) + \frac{1}{3}(x_1\cot(h) - x_0\csc(h))^4\tan(h) \\
&= \cot(h)\left(1 + \frac{\cot^2(h)}{3}\right)x_1^4 - \frac{4}{3}x_0x_1^3\csc^3(h) + 2x_0^2x_1^2\cot(h)\csc^2(h) - \frac{4}{3}x_0^3x_1\csc^3(h) \\
&\quad + \cot(h)\left(1 + \frac{\cot^2(h)}{3}\right)x_0^4\,.
\end{aligned}$$

Chapter 5

Inverse problem for constrained second order systems

In this chapter we will deal with mechanical systems with constraints. There is a crucial difference between a planar pendulum, which is restricted to move on S^1, a submanifold in $\mathbb{R}^2$ and a disk rolling without sliding on a plane, which can get to any configuration, but has restrictions on the velocities. The first one is an example of a holonomic system while the second is an example of nonholonomic system.

In Section 5.1 we will introduce nonholonomic systems, including some examples, for instance the rolling disk and Chaplygin systems. In Section 5.2 we briefly introduce vakonomic mechanics. In Section 5.3 we provide a definition of constrained variational SODE and show some consequences of this notion. A generalization of Crampin's Theorem holds and also the original system, if variational, can be regarded as a subsystem of a Lagrangian system. In Section 5.4 we need to adapt the definition to include systems with holonomic constraints. Then we show the relationship between the inverse problem for holonomic systems and the inverse problem without constraints. In Section 5.5 we further adapt the notion of variational systems, for time-dependent constrained systems, using now the notion of isotropic submanifold for a Poisson manifold and Crampin's Theorem follows similarly. Finally in Section 5.6 we discuss the inverse problem for discrete constrained systems. We obtain results similar to the continuous case and we also start some discussion on the possible advantages of using constrained variational integrators.

5.1 Nonholonomic systems

In this section, we will see one of the main examples where second order differential equations along submanifolds arise, namely the case of nonholonomic Lagrangian systems. There is considerable interest in the study of these systems since nonholonomic constraints are present in a great variety of mechanical systems in engineering and robotics. For instance, they describe the dynamics of wheeled vehicles, manipulation devices and locomotion systems (see [14, 18, 39, 40, 125, 126] and references therein).

Definition 5.1.1. *A nonholonomic Lagrangian system on a manifold Q consists of a pair (L, M) where $L : TQ \to \mathbb{R}$ is a Lagrangian function and M is a submanifold of TQ such that $\tau_Q(M) = Q$.*

In mechanical and real examples, the constraints are typically linear or affine in the velocities. Linear velocity constraints are constraints that are specified by a regular C^∞-distribution $\mathcal{D}$ on the configuration manifold Q, or equivalently, by a vector subbundle $\tau_{\mathcal{D}} : \mathcal{D} \to Q$ of the tangent bundle TQ with canonical inclusion $i_{\mathcal{D}} : \mathcal{D} \hookrightarrow TQ$. Therefore, we will say that a curve $\gamma : I \subseteq \mathbb{R} \to Q$ satisfies the constraints given by $\mathcal{D}$ if

$$\gamma'(t) = \frac{d\gamma}{dt}(t) \in \mathcal{D}_{\gamma(t)} \quad \text{for all} \quad t \in I\,. \tag{5.1}$$

A regular linear velocity constraint $\mathcal{D}$ is holonomic if $\mathcal{D}$ is integrable or involutive, that is, for any vector fields $X, Y \in \mathfrak{X}(Q)$ taking values on $\mathcal{D}$, it holds that the vector field $[X, Y]$ also takes values in $\mathcal{D}$. Such a constraint is nonholonomic if it is not holonomic. Observe that in the case of holonomic constraints all the curves through a point $q \in Q$ satisfying the constraints must lie on the maximal integral manifold for $\mathcal{D}$ through q.

From now on, we assume that $M = \mathcal{D}$ is a vector subbundle or an affine subbundle modelled on $\mathcal{D}$, and $\tau_Q(M) = Q$, avoiding the existence of holonomic constraints (see Section 5.4 for details on this case).

The existence of the constraints prescribed by M induces the introduction of reaction forces which restrict the motion to M. This forces are determined by the Lagrange-d'Alembert principle.

Define the set of admissible curves by

$$\mathcal{C}^2_M(q_0, q_1, [a,b]) = \left\{ c : [a,b] \subseteq \mathbb{R} \longrightarrow Q \;\middle|\; c \in \mathcal{C}^2(q_0, q_1, [a,b]),\ \dot{c}(t) \in M_{c(t)}\ \forall t \in [a,b] \right\},$$

and the set of possible virtual variations along c by

$$\mathcal{V}_c = \left\{ X : [a,b] \longrightarrow TQ \;\middle|\; X \in C^1, X(t) \in D_{c(t)}\ \forall t \in [a,b] \ \text{ and } X(a) = X(b) = 0 \right\}.$$

Definition 5.1.2. [Lagrange-d'Alembert's principle] *Let $c \in \mathcal{C}^2_M(q_0, q_1, [a,b])$, then c is a solution of the nonholonomic Lagrangian system (L, M) if*

$$\langle d\mathcal{J}(c), X \rangle = 0, \ \textit{for all } X \in \mathcal{V}_c.$$

Locally, if the submanifold M is determined by the vanishing of constraints $\phi^\alpha(q^i, \dot{q}^i) = 0$ (either linear or affine constraints), then the equations of motion of a nonholonomic Lagrangian system are

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} &= \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i}\,, \\ \phi^\alpha(q^i, \dot{q}^i) &= 0\,. \end{aligned} \tag{5.2}$$

If the constraints are written as $\phi^\alpha(q^i, \dot{q}^i) = \mu^\alpha_i(q)\dot{q}^i + \mu^\alpha_0(q)$, then the previous equations reduce to

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_\alpha \mu^\alpha_i(q)\,, \tag{5.3}$$

$$\mu^\alpha_i(q)\dot{q}^i + \mu^\alpha_0(q) = 0\,. \tag{5.4}$$

The right-hand side of Equation (5.3) represents the force induced by the constraints, while Equation (5.4) gives the constraints themselves.

It is important to stress that in Equations (5.3) it is necessary to use the Lagrangian defined on the full space TQ instead of working with the restriction of L to $\mathcal{D}$ (where we now consider $\mathcal{D}$ as a vector subbundle of TQ). Applying standard variational techniques and using $l = L|_{\mathcal{D}}$ we would derive a different set of equations than (5.3), which are not valid for nonholonomic mechanics. These other equations are called vakonomic equations, or variational constrained equations in the literature, see for instance [8]. They will be briefly recalled in Section 5.2.

If the Hessian matrix W of L with respect to the velocities is definite, then the matrix

$$\mathcal{C} = (C^{\alpha\beta}) \quad \text{with} \quad C^{\alpha\beta} = \mu_i^\alpha W^{ij} \mu_j^\beta$$

is regular, where (W^{ij}) is the inverse of the Hessian matrix $W_{ij} = \left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\right)$ and analogously $(C_{\alpha\beta})$ is the inverse of $\mathcal{C}$.

Observe that the definiteness condition is automatically satisfied for systems of mechanical type, that is, when the Lagrangian is given by $L = T - V$, where T is the kinetic energy associated to a Riemannian metric on Q and V, the potential energy, is a function on Q. It is easy to show that, under this condition, we can write the equations of motion of a nonholonomic system as a system of explicit second order differential equations on the constraint submanifold M. In fact, the Lagrange multipliers are determined univocally as

$$\lambda_\alpha(q, \dot{q}) \;\; = \;\; -C_{\alpha\beta} \left(\frac{\partial \mu_i^\beta}{\partial q^j} \dot{q}^i \dot{q}^j + \frac{\partial \mu_0^\beta}{\partial q^i} \dot{q}^i + \mu_i^\beta W^{ij} \left[\frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial \dot{q}^j \partial q^k} \dot{q}^k \right] \right),$$

and given an initial condition on M, $\dot{c}(0) \in M_{c(0)}$, the unique solution of the second order differential equation

$$\ddot{q}^i = W^{ij} \left[\lambda_\alpha(q, \dot{q}) \mu_j^\alpha(q) + \frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial \dot{q}^j \partial q^k} \dot{q}^k \right]$$

evolves on the constraint submanifold M, that is, $\dot{c}(t) \in M_{c(t)}$.

5.1.1 Examples

Some examples of nonholonomic systems are the nonholonomic particle, the vertical or falling rolling disk, the rolling ball, the knife edge, the Chaplygin sleigh, the snakeboard and the rattleback. A description of each of these examples and additional ones can be found in [14, 40, 126].

We will now see a detailed example, the vertical rolling disk, which will also appear in Section 5.3 and Chapter 6. Later we will also introduce the nonholonomic particle and the Chaplygin sleigh.

Example 5.1.3 (Rolling disk)**.** One of the simplest examples of a nonholonomic system is the unicycle, that is, a disk of radius r which rolls on a horizontal plane, and always remains exactly upright, see for instance [14]. The coordinates (x_1, x_2, θ, ϕ) describe the possible configurations of the system, where (x_1, x_2) are the coordinates of the contact point with the x_1x_2-plane, θ the heading angle and ϕ the self-rotation angle. The system is shown in Figure 5.1. Not all the velocities are admissible for this system, since the constraint that the disk roll without slipping is specified by the linear velocity constraints

$$\dot{x}_1 - r\dot{\phi}\cos\theta = 0\,, \qquad \dot{x}_2 - r\dot{\phi}\sin\theta = 0\,, \tag{5.5}$$

and therefore,

$$\mathcal{D} = \operatorname{span}\left\{ X_1 = r\cos\theta \frac{\partial}{\partial x_1} + r\sin\theta \frac{\partial}{\partial x_2} + \frac{\partial}{\partial \phi}\,, X_2 = \frac{\partial}{\partial \theta} \right\}.$$

The constraints are nonholonomic since the distribution $\mathcal{D}$ is not involutive. Indeed,

$$[X_1, X_2] = r\sin\theta \frac{\partial}{\partial x_1} - r\cos\theta \frac{\partial}{\partial x_2}\,,$$

and $[X_1, X_2](q) \notin \mathcal{D}_q$ for all $q \in Q$.

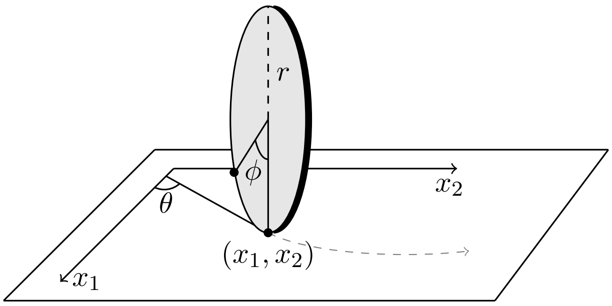


FIGURE 5.1: The geometry of the rolling disk.

In this case the Riemannian metric is

$$g = m\, dx_1 \otimes dx_1 + m\, dx_2 \otimes dx_2 + J_\theta\, d\theta \otimes d\theta + J_\phi\, d\phi \otimes d\phi\,,$$

where m is the mass of the disk and J_θ and J_ϕ are the moment of inertia about the θ and ϕ axis respectively. We assume that the Lagrangian is purely kinetic, i.e. $V = 0$, and thus

$$L(x_1, x_2, \theta, \phi, \dot{x}_1, \dot{x}_2, \dot{\theta}, \dot{\phi}) = \frac{1}{2}\left(m\,\dot{x}_1^2 + m\,\dot{x}_2^2 + J_\theta\,\dot{\theta}^2 + J_\phi\,\dot{\phi}^2 \right).$$

Therefore the nonholonomic equations (5.3) are

$$\begin{aligned}
m\ddot{x}_1 &= \lambda_1\,, \\
m\ddot{x}_2 &= \lambda_2\,, \\
J_\theta\ddot{\theta} &= -\lambda_1 r\cos(\phi) - \lambda_2 r\sin(\phi)\,, \\
J_\phi\ddot{\phi} &= 0\,, \\
\dot{x}_1 - r\dot{\phi}\cos\theta &= 0\,, \\
\dot{x}_2 - r\dot{\phi}\sin\theta &= 0\,,
\end{aligned}$$

which, substituting the values of λ_α and using the constraints, reduce to the system

$$\ddot{\theta} = 0\,, \qquad \ddot{\phi} = 0\,, \qquad \dot{x}_1 - r\dot{\phi}\cos\theta = 0\,, \qquad \dot{x}_2 - r\dot{\phi}\sin\theta = 0\,.$$

5.1.2 Nonholonomic Chaplygin systems

We will now consider nonholonomic Lagrangian systems with symmetry, that is, nonholonomic systems $(L, \mathcal{D})$, where $\mathcal{D}$ is a vector subbundle of TQ, with a Lie group action $\Psi : G \times Q \to Q$, such that both L and $\mathcal{D}$ are G-invariant with respect to the induced action on TG. The subclass of nonholonomic systems with symmetry such that

$$\mathcal{D}_q \oplus T_q\mathrm{Orb}(q) = T_qQ\,,$$

is known as the purely kinematic case, where the symmetry directions complement the constraints given by $\mathcal{D}$. Here $\mathrm{Orb}(q) = \{\bar{q} \in Q \;\big|\; \bar{q} = \Psi(g, q), \text{with } g \in G\}$ denotes the orbit of $q \in Q$.

Chaplygin systems are a particular type of nonholonomic systems with symmetry (see for instance [40]). In this case the configuration space is a principal G-bundle $\pi : Q \to Q/G$, associated with a free and proper action $\Psi : G \times Q \to Q$ such that L is G-invariant and D is determined by the horizontal distribution of a principal connection $\mathcal{A} : TQ \to \mathfrak{g}$. Remember that $\mathcal{A}(\xi_Q(q)) = \xi$, where

$$\xi_Q(q) = \frac{d}{dt}\Big|_{t=0} \Psi(\exp\,(t\xi), q), \text{ with } \xi \in \mathfrak{g}$$

and $\mathcal{A}(T\Psi_g(X)) = \mathrm{Ad}_g(\mathcal{A}(X))$, for all $X \in TQ$ where $\Psi_g(q) = \Psi(g, q)$. Observe that in this case

$$\mathcal{D}_q = \{v_q \in T_qQ \;\Big|\; \mathcal{A}(v_q) = 0\}\,,$$

that is, D_q is the horizontal subspace at q determined by the connection $\mathcal{A}$.

Therefore, for any $v_q \in T_qQ$ we have a unique decomposition $v_q = \mathrm{hor}_q v_q + \mathrm{ver}_q v_q$, where $\mathrm{ver}_q v_q = (\mathcal{A}(v_q))_Q(q)$ and then $\mathrm{hor}_q v_q = v_q - (\mathcal{A}(v_q))_Q(q) \in \mathcal{D}_q$. The projection map $\pi : Q \to Q/G$ induces an isomorphism from $\mathcal{D}_q$ to $T_{\pi(q)}(Q/G)$, and the inverse map is called the horizontal lift. Thus for any vector field $X \in \mathfrak{X}(Q/G)$ on the base space, we have a unique vector field X^h (the horizontal lift of X) that is horizontal and π-related to X. Consider a local trivialization $U \times G$ of π where now the action of G is given by left translation on the second factor and U is a neighborhood of Q/G. Take coordinates r^a on U and a basis $\{e_\alpha\}$ of $\mathfrak{g}$. Then any element $\xi \in \mathfrak{g}$ is written as $\xi = \xi^\alpha e_\alpha$. In this local trivialization we can write the connection $\mathcal{A}$ as

$$\mathcal{A}(r, g, \dot{r}, \dot{g}) = \mathrm{Ad}_g(g^{-1}\dot{g} + A_a^\alpha \dot{r}^a e_\alpha)$$

and the coefficients of the curvature, $\mathcal{B}(X, Y) = -\mathcal{A}([X^h, Y^h])$, $\qquad X, Y \in \mathfrak{X}(Q/G)$, are

$$\mathcal{B}_{ab}^\alpha = \frac{\partial A_a^\alpha}{\partial r^b} - \frac{\partial A_b^\alpha}{\partial r^a} - C_{\beta\gamma}^\alpha A_b^\beta A_a^\gamma\,, \quad \text{where} \quad \mathcal{B}\left(\frac{\partial}{\partial r^a}, \frac{\partial}{\partial r^b}\right) = \mathcal{B}_{ab}^\alpha e_\alpha\,.$$

In this case, the Lagrangian $L : TQ \to \mathbb{R}$ induces a Lagrangian $L^* : T(Q/G) \to \mathbb{R}$ by

$$L^*(X(\bar{q})) = L(X^h(q))\,.$$

Locally, $L^*(r^a, \dot{r}^a) = l(r^a, \dot{r}^a, -A_a^\alpha \dot{r}^a e_\alpha)$, where $l : TU \times \mathfrak{g} \to \mathbb{R}$ represents the reduction of $L : T(U \times G) \to \mathbb{R}$ to TQ/G.

After some computations, we can see that the reduced dynamics are given by the following system of equations on $T(Q/G)$:

$$\frac{d}{dt}\left(\frac{\partial L^*}{\partial \dot{r}^a}\right) - \frac{\partial L^*}{\partial r^a} = \Lambda_a, \quad \text{where} \quad \Lambda_a = -\left(\frac{\partial l}{\partial \xi^\alpha}\right)_{\mathbf{c}} \mathcal{B}^\alpha_{ab}\dot{r}^b$$

and the subindex "**c**" on the right-hand side indicates that, after computing the derivative of l with respect to ξ^a, one evaluates this partial derivative on $(r^a, \dot{r}^a, -A^\alpha_a \dot{r}^a e_\alpha)$.

Moreover, if L is regular, we have that L^* is also regular and we obtain the system of second-order differential equations now defined on the full space $T(Q/G)$

$$\frac{d^2 r^a}{dt^2} = \widehat{W}^{ab}\left(\frac{\partial L^*}{\partial r^b} - \dot{r}^c \frac{\partial^2 L^*}{\partial r^c \partial \dot{r}^b} + \Lambda_b\right), \tag{5.6}$$

where $(\widehat{W}^{ab})$ is the inverse of the Hessian matrix $\widehat{W}_{ab} = \left(\frac{\partial^2 L^*}{\partial \dot{r}^a \partial \dot{r}^b}\right)$.

5.1.3 Chaplygin Hamiltonization

As we have seen in Section 5.1.2, the equations of motion of a noholonomic Chaplygin system can be reduced to a second-order differential equation on Q/G. Then, we can apply the inverse problem of the calculus of variations in an attempt to find a Lagrangian $L : T(Q/G) \to \mathbb{R}$ such that equations (5.6) are equivalent to the Euler-Lagrange equations for the Lagrangian L.

Denote by Γ the SODE on $T(Q/G)$ in equations (5.6). By Theorem 1.5.2, Γ is equivalent to the Euler-Lagrange equations of a Lagrangian if there exists a fiber diffeomorphism $F : T(Q/G) \to T^*(Q/G)$ such that $\mathrm{Im}(\mu_{\Gamma,F})$ is a Lagrangian submanifold of $(T^*T(Q/G), \omega_{T(Q/G)})$.

Equivalently, in the case of Chaplygin systems we can use the reduced Lagrangian $L^* : T(Q/G) \to \mathbb{R}$ defined in Section 5.1.2 and its associated Legendre transformation

$$Leg_{L^*} : T(Q/G) \to T^*(Q/G) .$$

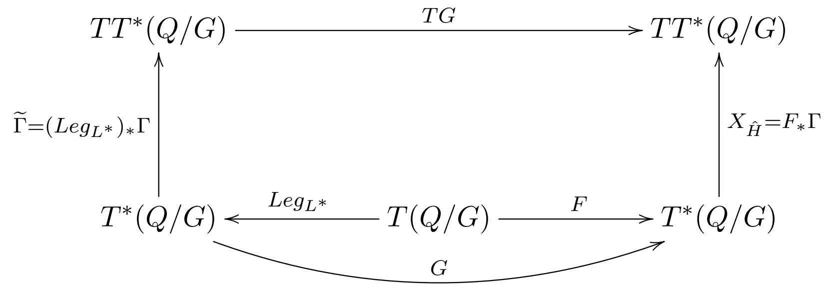
Then we can define the vector field $\widetilde{\Gamma} = (Leg_{L^*})_*\Gamma$ on $T^*(Q/G)$ representing the nonholonomic dynamics, now on the Hamiltonian side. But if there exists a solution $F : T(Q/G) \to T^*(Q/G)$ of the inverse problem of calculus of variations then the vector field $F_*\Gamma$ is locally Hamiltonian. That is, locally there exists a function $\widehat{H} : T^*(Q/G) \to \mathbb{R}$ such that

$$i_{F_*\Gamma}\omega_{Q/G} = d\hat{H}.$$

Therefore, if we consider the diffeomorphism $G : T^*(Q/G) \to T^*(Q/G)$ given by $G = F \circ (Leg_{L^*})^{-1}$ then it is clear by construction that $G_*\widetilde{\Gamma} = F_*\Gamma$ and

$$i_{\widetilde{\Gamma}}\Omega = d\hat{H}, \tag{5.7}$$

where $\Omega = G^*(\omega_{Q/G})$ and $\hat{H} = H \circ G$. Equation (5.7) corresponds to the standard notion of Hamiltonization of a Chaplygin system [10, 17].



5.2 Variational constrained equations (or vakonomic mechanics)

Now we will introduce a dynamical system described by the same pair (L, M), but using purely variational techniques [8]. As above, let us consider a regular Lagrangian $L : TQ \to \mathbb{R}$, and a set of constraints $\phi^\alpha(q^i, \dot{q}^i) = 0$, $1 \leq \alpha \leq m$ that determine a $2n - m$ dimensional submanifold $M \subset TQ$. Take the extended Lagrangian $\mathcal{L} = L + \lambda_\alpha \phi^\alpha$ which includes the Lagrange multipliers λ_α as new extra variables. The equations of motion for the constrained variational problem are the Euler-Lagrange equations for $\mathcal{L}$, that is,

$$
\begin{aligned}
&\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = -\dot{\lambda}_\alpha \frac{\partial \phi^\alpha}{\partial \dot{q}^i} - \lambda_\alpha \left[\frac{d}{dt}\left(\frac{\partial \phi^\alpha}{\partial \dot{q}^i}\right) - \frac{\partial \phi^\alpha}{\partial q^i}\right], \\
&\phi^\alpha(q^i, \dot{q}^i) = 0, \quad 1 \leq \alpha \leq m\,.
\end{aligned}
\tag{5.8}
$$

Observe that the equations of a variational constrained system are different from the equations of a nonholonomic system given in (5.2).

We can alternatively derive the vakonomic equations (5.8) by extremizing the action functional among all curves satisfying the constraints. More precisely, if we consider

$$
\begin{array}{rccl}
\tilde{\mathcal{J}}: & \mathcal{C}^2_M(q_0, q_1, [a, b]) & \longrightarrow & \mathbb{R} \\
& \tilde{c} & \longmapsto & \int_a^b L(\tilde{c}(t), \dot{\tilde{c}}(t))\, dt\,,
\end{array}
$$

that is, the action functional restricted to the set of admissible curves, then the variational principle

$$
d\tilde{\mathcal{J}}(\tilde{c})(\tilde{X}) = 0 \quad \text{for all} \quad \tilde{X} \in T_{\tilde{c}}\mathcal{C}^2_M(q_0, q_1, [a, b])
$$

provides again equations (5.8). Notice that in this case we only need L defined on M, in contrast with nonholonomic mechanics. The nonholonomic equations describe physical systems while the vakonomic ones appear in optimal control problems [102].

5.3 The inverse problem for constrained systems

The use of other distinguished submanifolds of symplectic manifolds, namely isotropic submanifolds, turns out to be suitable to characterize the inverse problem for constrained variational calculus. Moreover, using a standard construction in symplectic geometry (see Appendix A) we can extend these isotropic submanifolds to Lagrangian ones, allowing us to describe the constrained solutions as solutions of a variational problem without constraints. The solutions of the new variational problem with initial conditions verifying the constraints are precisely real solutions of the original constrained system. Our techniques are also related to classical results about the comparison of solutions of nonholonomic systems and constrained variational problems (see [16, 38, 61] and references therein).

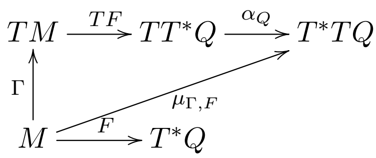
In this section, we will study the extension of the inverse problem of the calculus of variations to the case of constrained systems. Consider a submanifold M of TQ and a vector field Γ on M verifying the SODE condition, that is,

$$S_x(\Gamma(x)) = \Delta(x), \quad \text{for all} \quad x \in M\,.$$

Nonholonomic mechanics is an example of this situation if the Lagrangian is of mechanical type, as we have already seen in Section 5.1.

From now on, we assume that M projects over the whole configuration manifold Q. Inspired by Theorem 1.5.2 we give the following definition:

Definition 5.3.1. *A SODE Γ on the submanifold M of TQ is variational if there exists an immersion $F : M \to T^*Q$ over Q such that $\Sigma_{\Gamma,F} := Im(\mu_{\Gamma,F})$ is an isotropic submanifold of (T^*TQ, ω_{TQ}), where $\mu_{\Gamma,F} = \alpha_Q \circ TF \circ \Gamma$.*



Assume that M is locally determined by the constraints $\dot{q}^\alpha = \psi^\alpha(q^i, \dot{q}^a), 1 \leq \alpha \leq m$, so that $(q^i, \dot{q}^a)$ are local coordinates on M, $1 \leq a \leq n - m$, $n = \dim Q$. Then the solutions of the SODE Γ are now represented by the system of differential equations

$$\begin{aligned} \ddot{q}^a &= \Gamma^a(q^i, \dot{q}^a)\,, \\ \dot{q}^\alpha &= \psi^\alpha(q^i, \dot{q}^a)\,. \end{aligned}$$

For each map

$$\begin{array}{rccc} F: & M & \longrightarrow & T^*Q \\ & (q^i, \dot{q}^a) & \longmapsto & (q^i, F_j(q^i, \dot{q}^a)) \end{array}$$

satisfying that rank $\left(\frac{\partial F_i}{\partial \dot{q}^a}\right) = n - m$, the submanifold $\mathrm{Im}(\alpha_Q \circ TF \circ \Gamma) = \mathrm{Im}(\mu_{\Gamma,F})$ is given in coordinates by

$$\left(q^i, \dot{q}^a, \psi^\alpha, \frac{\partial F_i}{\partial q^a}\dot{q}^a + \frac{\partial F_i}{\partial q^\alpha}\psi^\alpha + \frac{\partial F_i}{\partial \dot{q}^a}\Gamma^a, F_i\right)\,.$$

We look for an immersion $F : M \to T^*Q$ such that $\mathrm{Im}(\mu_{\Gamma,F})$ is isotropic in (T^*TQ, ω_{TQ}), that is, such that the following conditions are satisfied:

$$0 = \frac{\partial F_a}{\partial \dot{q}^b} + \frac{\partial \psi^\alpha}{\partial \dot{q}^a}\frac{\partial F_\alpha}{\partial \dot{q}^b} - \frac{\partial F_b}{\partial \dot{q}^a} - \frac{\partial \psi^\alpha}{\partial \dot{q}^b}\frac{\partial F_\alpha}{\partial \dot{q}^a}\,, \tag{5.9}$$

$$0 = \frac{\partial^2 F_i}{\partial q^j \partial q^b}\dot{q}^b + \frac{\partial^2 F_i}{\partial q^j \partial q^\beta}\psi^\beta + \frac{\partial F_i}{\partial q^\beta}\frac{\partial \psi^\beta}{\partial q^j} + \frac{\partial^2 F_i}{\partial q^j \partial \dot{q}^b}\Gamma^b + \frac{\partial F_i}{\partial \dot{q}^b}\frac{\partial \Gamma^b}{\partial q^j} + \frac{\partial \psi^\alpha}{\partial q^i}\frac{\partial F_\alpha}{\partial q^j}$$

$$-\frac{\partial^2 F_j}{\partial q^i \partial q^b}\dot{q}^b - \frac{\partial^2 F_j}{\partial q^i \partial q^\beta}\psi^\beta - \frac{\partial F_j}{\partial q^\beta}\frac{\partial \psi^\beta}{\partial q^i} - \frac{\partial^2 F_j}{\partial q^i \partial \dot{q}^b}\Gamma^b - \frac{\partial F_j}{\partial \dot{q}^b}\frac{\partial \Gamma^b}{\partial q^i} - \frac{\partial \psi^\alpha}{\partial q^j}\frac{\partial F_\alpha}{\partial q^i}\,, \tag{5.10}$$

$$\begin{aligned} 0 \;=\; & \frac{\partial^2 F_i}{\partial \dot{q}^a \partial q^b}\dot{q}^b + \frac{\partial F_i}{\partial q^a} + \frac{\partial^2 F_i}{\partial \dot{q}^a \partial q^\beta}\psi^\beta + \frac{\partial F_i}{\partial q^\beta}\frac{\partial \psi^\beta}{\partial \dot{q}^a} + \frac{\partial^2 F_i}{\partial \dot{q}^a \partial \dot{q}^b}\Gamma^b + \frac{\partial F_i}{\partial \dot{q}^b}\frac{\partial \Gamma^b}{\partial \dot{q}^a} \\ & -\frac{\partial F_a}{\partial q^i} + \frac{\partial \psi^\alpha}{\partial q^i}\frac{\partial F_\alpha}{\partial \dot{q}^a} - \frac{\partial \psi^\alpha}{\partial \dot{q}^a}\frac{\partial F_\alpha}{\partial q^i}\,. \end{aligned} \tag{5.11}$$

We will refer to them as **constrained Helmholtz conditions**.

Now we will see the relationship between $\mathrm{Im}(\mu_{\Gamma,F})$ and the dynamics given by the SODE Γ on M. Take the submanifold $\alpha_Q^{-1}(\mathrm{Im}(\mu_{\Gamma,F})) = TF(\Gamma(M))$ of TT^*Q. Since TT^*Q is a tangent bundle, we have dynamics related to any submanifold. In our case $TF(\Gamma(M))$ is given by

$$\left(q^i, F_i(q^j, \dot{q}^b), \dot{q}^a, \psi^\alpha(q^j, \dot{q}^b), \frac{\partial F_i}{\partial q^a}\dot{q}^a + \frac{\partial F_i}{\partial q^\alpha}\psi^\alpha + \frac{\partial F_i}{\partial \dot{q}^a}\Gamma^a\right)$$

in the typical coordinates in TT^*Q. Tangent curves to this submanifold satisfy the equations

$$\dot{q}^\alpha = \psi^\alpha(q^j, \dot{q}^b) \quad \text{and} \quad \frac{d}{dt}F_i = \frac{\partial F_i}{\partial q^a}\dot{q}^a + \frac{\partial F_i}{\partial q^\alpha}\psi^\alpha + \frac{\partial F_i}{\partial \dot{q}^a}\Gamma^a.$$

Then

$$\frac{\partial F_i}{\partial \dot{q}^a}\left(\ddot{q}^a - \Gamma^a(q^j, \dot{q}^b)\right) = 0.$$

Since $\left(\frac{\partial F_i}{\partial \dot{q}^a}\right)$ is assumed to have maximal rank, we get $\ddot{q}^a = \Gamma^a(q^j, \dot{q}^b)$ and $\dot{q}^\alpha = \psi^\alpha(q^j, \dot{q}^b)$. In this case we have seen that the isotropic submanifold $TF(\Gamma(M)) = \alpha_Q^{-1}(\Sigma_{\Gamma,F})$ on TT^*Q carries the original dynamics defined by the SODE Γ on M.

Now we will generalize the characterization of Theorem 1.4.1 to the case of constrained systems.

Theorem 5.3.2. *A SODE Γ on M is variational if and only if there exists a two-form Ω on M satisfying*

(i) $d\Omega = 0$,

(ii) $\Omega(v_1, v_2) = 0$ *for all* $v_1, v_2 \in V(M)$,

(iii) $\mathcal{L}_\Gamma \Omega = 0$,

(iv) $\flat_\Omega|_{V(M)}$ *is injective.*

Proof. $\Rightarrow$ Assume that Γ is variational, that is, there exists an immersion $F : M \longrightarrow T^*Q$ such that $\Sigma_{\Gamma,F} = \mathrm{Im}(\mu_{\Gamma,F})$ is isotropic in (T^*TQ, ω_{TQ}). Then we define $\Omega = dF^*\theta_Q \in \Lambda^2(M)$. We first prove that $\Sigma_{\Gamma,F}$ is isotropic if and only if $\mathrm{Im}(\mathcal{L}_\Gamma F^*\theta_Q)$ is Lagrangian in (T^*M, ω_M), that is, $d(\mathcal{L}_\Gamma F^*\theta_Q) = \mathcal{L}_\Gamma \Omega = 0$. In local coordinates $(q^i, \dot{q}^a)$ on M,

$$\mathcal{L}_\Gamma F^*\theta_Q = \left(\Gamma(F_i) + \frac{\partial \psi^\alpha}{\partial q^i}F_\alpha\right)dq^i + \left(F_a + \frac{\partial \psi^\alpha}{\partial \dot{q}^a}F_\alpha\right)d\dot{q}^a.$$

On the other hand, $\Sigma_{\Gamma,F}$ is given by the following set of points of T^*TQ:

$$\left(q^i, \dot{q}^a, \psi^\alpha, \frac{\partial F_i}{\partial q^a}\dot{q}^a + \frac{\partial F_i}{\partial q^\alpha}\psi^\alpha + \frac{\partial F_i}{\partial \dot{q}^a}\Gamma^a, F_i\right).$$

If we denote by $i_{\Sigma_{\Gamma,F}}:\Sigma_{\Gamma,F}\longrightarrow T^*TQ$ the inclusion, then

$$i^*_{\Sigma_{\Gamma,F}}\theta_{TQ}=\Gamma(F_i)dq^i+F_a d\dot q^a+F_\alpha d\psi^\alpha=\left(\Gamma(F_i)+\frac{\partial\psi^\alpha}{\partial q^i}F_\alpha\right)dq^i+\left(F_a+\frac{\partial\psi^\alpha}{\partial\dot q^a}F_\alpha\right)d\dot q^a.$$

Now it is clear that the condition of isotropy, $i^*_{\Sigma_{\Gamma,F}}\omega_{TQ}=0$, is equivalent to the condition of $\mathrm{Im}(\mathcal{L}_\Gamma F^*\theta_Q)$ being Lagrangian in (T^*M,ω_M), in other words, $d\mathcal{L}_\Gamma F^*\theta_Q=\mathcal{L}_\Gamma\Omega=0$.

The first two properties in the statement of the theorem follow directly from the definition of Ω and the last one from F being an immersion. Indeed, $\Omega=\left(\frac{\partial F_i}{\partial q^j}\right)dq^j\wedge dq^i+\left(\frac{\partial F_i}{\partial\dot q^a}\right)d\dot q^a\wedge dq^i$, and for any v_1,v_2 in $V(M)$, $i_{v_1}\Omega-i_{v_2}\Omega=(v_1^a-v_2^a)\left(\frac{\partial F_i}{\partial\dot q^a}\right)dq^i=0$ for all $i=1,\dots,n$. As $\left(\frac{\partial F_i}{\partial\dot q^a}\right)$ has maximal rank, $v_1=v_2$.

$\Leftarrow$ Conversely, given a two-form on M satisfying the conditions in the statement, we construct an immersion that provides an isotropic submanifold $\Sigma_{\Gamma,F}$ of (T^*TQ,ω_{TQ}). Since $d\Omega=0$, locally we can write $\Omega=d\Theta$. Then using the second condition we get that there exists a locally defined function f on M such that $\Theta(v)=df(v)$ for each vertical vector $v\in V(M)$. We can define $\tilde\Theta=\Theta-df$ which is a semi-basic one-form on M, that is, it vanishes on vertical vectors and can be written in coordinates as $\tilde\Theta=\mu_i dq^i$, μ_i being functions on M. Moreover $d\tilde\Theta=\Omega$. Then we define $F\colon M\to T^*Q$ by

$$\langle F(m),v_q\rangle=\langle\tilde\Theta(m),w_m\rangle\,,$$

where $m\in M$ and w_m is any vector in T_mM satisfying $T_m\tau_Q|_M(w_m)=v_q$. This definition does not depend on the choice of w_m since $\tilde\Theta$ vanishes on vertical vectors and it gives $\tilde\Theta=F^*\theta_Q$.

Since the one-form $\mathcal{L}_\Gamma F^*\theta_Q\in\Lambda^1(M)$ is closed, then $\mathrm{Im}(\mathcal{L}_\Gamma F^*\theta_Q)$ is a Lagrangian submanifold of (T^*M,ω_M). Having Proposition A.2.3 in mind, we obtain from it a Lagrangian submanifold of (T^*TQ,ω_{TQ}),

$$\mathrm{Im}(\widetilde{\mathcal{L}_\Gamma F^*}\theta_Q)=\{\mu\in T^*TQ\mid i^*_M\mu\in\mathrm{Im}(\mathcal{L}_\Gamma F^*\theta_Q)\}\,,$$

where $i_M:M\hookrightarrow TQ$ is the canonical inclusion. In coordinates, $\mathrm{Im}(\widetilde{\mathcal{L}_\Gamma F^*}\theta_Q)$ is expressed as

$$\left(q^i,\dot q^a,\psi^\alpha,\Gamma(F_i)+\frac{\partial\psi^\alpha}{\partial q^i}F_\alpha-\frac{\partial\psi^\alpha}{\partial q^i}\tilde p_\alpha,F_a+\frac{\partial\psi^\alpha}{\partial\dot q^a}F_\alpha-\frac{\partial\psi^\alpha}{\partial\dot q^a}\tilde p_\alpha,\tilde p_\alpha\right).$$

In particular for $\tilde p_\alpha=F_\alpha$ we have

$$\mathrm{Im}(\mu_{\Gamma,F})\subset\mathrm{Im}(\widetilde{\mathcal{L}_\Gamma F^*}\theta_Q)\,.$$

As $\mathcal{L}_\Gamma\Omega=0$, we get that both $\mathrm{Im}(\mathcal{L}_\Gamma F^*\theta_Q)$ and $\mathrm{Im}(\widetilde{\mathcal{L}_\Gamma F^*}\theta_Q)$ need to be Lagrangian and therefore $\mathrm{Im}(\mu_{\Gamma,F})$ is isotropic in (T^*TQ,ω_{TQ}).

Finally since $\flat_\Omega|_{V(TM)}$ is injective and $d\tilde\Theta=\Omega$, $\left(\frac{\partial F_i}{\partial\dot q^a}\right)$ has maximal rank and F is an immersion. Now we conclude that Γ is variational according to Definition 5.3.1. ■

Remark 5.3.3. Note that in the proof above we have described a way to assign to each isotropic submanifold $\Sigma_{\Gamma,F}$ a Lagrangian submanifold that contains it and projects over the constraint submanifold, see Proposition A.2.3. From $\mathcal{L}_\Gamma\Omega=0$ we obtain a locally defined function $l:M\longrightarrow\mathbb{R}$

such that $\mathcal{L}_\Gamma F^*\theta_Q = dl$. Since $\mathrm{Im}(\widetilde{\mathcal{L}_\Gamma F^*\theta_Q})$ coincides with $\Sigma_l = \{\mu \in T^*TQ : i^*\mu = dl\} \subset T^*TQ$, the construction from Theorem A.2.4, it gives the constrained variational mechanics associated to l (see Section 5.2 and Appendix B). Summing up, given a variational SODE Γ on M, we can always find a local Lagrangian l on M such that the solutions of Γ are constrained variational trajectories for l.

Note that in this case we were not adressing the question of finding a Lagrangian $L : TQ \longrightarrow \mathbb{R}$ such that the solutions of the nonholonomic equations for L coincide with the solutions of Γ, but asking when the nonholonomic dynamics can be seen as constrained variational dynamics, see Sections 5.1 and 5.2.

Next we will study the problem of how to derive a description of the constrained dynamics in terms of a variational problem without constraints (see [17, 120]). We will need the following lemma.

Lemma 5.3.4. *Let P be a smooth manifold, $C \subset P$ a submanifold and γ a section of $T^*P|_C \longrightarrow C$, where $T^*P|_C = \{\mu \in T^*P : \pi_P(\mu) \in C\}$ and $\pi_P : T^*P \longrightarrow P$ denotes the projection over P. If $\gamma(C)$ is isotropic in (T^*P, ω_P), then there is a one-form $\tilde{\gamma}$ defined in a neighborhood of C such that*

- $\tilde{\gamma}|_C = \gamma$,
- $d\tilde{\gamma} = 0$.

Proof. Take adapted coordinates (x^i, y^a), $i = 1, \dots, n-m$, $a = 1, \dots, m$, on P such that C is given by $y^a = 0$ and denote the corresponding momenta coordinates by p_i and $\tilde{p}_a$. Then $\gamma(C)$ is given by

$$\left(x^i, 0, \gamma_i(x), \tilde{\gamma}_a(x)\right),$$

and it projects over C. The isotropy condition gives $\frac{\partial \gamma_i}{\partial x^j} = \frac{\partial \gamma_j}{\partial x^i}$. We want to see $\gamma(C)$ inside some submanifold N of T^*P of dimension $2n - m$ and then apply the construction at the end of Appendix A to extend it to a Lagrangian submanifold via the Hamiltonian vector fields corresponding to the constraints defining N. For that we have many options, for instance we can choose among the constraints

$$y^a = 0, \quad p_i - \gamma_i = 0, \quad \tilde{p}_a - \tilde{\gamma}_a = 0$$

and linear combinations of them. If we consider $\phi_a = \tilde{p}_a - \tilde{\gamma}_a$ the Hamiltonian vector field is given by

$$X_{\phi_a} = \frac{\partial}{\partial y^a} + \frac{\partial \tilde{\gamma}_a}{\partial x^j}\frac{\partial}{\partial p_j},$$

which satisfies $X_{\phi_a}(y_a) = 1$, so it is not tangent to $\gamma(C)$. Extending $\gamma(C)$ along the flows of X_{ϕ_a} we obtain

$$\left(x^i, y^a, \gamma_i(x) + \frac{\partial \tilde{\gamma}_a}{\partial x^i} y^a, \tilde{\gamma}_a(x)\right),$$

which is the image of $\tilde{\gamma} = dL$ with $L\colon P \to \mathbb{R}$, $L(x, y) = \tilde{\gamma}_a(x) y^a + f(x)$, not necessarily regular, and $\frac{\partial f}{\partial x^i} = \gamma_i$. The existence of such a function f on C is guaranteed by the isotropy condition. ∎

Remark 5.3.5. Note that there are many possible ways to choose the constraints and construct the corresponding Lagrangian functions. For instance, taking $\phi_a = y^a + \tilde{p}_a - \gamma_a$ we obtain $L = \gamma_a y^a + \gamma_i x^i - \frac{\partial \gamma_a}{\partial x^j}(y^a)^2$. On the other hand, if we take $\phi_a = y^a$, we obtain a Lagrangian submanifold projecting over M which corresponds to the constrained variational description.

As a consequence of Lemma 5.3.4, taking $\gamma(C) = \Sigma_{\Gamma,F}$ we obtain the following result.

Theorem 5.3.6. *If a SODE Γ on M is variational, then there exists a Lagrangian $L : TQ \longrightarrow \mathbb{R}$ such that the integral curves of Γ are the restriction of the solutions of the Euler-Lagrange equations of L to M.*

Example 5.3.7. Let $Q = \mathbb{R}^2$ with coordinates (x, y) and denote fibered coordinates on TQ and T^*TQ by $(x, y, \dot{x}, \dot{y})$ and $(x, y, \dot{x}, \dot{y}, \mu_x, \mu_y, \tilde{\mu}_x, \tilde{\mu}_y)$ respectively. Let $N = \{(x, y, \dot{x}, f(x, y, \dot{x}))\} \subset TQ$ be the constraint submanifold and the SODE Γ on N be given by $\ddot{x} = 0$. That is, we have the dynamics given by

$$\ddot{x} = 0, \quad \dot{y} = f(x, y, \dot{x}) \ .$$

We define $F : N \longrightarrow T^*Q$ by $F(x, y, \dot{x}) = (x, y, \dot{x} + y, x)$, which is an immersion. Then $\Sigma_{\Gamma,F} \subset T^*TQ$ is locally described by $(x, y, \dot{x}, f, f, \dot{x}, \dot{x} + y, x)$ and is an isotropic submanifold of dimension 3, for $dx \wedge df + dy \wedge d\dot{x} + d\dot{x} \wedge d(\dot{x} + y) + df \wedge dx = 0$. Note that

$$\mathcal{L}_\Gamma F^*\theta_Q = \left(f + x\frac{\partial f}{\partial x}\right) dx + \left(\dot{x} + x\frac{\partial f}{\partial y}\right) dy + \left(\dot{x} + y + x\frac{\partial f}{\partial \dot{x}}\right) d\dot{x} \, .$$

Therefore $\mathrm{Im}\widetilde{\mathcal{L}_\Gamma F^*}\theta_Q \subset T^*TQ$ is locally described by

$$\left(x, y, \dot{x}, f, f + x\frac{\partial f}{\partial x} - \frac{\partial f}{\partial x}\tilde{\mu}_y, \dot{x} + x\frac{\partial f}{\partial y} - \frac{\partial f}{\partial y}\tilde{\mu}_y, \dot{x} + y + x\frac{\partial f}{\partial \dot{x}} - \frac{\partial f}{\partial \dot{x}}\tilde{\mu}_y, \tilde{\mu}_y\right).$$

When $\tilde{\mu}_y = x$, $\Sigma_{\Gamma,F}$ is recovered. Since $d\mathcal{L}_\Gamma F^*\theta_Q = 0$, we have a local Lagrangian $l : N \longrightarrow \mathbb{R}$, $l = \frac{\dot{x}^2}{2} + \dot{x}y + xf(x, y, \dot{x})$, satisfying

$$\frac{\partial l}{\partial x} = f + x\frac{\partial f}{\partial x}, \; \frac{\partial l}{\partial y} = \dot{x} + x\frac{\partial f}{\partial y}, \; \frac{\partial l}{\partial \dot{x}} = \dot{x} + y + x\frac{\partial f}{\partial \dot{x}}.$$

Note that l is the restriction of the singular Lagrangian $L_1 = \frac{\dot{x}^2}{2} + \dot{x}y + x\dot{y}$ to $\dot{y} = f$.

Consider the constraint $\phi = \dot{y} - f + \tilde{\mu}_y - x$ and the corresponding Hamiltonian vector field for the symplectic structure ω_{TQ}, that is,

$$X_\phi = -\frac{\partial}{\partial \tilde{\mu}_y} + \frac{\partial f}{\partial \dot{x}}\frac{\partial}{\partial \tilde{\mu}_x} + \frac{\partial f}{\partial x}\frac{\partial}{\partial \mu_x} + \frac{\partial f}{\partial y}\frac{\partial}{\partial \mu_y} + \frac{\partial}{\partial \dot{y}} + \frac{\partial}{\partial \mu_x}.$$

If we extend the isotropic submanifold $\Sigma_{\Gamma,F}$ along its flow we obtain the Lagrangian submanifold

$$\left(x, y, \dot{x}, \dot{y}, \dot{y} + \frac{\partial f}{\partial x}\dot{y} - \frac{\partial f}{\partial x}f, \dot{x} + \frac{\partial f}{\partial y}\dot{y} - \frac{\partial f}{\partial y}f, \dot{x} + y + \frac{\partial f}{\partial \dot{x}}\dot{y} - \frac{\partial f}{\partial \dot{x}}f, x - \dot{y} + f\right),$$

which is the image of dL_2 with $L_2 = x\dot{y} - \frac{\dot{y}^2}{2} + f\dot{y} + \frac{\dot{x}^2}{2} + \dot{x}y - \frac{f^2}{2}$, another extension of l. However, this is a regular Lagrangian since $\det\left(\frac{\partial^2 L_2}{\partial \dot{q}^i \partial \dot{q}^j}\right) = -1 - \dot{y}\frac{\partial^2 f}{\partial \dot{x}^2} + f\frac{\partial^2 f}{\partial \dot{x}^2}$, which does not vanish in a neighborhood of $\Sigma_{\Gamma,F}$. It is possible to recover Γ by computing the corresponding Euler-Lagrange equations and restricting them to M.

Example 5.3.8 (Vertical rolling disk)**.** Consider the configuration space $Q = S^1 \times S^1 \times \mathbb{R}^2$ with coordinates (θ, φ, x, y), where θ denotes the angle of rotation, φ the angle between the direction in which the disk moves and the x-axis and (x, y) are the coordinates of the contact point. We consider the Lagrangian $L = \frac{1}{2}(\dot\theta^2 + \dot\varphi^2 + \dot x^2 + \dot y^2)$ and the constraints given by the condition of rolling without sliding are $\dot x = \cos(\varphi)\dot\theta$ and $\dot y = \sin(\varphi)\dot\theta$.

We know that for the rolling disk the nonholonomic equations are

$$\ddot\theta = 0, \quad \ddot\varphi = 0, \quad \dot x = \cos(\varphi)\dot\theta, \quad \dot y = \sin(\varphi)\dot\theta\,,$$

and the variational constrained ones are

$$2\ddot\theta = \dot\varphi(-A\sin(\varphi) + B\cos(\varphi)), \quad \ddot\varphi = \dot\theta(A\sin(\varphi) - B\cos(\varphi))\,,$$

$$\dot x = \cos(\varphi)\dot\theta, \quad \dot y = \sin(\varphi)\dot\theta\,,$$

where A and B are constants, see [16]. Taking $A = B = 0$ we see that the set of nonholonomic solutions is contained in the set of variational constrained ones. Now consider the constrained Lagrangian $l(\theta, \varphi, x, y, \dot\theta, \dot\varphi) = \dot\theta^2 + \frac{\dot\varphi^2}{2}$ and define F as the Legendre transformation associated to the extension $L(\theta, \varphi, x, y, \dot\theta, \dot\varphi, \dot x, \dot y) = \dot\theta^2 + \frac{\dot\varphi^2}{2}$, that is,

$$\begin{array}{rccc} F \equiv Leg_L : & M & \longrightarrow & T^*Q \\ & (\theta, \varphi, x, y, \dot\theta, \dot\varphi) & \longmapsto & (\theta, \varphi, x, y, 2\dot\theta, \dot\varphi, 0, 0)\,. \end{array}$$

As $\Gamma^1 = \Gamma^2 = 0$, the submanifold $\Sigma_{\Gamma,F} \subset T^*TQ$ can be locally described by

$$\Big(\theta, \varphi, x, y, \dot\theta, \dot\varphi, \cos(\varphi)\dot\theta, \sin(\varphi)\dot\theta, 0, 0, 0, 0, 2\dot\theta, \dot\varphi, 0, 0\Big)\,.$$

It is isotropic and has dimension 6, so we want to choose 2 constraint functions on T^*TQ satisfied by $\Sigma_{\Gamma,F}$ and extend it in the corresponding directions. First we take the constraints

$$\phi_1 = \dot x - \cos(\varphi)\dot\theta + \tilde\mu_x, \quad \phi_2 = \dot y - \sin(\varphi)\dot\theta + \tilde\mu_y\,,$$

with corresponding Hamiltonian vector fields

$$X_{\phi_1} = -\frac{\partial}{\partial\tilde\mu_x} + \cos(\varphi)\frac{\partial}{\partial\tilde\mu_\theta} - \sin(\varphi)\dot\theta\frac{\partial}{\partial\mu_\varphi} + \frac{\partial}{\partial\dot x}\,,$$

$$X_{\phi_2} = -\frac{\partial}{\partial\tilde\mu_y} + \sin(\varphi)\frac{\partial}{\partial\tilde\mu_\theta} + \cos(\varphi)\dot\theta\frac{\partial}{\partial\mu_\varphi} + \frac{\partial}{\partial\dot y}\,.$$

Extending $\Sigma_{\Gamma,F}$ along the flows of X_{ϕ_1} and X_{ϕ_2} we obtain the Lagrangian submanifold with local expression

$$\Big(\theta, \varphi, x, y, \dot\theta, \dot\varphi, \dot x, \dot y, 0, \dot\theta(\cos(\varphi)\dot y - \sin(\varphi)\dot x), 0, 0, \dot\theta + \cos(\varphi)\dot x + \sin(\varphi)\dot y, \dot\varphi, -\dot x + \cos(\varphi)\dot\theta, -\dot y + \sin(\varphi)\dot\theta\Big)$$

which is the image of $d\bar L$ with $\bar L = \frac{1}{2}(\dot\theta^2 + \dot\varphi^2 - \dot x^2 - \dot y^2) + \dot\theta(\cos(\varphi)\dot x + \sin(\varphi)\dot y)$. So we have obtained a regular Lagrangian whose unconstrained trajectories include the nonholonomic trajectories of the first Lagrangian. This is the same Lagrangian as the one obtained in [61].

If we take $\phi_1 = \tilde{\mu}_x,\ \phi_2 = \tilde{\mu}_y$ then we obtain the Lagrangian submanifold

$$(\theta, \varphi, x, y, \dot\theta, \dot\varphi, \dot x, \dot y, 0, 0, 0, 0, 2\dot\theta, \dot\varphi, 0, 0)$$

and recover the singular Lagrangian function $L = \dot\theta^2 + \frac{\dot\varphi^2}{2}$.

For $\phi_1 = \dot x - \cos(\varphi)\dot\theta,\quad \phi_2 = \dot y - \sin(\varphi)\dot\theta$ we get the Lagrangian submanifold

$$\Big(\theta, \varphi, x, y, \dot\theta, \dot\varphi, \cos(\varphi)\dot\theta, \sin(\varphi)\dot\theta, 0, \dot\theta(\tilde\mu_x \sin(\varphi) - \tilde\mu_y \cos(\varphi)), 0, 0,$$
$$2\dot\theta - \tilde\mu_x \cos(\varphi) - \tilde\mu_y \sin(\varphi), \dot\varphi, \tilde\mu_x, \tilde\mu_y\Big)\ ,$$

which coincides with $\mathrm{Im}(\widetilde{\mathcal{L}_\Gamma F^*}\theta_Q)$, for $\frac{\partial\psi^1}{\partial\dot\theta} = \cos(\varphi)$ and $\frac{\partial\psi^2}{\partial\dot\theta} = \sin(\varphi)$, where $\psi^1 = \cos(\varphi)\dot\theta,\ \psi^2 = \sin(\varphi)\dot\theta$. Therefore, we obtain the variational constrained equations for the constrained Lagrangian $l : M \to \mathbb{R}$.

Now we find another immersion $F : M \longrightarrow T^*Q$ that makes $\Sigma_{\Gamma,F}$ isotropic. After extending it we get new Lagrangian functions defined on TQ.

We make the following assumptions on the dependence of coordinates of F:

$$F_\theta(\dot\theta, \dot\varphi) = F_x(\dot\theta, \dot\varphi) = F_y(\dot\theta, \dot\varphi), \quad F_\varphi(\varphi, \dot\theta, \dot\varphi)\,.$$

Then the only constrained Helmholtz equations (5.9)-(5.11) that do not vanish identically are

$$\frac{\partial F_\varphi}{\partial\dot\theta} = (1 + \cos(\varphi) + \sin(\varphi))\frac{\partial F_\theta}{\partial\dot\varphi}\,, \tag{5.12}$$
$$0 = \dot\varphi\frac{\partial^2 F_\varphi}{\partial\dot\theta\partial\varphi} + \dot\theta(\cos(\varphi) - \sin(\varphi))\frac{\partial F_x}{\partial\dot\theta}\,, \tag{5.13}$$
$$\frac{\partial F_\varphi}{\partial\varphi} = \frac{\partial}{\partial\dot\varphi}\left(\frac{\partial F_\varphi}{\partial\varphi}\dot\varphi\right) + \dot\theta(\cos(\varphi) - \sin(\varphi))\frac{\partial F_x}{\partial\dot\varphi}\,, \tag{5.14}$$

and $F_\theta = F_x = F_y = \frac{\dot\theta}{\dot\varphi}$, $F_\varphi = \rho(\dot\varphi) - \frac{\dot\theta^2}{2\dot\varphi^2}(1 + \cos(\varphi) + \sin(\varphi))$ is a solution, where $\rho : \mathbb{R} \to \mathbb{R}$ is arbitrary.

Setting $\rho(\dot\varphi) = \dot\varphi$, define

$$\begin{array}{rccc} F: & M & \longrightarrow & T^*Q \\ & (\theta, \varphi, x, y, \dot\theta, \dot\varphi) & \longmapsto & \left(\theta, \varphi, x, y, \frac{\dot\theta}{\dot\varphi}, \dot\varphi - \frac{\dot\theta^2}{2\dot\varphi^2}\left(1 + \cos(\varphi) + \sin(\varphi)\right), \frac{\dot\theta}{\dot\varphi}, \frac{\dot\theta}{\dot\varphi}\right) \end{array}$$

to get $\Sigma_{\Gamma,F}$ given by

$$\left(\theta, \varphi, x, y, \cos(\varphi)\dot\theta, \sin(\varphi)\dot\theta, \dot\theta, \dot\varphi, 0, \frac{1}{2}\frac{\dot\theta^2}{\dot\varphi}(\sin(\varphi) - \cos(\varphi)), 0, 0, \frac{\dot\theta}{\dot\varphi}, \dot\varphi - \frac{\dot\theta^2}{2\dot\varphi^2}(1 + \cos(\varphi) + \sin(\varphi)), \frac{\dot\theta}{\dot\varphi}, \frac{\dot\theta}{\dot\varphi}\right),$$

which is isotropic of dimension 6 on (T^*TQ, ω_{TQ}) .

If we take $\phi_1 = \tilde\mu_x - \frac{\dot\theta}{\dot\varphi}$ and $\phi_2 = \tilde\mu_y - \frac{\dot\theta}{\dot\varphi}$, the corresponding Hamiltonian vector fields are

$$X_{\phi_1} = \frac{\partial}{\partial\dot x} + \frac{1}{\dot\varphi}\frac{\partial}{\partial\tilde\mu_\theta} - \frac{\dot\theta}{\dot\varphi^2}\frac{\partial}{\partial\tilde\mu_\varphi}\,,$$
$$X_{\phi_2} = \frac{\partial}{\partial\dot y} + \frac{1}{\dot\varphi}\frac{\partial}{\partial\tilde\mu_\theta} - \frac{\dot\theta}{\dot\varphi^2}\frac{\partial}{\partial\tilde\mu_\varphi}\,.$$

Extending $\Sigma_{\Gamma,F}$ along the flows of X_{ϕ_1} and X_{ϕ_2} we get

$$\left(\theta,\varphi,x,y,\dot\theta,\dot\varphi,\dot x,\dot y,0,\frac{1}{2}\frac{\dot\theta^2}{\dot\varphi}(\sin(\varphi)-\cos(\varphi)),0,0,\frac{\dot\theta}{\dot\varphi}(1-\sin(\varphi)-\cos(\varphi))+\frac{\dot x+\dot y}{\dot\varphi},\right.$$

$$\left.\dot\varphi-\frac{\dot\theta^2}{2\dot\varphi}(1-\sin(\varphi)-\cos(\varphi))-\frac{\dot\theta}{\dot\varphi^2}(\dot x+\dot y),\frac{\dot\theta}{\dot\varphi},\frac{\dot\theta}{\dot\varphi}\right),$$

which is the image of $d\bar L$ for the singular Lagrangian

$$\bar L=\frac{\dot\varphi^2}{2}+\frac{\dot\theta^2}{\dot\varphi}\left(\frac{1}{2}-\cos(\varphi)-\sin(\varphi)\right)+\frac{\dot\theta}{\dot\varphi}(\dot x+\dot y)\,.$$

Now we choose constraints $\phi_1=\dot x-\cos(\varphi)\dot\theta+\tilde\mu_x-\frac{\dot\theta}{\dot\varphi}$ and $\phi_2=\dot y-\sin(\varphi)\dot\theta+\tilde\mu_y-\frac{\dot\theta}{\dot\varphi}$ with Hamiltonian vector fields

$$\begin{aligned}X_{\phi_1}&=-\frac{\partial}{\partial\tilde\mu_x}+\left(\cos(\varphi)+\frac{1}{\dot\varphi}\right)\frac{\partial}{\partial\tilde\mu_\theta}-\frac{\dot\theta}{\dot\varphi^2}\frac{\partial}{\partial\tilde\mu_\varphi}-\dot\theta\sin(\varphi)\frac{\partial}{\partial\mu_\varphi}+\frac{\partial}{\partial\dot x},\\ X_{\phi_2}&=-\frac{\partial}{\partial\tilde\mu_y}+\left(\sin(\varphi)+\frac{1}{\dot\varphi}\right)\frac{\partial}{\partial\tilde\mu_\theta}-\frac{\dot\theta}{\dot\varphi^2}\frac{\partial}{\partial\tilde\mu_\varphi}+\dot\theta\cos(\varphi)\frac{\partial}{\partial\mu_\varphi}+\frac{\partial}{\partial\dot y}.\end{aligned}$$

Extending $\Sigma_{\Gamma,F}$ along their flows we obtain

$$\left(\theta,\varphi,x,y,\dot\theta,\dot\varphi,\dot x,\dot y,0,\frac{1}{2}\frac{\dot\theta^2}{\dot\varphi}(\sin(\varphi)-\cos(\varphi))-\dot x\dot\theta\sin(\varphi)+\dot\theta\dot y\cos(\varphi),0,0,\right.$$

$$\frac{\dot\theta}{\dot\varphi}(1-\cos(\varphi)-\sin(\varphi))+\dot x\cos(\varphi)-\dot\theta+\frac{\dot x+\dot y}{\dot\varphi}+\dot y\sin(\varphi),\dot\varphi-\frac{\dot\theta^2}{2\dot\varphi^2}(1-\cos(\varphi)-\sin(\varphi))-\frac{\dot\theta}{\dot\varphi^2}(\dot x+\dot y),$$

$$\left.\frac{\dot\theta}{\dot\varphi}-\dot x+\cos(\varphi)\dot\theta,\frac{\dot\theta}{\dot\varphi}-\dot y+\sin(\varphi)\dot\theta\right),$$

which is the image of $d\bar L$ for

$$\bar L=\frac{1}{2}\left(\dot\varphi^2-\dot\theta^2-\dot x^2-\dot y^2\right)+\frac{\dot\theta^2}{2\dot\varphi}\left(1-\cos(\varphi)-\sin(\varphi)\right)+\dot\theta\dot x\left(\cos(\varphi)+\frac{1}{\dot\varphi}\right)+\dot\theta\dot y\left(\sin(\varphi)+\frac{1}{\dot\varphi}\right).$$

As $\det\left(\frac{\partial^2 L}{\partial\dot q^i\partial\dot q^j}\right)=\frac{1}{\dot\varphi^5}\left(-2\dot\theta^2(1-\sin(\varphi)-\cos(\varphi))-\dot\theta^2\dot\varphi+2\dot\varphi^3+\dot\varphi^4(1+\sin(\varphi)+\cos(\varphi))\right)$, observe that this Lagrangian is regular except at a hypersurface of singular points.

Example 5.3.9 (Nonholonomic particle)**.** Consider the system defined by $Q=\mathbb{R}^3$, $L=\frac{1}{2}(\dot x^2+\dot y^2+\dot z^2)$ and constraint $\dot z=-x\dot y$. The nonholonomic SODE is given by $\Gamma^1=0,\Gamma^2=-\frac{x\dot x\dot y}{1+x^2}$. This SODE is variational as a constrained system as we will see. Indeed, in [17] the authors show that this system can be represented as the restriction of the Euler-Lagrange vector field associated to a Lagrangian defined on the full space TQ. In our framework, we define the map

$$\begin{array}{rccc}F: & M & \longrightarrow & T^*Q\\ & (x,y,z,\dot x,\dot y) & \longmapsto & \left(x,y,z,\dot x-\frac{\dot y^2}{2\dot x^2}\sqrt{1+x^2}\,(1+x)\,,\frac{\sqrt{1+x^2}\dot y}{\dot x},-\frac{\sqrt{1+x^2}\dot y}{\dot x}\right),\end{array}$$

and then $\Sigma_{\Gamma,F}$ is given by

$$\left(x,y,z,\dot x,\dot y,-x\dot y,-\frac{\dot y^2}{2\dot x}\frac{(1-x)}{\sqrt{1+x^2}},0,0,\dot x-\frac{\dot y^2}{2\dot x^2}\sqrt{1+x^2}\,(1+x)\,,\frac{\sqrt{1+x^2}\dot y}{\dot x},-\frac{\sqrt{1+x^2}\dot y}{\dot x}\right),$$

which is isotropic in (T^*TQ, ω_{TQ}). Also $\mathcal{L}_\Gamma F^*\theta_Q = dl$ for

$$l = \frac{\dot{x}^2}{2} + \frac{\dot{y}^2}{2}\frac{\sqrt{1+x^2}}{\dot{x}}(1+x).$$

Note that $l \neq L|_M = \frac{1}{2}(\dot{x}^2+\dot{y}^2(1+x^2))$. Since $\Sigma_{\Gamma,F} \subset \Sigma_l$, the solutions of Γ can be seen as constrained variational for l, although not for $L|_M$ (see [61]).

Now we look for a Lagrangian on TQ. Taking $\phi = \tilde{\mu}_z + \frac{\sqrt{1+x^2}\dot{y}}{\dot{x}}$ as a constraint and extending $\Sigma_{\Gamma,F}$ along the flow of

$$X_\phi = \frac{\partial}{\partial \dot{z}} - \frac{\sqrt{1+x^2}}{\dot{x}}\frac{\partial}{\partial \tilde{\mu}_y} + \frac{\sqrt{1+x^2}\dot{y}}{\dot{x}^2}\frac{\partial}{\partial \tilde{\mu}_x} - \frac{\dot{y}}{\dot{x}}\frac{x}{\sqrt{1+x^2}}\frac{\partial}{\partial \mu_x}$$

we get

$$\left(x, y, z, \dot{x}, \dot{y}, \dot{z}, -\frac{\dot{y}^2}{2\dot{x}}\frac{(1-x)}{\sqrt{1+x^2}} - \frac{\dot{y}}{\dot{x}}\frac{x}{\sqrt{1+x^2}}(\dot{z}+x\dot{y}), 0, 0, \dot{x} + \frac{\sqrt{1+x^2}}{2\dot{x}^2}\dot{y}^2(x-1),\right.$$
$$\left.\frac{\sqrt{1+x^2}}{\dot{x}}\dot{y}(1-x) - \frac{\sqrt{1+x^2}}{\dot{x}}\dot{z}, -\frac{\sqrt{1+x^2}}{\dot{x}}\dot{y}\right),$$

generated by the regular Lagrangian

$$\bar{L} = \frac{\dot{x}^2}{2} + \frac{(1-x)\sqrt{1+x^2}}{2\dot{x}}\dot{y}^2 - \frac{\sqrt{1+x^2}}{\dot{x}}\dot{z}\dot{y}.$$

Remark 5.3.10. If a SODE Γ on M is variational, from Theorem 5.3.6 we know that there exists a Lagrangian function such that its associated Euler-Lagrange vector field Γ_L verifies $(\Gamma_L)|_M = \Gamma$. Since $i_{\Gamma_L}\Omega_L = dE_L$ and $i_M\colon M \to TQ$, if we define the two-form $\Omega_M = i_M^*\Omega_L$ then $i_\Gamma\Omega_M = d(E_L)|_M$. As a result, the flow of Γ preserves the two-form Ω_M (this result is also a direct consequence of Theorem 5.3.2). Hence, $\mathcal{L}_\Gamma\Omega_M^k = 0$, for all k, giving information about the qualitative behavior of the flow of Γ.

Additionally, if we derive a constant of motion $I : TQ \to \mathbb{R}$ for Γ_L then the restriction of I to M is also a constant of motion of Γ. Thus, $(E_L)|_M$ is a constant of motion of Γ. We can also apply Noether's theorem to derive these constants of motion. For example, if we find a vector field $X \in \mathfrak{X}(Q)$ such that $X^c(L) = 0$ then $X^v(L)$ is a constant of motion of Γ_L and so is $X^v(L)|_M$ for Γ. Here X^c and X^v denote respectively the complete and vertical lift of X.

5.4 The inverse problem for holonomic constraints

A particular case of constrained systems is given by a submanifold M of TQ which is precisely a tangent bundle of a submanifold N of Q, that is $M = TN$, which is the case of **holonomic constraints**. In many cases of interest it is useful *to work extrinsically*, that is, on the manifold Q instead of *intrinsically*, that is, on N. As a result, the system on N is described in terms of a system on Q. Assume that TN is locally described by the vanishing of the constraints

$$\psi^\alpha(q^a, q^\beta) = 0 \quad \text{and} \quad \frac{\partial\psi^\alpha}{\partial q^a}\dot{q}^a + \frac{\partial\psi^\alpha}{\partial q^\beta}\dot{q}^\beta = 0, \quad 1 \leq \alpha \leq m\,.$$

For simplicity and without loss of generality, we consider the local coordinates on Q adapted to N and the corresponding local coordinates on TQ adapted to TN, so that

$$N = \{(q^a, q^\alpha) \in Q \mid q^\alpha = 0\} \quad \text{and} \quad TN = \{(q^a, q^\alpha, \dot{q}^a, \dot{q}^\alpha) \in TQ \mid q^\alpha = 0,\ \dot{q}^\alpha = 0\}\,,$$

where $a = 1, \dots, n-m$. The SODE Γ on TN is locally described by

$$\Gamma(q^a, \dot{q}^a) = (q^a, \dot{q}^a, \dot{q}^a, \Gamma^a(q^b, \dot{q}^b))\,.$$

The difference between holonomic dynamics and the nonholonomic one considered in Section 5.3 is that $M = TN$ does not project over the entire configuration manifold Q. Thus, the notion of variational SODE for constrained systems in Definition 5.3.1 must be adapted, because if M does not project over the entire manifold Q, then $F\colon M \to T^*Q$ might not be an immersion.

Definition 5.4.1. *Let Γ be a SODE along M and assume that $N = \tau_Q(M)$ is a submanifold so that we have the canonical inclusion $i_{TN}\colon TN \to TQ$. The SODE Γ is variational if there exists a function $F\colon M \to T^*Q$ such that the map $(i^*_{TN} \circ F)_{|M\cap TN}\colon M \cap TN \to T^*N$ is an immersion and $\Sigma_{\Gamma,F} = Im(\alpha_Q \circ TF \circ \Gamma)$ is an isotropic submanifold of (T^*TQ, ω_{TQ}), where i^*_{TN} is the transpose map of i_{TN} as defined below in* (5.15).

With this adapted notion of a variational SODE for holonomic constraints, Theorem 5.3.2 can also be proved similarly to the proof in Section 5.3 for the case when M projects onto the entire Q.

Our goal now is to establish a relationship between the inverse problem without constraints when we work intrinsically on TN and the inverse problem with holonomic constraints, when we work extrinsically on TQ.

$$\begin{array}{ccccccccc}
T^*TN & \xleftarrow{\alpha_N} & TT^*N & \xleftarrow{Tf} & TTN & \xrightarrow{TF} & TT^*Q & \xrightarrow{\alpha_Q} & T^*TQ \\
 & \nwarrow^{\mu_{\Gamma,f}} & & & \uparrow{\scriptstyle\Gamma} & & & \nearrow^{\mu_{\Gamma,F}} & \\
 & & T^*N & \xleftarrow[f]{} & TN & \xrightarrow[F]{} & T^*Q & &
\end{array}$$

Theorem 5.4.2. *A SODE Γ on TN is variational for the inverse problem of the calculus of variations without constraints if and only if it is variational along the submanifold TN of TQ in the inverse problem for constrained systems.*

Proof. $\Rightarrow$ If Γ is variational for the unconstrained system on TN, then there exists a regular Lagrangian $l\colon TN \to \mathbb{R}$ whose solutions of the Euler-Lagrange equations are also integral curves of the SODE Γ and vice versa. The function $f\colon TN \to T^*N$ in the above diagram is the Legendre transformation of l, that is, $f(q,\dot{q}) = \mathrm{Leg}_l(q,\dot{q}) = (q, \partial l/\partial \dot{q})$. Moreover, $\mathrm{Im}(\mu_{\Gamma,f})$ is a Lagrangian submanifold of (T^*TN, ω_N).

Let $i_{TN}\colon TN \to TQ$ be the inclusion and consider an arbitrary fiber function $F: TN \to T^*Q$ such that the diagram

$$\begin{array}{ccc}
TN & \xrightarrow{F} & T^*_N Q \\
{\scriptstyle \mathrm{Leg}_l}\downarrow & \swarrow_{i^*_{TN}} & \\
T^*N & &
\end{array}$$

is commutative, where i^*_{TN} is the transpose map of i_{TN} defined by

$$\langle i^*_{TN}(p_{\tilde{q}}), v_q\rangle = \langle p_{\tilde{q}}, i_{TN}(v_q)\rangle \,, \tag{5.15}$$

where $p_{\tilde{q}} \in T^*_N Q$, $v_q \in TN$, $(\tau_Q \circ i_{TN})(v_q) = \pi_Q(p_{\tilde{q}})$, $q \in N$, $\tilde{q} \in Q$, $\pi_N(q) = \tilde{q}$.

Since $l : TN \to \mathbb{R}$ is regular (that is, $\mathrm{Leg}_l : TN \to T^*N$ is a local diffeomorphism), it is easy to deduce that $F : TN \to T^*Q$ is an immersion. In local coordinates, the function F looks like

$$\begin{array}{rcl} F\colon TN & \longrightarrow & T^*Q \\ (q^a, \dot{q}^a) & \longmapsto & \left(q^a, 0, \frac{\partial l}{\partial \dot{q}^a}, F_\alpha(q^b, \dot{q}^b)\right) \,, \end{array}$$

where F_α are arbitrary functions on TN.

The local expression in adapted coordinates of the submanifold $\mathrm{Im}(\mu_{\Gamma,F})$ of T^*TQ is

$$\left(q^a, 0, \dot{q}^a, 0; \frac{\partial^2 l}{\partial \dot{q}^a \partial q^b}\dot{q}^b + \frac{\partial^2 l}{\partial \dot{q}^a \partial \dot{q}^b}\Gamma^b, \frac{\partial F_\alpha}{\partial q^b}\dot{q}^b + \frac{\partial F_\alpha}{\partial \dot{q}^b}\Gamma^b, \frac{\partial l}{\partial \dot{q}^a}, F_\alpha\right) .$$

This submanifold is isotropic if $(\mu_{\Gamma,F})^*(\omega_{TQ})$ vanishes, or equivalently if

$$d\left(\frac{\partial^2 l}{\partial \dot{q}^a \partial q^b}\dot{q}^b + \frac{\partial^2 l}{\partial \dot{q}^a \partial \dot{q}^b}\Gamma^b\right) \wedge dq^a + d\left(\frac{\partial l}{\partial \dot{q}^a}\right) \wedge d\dot{q}^a = d^2 l = 0$$

because Γ is the Euler-Lagrange vector field for $l : TN \to \mathbb{R}$, that is, locally

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{q}^a}\right) = \frac{\partial^2 l}{\partial \dot{q}^a \partial q^b}\dot{q}^b + \frac{\partial^2 l}{\partial \dot{q}^a \partial \dot{q}^b}\Gamma^b = \frac{\partial l}{\partial q^a} \,.$$

$\Leftarrow$ Assuming now that Γ is variational for the inverse problem with constraints, then there exists $F\colon TN \to T^*Q$ such that the map $(i^*_{TN} \circ F)\colon TN \to T^*N$ is an immersion and $\mathrm{Im}(\mu_{\Gamma,F})$ is isotropic in (T^*TQ, ω_{TQ}). Now we find a solution of the inverse problem of the calculus of variations (without constraints) by taking $f = i^*_{TN} \circ F : TN \to T^*N$. In coordinates, $f(q^a, \dot{q}^a) = (q^a, F_a(q^b, \dot{q}^b))$. Obviously, $\mathrm{Im}(\mu_{\Gamma,f})$ is Lagrangian in (T^*TN, ω_{TN}) and f is a local diffeomorphism.

This result can be also proved intrinsically because f and F must make the following diagram commutative:

$$\begin{array}{ccc} TN & \xrightarrow{F} & T^*_N Q \\ {\scriptstyle f}\downarrow & \swarrow {\scriptstyle i^*_{TN}} & \\ T^*N & & \end{array}$$

Note that the diagram is commutative if $F_a = f_a$, but the remaining F_α are arbitrary. It can be easily proved that $f^*\theta_N = F^*\theta_Q$. Then the two-form characterizing the inverse problem for the calculus of variations, Theorem 1.4.1, and the one characterizing the inverse problem for constrained systems, Theorem 5.3.2, coincide. This concludes the proof. ■

Let Γ be a SODE on TN which is the Euler-Lagrange vector field corresponding to a regular Lagrangian $l : TN \to \mathbb{R}$. Applying Theorem 5.4.2 we obtain an isotropic submanifold of (T^*TQ, ω_{TQ}) by simply taking $\mathrm{Im}(\mu_{\Gamma,F})$ for any map $F : M \to T^*Q$ verifying

$$i^*_{TN} \circ F = \mathrm{Leg}_l \,,$$

where $\mathrm{Leg}_l : TN \to T^*N$ is the Legendre transformation associated to $l : TN \to \mathbb{R}$.

Recall that in Section 5.3 for the case of a submanifold projecting over the entire Q, we saw that a constrained variational SODE could be seen as the restriction of a variational SODE on TQ, Theorem 5.3.6. In order to do this we just need to find a Lagrangian submanifold projecting over the entire TQ and containing $\mathrm{Im}(\mu_{\Gamma,F})$ which in this case has the expression

$$\left(q^a, 0, \dot{q}^a, 0; \frac{\partial^2 l}{\partial \dot{q}^a \partial q^b}\dot{q}^b + \frac{\partial^2 l}{\partial \dot{q}^a \partial \dot{q}^b}\Gamma^b, \frac{\partial F_\alpha}{\partial q^b}\dot{q}^b + \frac{\partial F_\alpha}{\partial \dot{q}^b}\Gamma^b, \frac{\partial l}{\partial \dot{q}^a}, F_\alpha\right).$$

If we take a Lagrangian $L : TQ \to \mathbb{R}$ such that $L_{|TN} = l$ and verifying

$$\frac{\partial L}{\partial q^\alpha} = \frac{\partial^2 L}{\partial q^a \partial \dot{q}^\alpha}\dot{q}^a + \frac{\partial^2 L}{\partial \dot{q}^a \partial \dot{q}^\alpha}\Gamma^a$$

on TN, then we can define $F = \mathrm{Leg}_L\big|_{TN} : TN \to T^*Q$ and get $\mathrm{Im}(\mu_{\Gamma,F}) \subset dL$.

For instance, in adapted coordinates to TN, we can take any Lagrangian $L\colon TQ \to \mathbb{R}$ of the form

$$L(q,\dot{q}) = l(q^a,\dot{q}^a) + \frac{1}{2}(\dot{q}^\alpha)^2 A_\alpha(q,\dot{q}) + \frac{1}{2}(q^\alpha)^2 B_\alpha(q,\dot{q})\,,$$

where $A_\alpha, B_\alpha \in C^\infty(TQ)$. Obviously

$$F(q^a,\dot{q}^a) = \left(q^a, 0, \frac{\partial l}{\partial \dot{q}^a}, 0\right)\,.$$

Therefore, we conclude that the solutions of the holonomic problem given by l are included in the solutions of L with initial conditions given on TN.

Example 5.4.3. Planar pendulum of length h with a particle of mass m. In this case $TN = TS^1$ and $TQ = T\mathbb{R}^2$. The local adapted coordinates are $(q^1, q^2) = (\theta, r-h)$. We consider the SODE Γ on TS^1 coming from the Lagrangian $l\colon TS^1 \to \mathbb{R}$,

$$l(\theta,\dot{\theta}) = \frac{1}{2}mh^2\dot{\theta}^2 - mgh\cos\theta\,.$$

In this case $f(\theta,\dot{\theta}) = (\theta, mh^2\dot{\theta})$ and we could take $F(\theta,\dot{\theta}) = \Big(\theta, 0, mh^2\dot{\theta}, F_2(\theta,0,\dot{\theta},0)\Big)$. Proposition 5.4.2 guarantees that $\mathrm{Im}(\mu_{\Gamma,F})$ is isotropic in (T^*TQ, ω_{TQ}). A choice of Lagrangian $L\colon TQ \to \mathbb{R}$ associated with that F is

$$L = \frac{1}{2}mh^2\dot{\theta}^2 - mgh\cos\theta + \frac{1}{2}\dot{r}^2 A(q,\dot{q}) + \frac{1}{2}(r-h)^2 B(q,\dot{q})$$

and a regular one is, for instance,

$$L = l + \frac{1}{2}\dot{r}^2 + B(q,\dot{q})(r-h)^2.$$

5.5 The inverse problem for time-dependent constrained systems

Now let us extend the time-dependent Helmholtz conditions reviewed in Section 1.5.2 to constrained systems. Let $M \subset TQ$ be a submanifold projecting over the whole configuration manifold Q, and Γ a SODE on $\mathbb{R} \times M$. If $(t, q^i, \dot q^a)$ denote coordinates on $\mathbb{R} \times M$, $i = 1, \ldots, n = \dim Q$, $a = 1, \ldots, m \leq n$, then the solutions of Γ are given by

$$\ddot q^a = \Gamma^a(t, q^j, \dot q^b), \quad \dot q^\alpha = \psi^\alpha(t, q^j, \dot q^b), \quad \alpha = 1, \ldots, n-m\,.$$

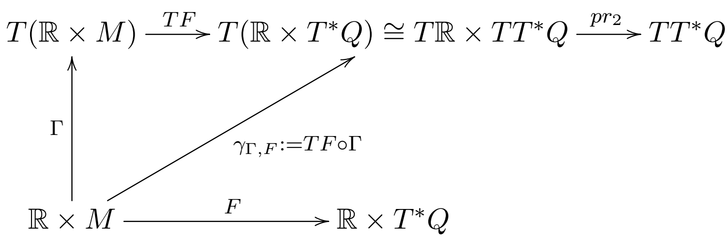
As in Section 5.3 we need to introduce the notion of isotropic submanifolds but now in the Poisson context (see Section 1.5.2).

Definition 5.5.1 ([155])**.** *Let $(P, \{\cdot,\cdot\})$ be a Poisson manifold and denote by $\sharp : T^*P \longrightarrow TP$ the morphism of vector bundles induced by the Poisson bivector. Let $N \subset P$ be a submanifold. We say that it is isotropic if*

$$\sharp(TN^\circ) \supseteq TN \cap \mathcal{C}\,.$$

Recall that $\mathcal{C} = Im(\sharp)$ denotes the characteristic distribution.

Definition 5.5.2. *We say that a SODE Γ on $\mathbb{R} \times M$ is variational if there is an immersion $F : \mathbb{R}\times M \longrightarrow \mathbb{R}\times T^*Q$ over $\mathbb{R}\times Q$ such that $Im(TF\circ\Gamma)$ is an isotropic submanifold of $(T(\mathbb{R}\times T^*Q), \{\cdot,\cdot\}^T)$.*



We will now impose the isotropy condition on $\mathrm{Im}(\gamma_{\Gamma,F})$ to obtain the time-dependent Helmholtz conditions for constrained systems. In local coordinates $\gamma_{\Gamma,F}$ is given by

$$\gamma_{\Gamma,F}(t, q^i, \dot q^a) = \left(t, q^i, F_i, 1, \dot q^a, \psi^\alpha, \Gamma(F_i) = \frac{\partial F_i}{\partial t} + \dot q^a \frac{\partial F_i}{\partial q^a} + \psi^\alpha \frac{\partial F_i}{\partial q^\alpha} + \Gamma^a \frac{\partial F_i}{\partial \dot q^a}\right).$$

We also have

$$\begin{aligned}
T(\mathrm{Im}(\gamma_{\Gamma,F})) \cap \mathcal{C} &= \mathrm{span}\left\{ V_i := \frac{\partial}{\partial q^i} + \frac{\partial F_j}{\partial q^i}\frac{\partial}{\partial p_j} + \frac{\partial \psi^\alpha}{\partial q^i}\frac{\partial}{\partial \dot q^\alpha} + \frac{\partial \Gamma(F_j)}{\partial q^i}\frac{\partial}{\partial \dot p_j}, \right.\\
&\qquad \left. W_a := \frac{\partial}{\partial \dot q^a} + \frac{\partial F_i}{\partial \dot q^a}\frac{\partial}{\partial p_i} + \frac{\partial \psi^\alpha}{\partial \dot q^a}\frac{\partial}{\partial \dot q^\alpha} + \frac{\partial \Gamma(F_i)}{\partial \dot q^a}\frac{\partial}{\partial \dot p_i} \right\},\\
\sharp(T\mathrm{Im}(\gamma_{\Gamma,F})^\circ) &= \mathrm{span}\left\{ A_i := \frac{\partial}{\partial \dot q^i} + \frac{\partial F_i}{\partial q^j}\frac{\partial}{\partial \dot p_j} + \frac{\partial F_i}{\partial \dot q^a}\frac{\partial}{\partial p_a}, B_i := \frac{\partial}{\partial q^i} + \frac{\partial \Gamma(F_i)}{\partial q^j}\frac{\partial}{\partial \dot p_j} + \frac{\partial \Gamma(F_i)}{\partial \dot q^a}\frac{\partial}{\partial p_a}, \right.\\
&\qquad \left. C^\alpha := -\frac{\partial}{\partial p_\alpha} + \frac{\partial \psi^\alpha}{\partial q^j}\frac{\partial}{\partial \dot p_j} + \frac{\partial \psi^\alpha}{\partial \dot q^a}\frac{\partial}{\partial p_a} \right\}.
\end{aligned}$$

Then the time-dependent Helmholtz conditions for constrained systems, obtained by imposing $T(\mathrm{Im}(\gamma_{\Gamma,F}))\cap\mathcal{C}\subset\sharp(T(\mathrm{Im}(\gamma_{\Gamma,F}))^\circ)$ are

$$\frac{\partial F_a}{\partial\dot q^b}+\frac{\partial\psi^\alpha}{\partial\dot q^a}\frac{\partial F_\alpha}{\partial\dot q^b} = \frac{\partial F_b}{\partial\dot q^a}+\frac{\partial\psi^\alpha}{\partial\dot q^b}\frac{\partial F_\alpha}{\partial\dot q^a}\,, \tag{5.16}$$

$$\frac{\partial\Gamma(F_i)}{\partial q^k}+\frac{\partial F_\alpha}{\partial q^k}\frac{\partial\psi^\alpha}{\partial q^i} = \frac{\partial\Gamma(F_k)}{\partial q^i}+\frac{\partial F_\alpha}{\partial q^i}\frac{\partial\psi^\alpha}{\partial q^k}\,, \tag{5.17}$$

$$\frac{\partial F_a}{\partial q^i}+\frac{\partial\psi^\alpha}{\partial\dot q^a}\frac{\partial F_\alpha}{\partial q^i} = \frac{\partial\Gamma(F_i)}{\partial\dot q^a}+\frac{\partial F_\alpha}{\partial\dot q^a}\frac{\partial\psi^\alpha}{\partial q^i}\,. \tag{5.18}$$

Equations (5.16) and (5.18) are obtained by imposing that W_a be in $\sharp(T\mathrm{Im}(\gamma_{\Gamma,F})^\circ)$, while (5.17) and (5.18) are the conditions that arise when imposing that V_i be in $\sharp(T\mathrm{Im}(\gamma_{\Gamma,F})^\circ)$.

Theorem 5.5.3. *A SODE Γ on $\mathbb{R}\times M$ is variational if and only if there is a two-form Ω on $\mathbb{R}\times M$ such that*

(i) $d\Omega=0$,

(ii) $\Omega(v_1,v_2)=0$, *for all vertical vectors* $v_1,v_2\in V(\mathbb{R}\times M)$,

(iii) $i_\Gamma\Omega=0$,

(iv) $\flat_\Omega|_{V(\mathbb{R}\times M)}$ *is injective.*

Proof. We can prove this result using Theorem 1.4.2.

$\Rightarrow$ If Γ is variational in the sense given in Definition 5.5.2, then we define a two-form on $\mathbb{R}\times M$ by

$$\Omega=-dF^*\theta_Q+di_\Gamma F^*\theta_Q\wedge dt-\mathcal{L}_\Gamma F^*\theta_Q\wedge dt\,.$$

Condition (ii) is readily satisfied and condition (iii) can also be checked without making use of the conditions on F, since

$$\begin{aligned} i_\Gamma\Omega &= -i_\Gamma dF^*\theta_Q+\overbrace{i_\Gamma(di_\Gamma F^*\theta_Q\wedge dt)}^{-di_\Gamma F^*\theta_Q+(i_\Gamma di_\Gamma F^*\theta_Q)dt}-\overbrace{i_\Gamma(\mathcal{L}_\Gamma F^*\theta_Q\wedge dt)}^{-\mathcal{L}_\Gamma F^*\theta_Q+i_\Gamma\mathcal{L}_\Gamma F^*\theta_Q\,dt}\\ &= -i_\Gamma dF^*\theta_Q-di_\Gamma F^*\theta_Q+(i_\Gamma di_\Gamma F^*\theta_Q)dt+i_\Gamma dF^*\theta_Q\\ &\quad +di_\Gamma F^*\theta_Q-i_\Gamma(i_\Gamma dF^*\theta_Q+di_\Gamma F^*\theta_Q)dt\\ &= -(i_\Gamma i_\Gamma dF^*\theta_Q)dt=0\,. \end{aligned}$$

Condition (i) is equivalent to $d(\mathcal{L}_\Gamma F^*\theta_Q\wedge dt)=0$, and this is guaranteed by equations (5.16), (5.17) and (5.18).

Finally condition (iv) is a consequence of F being an immersion. This can be checked using local coordinates as in Theorem 5.3.2. Now

$$\begin{aligned}\Omega &= -\frac{\partial F_i}{\partial q^j}dq^j\wedge dq^i-\frac{\partial F_i}{\partial\dot q^a}d\dot q^a\wedge dq^i\\ &\quad+\left(\frac{\partial F_a}{\partial q^i}\dot q^a+\frac{\partial F_\alpha}{\partial q^i}\psi^\alpha-\frac{\partial F_i}{\partial q^a}\dot q^a-\frac{\partial F_i}{\partial q^\alpha}\psi^\alpha-\frac{\partial F_i}{\partial\dot q^a}\Gamma^a\right)dq^i\wedge dt+\left(\frac{\partial F_a}{\partial\dot q^b}\dot q^a+\frac{\partial F_\alpha}{\partial\dot q^b}\psi^\alpha\right)d\dot q^b\wedge dt\end{aligned}$$

and therefore $i_{v_1}\Omega - i_{v_2}\Omega = -\frac{\partial F_i}{\partial \dot q^a}(v_1^a - v_2^a)dq^i + \left(\frac{\partial F_a}{\partial \dot q^b}\dot q^a + \frac{\partial F_\alpha}{\partial \dot q^b}\psi^\alpha\right)(v_1^b - v_2^b)dt$ for any v_1, v_2 in $V(\mathbb{R}\times M)$. Since $\left(\frac{\partial F_i}{\partial \dot q^a}\right)$ is assumed to have maximal rank, $i_{v_1}\Omega = i_{v_2}\Omega$ implies $v_1 = v_2$.

$\Leftarrow$ We proceed as in the proofs of Theorems 1.5.2 and 5.3.2 to get a local one-form $\tilde\Theta$ on $\mathbb{R}\times M$ such that $d\tilde\Theta = \Omega$ and $\tilde\Theta(v) = 0$ for all vertical vector fields v. We define

$$\begin{array}{rccc} F: & \mathbb{R}\times M & \longrightarrow & \mathbb{R}\times T^*Q \\ & (t, v_q) & \longmapsto & (t, \tilde F(t, v_q)) \end{array}$$

by

$$\langle \tilde F(t, v_q), w_q\rangle = \langle pr_2 \circ \tilde\Theta(t, v_q), W_{v_q}\rangle\,,$$

where $v_q \in M, w_q \in TQ, W_{v_q} \in TM$ and $T\tau_Q|_M(W_{v_q}) = w_q$.

$$\begin{array}{ccc} T^*(\mathbb{R}\times M) & \xrightarrow{pr_2} & T^*M \\ \uparrow{\scriptstyle\tilde\Theta} & & \\ \mathbb{R}\times M & \xrightarrow{\tilde F} & T^*Q \end{array}$$

We check that $\mathrm{Im}(\gamma_{\Gamma,F})$ is isotropic using local coordinates. As $\tilde\Theta$ vanishes on vertical vectors, we can write

$$\tilde\Theta = F_i dq^i + \mu_t dt\,.$$

Then

$$\begin{aligned} \Omega = -d\tilde\Theta &= -dF_i\wedge dq^i - d\mu_t\wedge dt \\ &= -\frac{\partial F_i}{\partial q^j}dq^j\wedge dq^i - \frac{\partial F_i}{\partial \dot q^a}d\dot q^a\wedge dq^i - \frac{\partial F_i}{\partial t}dt\wedge dq^i - \frac{\partial \mu_t}{\partial q^j}dq^j\wedge dt - \frac{\partial \mu_t}{\partial \dot q^a}d\dot q^a\wedge dt\,. \end{aligned}$$

By imposing the condition $i_\Gamma\Omega = 0$ we get

$$\begin{aligned} \frac{\partial \mu_t}{\partial \dot q^a} &= -\frac{\partial F_b}{\partial \dot q^a}\dot q^b - \frac{\partial F_\alpha}{\partial \dot q^a}\psi^\alpha\,, \\ \frac{\partial \mu_t}{\partial q^i} &= \Gamma(F_i) - \frac{\partial F_a}{\partial q^i}\dot q^a - \frac{\partial F_\alpha}{\partial q^i}\psi^\alpha\,, \end{aligned}$$

so we can write

$$\Omega = -dF_i\wedge dq^i - \left[\left(\Gamma(F_j) - \frac{\partial F_a}{\partial q^j}\dot q^a - \frac{\partial F_\alpha}{\partial q^j}\psi^\alpha\right)dq^j + \left(-\frac{\partial F_b}{\partial \dot q^a}\dot q^b - \frac{\partial F_\alpha}{\partial \dot q^a}\psi^\alpha\right)d\dot q^a\right]\wedge dt\,,$$

and now the closedness of the second factor gives equations (5.16), (5.17) and (5.18) for F.

Finally we see that F is an immersion. Condition (iv) states that

$$0 = i_{v_1}\Omega - i_{v_2}\Omega = -\frac{\partial F_i}{\partial \dot q^a}(v_1^a - v_2^a)dq^i - \left(-\frac{\partial F_\alpha}{\partial \dot q^a}\psi^\alpha - \frac{\partial F_b}{\partial \dot q^a}\dot q^b\right)(v_1^a - v_2^a)dt$$

is satisfied if and only if $v_1 = v_2$. Since $\frac{\partial F_i}{\partial \dot q^a}(v_1^a - v_2^a) = 0$ implies $\left(-\frac{\partial F_\alpha}{\partial \dot q^a}\psi^\alpha - \frac{\partial F_b}{\partial \dot q^a}\dot q^b\right)(v_1^a - v_2^a) = 0$, we have that $\frac{\partial F_i}{\partial \dot q^a}(v_1^a - v_2^a) = 0$ implies $i_{v_1}\Omega - i_{v_2}\Omega = 0$ and $v_1 = v_2$, that is, $\left(\frac{\partial F_i}{\partial \dot q^a}\right)$ has maximal rank and F is an immersion.

■

5.6 The inverse problem for discrete constrained systems

Now we will consider the case of constrained second order discrete systems and we will extend some results from Chapter 4 and Section 5.3 to this setting. A constrained second order discrete system is given by

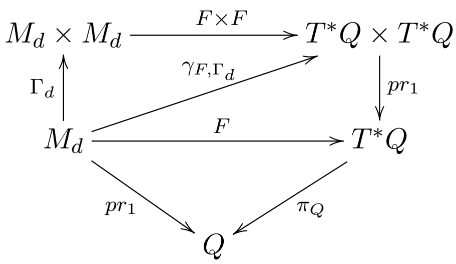
$$\begin{aligned} q^a_{k+1} &= \Gamma^a(q^j_{k-1}, q^b_k)\,, \\ q^\alpha_k &= \psi^\alpha_k(q^i_{k-1}, q^b_k)\,, \end{aligned}$$

where where $a, b = 1, \ldots, m < n$, $\alpha, \beta = m+1, \ldots, n$, and $i, j = 1, \ldots, n$.

More geometrically, let $M_d \subset Q \times Q$ be a submanifold defined by the discrete constraints $q^\alpha_k = \psi^\alpha_k(q^i_{k-1}, q^a_k)$ and let Γ_d be a second order difference equation on M_d regarded as a map $\Gamma_d : M_d \longrightarrow M_d \times M_d$ given by

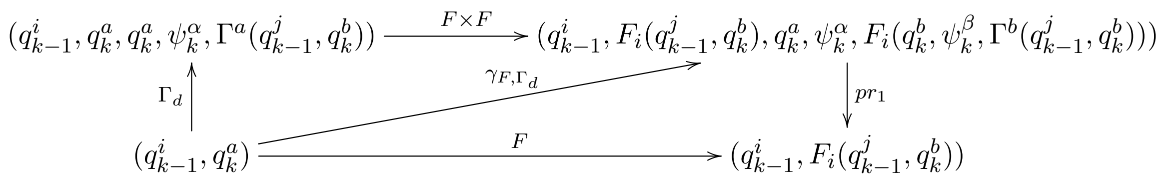
$$\Gamma_d(q^i_{k-1}, q^a_k) = (q^i_{k-1}, q^a_k, q^a_k, \psi^\alpha_k, q^a_{k+1} = \Gamma^a(q^j_{k-1}, q^b_k))\,.$$

We will also use the notation $q_{\tilde{k}} = (q^a_k)$, $q_{\bar{k}} = (q^\alpha_k)$. Given an immersion $F : M_d \longrightarrow T^*Q$ we define $\gamma_{F,\Gamma_d} := (F \times F) \circ \Gamma_d$, as shown in the following commutative diagram:



Definition 5.6.1. *A SOdE Γ_d on M_d is variational if there exists an immersion $F : M_d \longrightarrow T^*Q$ such that Im(γ_{F,Γ_d}) is an isotropic submanifold of $(T^*Q \times T^*Q, \Omega_Q)$.*

The above diagram in local coordinates becomes



Then the condition

$$d\left(F_a(q^b_k, \psi^\beta_k, \Gamma^b(q_{k-1}, q_{\tilde{k}}))dq^a_k + F_\alpha(q^b_k, \psi^\beta_k, \Gamma^b(q_{k-1}, q_{\tilde{k}}))d\psi^\alpha_k - F_i(q^j_{k-1}, q^b_k)dq^i_{k-1}\right) = 0$$

gives the discrete constraint Helmholtz conditions.

Some natural questions that immediately arise are the following:

1. Given a continuous variational SODE Γ on a submanifold $M \subset TQ$, find integrators Γ_d that are also variational in the sense of Definition 5.6.1.

2. From the existing integrators for nonholonomic systems [39, 63, 117], detect the ones that preserve the variational property.

One of the integrators mentioned in (ii) is the discrete Lagrange-d'Alembert (DLA) algorithm, derived from the so-called discrete Lagrange-d'Alembert principle [39]. Given a nonholonomic system, that is, a Lagrangian $L : TQ \to \mathbb{R}$ and a nonintegrable distribution $D \subset TQ$, it is necessary to choose a discrete Lagrangian L_d and a discrete constraint space $D_q \subset Q \times Q$, satisfying $\mathrm{diag}(Q\times Q) \subset D_d$ and $\dim D_d = \dim D$, and defined by the annihilation of functions $w_d^a : Q \times Q \to \mathbb{R}$, $a = 1\ldots, m$, regarded as discretizations of the constraint one-forms. As explained in [39], these discretizations should be chosen in a consistent way in order to get 'a desired order of accuracy'.

The DLA integrator is then given by

$$D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k) = \lambda_a w^a(q_k)\,, \tag{5.19}$$

$$w_d^a(q_k, q_{k+1}) = 0\,, \tag{5.20}$$

where λ_a are Lagrange multipliers, w^a are the constraint one-forms, L_d is a discrete Lagrangian and w_d^a is a discretization of the contraint one-forms.

Next we will study different choices of constraints and immersions $F : M_d \longrightarrow T^*Q$ for the example of the vertical rolling disk.

Example 5.6.2 (Vertical rolling disk)**.** The system represents a vertical disk rolling on a plane without sliding. It is defined on the configuration space $Q = S^1 \times S^1 \times \mathbb{R}^2$, with coordinates (θ, φ, x, y), where θ denotes the angle of self-rotation, φ the angle between the direction in which the disk moves and the x-axis and (x, y) are the coordinates of the contact point. The kinetic Lagrangian is given by $L = \frac{1}{2}(\dot\theta^2 + \dot\varphi^2 + \dot x^2 + \dot y^2)$, where all parameters are set to one, and the nonholonomic constraints of rolling without sliding are $\dot x = \cos(\varphi)\dot\theta$, $\dot y = \sin(\varphi)\dot\theta$, which define a submanifold $M \subset TQ$. Therefore the constraint one-forms are $w^1 = dx - \cos(\varphi)d\theta$ and $w^2 = dy - \sin(\varphi)d\theta$.

Recall that the immersion

$$\begin{array}{cccc} F_1: & M & \longrightarrow & T^*Q \\ & (\theta,\varphi,x,y,\dot\theta,\dot\varphi) & \longmapsto & (\theta,\varphi,x,y,2\dot\theta,\dot\varphi,0,0) \end{array}$$

provides an isotropic submanifold $\mathrm{Im}(TF_1 \circ \Gamma)$ of TT^*Q, and implies that Γ is variational in the sense of [11, Definition 5.1]. An alternative immersion is given by

$$\begin{array}{cccc} F_2: & M & \longrightarrow & T^*Q \\ & (\theta,\varphi,x,y,\dot\theta,\dot\varphi) & \longmapsto & \left(\theta,\varphi,x,y,\frac{\dot\theta}{\dot\varphi},\dot\varphi-\frac{\dot\theta^2}{2\dot\varphi^2}\left(1+\cos(\varphi)+\sin(\varphi)\right),\frac{\dot\theta}{\dot\varphi},\frac{\dot\theta}{\dot\varphi}\right)\,. \end{array}$$

which also provides an isotropic submanifold $\mathrm{Im}(TF_2 \circ \Gamma)$ of TT^*Q.

Now in order to derive a DLA integrator we can choose for instance the discretizations

$$\begin{aligned} L_d^{\frac{1}{2}}(q_k,q_{k+1}) &= \frac{1}{2}\left(\left(\frac{\theta_{k+1}-\theta_k}{h}\right)^2+\left(\frac{\varphi_{k+1}-\varphi_k}{h}\right)^2+\left(\frac{x_{k+1}-x_k}{h}\right)^2+\left(\frac{y_{k+1}-y_k}{h}\right)^2\right)\,, \\ w_d^1(q_k,q_{k+1}) &= \frac{x_{k+1}-x_k}{h}-\frac{\theta_{k+1}-\theta_k}{h}\cos\left(\frac{\varphi_k+\varphi_{k+1}}{2}\right)\,, \end{aligned}$$

$$w_d^2(q_k, q_{k+1}) = \frac{y_{k+1} - y_k}{h} - \frac{\theta_{k+1} - \theta_k}{h} \sin\left(\frac{\varphi_k + \varphi_{k+1}}{2}\right).$$

Equations (5.19) are then

$$\begin{aligned} -\frac{\theta_{k+1} - \theta_k}{h^2} + \frac{\theta_k - \theta_{k-1}}{h^2} &= -\lambda_1 \cos(\varphi_k) - \lambda_2 \sin(\varphi_k), \\ -\frac{\varphi_{k+1} - \varphi_k}{h^2} + \frac{\varphi_k - \varphi_{k-1}}{h^2} &= 0, \\ -\frac{x_{k+1} - x_k}{h^2} + \frac{x_k - x_{k-1}}{h^2} &= \lambda_1, \\ -\frac{y_{k+1} - y_k}{h^2} + \frac{y_k - y_{k-1}}{h^2} &= \lambda_2, \end{aligned} \tag{5.21}$$

from which we immediately obtain $\varphi_{k+1} = 2\varphi_k - \varphi_{k-1}$.

The discrete constraints chosen above yield the Lagrange multipliers

$$\begin{aligned} \lambda_1 &= -\frac{\theta_{k+1} - \theta_k}{h^2} \cos\left(\frac{\varphi_k + \varphi_{k+1}}{2}\right) + \frac{\theta_k - \theta_{k-1}}{h^2} \cos\left(\frac{\varphi_{k-1} + \varphi_k}{2}\right), \\ \lambda_2 &= -\frac{\theta_{k+1} - \theta_k}{h^2} \sin\left(\frac{\varphi_k + \varphi_{k+1}}{2}\right) + \frac{\theta_k - \theta_{k-1}}{h^2} \sin\left(\frac{\varphi_{k-1} + \varphi_k}{2}\right), \end{aligned}$$

and the substitution of them into (5.21) gives $\theta_{k+1} = 2\theta_k - \theta_{k-1}$ (as long as $\varphi_k - \varphi_{k-1} \neq 2(2n+1)\pi$, $n \in \mathbb{Z}$).

Therefore we have seen that Γ_d is given by

$$\begin{gathered} \Gamma_d(\theta_{k-1}, \varphi_{k-1}, x_{k-1}, y_{k-1}, \theta_k, \varphi_k) = \\ \left(\theta_{k-1}, \varphi_{k-1}, x_{k-1}, y_{k-1}, \theta_k, \varphi_k, \theta_k, \varphi_k, x_{k-1} + \cos\left(\frac{\varphi_{k-1} + \varphi_k}{2}\right)(\theta_k - \theta_{k-1}),\right. \\ \left. y_{k-1} + \sin\left(\frac{\varphi_{k-1} + \varphi_k}{2}\right)(\theta_k - \theta_{k-1}), 2\theta_k - \theta_{k-1}, 2\varphi_k - \varphi_{k-1}\right) \end{gathered}$$

If we define $F_{d1} : M_d \longrightarrow T^*Q$ in coordinates by

$$F_{d1}(\theta_{k-1}, \varphi_{k-1}, x_{k-1}, y_{k-1}, \theta_k, \varphi_k) = \left(\theta_{k-1}, \varphi_{k-1}, x_{k-1}, y_{k-1}, 2\frac{\theta_k - \theta_{k-1}}{h}, \frac{\varphi_k - \varphi_{k-1}}{h}, 0, 0\right),$$

which is a discretization of F_1 given above, then Im$((F_{d1} \times F_{d1}) \circ \Gamma_d)$=Im$(\gamma_{F_{d1},\Gamma_d})$ becomes

$$\begin{gathered} \left(\theta_{k-1}, \varphi_{k-1}, x_{k-1}, y_{k-1}, 2\frac{\theta_k - \theta_{k-1}}{h}, \frac{\varphi_k - \varphi_{k-1}}{h}, 0, 0,\right. \\ \theta_k, \varphi_k, x_{k-1} + \cos\left(\frac{\varphi_{k-1} + \varphi_k}{2}\right)(\theta_k - \theta_{k-1}), y_{k-1} + \sin\left(\frac{\varphi_{k-1} + \varphi_k}{2}\right)(\theta_k - \theta_{k-1}), \\ \left. 2\frac{\theta_k - \theta_{k-1}}{h}, \frac{\varphi_k - \varphi_{k-1}}{h}, 0, 0\right). \end{gathered}$$

Let $i : \text{Im}(\gamma_{F_{d1},\Gamma_d}) \hookrightarrow T^*Q \times T^*Q$ denote the inclusion. Then Im$(\gamma_{F_{d1},\Gamma_d})$ is an isotropic submanifold because

$$i^*\Omega_Q = 2d\left(\frac{\theta_k - \theta_{k-1}}{h}\right) \wedge d\theta_k + d\left(\frac{\varphi_k - \varphi_{k-1}}{h}\right) \wedge d\varphi_k$$

$$-2d\left(\frac{\theta_k-\theta_{k-1}}{h}\right)\wedge d\theta_{k-1}-d\left(\frac{\varphi_k-\varphi_{k-1}}{h}\right)\wedge d\varphi_{k-1}=0\,.$$

For the chosen discrete Lagrangian $L_d^{\frac{1}{2}}$ and F_{d1}, but with arbitrary constraints, the isotropy condition is equivalent to

$$2d\left(\frac{\theta_{k+1}-\theta_{k-1}}{h}\right)\wedge d\theta_k-2d\left(\frac{\theta_k-\theta_{k-1}}{h}\right)\wedge d\theta_{k-1}=0 \tag{5.22}$$

since the choice of constraints does not affect the evolution of φ, given by $\varphi_{k+1}=2\varphi_k-\varphi_{k-1}$. Thus we must necessarily have an evolution of the form $\theta_{k+1}=-\theta_{k-1}+f(\theta_k)$ in order to obtain an isotropic submanifold. For instance, if we choose the alternative constraints

$$\begin{aligned} w_d^1(q_k,q_{k+1}) &= \frac{x_{k+1}-x_k}{h}-\frac{\theta_{k+1}-\theta_k}{h}\cos\left(\varphi_k\right)\,, \\ w_d^2(q_k,q_{k+1}) &= \frac{y_{k+1}-y_k}{h}-\frac{\theta_{k+1}-\theta_k}{h}\sin\left(\varphi_k\right)\,, \end{aligned}$$

then we get the evolution $\theta_{k+1}=\theta_k+\left(\frac{\theta_k-\theta_{k-1}}{2}\right)(1+\cos(\varphi_k-\varphi_{k-1}))$ and therefore $\mathrm{Im}(F_{d1}\times F_{d1})\circ\Gamma_d$ is not an isotropic submanifold.

On the other hand, if we take the discrete constraints

$$\begin{aligned} w_d^1(q_k,q_{k+1}) &= \frac{x_{k+1}-x_k}{h}-\frac{\theta_{k+1}-\theta_k}{h}\left(\frac{1}{2}\cos\left((1-\alpha)\varphi_k+\alpha\varphi_{k+1}\right)+\frac{1}{2}\cos\left(\alpha\varphi_k+(1-\alpha)\varphi_{k+1}\right)\right)\,, \\ w_d^2(q_k,q_{k+1}) &= \frac{y_{k+1}-y_k}{h}-\frac{\theta_{k+1}-\theta_k}{h}\left(\frac{1}{2}\sin\left((1-\alpha)\varphi_k+\alpha\varphi_{k+1}\right)+\frac{1}{2}\sin\left(\alpha\varphi_k+(1-\alpha)\varphi_{k+1}\right)\right)\,, \end{aligned}$$

then we still get the dynamics $\theta_{k+1}=2\theta_k-\theta_{k-1}$ for any $\alpha\in[0,1]$, and therefore we obtain an isotropic submanifold $\mathrm{Im}(F_{d1}\times F_{d1})\circ\Gamma_d$.

Notice that if we take the map

$$\bar{F}_{d1}(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\theta_k,\varphi_k)=(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},2(\theta_k-\theta_{k-1}),\varphi_k-\varphi_{k-1},0,0)\,,$$

instead of F_{d1} then $\mathrm{Im}(\bar{F}_{d1}\times\bar{F}_{d1})\circ\Gamma_d$ is still an isotropic submanifold. This choice will appear in the next section.

Finally we consider the midpoint discretization of the constraints and the midpoint discretization of the alternative F_2 given above, that is

$$\begin{aligned} F_{d2}(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\theta_k,\varphi_k) = \Bigg(&\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}}, \\ &\frac{\varphi_k-\varphi_{k-1}}{h}-\frac{(\theta_k-\theta_{k-1})^2}{2(\varphi_k-\varphi_{k-1})^2}\left(1+\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)+\sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)\right), \\ &\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}},\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}}\Bigg)\,. \end{aligned}$$

Then $\mathrm{Im}(F_{d2}\times F_{d2})\circ\Gamma_d$ becomes

$$\left(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}},\frac{\varphi_k-\varphi_{k-1}}{h}\right.$$

$$-\frac{(\theta_k-\theta_{k-1})^2}{2(\varphi_k-\varphi_{k-1})^2}\left(1+\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)+\sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)\right),\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}},\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}},$$
$$\theta_k,\varphi_k,x_{k-1}+h\cos(\frac{\varphi_{k-1}+\varphi_k}{2})\frac{\theta_k-\theta_{k-1}}{h},y_{k-1}+h\sin(\frac{\varphi_{k-1}+\varphi_k}{2})\frac{\theta_k-\theta_{k-1}}{h},$$
$$\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}},\frac{\varphi_k-\varphi_{k-1}}{h}-\frac{(\theta_k-\theta_{k-1})^2}{2(\varphi_k-\varphi_{k-1})}\left(1+\cos\left(\frac{3\varphi_k-\varphi_{k-1}}{2}\right)+\sin\left(\frac{3\varphi_k-\varphi_{k-1}}{2}\right)\right),$$
$$\left.\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}},\frac{\theta_k-\theta_{k-1}}{\varphi_k-\varphi_{k-1}}\right),$$

which is not an isotropic submanifold.

5.6.1 Extension to a Lagrangian submanifold

As pointed out in Remark 4.2.3, Definition 5.6.1 can be equivalently given by substituting the statement "*Im*$(\gamma_{F,\Gamma})$ *is an isotropic submanifold of* $(T^*Q\times T^*Q,\Omega_Q)$" by "*Im*$(\Psi^{-1}\circ\gamma_{F,\Gamma})$ *is an isotropic submanifold of* $(T^*(Q\times Q),\omega_{Q\times Q})$".

Next we will show how to extend the isotropic submanifold $\mathrm{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma})$ in order to obtain a Lagrangian one. For that we will use Lemma 5.3.4.

Now if we take $P=Q\times Q$, $C=M$, and $\gamma=\Psi^{-1}\circ\gamma_{F,\Gamma}$, since $(\Psi^{-1}\circ\gamma_{F,\Gamma})(M)$ is isotropic in $(T^*(Q\times Q),\omega_{Q\times Q})$, then there is a one-form $\tilde\gamma$ defined in a neighborhood of M such that $\tilde\gamma|_M=\gamma$ and $d\tilde\gamma=0$. Then by the Poincaré Lemma there is a locally defined function $L_d:Q\times Q\to\mathbb{R}$ such that $\tilde\gamma=dL_d$.

Recall from Appendix A, that in order to obtain a Lagrangian submanifold we need to choose $\dim(P)-\dim(\mathrm{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma}))$ constraints that define a submanifold $N\subset T^*(Q\times Q)$ such that $\mathrm{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma})\subset N$. Next we compute the corresponding Hamiltonian vector fields (with respect to $\omega_{Q\times Q}$). If they are independent and not tangent to $\mathrm{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma})$, we can extend the original manifold along its flows and obtain a Lagrangian submanifold [156], which depends on the choice of constraints. This method provides a source of (possibly) alternative Lagrangians. Recall from Section 5.3 that corresponding to the immersion F_1 we can obtain the Lagrangian function

$$L_1=\frac{1}{2}\left(\dot\theta^2+\dot\varphi^2-\dot x^2-\dot y^2\right)+\dot\theta(\cos(\varphi)\dot x+\sin(\varphi)\dot y)\,,$$

while for F_2 we get

$$L_2=\frac{1}{2}\left(\dot\varphi^2-\dot\theta^2-\dot x^2-\dot y^2\right)+\frac{\dot\theta^2}{2\dot\varphi}(1-\cos(\varphi)-\sin(\varphi))+\dot\theta\dot x\left(\cos(\varphi)+\frac{1}{\dot\varphi}\right)+\dot\theta\dot y\left(\sin(\varphi)+\frac{1}{\dot\varphi}\right).$$

We will now see an example of this process in the discrete setting.

Example 5.6.3. Consider again the vertical rolling disk. With Γ_d and F_{d1} as in Example 5.6.2, we obtain the isotropic submanifold $\mathrm{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma})$ of $(T^*(Q\times Q),\omega_{Q\times Q})$ given by

$$\left(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\theta_k,\varphi_k,x_{k-1}+\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1}),\right.$$
$$\left.y_{k-1}+\sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1}),2\frac{\theta_{k-1}-\theta_k}{h},\frac{\varphi_{k-1}-\varphi_k}{h},0,0,2\frac{\theta_k-\theta_{k-1}}{h},\frac{\varphi_k-\varphi_{k-1}}{h},0,0\right),$$

where we denote coordinates on $(T^*(Q\times Q),\omega_{Q\times Q})$ by

$$(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\theta_k,\varphi_k,x_k,y_k,p_{\theta_{k-1}},p_{\varphi_{k-1}},p_{x_{k-1}},p_{y_{k-1}},p_{\theta_k},p_{\varphi_k},p_{x_k},p_{y_k}).$$

Now we can choose for instance the constraints

$$\begin{aligned}
\phi_1 &= x_k - x_{k-1} - \cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1}) + p_{x_k}\,,\\
\phi_2 &= y_k - y_{k-1} - \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1}) + p_{y_k}\,,
\end{aligned}$$

with corresponding Hamiltonian vector fields

$$\begin{aligned}
X_{\phi_1} &= \frac{\partial}{\partial p_{x_k}} - \frac{\partial}{\partial p_{x_{k-1}}} - \cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)\left(\frac{\partial}{\partial p_{\theta_k}} - \frac{\partial}{\partial p_{\theta_{k-1}}}\right)\\
&\quad + \frac{\theta_k-\theta_{k-1}}{2}\sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)\left(\frac{\partial}{\partial p_{\varphi_k}} + \frac{\partial}{\partial p_{\varphi_{k-1}}}\right) - \frac{\partial}{\partial x_k}\,,\\
X_{\phi_2} &= \frac{\partial}{\partial p_{y_k}} - \frac{\partial}{\partial p_{y_{k-1}}} - \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)\left(\frac{\partial}{\partial p_{\theta_k}} - \frac{\partial}{\partial p_{\theta_{k-1}}}\right)\\
&\quad - \frac{\theta_k-\theta_{k-1}}{2}\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)\left(\frac{\partial}{\partial p_{\varphi_k}} + \frac{\partial}{\partial p_{\varphi_{k-1}}}\right) - \frac{\partial}{\partial y_k}\,.
\end{aligned}$$

If we extend along the flows of X_{ϕ_1} and X_{ϕ_2} we obtain the Lagrangian submanifold

$$\begin{gathered}
(\theta_{k-1},\varphi_{k-1},x_{k-1},y_{k-1},\theta_k,\varphi_k,x_k,y_k,\\
\left(\frac{-2}{h}+1\right)(\theta_k-\theta_{k-1}) - \cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(x_k-x_{k-1}) - \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(y_k-y_{k-1}),\\
-\frac{\varphi_k-\varphi_{k-1}}{h} + \frac{\theta_k-\theta_{k-1}}{2}\left(\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(y_k-y_{k-1}) - \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(x_k-x_{k-1})\right),\\
x_k-x_{k-1}-\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1}),\, y_k-y_{k-1}-\sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1})\\
\left(\frac{2}{h}-1\right)(\theta_k-\theta_{k-1}) + \cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(x_k-x_{k-1}) + \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(y_k-y_{k-1}),\\
\frac{\varphi_k-\varphi_{k-1}}{h} + \frac{\theta_k-\theta_{k-1}}{2}\left(\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(y_k-y_{k-1}) - \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(x_k-x_{k-1})\right),\\
-\left(x_k-x_{k-1}-\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1})\right), -\left(y_k-y_{k-1}-\sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k-\theta_{k-1})\right)\Bigg)\,,
\end{gathered}$$

with corresponding discrete Lagrangian

$$\begin{aligned}
L_d &= -\frac{1}{2}(x_k-x_{k-1})^2 - \frac{1}{2}(y_k-y_{k-1})^2 + \left(\frac{1}{h}-\frac{1}{2}\right)(\theta_k-\theta_{k-1})^2 + \frac{1}{2h}(\varphi_k-\varphi_{k-1})^2\\
&\quad + (\theta_k-\theta_{k-1})\left(\cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(x_k-x_{k-1}) + \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(y_k-y_{k-1})\right)\,. \qquad (5.23)
\end{aligned}$$

The DEL equations corresponding to L_d are

$$x_0 - 2x_1 + x_2 + (\theta_1-\theta_0)\cos\left(\frac{\varphi_0+\varphi_1}{2}\right) - (\theta_2-\theta_1)\cos\left(\frac{\varphi_1+\varphi_2}{2}\right) = 0\,, \qquad (5.24)$$

$$y_0 - 2y_1 + y_2 + (\theta_1 - \theta_0)\sin\left(\frac{\varphi_0+\varphi_1}{2}\right) - (\theta_2 - \theta_1)\sin\left(\frac{\varphi_1+\varphi_2}{2}\right) = 0\,, \tag{5.25}$$

$$\begin{aligned}\left(\frac{-2+h}{h}\right)(\theta_0-\theta_1) - \left(\frac{-2+h}{h}\right)(\theta_1-\theta_2)&\\ +(x_1-x_0)\cos\left(\frac{\varphi_0+\varphi_1}{2}\right) + (x_1-x_2)\cos\left(\frac{\varphi_1+\varphi_2}{2}\right)&\\ +(y_1-y_0)\sin\left(\frac{\varphi_0+\varphi_1}{2}\right) + (y_1-y_2)\sin\left(\frac{\varphi_1+\varphi_2}{2}\right) &= 0\,,\end{aligned} \tag{5.26}$$

$$\begin{aligned}\frac{\varphi_1-\varphi_0}{h} + \frac{\varphi_1-\varphi_2}{h}&\\ +\frac{1}{2}(\theta_0-\theta_1)\left((y_0-y_1)\cos\left(\frac{\varphi_0+\varphi_1}{2}\right) + (x_1-x_0)\sin\left(\frac{\varphi_0+\varphi_1}{2}\right)\right)&\\ +\frac{1}{2}(\theta_1-\theta_2)\left((y_1-y_2)\cos\left(\frac{\varphi_1+\varphi_2}{2}\right) + (x_2-x_1)\sin\left(\frac{\varphi_1+\varphi_2}{2}\right)\right) &= 0\,.\end{aligned} \tag{5.27}$$

When restricted to the constraint submanifold given by

$$\begin{aligned} x_k &= x_{k-1} + \cos\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k - \theta_{k-1})\,,\\ y_k &= y_{k-1} + \sin\left(\frac{\varphi_{k-1}+\varphi_k}{2}\right)(\theta_k - \theta_{k-1})\,,\end{aligned}$$

Equations (5.24) and (5.25) identically vanish and Equations (5.26) and (5.27) become $\theta_2 = 2\theta_1 - \theta_0$ and $\varphi_2 = 2\varphi_1 - \varphi_0$ respectively. Hence we recover the SOdE in Example 5.6.2.

Remark 5.6.4. Consider the Lagrangian obtained in [11] by extension of an isotropic submanifold corresponding to $F(\theta, \varphi, x, y, \dot\theta, \dot\varphi) = (\theta, \varphi, x, y, 2\dot\theta, \dot\varphi, 0, 0)$, given by

$$L_1 = \frac{1}{2}\left(\dot\theta^2 + \dot\varphi^2 - \dot x^2 - \dot y^2\right) + \dot\theta(\cos(\varphi)\dot x + \sin(\varphi)\dot y)\,.$$

If we take the discretization

$$\begin{aligned} L_{d1} &= \frac{1}{2}\left(\left(\frac{\theta_k-\theta_{k-1}}{h}\right)^2 + \left(\frac{\varphi_k-\varphi_{k-1}}{h}\right)^2 - \left(\frac{x_k-x_{k-1}}{h}\right)^2 - \left(\frac{y_k-y_{k-1}}{h}\right)^2\right)\\ &\quad + \frac{\theta_k-\theta_{k-1}}{h}\left(\cos\left(\frac{\varphi_k+\varphi_{k-1}}{2}\right)\frac{x_k-x_{k-1}}{h} + \sin\left(\frac{\varphi_k+\varphi_{k-1}}{2}\right)\frac{y_k-y_{k-1}}{h}\right)\,,\end{aligned}$$

then the DEL equations are $\theta_{k+1} = 2\theta_k - \theta_{k-1}$ and $\varphi_{k+1} = 2\varphi_k - \varphi_{k-1}$ when restricted to the constraint submanifold.

If instead of F_{d1} we consider $\bar F_{d1}$ in Example 5.6.2, by choosing the same constraints ϕ_1 and ϕ_2 as in Example 5.6.3, we obtain the discrete Lagrangian

$$\begin{aligned} \bar L_d &= \frac{h^2}{2}\left(\left(\frac{\theta_k-\theta_{k-1}}{h}\right)^2 + \left(\frac{\varphi_k-\varphi_{k-1}}{h}\right)^2 - \left(\frac{x_k-x_{k-1}}{h}\right)^2 - \left(\frac{y_k-y_{k-1}}{h}\right)^2\right.\\ &\quad \left. + \frac{\theta_k-\theta_{k-1}}{h}\left(\cos\left(\frac{\varphi_k+\varphi_{k-1}}{2}\right)\frac{x_k-x_{k-1}}{h} + \sin\left(\frac{\varphi_k+\varphi_{k-1}}{2}\right)\frac{y_k-y_{k-1}}{h}\right)\right)\,.\end{aligned}$$

We have run simulations of the vertical rolling disk using the DLA integrator (5.19)-(5.20). We have used several alternative discretizations for defining the discrete constraints w_d^a:

- Midpoint rule: $w_d^a(q_k, q_{k+1}) = w^a\left(\frac{q_k+q_{k+1}}{2}\right)\left(\frac{q_{k+1}-q_k}{h}\right)$;
- Trapezoidal rule: $w_d^a(q_k, q_{k+1}) = \frac{1}{2}\left(w^a\left(q_k\right)\left(\frac{q_{k+1}-q_k}{h}\right) + w^a\left(q_{k+1}\right)\left(\frac{q_{k+1}-q_k}{h}\right)\right)$;
- α-trapezoidal rule:

$$\begin{aligned} w_d^a(q_k, q_{k+1}) &= \frac{1}{2}\left(w^a\left((1-\alpha)q_k + \alpha q_{k+1}\right)\left(\frac{q_{k+1}-q_k}{h}\right)\right. \\ &\quad \left. + w^a\left(\alpha q_k + (1-\alpha)q_{k+1}\right)\left(\frac{q_{k+1}-q_k}{h}\right)\right), \end{aligned}$$

  which reduces to the trapezoidal rule for $\alpha = 0$ and $\alpha = 1$, and to the midpoint rule for $\alpha = 1/2$;
- Euler A: $w_d^a(q_k, q_{k+1}) = w^a\left(q_k\right)\left(\frac{q_{k+1}-q_k}{h}\right)$;
- Euler B: $w_d^a(q_k, q_{k+1}) = w^a\left(q_{k+1}\right)\left(\frac{q_{k+1}-q_k}{h}\right)$.

For the following choices of a Lagrangian function L, we have computed numerically the values of the energy $K = (\partial L/\partial \dot{q})\dot{q} - L$ along the solutions:

- $L_1 = \frac{1}{2}\left(\dot{\theta}^2 + \dot{\varphi}^2 - \dot{x}^2 - \dot{y}^2\right) + \dot{\theta}(\cos(\varphi)\dot{x} + \sin(\varphi)\dot{y})$ (see Remark 5.6.4), which gives $K_1 = L_1$,
- $L_2 = \frac{1}{2}\left(-\dot{\theta}^2 + \dot{\varphi}^2 - \dot{x}^2 - \dot{y}^2\right) + \frac{\dot{\theta}^2}{2\dot{\varphi}}(1 - \cos(\varphi) - \sin(\varphi)) + \dot{\theta}\dot{x}\left(\cos(\varphi) + \frac{1}{\dot{\varphi}}\right) + \dot{\theta}\dot{y}\left(\sin(\varphi) + \frac{1}{\dot{\varphi}}\right)$ (see Example 5.6.2) which gives $K_3 = \frac{1}{2}\left(-\dot{\theta}^2 + \dot{\varphi}^2 - \dot{x}^2 - \dot{y}^2\right) + \dot{\theta}(\cos(\varphi)\dot{x} + \sin(\varphi)\dot{y})$,
- $L_3 = h^2\left(-\frac{1}{2}\dot{x}^2 - \frac{1}{2}\dot{y}^2 + \left(\frac{1}{h} - \frac{1}{2}\right)\dot{\theta}^2 + \frac{1}{2h}\dot{\varphi}^2 + \dot{\theta}(\cos(\varphi)\dot{x} + \sin(\varphi)\dot{y})\right)$, whose corresponding midpoint discretization is (5.23), which gives $K_2 = L_2$.

The energy functions K_1, K_2 and K_3 were discretized using the midpoint rule to obtain K_1^d, K_2^d, K_3^d. The results of the simulations for all methods, except for Euler A and B, preserved the energy functions, up to numerical truncation errors. For the α-trapezoidal discretization, all the values of α that we have used preserve the energy functions. This is expected because we already saw in Example 5.6.2 that for any α we obtain a variational SOdE.

Note that K_1 and K_2 only differ in the sign of the term $\frac{1}{2}\dot{\theta}^2$, whose discrete version is $\frac{1}{2h^2}(\theta_{k+1}-\theta_k)^2$. For all α, one of the discrete evolution equations is $\theta_{k+1} = 2\theta_k - \theta_{k-1}$, so $K_2^d - K_1^d$ is constant along solutions. Similarly, it is easy to show that the preservation of either K_2^d or K_3^d along solutions implies the preservation of the other one. Indeed,

$$K_2 - \frac{K_3}{h^2} = -\frac{\dot{\theta}^2}{h} + \frac{h-1}{2h}\dot{\varphi}^2,$$

and θ_k and φ_k evolve uniformly, that is, both $\theta_{k+1} - \theta_k$ and $\varphi_{k+1} - \varphi_k$ are constant. This implies that $K_2^d - K_3^d/h^2$ is constant.

The energy behavior of the Euler A and B discretizations is shown in Figure 5.2.

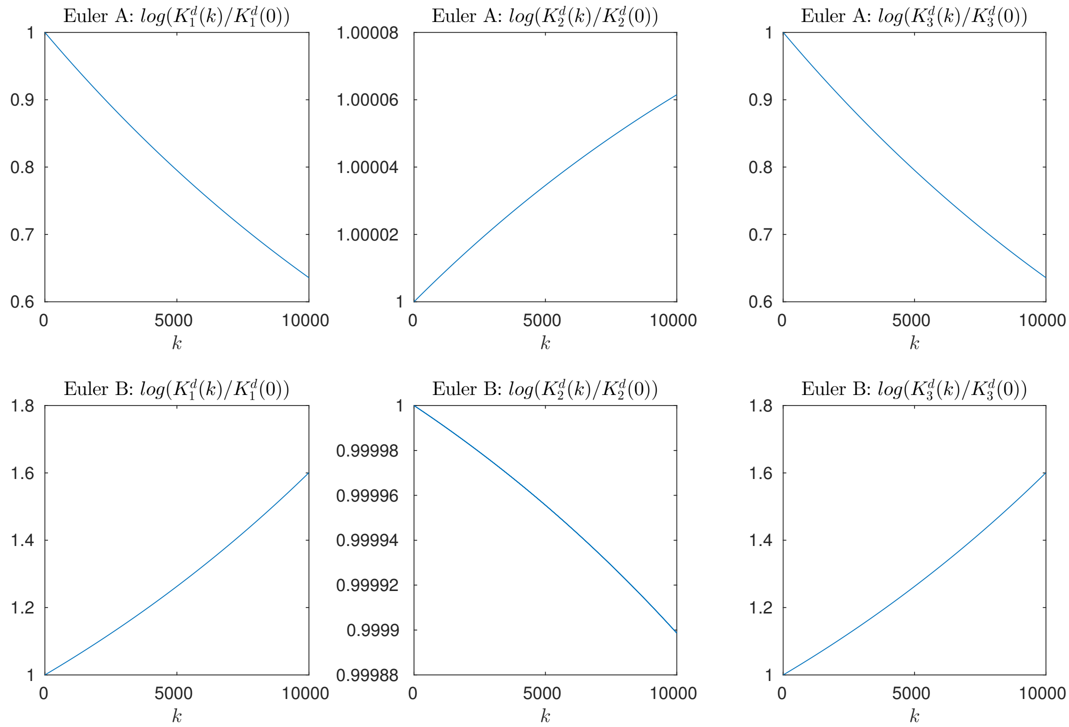


FIGURE 5.2: Energy behavior for Euler A and B, vertical rolling disk. $T = 500$, $h = 0.05$, $(x_0, y_0, \theta_0, \varphi_0) = (1, 1, 0.5, 0.3)$, $\theta_1 = 0.525$, $\varphi_1 = 0.31$; x_1 and y_1 satisfying the discrete constraints.

5.6.2 Crampin's Theorem with constraints

Now we will provide an extension of Theorem 4.2.7 to the discrete setting with constraints. We will need the following proposition from [76].

Proposition 5.6.5 ([76])**.** *Let $f : M \longrightarrow N$ be an immersion. For each Lagrangian submanifold $S \subset T^*M_d$ we can define a Lagrangian submanifold $\tilde{S} \subset T^*N$ by*

$$\tilde{S} = \{\mu \in T^*N : f^*\mu \in S\} .$$

Denote the flow of an explicit constrained second order difference equation $\Gamma_d : M_d \longrightarrow M_d \times M_d$ by $\Phi_{\Gamma_d} : M_d \longrightarrow M_d$ so that $\Phi_{\Gamma_d}(q_{k-1}, q_{\tilde{k}}) = (q_{\tilde{k}}, \psi_{\tilde{k}}, \Gamma_{\widetilde{k+1}})$.

Proposition 5.6.6. *An explicit constrained second order difference equation $\Gamma_d : M_d \longrightarrow M_d \times M_d$ is variational if and only if there is a nondegenerate two-form Ω_d on M_d such that*

(i) $\mathcal{L}^d_{\Gamma_d}\Omega_d = 0\,,$

(ii) $\Omega_d(v_1, v_2) = 0$ *for all* $v_1, v_2 \in Ker(Tpr_1)\,,$

(iii) $d\Omega_d = 0\,,$

(iv) $\flat_\Omega|_{Ker(Tpr_1)}$ *is injective,*

where $\mathcal{L}^d_{\Gamma_d}\Omega_d := (\Phi_{\Gamma_d})^*\Omega_d - \Omega_d$ *is regarded as a discrete analogue of the Lie derivative.*

Proof. The proof goes along the same lines as the analogue in [11].

If we assume that Γ_d is variational, we can define $\Omega_d = d(F^*\theta_Q)$ which clearly satisfies condition (iii). From the local expression

$$\Omega_d = \frac{\partial F_i}{\partial q^j_{k-1}} dq^j_{k-1} \wedge dq^i_{k-1} + \frac{\partial F_i}{\partial q^b_k} dq^b_k \wedge dq^i_{k-1}$$

condition (ii) is also clear since $Ker(Tpr_1) = \text{span}\left\{\frac{\partial}{\partial q^b_k}\right\}$. The requirement of F being an immersion implies that $\left(\frac{\partial F_i}{\partial q^b_k}\right)$ is of maximal rank. Thus, taking $v_1, v_2 \in Ker(Tpr_1)$, $v_1 = v^b_1 \frac{\partial}{\partial q^b_k}$, $v_2 = v^b_2 \frac{\partial}{\partial q^b_k}$ such that $i_{v_1}\Omega_d - i_{v_2}\Omega_d = (v^b_1 - v^b_2)\left(\frac{\partial F_i}{\partial q^b_k}\right) dq^i_{k-1} = 0$, we obtain $v_1 = v_2$ because of the rank condition. Therefore condition (iv) is satisfied.

Notice that

$$\mathcal{L}^d_{\Gamma_d}\Omega_d = \Phi^*_{\Gamma_d}\Omega_d - \Omega_d = d\Phi^*_{\Gamma_d}F^*\theta_Q - dF^*\theta_Q = d(\mathcal{L}^d_{\Gamma_d}F^*\theta_Q)\,.$$

In order to check condition (i) we locally compute $\mathcal{L}^d_{\Gamma_d}F^*\theta_Q = (F\circ\Phi_{\Gamma_d})^*\theta_Q - F^*\theta_Q$ to get

$$\mathcal{L}^d_{\Gamma_d}F^*\theta_Q = F_a(q_{\tilde{k}}, \psi_{\tilde{k}}, \Gamma_{\tilde{k}})dq^a_k + F_\alpha(q_{\tilde{k}}, \psi_{\tilde{k}}, \Gamma_{\tilde{k}})d\psi^a_k - F_i(q_{k-1}, q_{\tilde{k}})dq^i_{k-1}\,,$$

since $(F\circ\Phi_{\Gamma_d})(q_{k-1}, q_{\tilde{k}}) = (q_{\tilde{k}}, \psi_{\tilde{k}}, F_i(q_{\tilde{k}}, \psi_{\tilde{k}}, \Gamma_{\tilde{k}}))$. Note that the condition $d(\mathcal{L}^d_{\Gamma_d}F^*\theta_Q) = 0$ is exactly the same as requiring that $\text{Im}(\gamma_{F,\Gamma_d})$ be isotropic.

Conversely, let Ω_d be a two-form on M_d satisfying (i)-(iv). From (iii), locally $\Omega_d = d\Theta$ for a one-form Θ on M_d and from (ii) Θ has the local expression

$$\Theta = \alpha_i dq^i_{k-1} + \frac{\partial h}{\partial q^b_k}(q_{k-1}, q_{\tilde{k}})dq^b_k$$

for a locally defined map $h : M_d \longrightarrow \mathbb{R}$. Define $\bar{\Theta} = \Theta - dh$, which satisfies $\bar{\Theta}(V) = 0$ for all $V \in Ker(Tpr_1)$ and $d\bar{\Theta} = \Omega$. Then $F : M_d \longrightarrow T^*Q$ is given by

$$\langle F(q_{k-1}, q_{\tilde{k}}), v_{q_{k-1}}\rangle = \langle\bar{\Theta}(q_{k-1}, q_{\tilde{k}}), V_{v_{q_{k-1}}}\rangle \quad \text{for all } v_{q_{k-1}} \in TQ\,,$$

where $V_{v_{q_{k-1}}} \in TM_d$ is any vector satisfying $Tpr_1(V_{v_{q_{k-1}}}) = v_{q_{k-1}}$.

Since the one-form $\mathcal{L}^d_{\Gamma_d}\bar{\Theta} = \mathcal{L}^d_{\Gamma_d}F^*\theta_Q$ is closed, we obtain a Lagrangian submanifold $\text{Im}(\mathcal{L}^d_{\Gamma_d}F^*\theta_Q)$ of (T^*M_d, ω_{M_d}). Using Proposition A.2.3 (with $N = Q\times Q$) we obtain a Lagrangian submanifold of $(T^*(Q\times Q), \omega_{Q\times Q})$, described by

$$\text{Im}(\widetilde{\mathcal{L}^d_{\Gamma_d}F^*\theta_Q}) = \left\{\mu \in T^*(Q\times Q) : i^*_M\mu \in \text{Im}(\mathcal{L}^d_{\Gamma_d}F^*\theta_Q)\right\}\,,$$

where i_M denotes the inclusion. In coordinates $\text{Im}(\widetilde{\mathcal{L}^d_{\Gamma_d}F^*\theta_Q})$ is given by

$$\left(q^i_{k-1}, q_{\tilde{k}}, \psi_{\tilde{k}}, -F_i + F_\alpha\frac{\partial\psi^\alpha_k}{\partial q^i_{k-1}} - p_\alpha\frac{\partial\psi^\alpha_k}{\partial q^i_{k-1}}, F_a + F_\alpha\frac{\partial\psi^\alpha_k}{\partial q^a_k} - p_\alpha\frac{\partial\psi^\alpha_k}{\partial q^a_k}, p_\alpha\right)\,.$$

Since $\text{Im}(\Psi^{-1}\circ\gamma_{F,\Gamma_d}) \subset \text{Im}(\widetilde{\mathcal{L}^d_{\Gamma_d}F^*\theta_Q})$, $\text{Im}(\gamma_{F,\Gamma_d})$ is an isotropic submanifold of $(T^*Q\times T^*Q, \Omega_Q)$. Furthermore, condition (iv) implies that $\left(\frac{\partial F_i}{\partial q^b_k}\right)$ has maximal rank, that is, F is an immersion.

∎

Chapter 6

Energy-preserving integrators for nonholonomic systems

Geometric integrators are numerical methods for differential equations which preserve structural properties such as constants of the motion, symplectic or Poisson structures, phase-space volume, different symmetries of the system or isospectrality. Preservation of structural properties is often desirable in order to achieve correct qualitative behaviour and long time stability [77, 118, 139].

In this chapter we address the construction of geometric integrators for nonholonomic systems. In the unconstrained case, or when the constraints are holonomic, mechanical systems have many distinguishing geometric features. Among the most important are the preservation of energy, the symplectic form constructed from the Lagrangian (Poincaré-Cartan two-form) and the momentum map in the presence of symmetries according to the Noether Theorem. When we are dealing with nonholonomic constraints this symplectic form is no longer preserved, and the momentum map is not in general conserved in the presence of symmetries. However, the energy is still a conservation law for the system in the case of linear constraints. We therefore focus our attention on the exact preservation of energy, using geometric integrators, while writing the equations of motion in a format which ensures the nonholonomic constraints are satisfied.

The proposed approach is different from other recent approaches such as [40, 52, 58, 63, 98, 117], where the authors have introduced numerical integrators for nonholonomic systems with very good energy behavior, and properties such as the preservation of the discrete nonholonomic momentum map.

In this chapter we will consider a Lagrangian function $L : TQ \to \mathbb{R}$ of mechanical type, that is, kinetic minus potential energy, and a vector subbundle $\tau_{\mathcal{D}}$, which determines the nonholonomic constraints. To develop integrators we first introduce a Hamiltonian description of nonholonomic mechanics in terms of an almost-Poisson bracket. Using the Riemannian metric determined by the kinetic energy, and the standard symplectic structure on T^*Q, we can induce a linear almost-Poisson structure Π on the dual bundle $\pi_D : \mathcal{D}^* \to Q$. This so-called nonholonomic bracket is isomorphic to the nonholonomic bracket considered in [159].

Now, working on the "Hamiltonian system" determined by the triple given by (i) $\mathcal{D}^*$ as new phase space, (ii) the almost-Poisson bracket Π and (iii) the induced Hamiltonian function $\mathcal{H} : \mathcal{D}^* \to \mathbb{R}$, we apply energy-preserving integrators to simulate its dynamics. This is a coherent approach since the unique generic quantity preserved by the flow of the Hamiltonian vector field corresponding to

$\mathcal{H}$ is precisely the Hamiltonian function. The resulting integrators preserve by construction both the energy and nonholonomic constraints.

To approximate the solution while preserving the energy of the initial nonholonomic problem we use a class of geometric integrators called discrete gradient methods. Consider an ODE which can be written in skew-gradient form, i.e. $\dot{x} = \Pi(x)\nabla\mathcal{H}(x)$ with $x \in \mathbb{R}^N$ and $\Pi(x)$ a skew-symmetric matrix. In [119] it is shown that any ODE with a generic first integral $\mathcal{H}$ can be put into skew-gradient form. For a generalisation of these ideas to the case where the configuration space is a Lie group or a homogeneous manifold see [33].

Discrete gradient methods are based on the following construction. Let $x \approx x(nh)$ and $x' \approx x((n+1)h)$. Using a discrete gradient $\bar{\nabla}\mathcal{H}(x, x')$, which is an appropriate approximation of the gradient of $\mathcal{H}$ (see Section 6.2 for details), it is possible to define a class of integrators

$$\frac{x' - x}{h} = \tilde{\Pi}(x, x')\bar{\nabla}H(x, x'),$$

which preserve the first integral $\mathcal{H}$ exactly, i.e. $\mathcal{H}(x) = \mathcal{H}(x')$. Here $\tilde{\Pi}(x, x')$ is a skew-symmetric matrix approximating $\Pi(x)$. It can be shown that, in $\mathbb{R}^n$, any first integral-preserving (direct) integrator can be written as a discrete gradient method [68, 128, 132, 133]. In [69] discrete gradients were used to construct an energy preserving integrator for canonical mechanical systems with holonomic constraints.

For a given nonholonomic mechanical system, the equations of motion in canonical coordinates are generally assumed known. A potential obstacle in applying a discrete gradient method directly to the adapted coordinate system is the need for the user to analytically derive these equations. For this reason, we propose a reformulation of the methods using just information from the original system in canonical coordinates. With this approach the analytic reformulation of the system in adapted coordinates is avoided.

The outline of the chapter is as follows. In the next section we will recall the geometric framework for nonholonomic mechanics, to fix the notation. The main objective is to describe its dynamics as a Hamiltonian system on a vector bundle equipped with an almost-Poisson bracket. The resulting equations of motion in adapted coordinates are seen to be explicitly given in skew-gradient form. In Section 6.2 we apply discrete gradient integrators to the derived formulation to get energy-preserving integrators for nonholonomic systems. We then rewrite these integrators in an equivalent form by using only the information from the original nonholonomic system. Finally, in Section 6.3, we verify the properties and the performance of our integration techniques, applying them to several interesting examples: the chaotic quartic nonholonomic mechanical system, the Chaplygin sleigh system, the Suslov problem and the continuous gearbox driven by an asymmetric pendulum. Our methods are compared with other well-known numerical methods for nonholonomic mechanics.

6.1 Nonholonomic systems in adapted coordinates

Consider a nonholonomic system on a configuration manifold Q, of dimension n. Locally if (q^i) are coordinates on Q and $(q^i, \dot{q}^i)$ are the induced coordinates on TQ, the linear velocity constraints, specified by the regular C^∞-distribution $\mathcal{D} \subset TQ$, are written as

$$\mu_i^\alpha\,(q)\,\dot{q}^i = 0, \;\; m+1 \leq \alpha \leq n\,,$$

where $\operatorname{rank}(\mathcal{D}) = m \leq n$. The annihilator $\mathcal{D}^\circ$ is locally given by

$$\mathcal{D}^\circ = \operatorname{span}\left\{\mu^\alpha = \mu_i^\alpha(q)\,dq^i;\; m+1 \leq \alpha \leq n\right\}\,,$$

where the one-forms μ^α are independent. Equivalently, we can find independent vector fields $\{X_a\}$, $1 \leq a \leq m$ such that

$$\mathcal{D}_q = \operatorname{span}\{X_a\}\,.$$

Observe that $\mu^\alpha(X_a) = 0$, for all $m+1 \leq \alpha \leq n$ and $1 \leq a \leq m$.

6.1.1 Lagrangian equations for nonholonomic systems

Recall that, in addition to the constraints, the dynamics is specified by a Lagrangian function $L : TQ \to \mathbb{R}$, assumed to be of mechanical type, that is,

$$L\,(v_q) = \frac{1}{2}g\,(v_q, v_q) - V\,(q)\;, \quad v_q \in T_qQ\,,$$

where g is a Riemannian metric on the configuration space Q and $V : Q \to \mathbb{R}$ a potential function. The Lagrangian is written in coordinates $(q^i, \dot{q}^i)$ as

$$L(q^i, \dot{q}^i) = \frac{1}{2}\boldsymbol{g}_{ij}(q)\dot{q}^i\dot{q}^j - V(q)\,,$$

where $\boldsymbol{g}_{ij} = g\left(\partial/\partial q^i, \partial/\partial q^j\right)$, $1 \leq i, j \leq n$.

Recall that from the Lagrange-d'Alembert principle, Definition 5.1.2, we arrived at the well-known nonholonomic equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_\alpha \mu_i^\alpha\,, \tag{6.1a}$$

$$\mu_i^\alpha(q)\,\dot{q}^i = 0\,, \tag{6.1b}$$

where λ_α, $m+1 \leq \alpha \leq n$, is a set of Lagrange multipliers.

Adapted coordinates

Equations (6.1) are derived using a set of coordinates (q^i) on Q, and the induced coordinates $(q^i, \dot{q}^i)$ on TQ by the canonical coordinate frame $\left\{\frac{\partial}{\partial q^i}\right\}$, $1 \leq i \leq n$. Any element $v_q \in T_qQ$ can therefore be written without ambiguity as

$$v_q = \dot{q}^i \frac{\partial}{\partial q^i}\bigg|_q\,.$$

In the case of nonholonomic mechanics it can be useful to adapt the chosen frame to the linear velocity constraints. Specifically we consider a basis of vector fields $\{X_a, X_\alpha\}$, $1 \leq a \leq m$ and $m+1 \leq \alpha \leq n$, such that locally

$$\mathcal{D}_q = \operatorname{span}\{X_a(q)\} \quad \text{and} \quad \mathcal{D}_q^{\perp,g} = \operatorname{span}\{X_\alpha(q)\}\,,$$

where $\mathcal{D}_q^{\perp,g}$ is the Riemannian orthogonal to $\mathcal{D}$, that is,

$$g(X_a, X_\alpha) = 0 \quad \text{for all} \quad 1 \le a \le m \quad \text{and} \quad m+1 \le \alpha \le n\,.$$

Observe that $T_qQ = \mathcal{D}_q \oplus \mathcal{D}_q^{\perp,g}$.

The adapted basis $\{X_a, X_\alpha\}$ induces a new set of coordinates on the tangent bundle (q^i, y^a, y^α) (also called quasi-velocities) so that now

$$v_q = y^a X_a(q) + y^\alpha X_\alpha(q)\,.$$

Observe that the elements $v_q \in \mathcal{D}_q$ are distinguished by $y^\alpha = 0$. Therefore $y^\alpha = 0$ expresses the nonholonomic constraints in the adapted basis. Consequently $\mathcal{D}$ is completely described by coordinates (q^i, y^a).

Throughout this section and the next one, we will use the vertical rolling disk, introduced in Example 5.1.3, in order to illustrate how the nonholonomic equations can be written in almost-Poisson form and how to apply an energy-preserving integrator to these equations.

Example (Rolling disk)**.** We take an adapted basis $\{X_1, X_2, X_3, X_4\}$, where

$$\begin{aligned}
\mathcal{D} &= \text{span}\left\{ X_1 = r\cos\theta\frac{\partial}{\partial x_1} + r\sin\theta\frac{\partial}{\partial x_2} + \frac{\partial}{\partial\phi}, X_2 = \frac{\partial}{\partial\theta} \right\},\\
\mathcal{D}^{\perp,g} &= \text{span}\left\{ X_3 = \frac{1}{m}\frac{\partial}{\partial x_1} - \frac{r}{J_\phi}\cos\phi\frac{\partial}{\partial\phi}, X_4 = \frac{1}{m}\frac{\partial}{\partial x_2} - \frac{r}{J_\phi}\sin\phi\frac{\partial}{\partial\phi} \right\}.
\end{aligned}$$

This induces coordinates $(x_1, x_2, \theta, \phi, y^1, y^2, y^3, y^4)$ on TQ, which are related to the standard coordinates as follows:

$$\begin{aligned}
\dot{x}_1 &= ry^1\cos\theta + \frac{y^3}{m}\,,\\
\dot{x}_2 &= ry^1\sin\theta + \frac{y^4}{m}\,,\\
\dot{\theta} &= y^2\,,\\
\dot{\phi} &= y^1 - \frac{r}{J_\phi}y^3\cos\phi - \frac{r}{J_\phi}y^4\sin\phi\,.
\end{aligned}$$

Observe that the linear constraints have the simple form $y^3 = 0, y^4 = 0$ in the adapted basis.

Equations of motion in adapted coordinates

We now want to rewrite the equations of motion of the nonholonomic system in terms of the coordinates (q^i, y^a, y^α), instead of the canonical coordinates $(q^i, \dot{q}^i)$. Consider first Equation (6.1a). We can split it as the following system of equations:

$$0 = X_a^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - X_a^i\frac{\partial L}{\partial q^i} - \lambda_\beta\mu_i^\beta X_a^i = X_a^i\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - X_a^i\frac{\partial L}{\partial q^i}\,, \tag{6.2a}$$

$$0 = X_\alpha^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - X_\alpha^i\frac{\partial L}{\partial q^i} - \lambda_\beta\mu_i^\beta X_\alpha^i\,, \tag{6.2b}$$

with $1 \leq a \leq m$, $m+1 \leq \alpha, \beta \leq n$. In Equation (6.2a) we have used that $X_a(q) \in \mathcal{D}_q$. Observe that Equation (6.2b) uniquely gives information about the value of the Lagrange multipliers since g is a Riemannian metric and therefore $(\mu_i^\beta X_\alpha^i)$ is a regular matrix. As we are not interested in the evolution of the Lagrange multipliers we discard this second set of equations.

Define the Lagrangian in adapted coordinates as $\tilde{L}(q^i, y^a, y^\alpha) := L(q^i, X_a^i y^a + X_\alpha^i y^\alpha)$. We want to express Equation (6.2a) in terms of $\tilde{L}$. To this end observe that

$$\begin{aligned}
\frac{\partial \tilde{L}}{\partial y^a} &= X_a^i \frac{\partial L}{\partial \dot{q}^i}\,, \\
\frac{\partial \tilde{L}}{\partial y^\alpha} &= X_\alpha^i \frac{\partial L}{\partial \dot{q}^i}\,, \\
\frac{\partial \tilde{L}}{\partial q^j} &= \frac{\partial L}{\partial q^j} + \left(y^a \frac{\partial X_a^i}{\partial q^j} + y^\alpha \frac{\partial X_\alpha^i}{\partial q^j} \right) \frac{\partial L}{\partial \dot{q}^i}\,.
\end{aligned}$$

Now define the restricted Lagrangian $l : \mathcal{D} \to \mathbb{R}$ by $l := \tilde{L}\big|_{\mathcal{D}}$, that is, $l(q^i, y^a) := \tilde{L}(q^i, y^a, 0)$. It is interesting to note that

$$\tilde{L}(q^i, y^a, y^\alpha) = \frac{1}{2} g_{ab} y^a y^b + \frac{1}{2} g_{\alpha\beta} y^\alpha y^\beta - V(q)\,,$$

where $g_{ab} := g(X_a, X_b)$ and $g_{\alpha\beta} := g(X_\alpha, X_\beta)$, and thus

$$l(q^i, y^a) = \frac{1}{2} g_{ab} y^a y^b - V(q)\,.$$

We will make use of the fact that we can express the bracket $[X_a, X_b]$ in two ways using the different frames, concretely as

$$\begin{aligned}
[X_a, X_b] &= \left(\frac{\partial X_b^j}{\partial q^i} X_a^i - \frac{\partial X_a^j}{\partial q^i} X_b^i \right) \frac{\partial}{\partial q^j} = [X_a, X_b]^j \frac{\partial}{\partial q^j}\,, \\
[X_a, X_b] &= \mathcal{C}_{ab}^c X_c + \mathcal{C}_{ab}^\alpha X_\alpha\,.
\end{aligned}$$

Now, taking the restriction of Equation (6.2a) to $\mathcal{D}$, that is, using that $y^\alpha = 0$, we get

$$\begin{aligned}
0 &= X_a^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i} \right) - X_a^i \frac{\partial L}{\partial q^i} \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i} X_a^i \right) - \frac{dX_a^i}{dt} \frac{\partial L}{\partial \dot{q}^i} - X_a^i \frac{\partial L}{\partial q^i} \\
&= \frac{d}{dt}\left(\frac{\partial l}{\partial y^a} \right) + [X_a, X_b]^i y^b \frac{\partial L}{\partial \dot{q}^i} - X_a^i \frac{\partial l}{\partial q^i}\,.
\end{aligned}$$

The middle term of the last equation is

$$\begin{aligned}
[X_a, X_b]^i y^b \frac{\partial L}{\partial \dot{q}^i} &= (\mathcal{C}_{ab}^c X_c^i + \mathcal{C}_{ab}^\alpha X_\alpha^i) y^b \frac{\partial L}{\partial \dot{q}^i} \\
&= \mathcal{C}_{ab}^c y^b \frac{\partial l}{\partial y^c} + \mathcal{C}_{ab}^\alpha y^b \frac{\partial \tilde{L}}{\partial y^\alpha} \\
&= \mathcal{C}_{ab}^c y^b \frac{\partial l}{\partial y^c}\,,
\end{aligned}$$

since $\partial\tilde{L}/\partial y^\alpha = g_{\alpha\beta}y^\beta = 0$ because $y^\alpha = 0$. In conclusion we have

$$\begin{aligned} 0 &= X_a^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - X_a^i \frac{\partial L}{\partial q^i} \\ &= \frac{d}{dt}\left(\frac{\partial l}{\partial y^a}\right) + C_{ab}^c y^b \frac{\partial l}{\partial y^c} - X_a^i \frac{\partial l}{\partial q^i}\,. \end{aligned}$$

Therefore, the equations of motion of the nonholonomic system are rewritten in terms of the restricted Lagrangian l as

$$\frac{d}{dt}\left(\frac{\partial l}{\partial y^a}\right) + C_{ab}^c y^b \frac{\partial l}{\partial y^c} - X_a^i \frac{\partial l}{\partial q^i} = 0\,, \tag{6.3a}$$

$$\dot{q}^i = X_a^i(q) y^a\,, \tag{6.3b}$$

see for instance [72, 126].

Example (Rolling disk, continued)**.** We have the restricted Lagrangian

$$l(x_1, x_2, \theta, \phi, y^1, y^2) = \frac{1}{2}\left[(mr^2 + J_\phi)\,(y^1)^2 + J_\theta (y^2)^2\right]\,.$$

Now observe that

$$\begin{aligned} [X_1, X_2] &= r\sin\theta \frac{\partial}{\partial x_1} - r\cos\theta\frac{\partial}{\partial x_2} \\ &= mr\sin\theta\, X_3 - mr\cos\theta\, X_4\,. \end{aligned}$$

Therefore in this simple example we have $C_{ab}^c = 0$ for all $1 \le a, b, c \le 2$. The equations of motion (6.3) for this nonholonomic system are

$$\begin{aligned} \dot{x}_1 &= ry^1\cos\theta\,, \quad \dot{\theta} = y^2\,, \quad \dot{y}^1 = 0\,, \\ \dot{x}_2 &= ry^1\sin\theta\,, \quad \dot{\phi} = y^1\,, \quad \dot{y}^2 = 0\,, \end{aligned}$$

which are immediately explicitly integrated.

6.1.2 "Hamiltonian equations" for nonholonomic systems

On the cotangent bundle T^*Q the Lagrangian is replaced by the corresponding Hamiltonian H. We still assume a mechanical system, and let (q^i, p_i), $1 \le i \le n$, give local canonical coordinates on T^*Q through the Legendre transformation $\mathcal{F}L : TQ \to T^*Q$, i.e.

$$\mathcal{F}L : \,(q^i, \dot{q}^i) \longmapsto (q^i, p_i = \partial L/\partial \dot{q}^i)\,.$$

Then H is locally given by

$$H(q^i, p_i) = \frac{1}{2} p_i \boldsymbol{g}^{ij} p_j + V(q)\,,$$

where $(\boldsymbol{g}^{ij})$ is the inverse matrix of $(\boldsymbol{g}_{ij})$.

The Hamiltonian form of the nonholonomic equations (6.1) is then

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = J\begin{pmatrix} \frac{\partial H}{\partial q}(q,p) \\ \frac{\partial H}{\partial p}(q,p) \end{pmatrix} + \lambda_\alpha \begin{pmatrix} 0 \\ \mu^\alpha(q) \end{pmatrix}\,, \tag{6.4a}$$

$$\mu_i^\alpha(q)\frac{\partial H}{\partial p_i}(q,p) = \mu_i^\alpha \boldsymbol{g}^{ik} p_k = 0\,, \tag{6.4b}$$

where $m+1 \leq \alpha \leq n$ and $J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}$, see for instance [159].

Equations of motion in adapted coordinates

Now we will rewrite the restricted nonholonomic equations in a Hamilton-like way on D^* (see [9, 50, 72]). More precisely, consider the Legendre transformation $\mathcal{F}l : \mathcal{D} \to \mathcal{D}^*$, locally given by

$$\mathcal{F}l : \, (q^i, y^a) \longmapsto \left(q^i, \rho_a = \frac{\partial l}{\partial y^a}\right) .$$

From the Legendre transformation we can define the Hamiltonian function $\mathcal{H} : \mathcal{D}^* \to \mathbb{R}$, which in local coordinates becomes

$$\mathcal{H}(q^i, \rho_a) = \frac{1}{2} g^{ab} \rho_a \rho_b + V(q)\,.$$

Then upon changing coordinates in (6.3) using the Legendre transformation and $\mathcal{H}$, the equations of motion of a nonholonomic system are equivalently rewritten as

$$\dot{q}^i = X_b^i \frac{\partial \mathcal{H}}{\partial \rho_b}\,, \tag{6.5a}$$

$$\dot{\rho}_a = -\mathcal{C}_{ab}^c \rho_c \frac{\partial \mathcal{H}}{\partial \rho_b} - X_a^i \frac{\partial \mathcal{H}}{\partial q^i}\,. \tag{6.5b}$$

If we define the skew-symmetric matrix

$$\Pi(q,\rho) = \begin{pmatrix} 0 & X_b^i \\ -(X_a^j)^T & -\mathcal{C}_{ab}^c \rho_c \end{pmatrix} \tag{6.6}$$

then the Equations (6.5) will be given by

$$\dot{\zeta} = \Pi(\zeta)\nabla\mathcal{H}(\zeta)\,, \tag{6.7}$$

where $\zeta = (q^i, \rho_a)$ are coordinates on $\mathcal{D}^*$. This skew gradient format will allow the use of discrete gradient methods, as we will see in the next section.

Remark 6.1.1. It is possible to give a more intrinsic definition of these objects, as in [50]. Denote by $\{\cdot,\cdot\}$ the canonical bracket of the cotangent bundle T^*Q. Define a bracket of functions $\{\cdot,\cdot\}_{\mathcal{D}^*}$ on D^* by

$$\{f,g\}_{\mathcal{D}^*} = \{f \circ i_{\mathcal{D}}^*, g \circ i_{\mathcal{D}}^*\} \circ P^*\,,$$

for $f, g \in C^\infty(\mathcal{D}^*)$ where $i_{\mathcal{D}}^* : T^*Q \to \mathcal{D}^*$ and $P^* : \mathcal{D}^* \to T^*Q$ are the dual maps of the monomorphisms $i_{\mathcal{D}} : \mathcal{D} \to TQ$ and the projector $P : TQ \to \mathcal{D}$, respectively. Then the bivector field Π is given by

$$\Pi(df, dg) = \{f,g\}_{\mathcal{D}^*}\,.$$

This bracket does not in general satisfy the Jacobi identity, that is

$$\{f,\{g,h\}_{\mathcal{D}^*}\}_{\mathcal{D}^*} + \{g,\{h,f\}_{\mathcal{D}^*}\}_{\mathcal{D}^*} + \{h,\{f,g\}_{\mathcal{D}^*}\}_{\mathcal{D}^*} \neq 0\,.$$

By using this bracket, Equation (6.7) will be more appropriately written as

$$\dot{f} = \{f, \mathcal{H}\}_{\mathcal{D}^*} \text{ for all } f \in C^\infty(\mathcal{D}^*)\,.$$

Example (Rolling disk, continued)**.** We have the Hamiltonian function

$$\mathcal{H}(x_1, x_2, \theta, \phi, \rho_1, \rho_2) = \frac{1}{2}\left[\frac{\rho_1^2}{mr^2 + J_\phi} + \frac{\rho_2^2}{J_\theta}\right],$$

where $\rho_1 = (mr^2 + J_\phi)y^1$ and $\rho_2 = J_\theta y^2$. The skew-symmetric matrix (almost-Poisson structure) (6.6) is given by

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 0 & r\cos\theta & 0 \\ 0 & 0 & 0 & 0 & r\sin\theta & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -r\cos\theta & -r\sin\theta & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix},$$

and the equations of motion (6.7) are

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{\theta} \\ \dot{\phi} \\ \dot{\rho}_1 \\ \dot{\rho}_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & r\cos\theta & 0 \\ 0 & 0 & 0 & 0 & r\sin\theta & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -r\cos\theta & -r\sin\theta & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \dfrac{\rho_1}{mr^2 + J_\phi} \\ \dfrac{\rho_2}{J_\theta} \end{pmatrix}.$$

6.2 Energy-preserving integrators based on discrete gradients

In the previous section, we reduced the study of the nonholonomic dynamics to a system of differential equations

$$\dot{\zeta} = \Pi(\zeta)\nabla\mathcal{H}(\zeta)$$

on $\mathcal{D}^*$. In this section we will assume that Q is a real vector space, therefore $\mathcal{D}^* \cong \mathbb{R}^N$ where $n + m = N$. For a generalisation to the case of Lie groups and homogeneous manifolds see [33].

Since nonholonomic dynamics does not preserve the almost-Poisson structure Π in general, we will focus on the preservation of the energy using geometric integrators which preserve exactly this quantity. In particular, we will use discrete analogues of the gradient of the Hamiltonian function [119].

6.2.1 Discrete gradients

For ODEs in skew-gradient form, i.e. $\dot{x} = \Pi(x)\nabla H(x)$ with $x \in \mathbb{R}^N$ and $\Pi(x)$ a skew-symmetric matrix, it is immediate to check that H is a first integral. Indeed

$$\dot{H} = \nabla H(x)^T\dot{x} = \nabla H(x)^T\Pi(x)\nabla H(x) = 0,$$

due to the skew-symmetry of Π. Using discretizations of the gradient $\nabla H(x)$ it is possible to define a class of integrators which preserve the first integral H exactly.

Definition 6.2.1 ([68]). *Let $H : \mathbb{R}^N \longrightarrow \mathbb{R}$ be a differentiable function. Then $\bar{\nabla} H : \mathbb{R}^{2N} \longrightarrow \mathbb{R}^N$ is a discrete gradient of H if it is continuous and satisfies*

$$\bar{\nabla} H(x, x')^T (x' - x) = H(x') - H(x)\,, \qquad \text{for all } x, x' \in \mathbb{R}^N\,, \tag{6.8a}$$

$$\bar{\nabla} H(x, x) = \nabla H(x)\,, \qquad \text{for all } x \in \mathbb{R}^N\,. \tag{6.8b}$$

Some well-known examples of discrete gradients are:

- The mean value (or averaged vector field) discrete gradient introduced in [78] and given by

$$\bar{\nabla}_1 H(x, x') := \int_0^1 \nabla H((1-\xi)x + \xi x')d\xi\,, \qquad \text{for } x' \neq x\,. \tag{6.9}$$

- The midpoint (or Gonzalez) discrete gradient, introduced in [68] and given by

$$\bar{\nabla}_2 H(x, x') := \nabla H\left(\frac{1}{2}(x' + x)\right) + \frac{H(x') - H(x) - \nabla H\left(\frac{1}{2}(x'+x)\right)^T (x' - x)}{|x' - x|^2}(x' - x)\,, \tag{6.10}$$
for $x' \neq x$.

- The coordinate increment discrete gradient, introduced in [93], with each component given by

$$\bar{\nabla}_3 H(x, x')_i = \frac{H(x_1', \ldots, x_i', x_{i+1}, \ldots, x_n) - H(x_1', \ldots, x_{i-1}', x_i, \ldots, x_n)}{x_i' - x_i}\,, \quad 1 \leq i \leq N\,, \tag{6.11}$$

 when $x_i' \neq x_i$, and $\bar{\nabla}_3 H(x, x')_i = \frac{\partial H}{\partial x_i}(x_1', \ldots, x_{i-1}', x_i' = x_i, x_{i+1}, \ldots, x_n)$ otherwise.

It can be easily checked that these are indeed discrete gradients, see [68], [93] and [119].

6.2.2 Integrators based on discrete gradients

Once a discrete gradient $\bar{\nabla} H$ has been chosen, it is straightforward to define an energy-preserving integrator by

$$\frac{x' - x}{h} = \tilde{\Pi}(x, x', h)\bar{\nabla} H(x, x')\,, \tag{6.12}$$

where $\tilde{\Pi}$ is a differentiable skew-symmetric matrix approximating Π, that is, it satisfies $\tilde{\Pi}(x, x, 0) = \Pi(x)$. As in the continuous case, it is immediate to check that H is exactly preserved, since

$$H(x') - H(x) = \bar{\nabla} H(x, x')^T (x' - x) = h\bar{\nabla} H(x, x')^T \tilde{\Pi}(x, x', h)\bar{\nabla} H(x, x') = 0\,.$$

If we further wish to get a second order method then it is sufficient to choose $\tilde{\Pi}$ such that $\tilde{\Pi}(x, x', h) = \tilde{\Pi}(x', x, -h)$, and a differentiable discrete gradient such that $\bar{\nabla} H(x, x') = \bar{\nabla} H(x', x)$. This guarantees that the integration method (6.12) is time-symmetric and therefore second order accurate, see [132]. For instance it is enough to choose $\tilde{\Pi}(x, x', h) = \Pi\left(\frac{x+x'}{2}\right)$ and take the mean value discrete gradient or the midpoint discrete gradient. Higher order energy-preserving methods, which generalize the mean value discrete gradient (6.9), can be obtained by collocation methods as in [36].

Remark 6.2.2. If the Hamiltonian is quadratic then

$$\bar{\nabla}_1 H(x,x') = \bar{\nabla}_2 H(x,x') = \nabla H\left(\frac{1}{2}(x'+x)\right),$$

that is, the mean value discrete gradient (6.9) and the Gonzalez discrete gradient (6.10) coincide with the continuous gradient evaluated at the midpoint. Then if we choose $\tilde{\Pi}(x,x',h) = \Pi\left(\frac{x+x'}{2}\right)$ the method (6.12) reduces to the implicit midpoint rule. If the Hamiltonian is of the form $H(x) = \sum_{j=i}^{N} a_j x_j^2$, then $(\nabla H)_j = 2a_j x_j$ and

$$\bar{\nabla}_1 H(x,x')_i = \bar{\nabla}_2 H(x,x')_i = \bar{\nabla}_3 H(x,x')_i = a_i(x_i'+x_i)\,, \quad 1 \le i \le N\,,$$

that is, all three discrete gradients introduced above coincide with $\nabla H\left(\frac{1}{2}(x'+x)\right)$.

Remark 6.2.3. Preservation of the nonholonomic constraints. Going back to the case of nonholonomic systems, we can now apply an energy-preserving method (6.12) to Equation (6.7). Notice that if we take the approximation $\tilde{\Pi}(\zeta,\zeta',h)$ to be $\Pi(\bar{\zeta})$ for some $\bar{\zeta}(\zeta,\zeta') \in D^*$, and let $\partial\bar{\mathcal{H}}/\partial\rho_a$ be the discrete gradient component that approximates $\partial\mathcal{H}/\partial\rho_a$, then a discrete gradient method (6.12) gives

$$\frac{q'^j - q^j}{h} = X_a^j(\bar{q})\frac{\partial\bar{\mathcal{H}}}{\partial\rho_a}(\zeta,\zeta')\,,$$

where $\bar{q} = \pi_Q(\bar{\zeta})$. When applying $\mu^\alpha \in D^\circ$ we obtain

$$\mu_j^\alpha\left(\frac{q'^j - q^j}{h}\right) = \mu_j^\alpha X_a^j(\bar{q})\frac{\partial\bar{\mathcal{H}}}{\partial\rho_a}(\zeta,\zeta') = 0\,,$$

since $\mu_j^\alpha X_a^j = 0$ for all $1 \le a \le m$, $m+1 \le \alpha \le n$. All the nonholonomic constraints are thus preserved by the method.

Example (Rolling disk, continued)**.** Using any of the three discrete gradients introduced in Section 6.2.1 and a midpoint approximation of Π, we get the following energy preserving integrator, which is precisely the implicit midpoint rule:

$$x_1' = x_1 + hr\cos\left(\frac{\theta+\theta'}{2}\right)\frac{\rho_1}{mr^2+J_\phi}\,, \quad \theta' = \theta + h\frac{\rho_2}{J_\theta}\,, \quad \rho_1' = \rho_1\,,$$
$$x_2' = x_2 + hr\sin\left(\frac{\theta+\theta'}{2}\right)\frac{\rho_1}{mr^2+J_\phi}\,, \quad \phi' = \phi + h\frac{\rho_1}{mr^2+J_\phi}\,, \quad \rho_2' = \rho_2\,.$$

Observe that, as a consequence, we deduce the preservation of the constraints

$$\frac{x_1'-x_1}{h} - r\left(\frac{\phi'-\phi}{h}\right)\cos\left(\frac{\theta+\theta'}{2}\right) = 0\,, \quad \frac{x_2'-x_2}{h} - r\left(\frac{\phi'-\phi}{h}\right)\sin\left(\frac{\theta+\theta'}{2}\right) = 0\,,$$

which are discretizations of the nonholonomic constraints (5.5).

6.2.3 Integrators on T^*Q

The equations of motion in adapted coordinates for a given nonholonomic system are usually not known initially. A potential obstacle in applying a discrete gradient method to the equations in adapted coordinates (6.7) is therefore that the user must analytically obtain these equations. In this section we formulate the proposed schemes directly on the Hamiltonian equations of motion in canonical coordinates (6.4), and achieve preservation of energy and the nonholonomic constraints without explicitly deriving and using the formulation in adapted coordinates.

As a first attempt at an energy preserving method, we can define a numerical integrator for (6.4) directly on T^*Q by

$$\frac{z'-z}{h} = J\bar{\nabla}H(z,z') + \lambda_\alpha \begin{pmatrix} 0 \\ \mu^\alpha(\bar{q}) \end{pmatrix} , \quad \bar{\nabla}H(z,z')^T \begin{pmatrix} 0 \\ \mu^\alpha(\bar{q}) \end{pmatrix} = 0 , \tag{6.13}$$

where $\bar{\nabla}H$ is a discrete gradient, $z = (q,p)$ and $z' = (q',p')$. Notice that this method is energy-preserving, since

$$\begin{aligned} H(z') - H(z) &= \bar{\nabla}H(z,z')^T (z'-z) \\ &= h\bar{\nabla}H(z,z')^T J\bar{\nabla}H(z,z') + h\lambda_\alpha \bar{\nabla}H(z,z')^T \begin{pmatrix} 0 \\ \mu^\alpha(\bar{q}) \end{pmatrix} = 0 . \end{aligned}$$

However the constraints (6.4b) will in general only be approximately satisfied at the solution points by such a method.

To achieve exact preservation of both the energy and the nonholonomic constraints (6.4b), we use the restricted equations (6.7) on $\mathcal{D}^*$.

The method to step from (q,p) to (q',p') can be summarized as:

(i) Change coordinates from (q,p) to (q,ρ).

(ii) Step from (q,ρ) to (q',ρ') using a discrete gradient method (6.12) applied to (6.7).

(iii) Change coordinates from (q',ρ') to (q',p').

For step (i) and (iii) we make use of the following relations between the coordinates ρ_b on $\mathcal{D}^*$ and p_i on $\mathcal{F}L(\mathcal{D})$

$$p_i = \boldsymbol{g}_{ij} X_a^j g^{ab} \rho_b , \quad 1 \le i \le n , \tag{6.14a}$$

$$\rho_b = X_b^i p_i , \quad 1 \le b \le m . \tag{6.14b}$$

The challenge is performing step (ii) without any explicit knowledge of the equations in adapted coordinates. Specifically we need to evaluate $\nabla\mathcal{H}$ and Π in (6.7) for any $\zeta \in \mathcal{D}^*$. Let us therefore rewrite these expressions in a suitable format. Here p is considered as a dependent variable of (q,ρ) through (6.14a).

First observe that the skew symmetric matrix $-C_{ab}^c \rho_c$ in (6.6) may be written as

$$-C_{ab}^c \rho_c = -C_{ab}^c X_c^j p_j = -[X_a, X_b]^j p_j = \left(\frac{\partial X_a^j}{\partial q^i} X_b^i - \frac{\partial X_b^j}{\partial q^i} X_a^i \right) p_j . \tag{6.15}$$

Second we can write the partial derivatives of $\mathcal{H}$ as

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial q^i} &= \frac{1}{2} p_j \frac{\partial \boldsymbol{g}^{jk}}{\partial q^i} p_k + \frac{\partial p_j}{\partial q^i} \boldsymbol{g}^{jk} p_k + \frac{\partial V}{\partial q^i} = \frac{1}{2} p_j \frac{\partial \boldsymbol{g}^{jk}}{\partial q^i} p_k + \frac{\partial p_j}{\partial q^i} X_a^j g^{ab} \rho_b + \frac{\partial V}{\partial q^i} \\ &= \frac{1}{2} p_j \frac{\partial \boldsymbol{g}^{jk}}{\partial q^i} p_k - p_j \frac{\partial X_a^j}{\partial q^i} g^{ab} \rho_b + \frac{\partial V}{\partial q^i}\,, \qquad (6.16a)\\ \frac{\partial \mathcal{H}}{\partial \rho_a} &= g^{ab} \rho_b = g^{ab} X_b^i p_i\,. \qquad (6.16b)\end{aligned}$$

Expressing (6.7) using (6.15) and (6.16), the remaining issue is that we don't have explicit knowledge of a basis $X_a(q)$ for the distribution $\mathcal{D}$ or of the partial derivatives $\partial X_a(q)/\partial q^i$. For an arbitrary point q we generate $X_a(q)$ by computing the QR-factorization of the constraint matrix $(\mu_i^\alpha(q))$ using Householder reflections, see for instance [67]. The last m columns of the Q matrix can then be taken as $X_a(q)$, $1 \leq a \leq m$. Householder reflections were chosen because they are numerically stable and efficient.

We now make the assumption that the partial derivatives of $(\mu_i^\alpha(q))$ are either known or easily derived, which is usually the case. Then $\partial X_a(q)/\partial q^i$ can be calculated by augmenting the QR-factorization algorithm with corresponding steps for the partial derivatives. The specific procedure is given in Algorithm 1 in Appendix D for the matrix $A(q) := (\mu_i^\alpha(q))$.

To ensure we are sampling the same basis vector fields at different points in a given step when using Algorithm 1, it is sufficient to make sure the vector of sign choices s remains fixed for all factorizations in a given integration step. Because we only suppose knowledge of the full system (6.4), we transform back to canonical coordinates (q, p) after each step.

In theory this implementation can be combined with any discrete gradient method. However, since it is desirable to minimize the number of QR-factorizations per time step, this approach is best suited when used together with the Gonzalez discrete gradient and a midpoint approximation of Π. We shall refer to this specific method later as GONZALEZ-R.

Remark 6.2.4. Computational cost. For the initial direct method (6.13) it is necessary to evaluate the Lagrange multipliers. Moreover it is necessary to implement the constraint equations in each step of the algorithm. Applying a discrete gradient method (6.12) directly to the reduced system (6.7) simplifies the computational cost. This is so because the constraints are preserved automatically, and it is not necessary to compute the Lagrange multipliers as additional variables. Specifically, with the method (6.13) it is necessary to solve $3n - m$ variables while using (6.12) on (6.7) it is only necessary to compute $n + m$ variables.

Integrating the full system using the equations in adapted coordinates and the QR-factorization approach, we avoid the problem with Lagrange multipliers, but still see a rise in computational cost due to the necessity of moving between coordinate systems, and the general added cost in evaluating $\nabla \mathcal{H}$ and Π.

Thus the trade-off in not requiring knowledge of the reduced system is an increase in computational cost. It is therefore generally more efficient to analytically derive (6.7) and apply a discrete gradient method (6.12).

Remark 6.2.5. Implementation using finite differences. It is remarked in [117] that the condition (6.8b) for discrete gradients is only required to ensure consistency. Suppose this condition is relaxed slightly to

$$\bar{\nabla}H(x,x,h) = \nabla H(x) + \mathcal{O}(h^r)\,,$$

where r should at least match the order of the method. This is sufficient for the consistency of an integrator (6.12), and indeed for the method to have order r. We can use this to avoid having to evaluate $\partial X_a(q)/\partial q^i$ at the midpoint in GONZALEZ-R by replacing it with an appropriate finite difference approximation, for instance the central difference approximation

$$\frac{\bar{\partial X_a^j}}{\partial q^i}(q) := \frac{X_a^j(q+he_i) - X_a^j(q-he_i)}{2h} = \frac{\partial X_a^j}{\partial q^i}(q) + \mathcal{O}(h^2)\,,$$

where e_i is the canonical unit vector i. The resulting method retains second order and still preserves energy and the nonholonomic constraints.

6.3 Examples and numerical simulations

In this section we apply discrete gradient methods to some illustrative examples of nonholonomic systems. In the first three examples we will derive Equations (6.7) analytically. In the last one we will compare the strategies proposed in Section 6.2.3.

6.3.1 A fully chaotic nonholonomic system

In [117] the authors remark that the key geometric properties for nonholonomic dynamics are not known for general nonintegrable systems. To compare integration methods for such systems, they focus on energy preservation, looking at the following chaotic quartic mechanical system on $Q = \mathbb{R}^{2n+1}$ with coordinates $q = (q^1, q^2, \dots, q^{2n+1})^T := (x, w_1, \dots, w_n, z_1, \dots, z_n)^T$, which is defined by the Lagrangian

$$L(q,\dot{q}) = K(q,\dot{q}) - V(q), \quad \text{where } K = \frac{1}{2}\|\dot{q}\|_2^2, \qquad V = \frac{1}{2}\left(\|q\|_2^2 + z_1^2 z_2^2 + \sum_{i=1}^{n} w_i^2 z_i^2\right), \tag{6.17}$$

with the single nonholonomic linear velocity constraint

$$\dot{x} + \sum_{i=1}^{n} w_i \dot{z}_i = 0\,. \tag{6.18}$$

This system is reversible and preserves energy, i.e. $\dot{H} = 0$.

To derive the equations of motion in adapted coordinates, first note that this Lagrangian is of mechanical type. We can write the kinetic energy K as

$$K(q,\dot{q}) = \frac{1}{2} g_q(\dot{q},\dot{q})\,,$$

with the Euclidean Riemannian metric

$$g = \sum_{i=1}^{2n+1} dq^i \otimes dq^i\,,$$

which does not depend on q.

The distribution $\mathcal{D}$ and its orthogonal complement $\mathcal{D}^{\perp,g}$ are given by

$$\mathcal{D} = \operatorname{span}\left\{X_i := \frac{\partial}{\partial w_i},\, X_{n+i} := w_i\frac{\partial}{\partial x} - \frac{\partial}{\partial z_i},\, i=1,\ldots,n\right\},$$
$$\mathcal{D}^{\perp,g} = \operatorname{span}\left\{X_{2n+1} := \frac{\partial}{\partial x} + \sum_{i=1}^{n} w_i\frac{\partial}{\partial z_i}\right\}.$$

The adapted basis $\{X_1, X_2, \ldots, X_{2n+1}\}$, induces new coordinates (q^i, y^a, y^{2n+1}), $1 \le i \le 2n+1$, $1 \le a \le 2n$ on TQ for which the nonholonomic constraint reduces to $y^{2n+1} = 0$. The restricted Lagrangian $l : \mathcal{D} \to \mathbb{R}$ for this system is

$$l(q^i, y^a) = \frac{1}{2}\left(g_{ab}y^a y^b - V(q)\right),$$

where

$$(g_{ab}) = \begin{pmatrix} I_n & 0_n \\ 0_n & I_n + ww^T \end{pmatrix} \quad \text{and} \quad w = (w_1, w_2, \ldots, w_n)^T.$$

Moving to the Hamiltonian side we replace the velocities in adapted coordinates y^a with the momenta

$$\rho_a = \frac{\partial l}{\partial y^a} = g_{ab}y^b, \quad a = 1, \ldots, 2n.$$

The restricted Hamiltonian $\mathcal{H} : \mathcal{D}^* \to \mathbb{R}$ is then

$$\mathcal{H}(q^i, \rho_a) = \frac{1}{2}g^{ab}\rho_a\rho_b + V(q),$$

where (g^{ab}) is the inverse of (g_{ab}), i.e.

$$(g^{ab}) = \begin{pmatrix} I_n & 0_n \\ 0_n & I_n - \frac{ww^T}{1+w^T w} \end{pmatrix}.$$

The equations of motion on the Hamiltonian side are then given by (6.7), where

$$\Pi(q,\rho) = \begin{pmatrix} & & & 0_{1\times n} & w^T \\ & 0_{2n+1} & & I_n & 0_n \\ & & & 0_n & -I_n \\ 0_{n\times 1} & -I_n & 0_n & 0_n & -\kappa I_n \\ -w & 0_n & I_n & \kappa I_n & 0_n \end{pmatrix}, \quad \nabla\mathcal{H}(q,\rho) = \begin{pmatrix} x \\ w_1 + w_1 z_1^2 - \kappa\rho_{n+1} + \kappa^2 w_1 \\ \vdots \\ w_n + w_n z_n^2 - \kappa\rho_{n+n} + \kappa^2 w_n \\ z_1 + y_1^2 z_1 + z_1 z_2^2 \\ z_2 + y_2^2 z_2 + z_1^2 z_2 \\ z_3 + y_3^2 z_3 \\ \vdots \\ z_n + y_n^2 z_n \\ g^{ab}\rho_b \end{pmatrix}, \tag{6.19}$$

$\kappa := \frac{w^T\eta}{1+w^T w}$ and $\eta := (\rho_{n+1}, \ldots, \rho_{2n})^T$.

Numerical simulations

We follow the approach in [117], and integrate this system, with $n = 3$, from a random initial state with energy $H = 3.06$. We here compare five different methods. The first two are variational integrators based on the discrete Lagrange d'Alembert (DLA) for the full system (6.17) and (6.18). The semi implicit reversible DLA variational integrator proposed in [117] (SI-DLA), and the implicit reversible DLA variational integrator based on a midpoint discrete Lagrangian (I-DLA) which is also described in [117] among others. The third method is the 2-stage Lobatto IIIA-B-C-C*-D SPARK method described in [94] for index 2 DAEs (SPARK), which again discretize the equations of motion of the full system. For the last two methods, we integrate the reduced system, (6.7) with (6.19), using a discrete gradient method (6.12), with two different discrete gradients: The averaged vector field discrete gradient (AVF) (6.9), and the Gonzalez discrete gradient (GONZALEZ) (6.10).

As seen in Figure 6.1, while all five methods are known to be second order accurate and respect the constraint, only the discrete gradient methods conserve the energy up to round off error. In [117] it is shown that the energy error for SI-DLA closely follows a random walk with the variance $\sigma^2 = 10^{-4}h^4t$. In Figure 6.1 we also show that I-DLA and SPARK behaves similarly, with all comparison methods exhibiting similar linear time growth. As expected, since the discrete gradient methods have no energy error, they also have zero variance up to round off error.

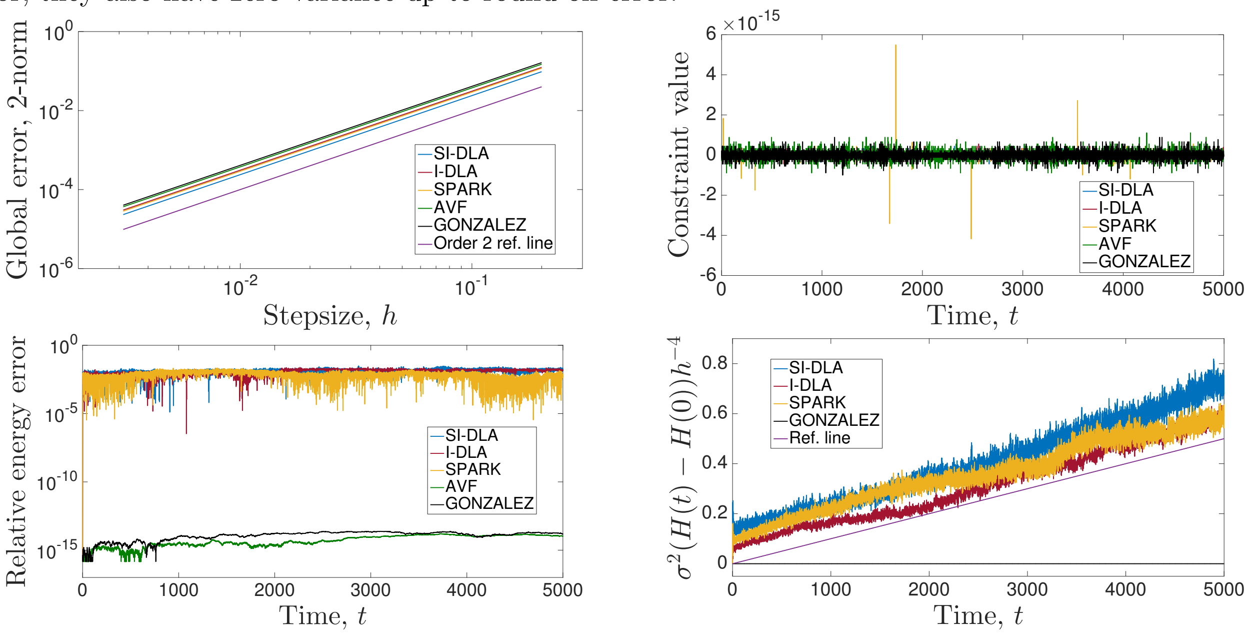


FIGURE 6.1: Comparison of the different methods for the fully chaotic system (6.17)-(6.18). **Top left**: Order plot, integrating up to $t = 10$. All methods are seen to be second order. **Top right**: Value of the left hand side of the constraint expression (6.18) for a sample trajectory with random initial conditions and step size $h = 0.2$. The methods all respect the constraint up to machine precision. **Bottom left**: Relative energy error, i.e. $|H(t) - H(0)/H(0)|$, for the same trajectory. Only the discrete gradient methods conserve the energy up to machine precision. **Bottom right**: The variance of the energy errors $\sigma^2(H(t) - H(0))$ for 200 different initial conditions scaled by their expected h^4 dependence on the time step. Again $h = 0.2$. The reference line $10^{-4}t$ is included for comparison. All comparison methods exhibit similar linear time growth in accordance with the reference line, while the discrete gradient method GONZALEZ has zero variance up to machine precision as expected. AVF is not shown since it was indistinguishable from GONZALEZ.

6.3.2 The Chaplygin sleigh

In this example we will see that the transformation of the systems into adapted coordinates can give rise to some additional numerical advantages apart from the possibility of achieving energy preservation.

The Chaplygin sleigh is a rigid body moving on a horizontal plane with three contact points, two of which slide freely without friction. The third one is a knife edge, which imposes the nonholonomic constraint of no motion perpendicular to the direction of the blade. The configuration space is $Q = SE(2)$, with coordinates (x_1, x_2, θ). The coordinates (x_1, x_2) denote the contact point of the blade with the plane and θ the orientation of the blade. The Lagrangian is of kinetic type and if we assume that the center of mass lies in the line through the blade then it is given by

$$L = \frac{1}{2}\left((J+ma^2)\dot{\theta}^2 + m\left(\dot{x}_1^2 + \dot{x}_2^2 + 2a\dot{\theta}(-\dot{x}_1\sin(\theta) + \dot{x}_2\cos(\theta))\right)\right),$$

where m denotes the mass of the body, J the moment of inertia relative to the center of mass and a the distance between the center of mass and the contact point of the blade. The matrix of the metric defining the kinetic Lagrangian is then given by

$$\begin{pmatrix} m & 0 & -ma\sin(\theta) \\ 0 & m & ma\cos(\theta) \\ -ma\sin(\theta) & ma\cos(\theta) & J+ma^2 \end{pmatrix}.$$

The nonholonomic constraint is $-\dot{x}_1\sin(\theta) + \dot{x}_2\cos(\theta) = 0$, which defines the non-integrable distribution

$$\mathcal{D} = \operatorname{span}\left\{\frac{\partial}{\partial\theta}, \cos(\theta)\frac{\partial}{\partial x_1} + \sin(\theta)\frac{\partial}{\partial x_2}\right\}.$$

For more details on this system, see [126].

In [59] there is a qualitative study of the DLA method when applied to the Chaplygin sleigh. More precisely, it is shown that the discrete momentum dynamics reproduces the same qualitative behaviour as the continuous momentum dynamics, as long as $\mid \theta' - \theta \mid < 2\pi$ and the momentum variable ρ_2 satisfies some bound. In the present example we examine the same issue using a discrete gradient method to the equations in adapted coordinates. We will obtain a bound on h but no bound on the momentum variables.

To derive the equations in adapted coordinates, we choose the following orthonormal basis adapted to D and $D^{\perp,g}$:

$$\mathcal{D} = \operatorname{span}\left\{X_1 = \frac{1}{\sqrt{J+ma^2}}\frac{\partial}{\partial\theta}\,,\, X_2 = \frac{1}{\sqrt{m}}\left(\cos(\theta)\frac{\partial}{\partial x_1} + \sin(\theta)\frac{\partial}{\partial x_2}\right)\right\},$$

$$\mathcal{D}^{\perp,g} = \operatorname{span}\left\{X_3 = \frac{1}{\sqrt{\frac{(J+ma^2)^2}{ma^2} - (J+ma^2)}}\left(\frac{(J+ma^2)}{ma}\sin(\theta)\frac{\partial}{\partial x_1} - \frac{(J+ma^2)}{ma}\cos(\theta)\frac{\partial}{\partial x_2} + \frac{\partial}{\partial\theta}\right)\right\},$$

and denote the induced coordinates on TQ by $(x_1, x_2, \theta, y^1, y^2, y^3)$. In these coordinates the restricted Lagrangian $l : \mathcal{D} \longrightarrow \mathbb{R}$ is given by $l(q^i, y^a) = \frac{1}{2}((y^1)^2 + (y^2)^2)$, and the nonholonomic constraint by $y^3 = 0$.

Since we have chosen an orthonormal basis we have the restricted Hamiltonian $\mathcal{H}(q^i,\rho_1,\rho_2) = \frac{1}{2}\left(\rho_1^2+\rho_2^2\right)$, where $\rho_1 = \frac{\partial l}{\partial y^1} = y^1, \rho_2 = \frac{\partial l}{\partial y^2} = y^2$. Then

$$\Pi(\theta,\rho_1) = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{\cos(\theta)}{\sqrt{m}} \\ 0 & 0 & 0 & 0 & \frac{\sin(\theta)}{\sqrt{m}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{J+ma^2}} & 0 \\ 0 & 0 & -\frac{1}{\sqrt{J+ma^2}} & 0 & -\frac{a\sqrt{m}}{J+ma^2}\rho_1 \\ -\frac{\cos(\theta)}{\sqrt{m}} & -\frac{\sin(\theta)}{\sqrt{m}} & 0 & \frac{a\sqrt{m}}{J+ma^2}\rho_1 & 0 \end{pmatrix}, \qquad \nabla\mathcal{H}(\zeta) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \rho_1 \\ \rho_2 \end{pmatrix},$$

and the equations of motion (6.7) are for the position

$$\dot{x}_1 = \frac{\cos(\theta)}{\sqrt{m}}\rho_2, \quad \dot{x}_2 = \frac{\sin(\theta)}{\sqrt{m}}\rho_2, \quad \dot{\theta} = \frac{1}{\sqrt{J+ma^2}}\rho_1\,,$$

and for the momenta

$$\dot{\rho_1} = -\frac{a\sqrt{m}}{J+ma^2}\rho_1\rho_2, \quad \dot{\rho_2} = \frac{a\sqrt{m}}{J+ma^2}\rho_1^2\,. \tag{6.20}$$

The obtained equations are rather simple, since we have a quadratic vector field, a quadratic Hamiltonian and no constraints.

The mean value discrete gradient (6.9), the midpoint discrete gradient (6.10) and the coordinate increment discrete gradient (6.11) all coincide and give $\bar{\nabla}\mathcal{H}(\zeta,\zeta') = \left(0,0,0,\frac{\rho_1+\rho_1'}{2},\frac{\rho_2+\rho_2'}{2}\right)^T$. As an approximation to the matrix Π we have chosen the midpoint value $\tilde{\Pi}(\zeta,\zeta') = \Pi\left(\frac{\zeta+\zeta'}{2}\right)$. Recalling Remark 6.2.2, the energy-preserving integrator (6.12) with any of these discrete gradients then collapses to the implicit midpoint rule, and is consequently given by

$$x_1' = x_1 + \frac{h}{2\sqrt{m}}\cos\left(\frac{\theta+\theta'}{2}\right)(\rho_2+\rho_2')\,, \tag{6.21a}$$

$$x_2 = x_2 + \frac{h}{2\sqrt{m}}\sin\left(\frac{\theta+\theta'}{2}\right)(\rho_2+\rho_2')\,, \tag{6.21b}$$

$$\theta' = \theta + \frac{h}{2}\frac{1}{\sqrt{J+ma^2}}(\rho_1+\rho_1')\,, \tag{6.21c}$$

$$\rho_1' = \rho_1 - \frac{h}{4}\frac{a\sqrt{m}}{J+ma^2}(\rho_1+\rho_1')(\rho_2+\rho_2')\,, \tag{6.21d}$$

$$\rho_2' = \rho_2 + \frac{h}{4}\frac{a\sqrt{m}}{J+ma^2}(\rho_1+\rho_1')^2\,. \tag{6.21e}$$

We will write the equations (6.21d) and (6.21e) as

$$F(\rho_1,\rho_2,\rho_1',\rho_2',h) := \rho_1' - \rho_1 + \frac{h}{4}C_{12}^1(\rho_1+\rho_1')(\rho_2+\rho_2') = 0\,,$$
$$G(\rho_1,\rho_2,\rho_1',\rho_2',h) := \rho_2' - \rho_2 + \frac{h}{4}C_{21}^1(\rho_1+\rho_1')^2 = 0\,.$$

Notice that $F(\rho_1,\rho_2,\rho_1,\rho_2,0) = G(\rho_1,\rho_2,\rho_1,\rho_2,0) = 0$ and compute

$$\left.\begin{pmatrix} \frac{\partial F}{\partial \rho_1'} & \frac{\partial F}{\partial \rho_2'} \\ \frac{\partial G}{\partial \rho_1'} & \frac{\partial G}{\partial \rho_2'} \end{pmatrix}\right|_{(\rho_1,\rho_2,\rho_1,\rho_2,0)} = \left.\begin{pmatrix} 1+\frac{h}{4}C_{12}^1(\rho_2+\rho_2') & \frac{h}{4}C_{12}^1(\rho_1+\rho_1') \\ \frac{h}{2}C_{21}^1(\rho_1+\rho_1') & 1 \end{pmatrix}\right|_{(\rho_1,\rho_2,\rho_1,\rho_2,0)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

By the implicit function theorem we can write $\rho_1' = f(\rho_1,\rho_2,h)$ and $\rho_2' = g(\rho_1,\rho_2,h)$ in a neighbourhood of $(\rho_1,\rho_2,0)$, with (ρ_1',ρ_2') also in a neighbourhood of (ρ_1,ρ_2).

The continuous system has certain qualitative characteristics. Specifically, for the continuous system we have from (6.20), in the case $a \neq 0$, a one-dimensional manifold of equilibria $\{\rho_1 = 0\}$. These equilibria are stable and asymptotically stable with respect to ρ_1 if $\rho_2 > 0$ and unstable if $\rho_2 < 0$. We will now study how the qualitative behaviour of (6.21d)-(6.21e) compares, as in [59].

Equilibria: If $h \neq 0$ then $F(\rho_1, \rho_2, \rho_1, \rho_2, h) = 0$ and $G(\rho_1, \rho_2, \rho_1, \rho_2, h) = 0$ imply $\rho_1 = 0$. Then the set $\{\rho_1 = 0\}$ is a one-dimensional manifold of equilibria.

Stability: Now we study the linearization of (f, g) at the equilibrium points $eq = (0, \rho_2, 0, \rho_2, h)$. Assuming that $\rho_2 \neq 0$ and $h < \left|\frac{2}{C^1_{12}\rho_2}\right|$ we compute

$$\left(\begin{array}{cc} \frac{\partial f}{\partial \rho_1} & \frac{\partial f}{\partial \rho_2} \\ \frac{\partial g}{\partial \rho_1} & \frac{\partial g}{\partial \rho_2} \end{array}\right)\Bigg|_{eq} = \left(\begin{array}{cc} \frac{2-hC^1_{12}\rho_2}{2+hC^1_{12}\rho_2} & 0 \\ 0 & 1 \end{array}\right),$$

with eigenvalues

$$\lambda_1 = \frac{2 - hC^1_{12}\rho_2}{2 + hC^1_{12}\rho_2} = \frac{2(J + ma^2) - ha\sqrt{m}\rho_2}{2(J + ma^2) + ha\sqrt{m}\rho_2}, \quad \lambda_2 = 1\,.$$

Since $0 < h < \left|\frac{2(J+ma^2)}{a\sqrt{m}\rho_2}\right| = \left|\frac{2}{C^1_{12}\rho_2}\right|$, we have $\lambda_1 > 0$, regardless of $\rho_2 \neq 0$. Further if $\rho_2 > 0$ then $\lambda_1 < 1$ and hence the equilibrium is stable and asymptotically stable with respect to ρ_1. On the other hand if $\rho_2 < 0$ then $\lambda_1 > 1$ and hence the equilibrium is unstable. Therefore the proposed discrete method reproduces the same qualitative behaviour as the continuous system. This is not guaranteed when applying the midpoint rule to the Chaplygin sleigh system in the original coordinates.

Proposition 6.3.1. *The energy-preserving method (6.21d)-(6.21e) has a one-dimensional manifold of equilibria* $\{\rho_1 = 0\}$*. Assuming that* $h < \left|\frac{2}{C^1_{12}\rho_2}\right|$*, the equilibria* $(0, \rho_2)$ *are stable and asymptotically stable with respect to* ρ_1 *if* $\rho_2 > 0$ *and are unstable if* $\rho_2 < 0$*.*

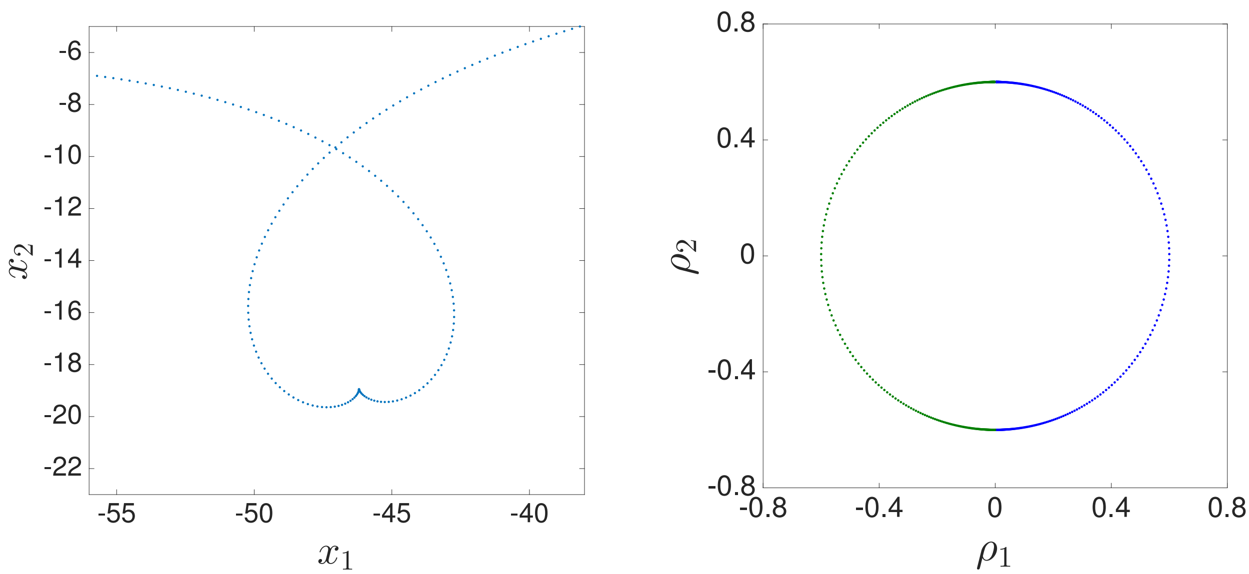


FIGURE 6.2: Integration results for the sleigh, using the method (6.21).

In Figure 6.2 we see an example of how the method exhibits correct behaviour by converging towards a stable equilibrium point when starting very close to an unstable one. The parameters are set to $J = 8$,

$a = m = 1$, step-size $h = 0.5$ and initial values $x_1 = -5$, $x_2 = 0$, $\theta = 0.1$, $\rho_1 \in \{-0.001, 0.001\}$, $\rho_2 = -0.6$. On the left we see a partial x_1x_2 trajectory, while in the right there are two $\rho_1\rho_2$ trajectories.

Remark 6.3.2. Similarly it is possible to show that any convergent Runge-Kutta method will give the correct behaviour for h small enough, when applied to the equations (6.20). For example, applying the explicit Euler method to these equations, we obtain the same conclusion as in Proposition 6.3.1 if we assume $h < \left|\frac{1}{C_{12}^1\rho_2}\right|$. This confirms the fact that the illustrated good qualitative behaviour with respect to stability of equilibria is more an effect of the choice of coordinates rather than of the choice of the method applied in those coordinates.

6.3.3 Euler-Poincaré-Suslov problem on $\mathfrak{so}(3)$

In this example we show that the approach that we have presented is also valid for nonholonomic systems defined on a Lie algebra (and more generally on a Lie algebroid [37]).
Let $\{e_1, e_2, e_3\}$ be a basis of the Lie algebra $\mathfrak{so}(3) \cong \mathbb{R}^3$ and denote the corresponding coordinates by $(\omega_1, \omega_2, \omega_3)$. Consider a kinetic Lagrangian on $\mathfrak{so}(3)$ defined by the matrix

$$(\boldsymbol{g}_{ij}) = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix},$$

and introduce the nonholonomic constraints given by $\sum a_i\omega_i = 0$, where $a \in \mathfrak{so}(3)$ is a fixed element. We can choose the frame $\{e_1, e_2, e_3\}$ in such a way that $I_{12} = 0$ and $a = e_3$. Then the Lagrangian is given by

$$L = \frac{1}{2}(I_{11}\omega_1^2 + I_{22}\omega_2^2 + I_{33}\omega_3^2 + 2I_{13}\omega_1\omega_3 + 2I_{23}\omega_2\omega_3),$$

and the constraint reduces to $\omega_3 = 0$. This defines the distribution

$$\mathcal{D} = \text{span}\{X_1 := (1,0,0),\ X_2 := (0,1,0)\}.$$

Since the bracket on $\mathfrak{so}(3)$ is given by the cross product, it is immediate that $\mathcal{D}$ is not involutive, and hence the constraint $\omega_3 = 0$ is nonholonomic. On the other hand,

$$\mathcal{D}^{\perp,g} = \text{span}\{X_3 := (I_{22}I_{13}, I_{11}I_{23}, -I_{11}I_{22})\}.$$

The Lie bracket of X_1 and X_2 is expressed in terms of the adapted basis X_1, X_2, X_3 as

$$[X_1, X_2] = (1,0,0) \times (0,1,0) = (0,0,1) = \frac{I_{13}}{I_{11}}X_1 + \frac{I_{23}}{I_{22}}X_2 - \frac{1}{I_{11}I_{22}}X_3.$$

Thus the nonvanishing structure constants of the projected bracket are

$$C_{12}^1 = \frac{I_{13}}{I_{11}} \quad \text{and} \quad C_{12}^2 = \frac{I_{23}}{I_{22}}.$$

If we denote by (y^1, y^2, y^3) the coordinates corresponding to the adapted basis $\{X_1, X_2, X_3\}$, the change of coordinates is given by

$$\omega_1 = y^1 + I_{22}I_{13}y^3, \quad \omega_2 = y^2 + I_{11}I_{23}y^3, \quad \omega_3 = -I_{11}I_{22}y^3.$$

Then the restricted Lagrangian becomes $l = \frac{1}{2}(I_{11}(y^1)^2 + I_{22}(y^2)^2)$ and the nonholonomic constraint is $y^3 = 0$.

In this example, since there are no (q^i) variables, the equations of motion (6.5a)-(6.5b) reduce to $\dot{\rho}_a = -C^c_{ab}\rho_c \frac{\partial \mathcal{H}}{\partial \rho_b}$, where $\mathcal{H} = \frac{1}{2}\left(\frac{1}{I_{11}}\rho_1^2 + \frac{1}{I_{22}}\rho_2^2\right)$ and $\rho_i = \frac{\partial l}{\partial y^i} = I_{ii}y^i$, that is

$$\dot{\rho}_1 = -\frac{1}{I_{22}}(C^1_{12}\rho_1 + C^2_{12}\rho_2)\rho_2 \quad \text{and} \quad \dot{\rho}_2 = -\frac{1}{I_{11}}(C^1_{21}\rho_1 + C^2_{21}\rho_2)\rho_1 \,.$$

In matrix form, using $C^c_{ab} = -C^c_{ba}$, we get

$$\begin{pmatrix} \dot{\rho}_1 \\ \dot{\rho}_2 \end{pmatrix} = -\begin{pmatrix} 0 & C^1_{12}\rho_1 + C^2_{12}\rho_2 \\ -(C^1_{12}\rho_1 + C^2_{12}\rho_2) & 0 \end{pmatrix}\begin{pmatrix} \frac{\rho_1}{I_{11}} \\ \frac{\rho_2}{I_{22}} \end{pmatrix}.$$

We apply the same discrete gradient method as in the previous example to get the integrator

$$\begin{pmatrix} \frac{\rho_1'-\rho_1}{h} \\ \frac{\rho_2'-\rho_2}{h} \end{pmatrix} = -\begin{pmatrix} 0 & C^1_{12}\frac{\rho_1'+\rho_1}{2} + C^2_{12}\frac{\rho_2'+\rho_2}{2} \\ -\left(C^1_{12}\frac{\rho_1'+\rho_1}{2} + C^2_{12}\frac{\rho_2'+\rho_2}{2}\right) & 0 \end{pmatrix}\begin{pmatrix} \frac{\rho_1'+\rho_1}{2I_{11}} \\ \frac{\rho_2'+\rho_2}{2I_{22}} \end{pmatrix}.$$

As mentioned previously the integrator is here equivalent to the implicit midpoint method.

6.3.4 Continuous gearbox driven by an asymmetric pendulum

In this final example we compare the performance of integrators applied directly to the formulation of the system in canonical coordinates. We consider a continuous gearbox driven by an asymmetric pendulum. This system is a special case of the continuous gearbox system discussed in [122]. Here $Q = \mathbb{R}^3$ with Hamiltonian

$$H(q^i, p_i) = \frac{1}{2}\left(p_1^2 + p_2^2 + p_3^2\right) - V(q^i)\,, \tag{6.22a}$$

$$V(q^i) = \frac{1}{2}\left((q^1)^2 + (q^2)^2\right) + \cos(q^3) - \frac{1}{5}\sin(2q^3)\,. \tag{6.22b}$$

The single nonholonomic linear velocity constraint is

$$\dot{q}^1 + \sin(q^3)\dot{q}^2 = p_1 + \sin(q^3)p_2 = 0\,,$$

since clearly $p_i = \dot{q}^i$, $1 \le i \le 3$ because $(\boldsymbol{g}_{ij}) = I_3$.

Again comparing with the semi implicit reversible DLA variational integrator (SI-DLA) proposed in [117], we consider from Section 6.2.3 the initial method (6.13) using the Gonzalez midpoint discrete gradient (GONZALEZ-F) and the canonical coordinate implementation of the Gonzalez midpoint discrete gradient method for the reduced system (GONZALEZ-R). In Figure 6.3 we compare the methods for a long time simulation $t \in [0, 100000]$ with random initial values chosen to ensure non-periodic behaviour. On the left we plot the relative energy error, i.e. $|(H - H_0)/H_0|$, where $H_0 := H(0)$ is the initial energy. On the right we plot a relative error in the nonholonomic constraint

$$\frac{|p_1 + \sin(q^3)p_2|}{\|(p_1, \sin(q^3)p_2, \cos(q^3)q^3p_2)\|_\infty}. \tag{6.23}$$

The denominator in (6.23) accounts for the dependence of the round off error, when computing the constraint error $|p_1 + \sin(q^3)p_2|$, on the size of the solution components, since the components for

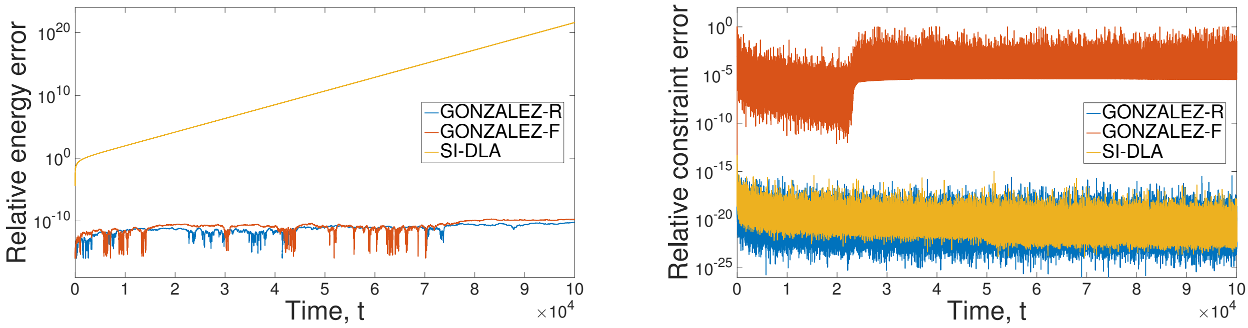


FIGURE 6.3: Integration results for methods GONZALEZ-R, GONZALEZ-F and SI-DLA, applied to the system (6.22) with random initial values, $h = 0.1$, and $t \in [0, 100000]$.

SI-DLA grow very large. A relative constraint error at the level of machine precision implies that the constraint error is due to round off.

For SI-DLA we observe an exponential growth in the relative energy error, while both Gonzalez methods preserve the energy to machine precision. SI-DLA and GONZALEZ-R both preserve the nonholonomic constraint to machine precision. GONZALEZ-F does not respect the nonholonomic constraint. The results are thus as expected.

Chapter 7

Future lines of research

In this last chapter we propose some future projects which are a natural continuation of the work presented in this manuscript.

Matching via the inverse problem II

Matching conditions have been derived for various contexts, including Euler-Poincaré mechanical systems [22] and discrete mechanical systems [20], for which we have Helmholtz type conditions available, derived for instance in Chapters 3 and 4. One of the projects I want to work on is to use these Helmholtz conditions, along the lines of the current project, described in Section 2.4, in order to also derive new matching conditions. In particular I would try to use the Helmholtz conditions derived in Chapter 4 for discrete systems in order to obtain discrete matching conditions, and compare them with the results in [20].

Extension of paper [57] to higher dimensions

As explained in Chapter 2, the results in [57] are restricted to two-dimensional systems since this is the only case, apart from dimension one, for which a classification is available. Nevertheless, as pointed out before, some of the cases of Douglas' classification have been generalized to arbitrary dimension, and one of them is precisely Case IIa1, the case we use in [57] to make systems variational. This case has been studied by M. Crampin, G. E. Prince, W. Sarlet, and G. Thompson in [47], showing that it is always variational. One project I would like to continue is to study the applicability of these results to tackle higher dimensional systems along the lines of [57]. It would also be interesting to study whether the condition of a cyclic variable can be removed using similar techniques. A paper in that direction is [15], where the method of controlled Lagrangians is extended to systems with a symmetry in the kinetic energy, but with a potential that breaks the symmetry of the Lagrangian.

Matching tecniques in one step

In Section 2.3.5 we have used two steps in order to achieve asymptotic stability. First we have assumed that we could add a control in such a way that the system is variational, second we have added an extra control to make the system dissipative, but with the same Lagrangian as the one of the first step. The first step is based on an analysis of the solution space of the Helmholtz conditions (1.14)-(1.16). However, also for dissipative systems, there exist Helmholtz type conditions. That is to say, given a

SODE in any dimension n one could wonder under what conditions there exist a multiplier matrix (g_{ij}) and functions L and D such that

$$g_{ij}\left(\ddot{q}^j - \Gamma^j\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} - \frac{\partial D}{\partial \dot{q}^i}\,.$$

As it turns out, in [46, 121] it is shown that necessary and sufficient conditions for this to occur can also be written entirely in terms of the multipliers g_{ij}, without having to make reference to the sought functions L and D. One of the conditions is of algebraic type and of the form

$$\sum_{X,Y,Z} g(R(X,Y),Z) = 0\,,$$

where R stands for the curvature of the nonlinear connection that can be associated to a SODE. A possible classification of such dissipative SODEs would be based on properties of this curvature and its derivatives (much like the Douglas' classification is based on Φ and its corresponding Helmholtz condition). In dimension $n = 2$, however, the curvature condition is automatically satisfied and in [27] it is even shown that every two-dimensional system of second order differential equations is dissipative. It would be an interesting path to investigate whether, based on these ideas, one may find assumptions under which one may asymptotically stabilize a two-dimensional mechanical system.

Use of exterior differential systems theory in the constrained inverse problem

Exterior differential systems theory [26] (EDS) has been applied successfully to the inverse problem of the calculus of variations via Theorems 1.4.1 and 1.4.2, which give characterizations of the inverse problem in terms of the existence of a closed two-form with further properties. Using EDS some of the variational cases that appear in Douglas' classification have been generalized to arbitrary dimension [3]. These techniques have also enabled a better understanding of Douglas' classification, as well as a proposal for a new classification scheme, yet to be exploited in general [54].
One of my goals in view of the results for the classical inverse problem, is to try to use EDS for the extension of the inverse problem to constrained systems presented in Chapter 5, since a characterization in terms of the existence of a closed two-form is also available.

Geometric integration of differential-algebraic equations

In Chapter 6 we have not addressed the case when Q is a differentiable manifold. In a future paper, we will propose to adapt the discrete gradient approach taking the geometry of the configuration space into account, see for instance the methods in reference [33]. In order to adapt the ideas in [33] to a general differentiable manifold Q, we will need to introduce a finite difference map or retraction map $\Phi_h : U \subset \mathcal{D}^* \times \mathcal{D}^* \to T\mathcal{D}^*$ (see [117]) from a finite difference map initially defined on Q. In this case we will define a discrete gradient as a map $\bar{\nabla}\mathcal{H} : \mathcal{D}^* \times \mathcal{D}^* \longrightarrow T^*\mathcal{D}^*$ verifying similar properties to Definition 6.2.1 (see [33]).

$$\begin{array}{ccc} \mathcal{D}^* \times \mathcal{D}^* & \xrightarrow{\ \bar{\nabla}\mathcal{H}\ } & T^*\mathcal{D}^* \\ \Big\downarrow{\scriptstyle \Phi_h} & & \Big\downarrow{\scriptstyle \pi_{\mathcal{D}^*}} \\ T\mathcal{D}^* & \xrightarrow{\ \tau_{\mathcal{D}^*}\ } & \mathcal{D}^* \end{array}$$

In this case, an energy preserving integrator for Equation (6.7) would be

$$\Phi_h(\zeta, \zeta') = \Pi(\bar{\zeta})\bar{\nabla}\mathcal{H}(\zeta, \zeta')$$

with $\bar{\zeta} = \tau_{\mathcal{D}}^*(\Phi_h(\zeta, \zeta'))$. We will explore this possibility in a future paper since in many examples of nonholonomic systems the configuration space is a nonlinear space such as, for instance, a Lie group. Moreover, it would be interesting to compare the discrete gradient method approach introduced in [32] with other methods designed for nonholonomic systems. For instance, the Chaplygin case is given by a Lagrangian system with forces on the tangent space of a reduced space and then it is possible to use directly discrete variational integrators based on forced Lagrangian systems (see [39, 40]). Other interesting possibilities to compare our methods with are variational integrators from Hamiltonizable nonholonomic systems [62] or the geometric nonholonomic integrator [63].
Additionally, it would be interesting to study the possible applications of energy-preserving integrators to other situations, for instance, to the case of interconnection of port-Hamiltonian systems where the total energy is preserved [158].

Constrained version for Lie algebroids and Lie groupoids

In Chapter 5 we dicussed extensions of the inverse problem to constrained systems, both continuous systems on TQ and discrete systems on $Q \times Q$, but not for constrained systems on Lie algebroids. So it remains to study how our construction can be adapted to tackle the inverse problem for nonholonomic systems on Lie algebroids, using isotropic submanifolds similarly to the description on the tangent bundle given in [11]. Another possibility is to extend our technique to the context of Lie groupoids (see [165]) in order to be able to study for instance the variationality of discretizations of the Euler-Poincaré equations [109].

Integrators for Chaplygin systems

Given a Chaplygin system, the DLA algorithm can be reduced to an algorithm on Q/G called RDLA, provided that we choose the discrete Lagrangian L_d and discrete constraint space D_d to be invariant under the diagonal action of G on $Q \times Q$ [39]. It is in general of the form

$$D_1 L_d^*(r_k, r_{k+1}) + D_2 L_d^*(r_{k-1}, r_k) = F^-(r_k, r_{k+1}) + F^+(r_{k-1}, r_k)\,.$$

Under some extra assumptions we get $F^-(r_k, r_{k+1}) = F^+(r_{k-1}, r_k) = 0$, and therefore the RDLA gives a variational integrator on Q/G, but this is not generally the case. It would be interesting to check if it is possible to find an alternative Lagrangian for which the RDLA integrator becomes variational, and also for the methods proposed in [62].

From the discrete to the continuous setting

In Chapter 4 we have shown how to associate a discrete variational SOdE to a continuous variational SODE. We will complete the results in Section 4.3 with the opposite direction. Assume we have a family of discrete variational SOdEs, parametrized by the time step h and corresponding discrete Lagrangians L_d^h. We will show how to construct, using methods of backward error analysis [77], a family of continuous Lagrangians depending on h with corresponding exact discrete Lagrangians which are close to the original dicrete Lagrangian L_d^h for small h.

Appendix A

Isotropic and Lagrangian submanifolds

We briefly introduce some basic definitions and results from symplectic geometry that have been used throughout the manuscript. More details can be found in [1] and [104].

A.1 Symplectic vector spaces

Recall that a symplectic vector space is a pair (E, Ω) where E is a vector space and $\Omega\colon E \times E \to \mathbb{R}$ is a skew-symmetric bilinear map of maximal rank. See [76, 53, 104, 162] for more details.

Definition A.1.1. *Let (E, Ω) be a symplectic vector space and $F \subset E$ a subspace. The Ω-**orthogonal complement of** F is the subspace defined by*

$$F^{\perp} = \{e \in E \mid \Omega(e, e') = 0 \textit{ for all } e' \in F\} .$$

The subspace F is said to be

1. ***isotropic** if $F \subseteq F^{\perp}$, that is, $\Omega(e, e') = 0$ for all $e, e' \in F$.*
2. ***Lagrangian** if F is isotropic and has an isotropic complement, that is, $E = F \oplus F'$, where F' is isotropic.*

A well-known characterization of Lagrangian subspaces of finite dimensional symplectic vector spaces is summarized in the following result:

Proposition A.1.2. *Let (E, Ω) be a finite dimensional symplectic vector space and $F \subset E$ a subspace. The following assertions are equivalent:*

1. *F is Lagrangian,*
2. *$F = F^{\perp}$,*
3. *F is isotropic and $\dim F = \frac{1}{2}\dim E$.*

As a consequence, we can characterize a Lagrangian subspace F of (E, Ω) by checking if it has half the dimension of E and if the restriction of Ω to F vanishes, that is, $\Omega_{|F} = 0$.

A.2 Symplectic manifolds

A symplectic manifold is a pair (M, ω), where M is a smooth manifold and ω is a nondegenerate closed two-form on M. Therefore, for each $x \in M$, (T_xM, ω_x) is a symplectic vector space. A symplectic manifold has even dimension.
The notion of Lagrangian subspace can be transferred to submanifolds by requiring that the tangent space of the submanifold is a Lagrangian subspace for every point in the submanifold of a symplectic manifold.

Definition A.2.1. *Let (M, ω) be a symplectic manifold, $i\colon N \to M$ be an immersion and $T_xi\colon T_xN \to T_{i(x)}M$ be the tangent map of i. It is said that N is an* ***isotropic immersed submanifold*** *of (M, ω) if $Ti(T_xN) \subset T_{i(x)}M$ is an isotropic subspace for each $x \in N$. A submanifold $N \subset M$ is called* ***Lagrangian*** *if it is isotropic and there is an isotropic subbundle $P \subset TM|_N$ such that $TM|_N = TN \oplus P$.*

Note that $i\colon N \to M$ is an isotropic immersed submanifold if and only if $i^*\omega = 0$, that is, $\omega(T_xi(v_x), T_xi(u_x)) = 0$ for every $u_x, v_x \in T_xN$ and for every $x \in N$.
The canonical model of a symplectic manifold is the cotangent bundle T^*Q of an arbitrary manifold Q, which is the dual bundle of $\tau_Q : TQ \to Q$. Denote by $\pi_Q\colon T^*Q \to Q$ the canonical projection and define a canonical one-form θ_Q on T^*Q by

$$(\theta_Q)_{\alpha_q}(X_{\alpha_q}) = \langle \alpha_q, T_{\alpha_q}\pi_Q(X_{\alpha_q})\rangle\,, \tag{A.1}$$

where $X_{\alpha_q} \in T_{\alpha_q}T^*Q$, $\alpha_q \in T^*Q$ and $q \in Q$. If we consider bundle coordinates (q^i, p_i) on T^*Q such that $\pi_Q(q^i, p_i) = q^i$, then

$$\theta_Q = p_i dq^i\,.$$

The two-form $\omega_Q = -d\theta_Q$ is a symplectic form on T^*Q with local expression

$$\omega_Q = dq^i \wedge dp_i\,.$$

The Darboux Theorem states that this is the local model for an arbitrary symplectic manifold (M, ω). In other words, there always exist local coordinates (q^i, p_i) in a neighbourhood of each point in M such that $\omega = dq^i \wedge dp_i$.
Note that the canonical one-form θ_Q verifies that $\gamma^*(\theta_Q) = \gamma$ for any one-form γ on Q. Hence $\gamma^*(\omega_Q) = -d\gamma$.
A relevant example of a Lagrangian submanifold of the cotangent bundle is the following one.

Proposition A.2.2 ([104])**.** *Let γ be a one-form on Q and $\mathcal{L} = Im\,\gamma \subset T^*Q$. The submanifold $\mathcal{L}$ of T^*Q is Lagrangian if and only if γ is closed.*

The result follows because $\dim \mathcal{L} = \dim Q$ and $\gamma^*(\omega_Q) = -d\gamma$.
A useful extension of the previous construction is the following one:

Proposition A.2.3 ([76]). *Let $i : N \longrightarrow TQ$ be an immersion. For each Lagrangian submanifold $S \subset T^*N$ we can define a Lagrangian submanifold $\tilde{S} \subset T^*TQ$ by $\tilde{S} = \{\mu \in T^*TQ : i^*\mu \in S\}$.*

In the above proposition, if N is a submanifold and $S = \mathrm{Im}(df)$ for some $f : N \longrightarrow \mathbb{R}$, then we recover the following result:

Theorem A.2.4 ([153, 154]). *Let Q be a smooth manifold, $\tau_Q : TQ \to Q$ its tangent bundle projection, $N \subset Q$ a submanifold, and $f \colon N \to \mathbb{R}$ a smooth map. Then*

$$\Sigma_f = \left\{p \in T^*Q \mid \pi_Q(p) \in N \textit{ and } \langle p, v\rangle = \langle df, v\rangle \textit{ for all } v \in TN \subset TQ \textit{ such that } \tau_Q(v) = \pi_Q(p)\right\}$$

is a Lagrangian submanifold of T^*Q.

Given a symplectic manifold (M, ω), $\dim M = 2n$, it is well known that its tangent bundle TM is equipped with a symplectic structure, the tangent lift of ω to TM, denoted by $d_T\omega$. If we take Darboux coordinates (q^i, p_i) on M, that is, $\omega = dq^i \wedge dp_i$, then $d_T\omega = d\dot{q}^i \wedge dp_i + dq^i \wedge d\dot{p}_i$, where $(q^i, p_i, \dot{q}^i, \dot{p}_i)$ are the induced coordinates on TM. If we denote the bundle coordinates on T^*M by (q^i, p_i, a_i, b^i), then $\omega_M = dq^i \wedge da_i + dp_i \wedge db^i$. If we denote by $\flat_\omega : TM \to T^*M$ the isomorphism defined by ω, that is, $\flat_\omega(v) = i_v\,\omega$, then we have $\flat_\omega(q^i, p_i, \dot{q}^i, \dot{p}_i) = (q^i, p_i, -\dot{p}_i, \dot{q}^i)$. This isomorphim plays an important role in the description of the dynamics of Lagrangian and Hamiltonian systems as summarized latter in Appendix B (more details can be found in [153]).
Given a function $H : M \to \mathbb{R}$, and its associated Hamiltonian vector field X_H, that is, the unique vector field satisfying $i_{X_H}\omega = dH$, the image of X_H is a Lagrangian submanifold of $(TM, d_T\omega)$.

A.3 From isotropic to Lagrangian submanifolds

The following construction can be found in [156] and is used in Chapter 5. Assume we have a submanifold N of a symplectic manifold (M, ω) such that for a neighborhood U_p of a point p in M we can write

$$U_p \cap N = \{x \in M \mid \phi_1(x) = 0, \ldots, \phi_k(x) = 0\}\,.$$

If we have an isotropic submanifold $N_0 \subset N$ with $p \in N_0$, $\dim(N_0) = \frac{\dim(N)-k}{2}$ and the Hamiltonian vector fields $X_{\phi_1}, \ldots, X_{\phi_k}$ of $\phi_1, \ldots, \phi_k$ satisfy that

- $\exists\ \epsilon > 0$ such that the flows of X_{ϕ_i} are defined for all $|t| < \epsilon$,
- $X_{\phi_i}(p) \notin T_pN_0$, for all $i = 1, \ldots, k$ and $p \in N_0$,
- $X_{\phi_i}(p)$ are linearly independent for all $p \in N_0$,

then we can extend it to a Lagrangian submanifold transporting N_0 along the flows of the Hamiltonian vector fields $X_{\phi_1}, \ldots, X_{\phi_k}$.

We will illustrate the construction for the case $k = 1$ and rename ϕ_1 by ϕ. Since X_ϕ is transverse to N_0, there exists an open interval I about 0 in $\mathbb{R}$ such that $\exp\,(tX_\phi(\tilde{p}))$ is defined for all $t \in I$ and $\tilde{p} \in N_0 \cap U_p$. Therefore the map

$$\begin{array}{rrcl} j: & N_0 \times I & \longrightarrow & M \\ & (\tilde{p}, t) & \longmapsto & \exp\,(tX_\phi(\tilde{p})) \end{array}$$

allows us to realize locally $N_0 \times I$ as a submanifold Z of M whose tangent space is

$$T_{\exp\,(tX_\phi(\tilde{p}))} Z = (\exp\,(tX_\phi))_*(T_{\tilde{p}} N_0) \oplus \operatorname{span}\{X_\phi(\exp\,(tX_\phi(\tilde{p})))\}\ ,$$

where $(\exp\,(tX_\phi))_*$ is the pushforward of $\exp\,(tX_\phi)$. Obviously $\dim Z = \dim N_0 + 1$ and Z is also isotropic because, first, for any two vectors in $(\exp\,(tX_\phi))_*(T_{\tilde{p}} N_0)$ we have that

$$\omega((\exp\,(tX_\phi))_* v_1, (\exp\,(tX_\phi))_* v_2) = ((\exp\,(tX_\phi))^* \omega)(v_1, v_2) = \omega(v_1, v_2) = 0\,,$$

since $(\exp\,(tX_\phi))_*$ is a symplectomorphism and $v_1, v_2 \in T_{\tilde{p}} N_0$.
Second, it must be checked that the two-form ω also vanishes for a vector in $(\exp\,(tX_\phi))_*(T_{\tilde{p}} N_0)$ and one in $X_\phi(\exp\,(tX_\phi(\tilde{p})))$. Note that

$$\omega((\exp\,(tX_\phi))_* v, X_\phi(\exp\,(tX_\phi(\tilde{p})))) = d\phi(\tilde{p})(v) = 0\,,$$

because ϕ vanishes on N_0 and $v \in T_{\tilde{p}} N_0$.

Appendix B

Lagrangian mechanics using the Tulczyjew's triple

The theory of Lagrangian submanifolds gives an intrinsic geometric description of Lagrangian and Hamiltonian dynamics [153, 154]. Moreover, it allows us to relate the Lagrangian and Hamiltonian formalisms using the so-called Tulczyjew's triple

$$T^*TQ \xleftarrow{\alpha_Q} TT^*Q \xrightarrow{\beta_Q} T^*T^*Q\,.$$

The Tulczyjew map α_Q is an isomorphism between TT^*Q and T^*TQ. Besides, it is also a symplectomorphism between these vector bundles considered as symplectic manifolds, i.e. $(TT^*Q, d_T\,\omega_Q)$, where $d_T\,\omega_Q$ is the tangent lift of ω_Q, and (T^*TQ, ω_{TQ}). For completeness, we recall the construction of the symplectomorphism α_Q. To do this, it is necessary to introduce the canonical involution κ_Q on TTQ

$$\begin{array}{ccc} TTQ & \xrightarrow{\kappa_Q} & TTQ \\ \downarrow{\scriptstyle \tau_{TQ}} & & \downarrow{\scriptstyle T\tau_Q} \\ TQ & \xrightarrow[\mathrm{Id}]{} & TQ, \end{array}$$

defined by

$$\kappa_Q\left(\left.\frac{d}{ds}\right|_{s=0}\left.\frac{d}{dt}\right|_{t=0}\chi\left(s,t\right)\right) = \left.\frac{d}{ds}\right|_{s=0}\left.\frac{d}{dt}\right|_{t=0}\tilde{\chi}\left(s,t\right),$$

where $\chi : \mathbb{R}^2 \to Q$ and $\tilde{\chi} : \mathbb{R}^2 \to Q$ are related by $\tilde{\chi}\left(s,t\right) = \chi\left(t,s\right)$. If $\left(q^i\right)$ are the local coordinates for Q, $\left(q^i, v^i\right)$ for TQ and $\left(q^i, v^i, \dot{q}^i, \dot{v}^i\right)$ for TTQ, then the canonical involution is locally given by $\kappa_Q\left(q^i, v^i, \dot{q}^i, \dot{v}^i\right) = \left(q^i, \dot{q}^i, v^i, \dot{v}^i\right)$.

In order to describe α_Q it is also necessary to define a tangent pairing. Given two manifolds M and N, and a pairing $\langle\cdot,\cdot\rangle : M \times N \to \mathbb{R}$ between them, the tangent pairing $\langle\cdot,\cdot\rangle^T : TM \times TN \to \mathbb{R}$ is determined by

$$\left\langle \left.\frac{d}{dt}\right|_{t=0}\gamma\left(t\right), \left.\frac{d}{dt}\right|_{t=0}\delta\left(t\right)\right\rangle^T = \left.\frac{d}{dt}\right|_{t=0}\left\langle\gamma\left(t\right),\delta\left(t\right)\right\rangle$$

where $\gamma : \mathbb{R} \to M$ and $\delta : \mathbb{R} \to N$.

Finally, we can define α_Q as $\langle\alpha_Q\left(z\right), w\rangle = \langle z, \kappa_Q\left(w\right)\rangle^T$, where $z \in TT^*Q$ and $w \in TTQ$. In local coordinates $\left(q^i, p_i\right)$ for T^*Q and $\left(q^i, p_i, \dot{q}^i, \dot{p}_i\right)$ for TT^*Q, we have

$$\alpha_Q\left(q^i, p_i, \dot{q}^i, \dot{p}_i\right) = \left(q^i, \dot{q}^i, \dot{p}_i, p_i\right).$$

The isomorphism $\beta_Q : TT^*Q \to T^*T^*Q$ is just given by $\beta_Q = \flat_{\omega_Q}$, that is, $\flat_{\omega_Q}(v) = i_v\omega_Q$.
The Lagrangian dynamics is described by the Lagrangian submanifold $dL(TQ)$ of T^*TQ, where $L\colon TQ \to \mathbb{R}$ is the Lagrangian function, while the Hamiltonian formalism is described by the Lagrangian submanifold $dH(T^*Q)$ of T^*T^*Q, where $H\colon T^*Q \to \mathbb{R}$ is the corresponding Hamiltonian function. The solutions of the dynamics are curves $\gamma : I \subset \mathbb{R} \to T^*Q$ such that $\frac{d\gamma}{dt} : I \subset \mathbb{R} \to TT^*Q$ verifies that $\frac{d\gamma}{dt}(I) \subset \alpha_Q^{-1}(dL(TQ))$ in the Lagrangian description and $\frac{d\gamma}{dt}(I) \subset \beta_Q^{-1}(dH(T^*Q))$ in the Hamiltonian case.
Variationally constrained problems described in Section 5.2 are determined by a pair (M, l), where M is a submanifold of TQ, with inclusion $i_M : M \hookrightarrow TQ$, and $l : M \to \mathbb{R}$ is a Lagrangian function restricted to M. The submanifold Σ_l is a Lagrangian submanifold of (T^*TQ, ω_Q), see Theorem A.2.4. Now using the Tulczyjew's symplectomorphism α_Q, we induce a new Lagrangian submanifold $\alpha_Q^{-1}(\Sigma_l)$ of $(TT^*Q, d_T\omega_Q)$, which completely determines the constrained variational mechanics. Now we will see that this procedure gives the correct equations for the constrained variational mechanics. Take an arbitrary extension $L : TQ \to \mathbb{R}$ of $l : M \to \mathbb{R}$, that is, $L \circ i_M = l$. Locally,

$$\begin{aligned}\Sigma_l \;=\; & \{(q^i, \dot q^i, \mu_i, \tilde\mu_i) \in T^*TQ \mid \mu_i = \frac{\partial L}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial q^i}, \\ & \tilde\mu_i = \frac{\partial L}{\partial \dot q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot q^i}, \quad \phi^\alpha(q, \dot q) = 0, \;\; 1 \le \alpha \le m\}\,.\end{aligned}$$

Therefore,

$$\begin{aligned}\alpha_Q^{-1}(\Sigma_l) \;=\; & \{(q^i, p_i, \dot q^i, \dot p_i) \in TT^*Q \mid p_i = \frac{\partial L}{\partial \dot q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot q^i}, \\ & \dot p_i = \frac{\partial L}{\partial q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial q^i}, \; \phi^\alpha(q, \dot q) = 0, \;\; 1 \le \alpha \le m\}\,.\end{aligned}$$

The solutions for the dynamics given by $\alpha_Q^{-1}(\Sigma_l) \subset TT^*Q$ are curves $\gamma : I \subset \mathbb{R} \to T^*Q$ such that $\frac{d\gamma}{dt} : I \subset \mathbb{R} \to TT^*Q$ verifies that $\frac{d\gamma}{dt}(I) \subset \alpha_Q^{-1}(\Sigma_l)$. Locally, if $\gamma(t) = (q^i(t), p_i(t))$ then it must verify the following set of differential equations:

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q^i} + \lambda_\alpha \frac{\partial \phi^\alpha}{\partial \dot q^i}\right) - \frac{\partial L}{\partial q^i} - \lambda_\alpha \frac{\partial \phi^\alpha}{\partial q^i} &= 0\,, \\ \phi^\alpha(q^i, \dot q^i) &= 0\,,\end{aligned}$$

which coincide with Equations (5.8).

Appendix C

Equivalence between Helmholtz conditions

We will establish the equivalence between the equations for $\Sigma_{\Gamma,F} \subset T^*TQ$ to be Lagrangian and the Helmholtz conditions (1.14)-(1.16) for $g_{ij} = \frac{\partial F_i}{\partial \dot{q}^j}$ that appear in Section 1.2.
The equations we obtain by imposing that the submanifold $\Sigma_{F,\Gamma}$ be Lagrangian, that is,

$$d\left(\left(\frac{\partial F_i}{\partial q^j}\dot{q}^j + \frac{\partial F_i}{\partial \dot{q}^j}\Gamma^j\right)dq^i + F_i d\dot{q}^i\right) \quad = \quad 0\,,$$

are the following:

$$\frac{\partial F_i}{\partial \dot{q}^j} \quad = \quad \frac{\partial F_j}{\partial \dot{q}^i}\,, \tag{C.1}$$

$$\frac{\partial^2 F_i}{\partial q^j \partial q^k}\dot{q}^k + \frac{\partial^2 F_i}{\partial q^j \partial \dot{q}^k}\Gamma^k + \frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial q^j} \quad = \quad \frac{\partial^2 F_j}{\partial q^i \partial q^k}\dot{q}^k + \frac{\partial^2 F_j}{\partial q^i \partial \dot{q}^k}\Gamma^k + \frac{\partial F_j}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial q^i}\,, \tag{C.2}$$

$$\frac{\partial^2 F_i}{\partial \dot{q}^j \partial q^k}\dot{q}^k + \frac{\partial F_i}{\partial q^j} + \frac{\partial^2 F_i}{\partial \dot{q}^j \partial \dot{q}^k}\Gamma^k + \frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j} - \frac{\partial F_j}{\partial q^i} \quad = \quad 0\,. \tag{C.3}$$

Assume F is a local diffeomorphism that satisfies (C.1), (C.2) and (C.3). The first three sets of Helmholtz conditions (1.14), that is, $det(g_{ij}) \neq 0$, $g_{ij} = g_{ji}$ and $\frac{\partial g_{ij}}{\partial \dot{q}^k} = \frac{\partial g_{ik}}{\partial \dot{q}^j}$, are readily satisfied by $g_{ij} = \left(\frac{\partial F_i}{\partial \dot{q}^j}\right)$.
Taking the difference $(C.3)_{ij} - (C.3)_{ji} = 0$ we get $a_{ij} = a_{ji}$, where $a_{ij} = \frac{\partial F_i}{\partial q^j} + \frac{1}{2}\frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j}$. Then,

$$\begin{aligned}
\Gamma(g_{ij}) - \nabla_j^k g_{ik} - \nabla_i^k g_{kj} - (C.3) \quad = \quad & \frac{\partial^2 F_i}{\partial q^k \partial \dot{q}^j}\dot{q}^k + \frac{\partial^2 F_i}{\partial \dot{q}^k \partial \dot{q}^j}\Gamma^k + \frac{1}{2}\frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j} + \frac{1}{2}\frac{\partial F_k}{\partial \dot{q}^j}\frac{\partial \Gamma^k}{\partial \dot{q}^i} \\
& - \frac{\partial^2 F_i}{\partial \dot{q}^j \partial q^k}\dot{q}^k - \frac{\partial F_i}{\partial q^j} - \frac{\partial^2 F_i}{\partial \dot{q}^j \partial \dot{q}^k}\Gamma^k - \frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j} + \frac{\partial F_j}{\partial q^i} \\
= \quad & - \frac{\partial F_i}{\partial q^j} - \frac{1}{2}\frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j} + \frac{\partial F_j}{\partial q^i} + \frac{1}{2}\frac{\partial F_j}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^i} = a_{ji} - a_{ij} = 0\,.
\end{aligned}$$

Thus the ∇ condition (1.15) is satisfied.
Now we check that the Φ condition (1.16) is satisfied using $(C.2)$ and the condition $a_{ij} = a_{ji}$. From the latter, we have

$$\frac{\partial F_i}{\partial q^j} = \frac{\partial F_j}{\partial q^i} + \frac{1}{2}\frac{\partial F_j}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^i} - \frac{1}{2}\frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j}\,.$$

Substituting this on the left-hand side of $(C.2)$ we get

$$\begin{aligned}
& g_{jl}\left[\frac{1}{2}\frac{\partial^2 \Gamma^l}{\partial q^k \partial \dot{q}^i}\dot{q}^k + \frac{1}{2}\frac{\partial^2 \Gamma^l}{\partial \dot{q}^k \partial \dot{q}^i}\Gamma^k - \frac{\partial \Gamma^l}{\partial q^i}\right] + \frac{1}{2}\frac{\partial \Gamma^l}{\partial \dot{q}^i}\left(\frac{\partial^2 F_j}{\partial q^k \partial \dot{q}^l}\dot{q}^k + \frac{\partial^2 F_j}{\partial \dot{q}^k \partial \dot{q}^l}\Gamma^k\right) \\
& = g_{il}\left[\frac{1}{2}\frac{\partial^2 \Gamma^l}{\partial q^k \partial \dot{q}^j}\dot{q}^k + \frac{1}{2}\frac{\partial^2 \Gamma^l}{\partial \dot{q}^k \partial \dot{q}^j}\Gamma^k - \frac{\partial \Gamma^l}{\partial q^j}\right] + \frac{1}{2}\frac{\partial \Gamma^l}{\partial \dot{q}^j}\left(\frac{\partial^2 F_i}{\partial q^k \partial \dot{q}^l}\dot{q}^k + \frac{\partial^2 F_i}{\partial \dot{q}^k \partial \dot{q}^l}\Gamma^k\right)\,.
\end{aligned} \tag{C.4}$$

Using $(C.3)$ and $a_{ij} = a_{ji}$ again we get

$$\frac{1}{2}\frac{\partial \Gamma^l}{\partial \dot{q}^i}\left(\frac{\partial^2 F_j}{\partial q^k \partial \dot{q}^l}\dot{q}^k + \frac{\partial^2 F_j}{\partial \dot{q}^k \partial \dot{q}^l}\Gamma^k\right) = \frac{1}{2}\frac{\partial \Gamma^l}{\partial \dot{q}^i}\left(\frac{\partial F_l}{\partial q^j} - \frac{\partial F_j}{\partial q^l} - \frac{\partial F_j}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^l}\right)$$
$$= \frac{1}{2}\frac{\partial \Gamma^l}{\partial \dot{q}^i}\left(-\frac{1}{2}\frac{\partial F_l}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j} - \frac{1}{2}\frac{\partial F_j}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^l}\right) = -\frac{1}{4}g_{jk}\frac{\partial \Gamma^k}{\partial \dot{q}^l}\frac{\partial \Gamma^l}{\partial \dot{q}^i} - \frac{1}{4}\frac{\partial \Gamma^l}{\partial \dot{q}^i}\frac{\partial F_l}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j}\,.$$

Since the last term is equal on both sides of (C.4), that is,

$$\frac{\partial \Gamma^l}{\partial \dot{q}^i}\frac{\partial F_l}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j} = \frac{\partial \Gamma^l}{\partial \dot{q}^j}\frac{\partial F_l}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^i}\,,$$

we obtain the Φ condition

$$g_{jl}\Phi^l_i = g_{il}\Phi^l_j, \quad \text{where} \quad \Phi^l_j = \frac{\partial^2 \Gamma^l}{\partial q^k \partial \dot{q}^j}\dot{q}^k + \frac{\partial^2 \Gamma^l}{\partial \dot{q}^k \partial \dot{q}^j}\Gamma^k - 2\frac{\partial \Gamma^l}{\partial q^j} - \frac{1}{2}\frac{\partial \Gamma^l}{\partial \dot{q}^r}\frac{\partial \Gamma^r}{\partial \dot{q}^j}\,.$$

On the other hand, if we assume that the Helmholtz conditions are satisfied by g_{ij} then there exists a local diffeomorphism $F(q,\dot{q}) = (q^i, F_i(q,\dot{q}))$ such that $g_{ij} = \frac{\partial F_i}{\partial \dot{q}^j}$. Then $\Omega = -d(F^*\Theta_Q)$ satisfies the conditions in Theorem 1.4.1. According to [43], Ω is given by

$$\Omega = g_{ij} dq^i \wedge \nu^j\,,$$

where $\left\{dq^j, \nu^j = d\dot{q}^j - \frac{1}{2}\frac{\partial \Gamma^j}{\partial \dot{q}^k}dq^k\right\}$ is the dual basis to $\left\{\frac{\partial}{\partial q^i}, H_i = \frac{\partial}{\partial q^i} + \frac{1}{2}\frac{\partial \Gamma^k}{\partial \dot{q}^i}\frac{\partial}{\partial \dot{q}^k}\right\}$ and H_i is the horizontal lift of $\frac{\partial}{\partial q^i}$ with respect to the connection defined by Γ. Since

$$\Omega = -d(F_i dq^i) = -\frac{\partial F_i}{\partial q^j}dq^j \wedge dq^i - \frac{\partial F_i}{\partial \dot{q}^k}\left(\nu^k + \frac{1}{2}\frac{\partial \Gamma^k}{\partial \dot{q}^j}dq^j\right) \wedge dq^i$$
$$= \underbrace{\left(-\frac{\partial F_i}{\partial q^j} - \frac{1}{2}\frac{\partial F_i}{\partial \dot{q}^k}\frac{\partial \Gamma^k}{\partial \dot{q}^j}\right)}_{-a_{ij}} dq^j \wedge dq^i + \underbrace{\frac{\partial F_i}{\partial \dot{q}^k}}_{g_{ik}} dq^i \wedge \nu^k\,,$$

we obtain $a_{ij} = a_{ji}$ and we can reverse the calculations in the above implication.

Analogous computations can be carried out for the equations in the time-dependent case, now using the local expression for Ω in [48].

Appendix D

Algorithm for nonholonomic systems

We show here an algorithm that corresponds to the method presented in Section 6.2.3 to use energy-preserving integrators for nonholonomic systems, without explicitly writing the system in almost-Poisson form.

Algorithm 1 (QR factorization procedure with differentiation using Householder reflections). Computes the QR factorization of a differentiable matrix $A(q) \in \mathbb{R}^{n,n-m}$ as well as all partial derivatives $\partial Q/\partial q^i$, $1 \leq i \leq n$, for any $q \in \mathbb{R}^n$. Let $\partial_i := \partial/\partial q^i$ while $B_{i:j,k:l}$ with $i \leq j$ and $k \leq l$ denotes the submatrix containing rows i to j and columns k to l of a matrix B, with the shorthand $i := i : i$. ∂B denotes the tensor containing all partial derivatives of B at q. $s \in \{+1, -1\}^{n-m}$ is a vector of sign choices.

procedure QRdiff($A, \partial A, n, m, s$)
- $Q^{(0)} \leftarrow I_n$
- $R^{(0)} \leftarrow A$
- **for** $i = 1, 2, \ldots, n$ **do**
  - $\partial_i Q^{(0)} \leftarrow 0_n$
  - $\partial_i R^{(0)} \leftarrow \partial_i A$
- **end for**
- **for** $k = 0, 1, \ldots, n-m-1$ **do**
  - **for** $j = 1, 2, \ldots, n$ **do**
    - **if** $j \leq k$ **then**
      - $\tilde{w}_j^{(k)} \leftarrow 0$
    - **else if** $j = k+1$ **then**
      - $\tilde{w}_j^{(k)} \leftarrow R_{jk}^{(k)} + s_k \|R_{k:n,k}^{(k)}\|_2$
    - **else**
      - $\tilde{w}_j^{(k)} \leftarrow R_{jk}^{(k)}$
    - **end if**
  - **end for**
  - $\|\tilde{w}^{(k)}\|_2 \leftarrow \sqrt{2\left(\|R_{k:n,k}^{(k)}\|_2^2 + s_k R_{kk}^{(k)} \|R_{k:n,k}^{(k)}\|_2\right)}$
  - $w^{(k)} \leftarrow \frac{\tilde{w}^{(k)}}{\|\tilde{w}^{(k)}\|_2}$
  - $u^{(k)} \leftarrow {w^{(k)}}^T R^{(k)}$
  - $R^{(k+1)} \leftarrow R^{(k)} - 2 w^{(k)} u^{(k)}$
  - $Q^{(k+1)} \leftarrow Q^{(k)} - 2Q^{(k)} \left(w^{(k)} {w^{(k)}}^T\right)$
  - **for** $i = 1, 2, \ldots, n$ **do**
    - $\partial_i \|R_{k:n,k}^{(k)}\|_2 \leftarrow \frac{\left(\partial_i R_{k:n,k}^{(k)}\right)^T R_{k:n,k}^{(k)}}{\|R_{k:n,k}^{(k)}\|_2}$
    - **for** $j = 1, 2, \ldots, n$ **do**
      - **if** $j \leq k$ **then**
        - $\partial_i \tilde{w}_j^{(k)} \leftarrow 0$
      - **else if** $j = k+1$ **then**
        - $\partial_i \tilde{w}_j^{(k)} \leftarrow \partial_i R_{jk}^{(k)} + s_k \partial_i \|R_{k:n,k}^{(k)}\|_2$
      - **else**
        - $\partial_i \tilde{w}_j^{(k)} \leftarrow \partial_i R_{jk}^{(k)}$
      - **end if**
    - **end for**
    - $\partial_i \|\tilde{w}^{(k)}\|_2 \leftarrow \frac{\left(\partial_i \tilde{w}^{(k)}\right)^T \tilde{w}^{(k)}}{\|\tilde{w}^{(k)}\|_2}$
    - $\partial_i w^{(k)} \leftarrow \frac{\partial_i \tilde{w}^{(k)} \|\tilde{w}^{(k)}\|_2 - \tilde{w}^{(k)} \partial_i \|\tilde{w}^{(k)}\|_2}{\|\tilde{w}^{(k)}\|_2^2}$
    - $\partial_i u^{(k)} \leftarrow \partial_i {w^{(k)}}^T R^{(k)} + {w^{(k)}}^T \partial_i R^{(k)}$
    - $\partial_i R^{(k+1)} \leftarrow \partial_i R^{(k)} - 2\left(\partial_i w^{(k)} u^{(k)} + w^{(k)} \partial_i u^{(k)}\right)$
    - $\partial_i Q^{(k+1)} \leftarrow \partial_i Q^{(k)} - 2\partial_i Q^{(k)} \left(w^{(k)} {w^{(k)}}^T\right) + 2Q^{(k)} \left(\partial_i w^{(k)} {w^{(k)}}^T + w^{(k)} \partial_i {w^{(k)}}^T\right)$
  - **end for**
- **end for**
- **return** $Q^{(n-m)}, \partial Q^{(n-m)}$

end procedure

References

[1] R. Abraham and J. E. Marsden. *Foundations of mechanics.* Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978. Second edition, revised and enlarged, With the assistance of Tudor Ratiu and Richard Cushman.

[2] J. E. Aldridge, G. E. Prince, W. Sarlet, and G. Thompson. An EDS approach to the inverse problem in the calculus of variations. *J. Math. Phys.*, 47(10):103508, 22, 2006.

[3] I. Anderson and G. Thompson. The inverse problem of the calculus of variations for ordinary differential equations. *Mem. Amer. Math. Soc.*, 98(473):vi+110, 1992.

[4] I. M. Anderson. Aspects of the inverse problem to the calculus of variations. *Arch. Math. (Brno)*, 24(4):181–202, 1988.

[5] I. M. Anderson and T. Duchamp. On the existence of global variational principles. *Amer. J. Math.*, 102(5):781–868, 1980.

[6] I. M. Anderson and T. E. Duchamp. Variational principles for second-order quasilinear scalar equations. *J. Differential Equations*, 51(1):1–47, 1984.

[7] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein.

[8] V. I. Arnold, V. V. Kozlov, and A. I. Neĭshtadt. *Dynamical systems. III*, volume 3 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, Berlin, 1988. Translated from the Russian by A. Iacob.

[9] P. Balseiro, M. de León, J. C. Marrero, and D. Martín de Diego. The ubiquity of the symplectic Hamiltonian equations in mechanics. *J. Geom. Mech.*, 1(1):1–34, 2009.

[10] P. Balseiro and L. C. García-Naranjo. Gauge transformations, twisted Poisson brackets and Hamiltonization of nonholonomic systems. *Arch. Ration. Mech. Anal.*, 205(1):267–310, 2012.

[11] M. Barbero-Liñán, M. Farré Puiggalí, and D. Martín de Diego. Isotropic submanifolds and the inverse problem for mechanical constrained systems. *J. Phys. A*, 48:045210 (35 pp), 2015.

[12] M. Barbero-Liñán, M. Farré Puiggalí, and D. Martín de Diego. Inverse problem for Lagrangian systems on Lie algebroids and applications to reduction by symmetries. *Monatsh. Math.*, 180(4):665–691, 2016.

[13] G. Bhattarai. *Inverse Problem of Lagrangian Mechanics In Dimension Three.* ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–The University of Toledo.

[14] A. M. Bloch. *Nonholonomic mechanics and control*, volume 24 of *Interdisciplinary Applied Mathematics.* Springer Verlag, 2003.

[15] A. M. Bloch, D. E. Chang, N. E. Leonard, and J. E. Marsden. Controlled Lagrangians and the stabilization of mechanical systems. II. Potential shaping. *IEEE Trans. Automat. Control*, 46(10):1556–1571, 2001.

[16] A. M. Bloch and P. E. Crouch. Nonholonomic and vakonomic control systems on Riemannian manifolds. In *Dynamics and control of mechanical systems (Waterloo, ON, 1992)*, volume 1 of *Fields Inst. Commun.*, pages 25–52. Amer. Math. Soc., Providence, RI, 1993.

[17] A. M. Bloch, O. E. Fernandez, and T. Mestdag. Hamiltonization of nonholonomic systems and the inverse problem of the calculus of variations. *Rep. Math. Phys.*, 63(2):225–249, 2009.

[18] A. M. Bloch, J. E. Krishnaprasad, J. E. Marsden, and R. Murray. Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.*, 136:21–99, 1996.

[19] A. M. Bloch, D. Krupka, and D. V. Zenkov. The Helmholtz conditions and the method of controlled Lagrangians. In *The inverse problem of the calculus of variations*, volume 2 of *Atlantis Stud. Var. Geom.*, pages 1–29. Atlantis Press, Paris, 2015.

[20] A. M. Bloch, M. Leok, J. E. Marsden, and D. V. Zenkov. Controlled Lagrangians and stabilization of discrete mechanical systems. *Discrete Contin. Dyn. Syst. Ser. S*, 3(1):19–36, 2010.

[21] A. M. Bloch, N. E. Leonard, and J. E. Marsden. Controlled Lagrangians and the stabilization of mechanical systems. I. The first matching theorem. *IEEE Trans. Automat. Control*, 45(12):2253–2270, 2000.

[22] A. M. Bloch, N. E. Leonard, and J. E. Marsden. Controlled Lagrangians and the stabilization of Euler-Poincaré mechanical systems. *Internat. J. Robust Nonlinear Control*, 11(3):191–214, 2001.

[23] L. Bourdin and J. Cresson. Helmholtz's inverse problem of the discrete calculus of variations. *J. Difference Equ. Appl.*, 19(9):1417–1436, 2013.

[24] U. Brechtken-Manderscheid. *Introduction to the calculus of variations.* Chapman and Hall Mathematics Series. Chapman & Hall, London, 1991. Translated from the German by P. G. Engstrom.

[25] R. Brockett. The early days of geometric nonlinear control. *Automatica J. IFAC*, 50(9):2203–2224, 2014.

[26] R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, and P. A. Griffiths. *Exterior differential systems*, volume 18 of *Mathematical Sciences Research Institute Publications*. Springer-Verlag, New York, 1991.

[27] I. Bucataru and O. Constantinescu. Generalized Helmholtz conditions for non-conservative Lagrangian systems. *Math. Phys. Anal. Geom.*, 18(1):Art. 25, 24, 2015.

[28] I. Bucataru and M. F. Dahl. Semi-basic 1-forms and Helmholtz conditions for the inverse problem of the calculus of variations. *J. Geom. Mech.*, 1(2):159–180, 2009.

[29] F. Bullo and A. D. Lewis. *Geometric control of mechanical systems*, volume 49 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 2005. Modeling, analysis, and design for simple mechanical control systems.

[30] A. Cannas da Silva and A. Weinstein. *Geometric models for noncommutative algebras*, volume 10 of *Berkeley Mathematics Lecture Notes*. American Mathematical Society, Providence, RI; Berkeley Center for Pure and Applied Mathematics, Berkeley, CA, 1999.

[31] F. Cantrijn, M. de León, and E. A. Lacomba. Gradient vector fields on cosymplectic manifolds. *J. Phys. A*, 25(1):175–188, 1992.

[32] E. Celledoni, M. Farré Puiggalí, E. H. Høiseth, and D. Martín de Diego. Energy-preserving integrators applied to nonholonomic systems. arXiv:1605.02845, 2016.

[33] E. Celledoni and B. Owren. Preserving first integrals with symmetric Lie group methods. *Discrete Contin. Dyn. Syst.*, 34(3):977–990, 2014.

[34] H. Cendra, J. E. Marsden, and T. S. Ratiu. Lagrangian reduction by stages. *Mem. Amer. Math. Soc.*, 152(722):x+108, 2001.

[35] D. E. Chang. Stabilizability of controlled Lagrangian systems of two degrees of freedom and one degree of under-actuation by the energy-shaping method. *IEEE Trans. Automat. Control*, 55(8):1888–1893, 2010.

[36] D. Cohen and E. Hairer. Linear energy-preserving integrators for poisson systems. *BIT Numer Math*, 51:91–101, 2011.

[37] J. Cortés, M. de León, J. C. Marrero, and E. Martínez. Nonholonomic Lagrangian systems on Lie algebroids. *Discrete Contin. Dyn. Syst.*, 24(2):213–271, 2009.

[38] J. Cortés, M. de León, D. Martín de Diego, and S. Martínez. Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions. *SIAM J. Control Optim.*, 41(5):1389–1412, 2002.

[39] J. Cortés and S. Martínez. Non-holonomic integrators. *Nonlinearity*, 14(5):1365–1392, 2001.

[40] J. Cortés Monforte. *Geometric, control and numerical aspects of nonholonomic systems*, volume 1793 of *Lecture Notes in Mathematics*. Springer, 2002.

[41] T. Courant. Tangent Dirac structures. *J. Phys. A*, 23(22):5153–5168, 1990.

[42] D. Crăciun and D. Opriş. The Helmholtz conditions for the difference equations systems. *Balkan J. Geom. Appl.*, 1(2):21–30, 1996.

[43] M. Crampin. On the differential geometry of the Euler-Lagrange equations, and the inverse problem of Lagrangian dynamics. *J. Phys. A*, 14(10):2567–2575, 1981.

[44] M. Crampin. A note on non-Noether constants of motion. *Phys. Lett. A*, 95(5):209–212, 1983.

[45] M. Crampin and T. Mestdag. The inverse problem for invariant Lagrangians on a Lie group. *J. Lie Theory*, 18(2):471–502, 2008.

[46] M. Crampin, T. Mestdag, and W. Sarlet. On the generalized Helmholtz conditions for Lagrangian systems with dissipative forces. *ZAMM Z. Angew. Math. Mech.*, 90(6):502–508, 2010.

[47] M. Crampin, G. E. Prince, W. Sarlet, and G. Thompson. The inverse problem of the calculus of variations: separable systems. *Acta Appl. Math.*, 57(3):239–254, 1999.

[48] M. Crampin, G. E. Prince, and G. Thompson. A geometrical version of the Helmholtz conditions in time-dependent Lagrangian dynamics. *J. Phys. A*, 17(7):1437–1447, 1984.

[49] M. Crampin, W. Sarlet, E. Martínez, G. B. Byrnes, and G. E. Prince. Towards a geometrical understanding of Douglas' solution of the inverse problem of the calculus of variations. *Inverse Problems*, 10(2):245–260, 1994.

[50] M. de León, J. C. Marrero, and D. Martín de Diego. Linear almost Poisson structures and Hamilton-Jacobi equation. Applications to nonholonomic mechanics. *J. Geom. Mech.*, 2(2):159–198, 2010.

[51] M. de León, J.C. Marrero, and E. Martínez. Lagrangian submanifolds and dynamics on Lie algebroids. *J. Phys. A*, 38(24):R241–R308, 2005.

[52] M. de León, D. Martín de Diego, and A. Santamaria-Merino. Geometric integrators and nonholonomic mechanics. *J. Math. Phys.*, 45:1042–1064, 2004.

[53] M. de León and P. R. Rodrigues. *Methods of differential geometry in analytical mechanics*, volume 158 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1989.

[54] T. Do and G. Prince. New progress in the inverse problem in the calculus of variations. *Differential Geom. Appl.*, 45:148–179, 2016.

[55] J. Douglas. Solution of the inverse problem of the calculus of variations. *Trans. Amer. Math. Soc.*, 50:71–128, 1941.

[56] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[57] M. Farré Puiggalí and T. Mestdag. The inverse problem of the calculus of variations and the stabilization of controlled lagrangian systems. *SIAM J. Control Optim.*, 54(6):3297–3318, 2016.

[58] Y. N. Fedorov and D. V. Zenkov. Discrete nonholonomic LL systems on Lie groups. *Nonlinearity*, 18(5):2211–2241, 2005.

[59] Y. N. Fedorov and D. V. Zenkov. Dynamics of the discrete Chaplygin sleigh. *Discrete Contin. Dyn. Syst.*, (suppl.):258–267, 2005.

[60] R. L. Fernandes. Lie algebroids, holonomy and characteristic classes. *Adv. Math.*, 170(1):119–179, 2002.

[61] O. E. Fernandez and A. M. Bloch. Equivalence of the dynamics of nonholonomic and variational nonholonomic systems for certain initial data. *J. Phys. A*, 41(34):344005, 20, 2008.

[62] O. E. Fernandez, A. M. Bloch, and P. J. Olver. Variational integrators from Hamiltonizable nonholonomic systems. *J. Geom. Mech.*, 4(2):137–163, 2012.

[63] S. Ferraro, D. Iglesias, and D. Martín de Diego. Momentum and energy preserving integrators for nonholonomic dynamics. *Nonlinearity*, 21(8):1911–1928, 2008.

[64] P. L. García. The Poincaré-Cartan invariant in the calculus of variations. In *Symposia Mathematica, Vol. XIV (Convegno di Geometria Simplettica e Fisica Matematica, INDAM, Rome, 1973)*, pages 219–246. Academic Press, London, 1974.

[65] L. García-Naranjo. Reduction of almost Poisson brackets and Hamiltonization of the Chaplygin sphere. *Discrete Contin. Dyn. Syst. Ser. S*, 3(1):37–60, 2010.

[66] I. M. Gelfand and S. V. Fomin. *Calculus of variations.* Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[67] G. H. Golub and C. F. Van Loan. *Matrix computations.* The John Hopkins University Press, Baltimore, Maryland and London, England, third edition, 1996.

[68] O. Gonzalez. Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.*, 6(5):449–467, 1996.

[69] O. Gonzalez. Mechanical systems subject to holonomic constraints: Differential-algebraic formulations and conservative integration. *Physica D*, 132(12):165 – 174, 1999.

[70] M. J. Gotay, J. Isenberg, J. E. Marsden, and R. Montgomery. Momentum maps and classical relativistic fields. Part I: Covariant field theory. arXiv:physics/9801019v2 [math-ph].

[71] K. Grabowska, J. Grabowski, and P. Urbański. AV-differential geometry: Poisson and Jacobi structures. *J. Geom. Phys.*, 52(4):398–446, 2004.

[72] J. Grabowski, M. de León, J. C. Marrero, and D. Martín de Diego. Nonholonomic constraints: a new viewpoint. *J. Math. Phys.*, 50(1):013520, 17, 2009.

[73] J. Grabowski and P. Urbański. Tangent lifts of Poisson and related structures. *J. Phys. A*, 28(23):6743–6777, 1995.

[74] J. Grifone and Z. Muzsnay. Sur le problème inverse du calcul des variations: existence de lagrangiens associés à un spray dans le cas isotrope. *Ann. Inst. Fourier (Grenoble)*, 49(4):1387–1421, 1999.

[75] S. Grillo, J. Marsden, and S. Nair. Lyapunov constraints and global asymptotic stabilization. *J. Geom. Mech.*, 3(2):145–196, 2011.

[76] V. Guillemin and S. Sternberg. *Geometric asymptotics*. American Mathematical Society, Providence, R.I., 1977. Mathematical Surveys, No. 14.

[77] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.

[78] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.*, 25(1):35–61, 1983.

[79] P. Hartman. *Ordinary differential equations*, volume 38 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Corrected reprint of the second (1982) edition [Birkhäuser, Boston, MA; MR0658490 (83e:34002)], With a foreword by Peter Bates.

[80] M. Henneaux. On the inverse problem of the calculus of variations. *J. Phys. A*, 15(3):L93–L96, 1982.

[81] M. Henneaux. On the inverse problem of the calculus of variations in field theory. *J. Phys. A*, 17(1):75–85, 1984.

[82] M. Henneaux and C. Teitelboim. *Quantization of gauge systems*. Princeton University Press, Princeton, NJ, 1992.

[83] P. J. Higgins and K. C. H. Mackenzie. Algebraic constructions in the category of Lie algebroids. *J. Algebra*, 129(1):194–230, 1990.

[84] A. Hirsch. Die Existenzbedingungen des verallgemeinerten kinetischen Potentials. *Math. Ann.*, 50(2-3):429–441, 1898.

[85] D. D. Holm. *Geometric mechanics. Part I.* Imperial College Press, London, second edition, 2011. Dynamics and symmetry.

[86] D. D. Holm. *Geometric mechanics. Part II. Rotating, translating and rolling.* Imperial College Press, London, second edition, 2011.

[87] D. D. Holm, T. Schmah, and C. Stoica. *Geometric mechanics and symmetry*, volume 12 of *Oxford Texts in Applied and Engineering Mathematics.* Oxford University Press, Oxford, 2009. From finite to infinite dimensions, With solutions to selected exercises by David C. P. Ellis.

[88] P. E. Hydon and E. L. Mansfield. A variational complex for difference equations. *Found. Comput. Math.*, 4(2):187–217, 2004.

[89] L. A. Ibort and J. Marín-Solano. On the inverse problem of the calculus of variations for a class of coupled dynamical systems. *Inverse Problems*, 7(5):713–725, 1991.

[90] D. Iglesias, J. C. Marrero, D. Martín de Diego, E. Martínez, and Edith Padrón. Reduction of symplectic Lie algebroids by a Lie subalgebroid and a symmetry Lie group. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 3:049 (28 pages), 2007.

[91] D. Iglesias-Ponte, J. C. Marrero, D. Martín de Diego, and E. Padrón. Discrete dynamics in implicit form. *Discrete Contin. Dyn. Syst.*, 33(3):1117–1135, 2013.

[92] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta numerica, 2000*, volume 9 of *Acta Numer.*, pages 215–365. Cambridge Univ. Press, Cambridge, 2000.

[93] T. Itoh and K. Abe. Hamiltonian-conserving discrete canonical equations based on variational difference quotients. *J. Comput. Phys.*, 76(1):85–102, 1988.

[94] L. O. Jay. Solution of index 2 implicit differential-algebraic equations by Lobatto Runge-Kutta methods. *BIT*, 43(1):93–106, 2003.

[95] V. Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1997.

[96] H. K. Khalil. *Nonlinear systems.* Macmillan Publishing Company, New York, 1992.

[97] J. Klein. Espaces variationnels et mécanique. *Ann. Inst. Fourier (Grenoble)*, 12:1–124, 1962.

[98] M. Kobilarov, D. Martín de Diego, and S. Ferraro. Simulating nonholonomic dynamics. *Bol. Soc. Esp. Mat. Apl. S$\vec{e}$MA*, (50):61–81, 2010.

[99] D. Krupka. Global variational theory in fibred spaces. In *Handbook of global analysis*, pages 773–836, 1215. Elsevier Sci. B. V., Amsterdam, 2008.

[100] D. Krupka. *The Inverse Problem of the Calculus of Variations, An Introduction.* LRI Monograph Series, No. 2, 2013.

[101] O. Krupková and G. E. Prince. Second order ordinary differential equations in jet bundles and the inverse problem of the calculus of variations. In *Handbook of global analysis*, pages 837–904, 1215. Elsevier Sci. B. V., Amsterdam, 2008.

[102] A. D. Lewis and R. M. Murray. Variational principles for constrained systems: theory and experiment. *Internat. J. Non-Linear Mech.*, 30(6):793–815, 1995.

[103] P. Libermann. Lie algebroids and mechanics. *Arch. Math. (Brno)*, 32(3):147–162, 1996.

[104] P. Libermann and C.-M. Marle. *Symplectic geometry and analytical mechanics*, volume 35 of *Mathematics and its Applications.* D. Reidel Publishing Co., Dordrecht, 1987. Translated from the French by Bertram Eugene Schwarzbach.

[105] C. M. Campos. *Geometric Methods in Classical Field Theory and Continuous Media.* 2010.

[106] K. C. H. Mackenzie. *General theory of Lie groupoids and Lie algebroids*, volume 213 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 2005.

[107] G. Marmo and C. Rubano. Equivalent Lagrangians and Lax representations. *Nuovo Cimento B (11)*, 78(1):70–84, 1983.

[108] G. Marmo and E. J. Saletan. Ambiguities in the Lagrangian and Hamiltonian formalism: transformation properties. *Nuovo Cimento B (11)*, 40(1):67–89, 1977.

[109] J. C. Marrero, D. Martín de Diego, and E. Martínez. Discrete Lagrangian and Hamiltonian mechanics on Lie groupoids. *Nonlinearity*, 19(6):1313–1348, 2006.

[110] J. C. Marrero, D. Martín de Diego, and E. Martínez. On the exact discrete lagrangian function for variational integrators: theory and applications. arXiv:1608.01586, 2016.

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

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

[113] E. Martínez. Lagrangian mechanics on Lie algebroids. *Acta Appl. Math.*, 67(3):295–320, 2001.

[114] E. Martínez, J. F. Cariñena, and W. Sarlet. Derivations of differential forms along the tangent bundle projection. *Differential Geom. Appl.*, 2(1):17–43, 1992.

[115] E. Martínez, J. F. Cariñena, and W. Sarlet. Geometric characterization of separable second-order differential equations. *Math. Proc. Cambridge Philos. Soc.*, 113(1):205–224, 1993.

[116] A. Mayer. Die existenzbedingungen eines kinetishen potentiales. *Berich. Verh. Konig. Sach. Gesell. Wissen. Leipzig, Math. Phys. Kl.*, 84:519–529, 1896.

[117] R. McLachlan and M. Perlmutter. Integrators for nonholonomic mechanical systems. *J. Nonlinear Sci.*, 16(4):283–328, 2006.

[118] R. I. McLachlan and G. R. W. Quispel. Geometric integrators for ODEs. *J. Phys. A*, 39(19):5251–5285, 2006.

[119] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357(1754):1021–1045, 1999.

[120] T. Mestdag and M. Crampin. Nonholonomic systems as restricted Euler-Lagrange systems. *Balkan J. Geom. Appl.*, 15(2):78–89, 2010.

[121] T. Mestdag, W. Sarlet, and M. Crampin. The inverse problem for Lagrangian systems with certain non-conservative forces. *Differential Geom. Appl.*, 29(1):55–72, 2011.

[122] K. Modin and O. Verdier. Integrability of nonholonomically coupled oscillators. *Discrete Contin. Dyn. Syst.*, 34:1121–1130, 2014.

[123] B. L. Moiseiwitsch. *Variational principles.* Interscience Monographs and Texts in Physics and Astronomy, Vol. XX. Interscience Publishers John Wiley & Sons, London-New York-Sydney, 1966.

[124] G. Morandi, C. Ferrario, G. Lo Vecchio, G. Marmo, and C. Rubano. The inverse problem in the calculus of variations and the geometry of the tangent bundle. *Phys. Rep.*, 188(3-4):147–284, 1990.

[125] R. M. Murray and S. S. Sastry. Nonholonomic motion planning: steering using sinusoids. *IEEE Trans. Automat. Control*, 38(5):700–716, 1993.

[126] J. I. Neimark and N. A. Fufaev. *Dynamics of nonholonomic systems.* Translations of the American Mathematical Society, Providence, Rhode Island, 1972.

[127] H. Nijmeijer and A. van der Schaft. *Nonlinear dynamical control systems.* Springer-Verlag, New York, 1990.

[128] R. A. Norton, D. I. McLaren, G. R. W. Quispel, A. Stern, and A. Zanna. Projection methods and discrete gradient methods for preserving first integrals of ODEs. *Discrete Contin. Dyn. Syst.*, 35(5):2079–2098, 2015.

[129] G. W. Patrick and C. Cuell. Error analysis of variational integrators of unconstrained Lagrangian systems. *Numer. Math.*, 113(2):243–264, 2009.

[130] Y. Pinchover and J. Rubinstein. *An introduction to partial differential equations.* Cambridge University Press, Cambridge, 2005.

[131] L. Popescu. Metric nonlinear connections on Lie algebroids. *Balkan J. Geom. Appl.*, 16(1):111–121, 2011.

[132] G. R. W. Quispel and H. W. Capel. Solving ODEs numerically while preserving a first integral. *Phys. Lett. A*, 218(3-6):223–228, 1996.

[133] G. R. W. Quispel and G. S. Turner. Discrete gradient methods for solving ODEs numerically while preserving a first integral. *J. Phys. A*, 29(13):L341–L349, 1996.

[134] M. F. Rañada. Time-dependent Lagrangian systems: a geometric approach using semibasic forms. *J. Phys. A*, 23(15):3475–3482, 1990.

[135] O. Rossi and J. Musilová. On the inverse variational problem in nonholonomic mechanics. *Commun. Math.*, 20(1):41–62, 2012.

[136] S. V. Sabau. Some remarks on jacobi stability. *Nonlinear Analysis*, 63:e143–e153, 2005.

[137] H. Sagan. *Introduction to the calculus of variations.* Dover Publications, Inc., New York, 1992. Corrected reprint of the 1969 original.

[138] R. M. Santilli. *Foundations of theoretical mechanics. I.* Springer-Verlag New York, New York-Heidelberg, 1978. The inverse problem in Newtonian mechanics, Texts and Monographs in Physics.

[139] J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian problems*, volume 7 of *Applied Mathematics and Mathematical Computation.* Chapman & Hall, London, 1994.

[140] W. Sarlet. The Helmholtz conditions revisited. A new approach to the inverse problem of Lagrangian dynamics. *J. Phys. A*, 15(5):1503–1517, 1982.

[141] W. Sarlet. Geometric calculus for second-order differential equations and generalizations of the inverse problem of lagrangian mechanics. *International Journal of Non-Linear Mechanics*, 47:1132–1140, 2012.

[142] W. Sarlet, F. Cantrijn, and E. Martínez. Application of the theory of separability of second-order differential equations. *J. Nonlinear Math. Phys.*, 1(1):5–24, 1994.

[143] W. Sarlet, M. Crampin, and E. Martínez. The integrability conditions in the inverse problem of the calculus of variations for second-order ordinary differential equations. *Acta Appl. Math.*, 54(3):233–273, 1998.

[144] W. Sarlet, G. Thompson, and G. E. Prince. The inverse problem of the calculus of variations: the use of geometrical calculus in Douglas's analysis. *Trans. Amer. Math. Soc.*, 354(7):2897–2919 (electronic), 2002.

[145] W. Sarlet, A. Vandecasteele, F. Cantrijn, and E. Martínez. Derivations of forms along a map: the framework for time-dependent second-order equations. *Differential Geom. Appl.*, 5(2):171–203, 1995.

[146] D. J. Saunders. *The geometry of jet bundles*, volume 142 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1989.

[147] D. J. Saunders. Thirty years of the inverse problem in the calculus of variations. *Rep. Math. Phys.*, 66(1):43–53, 2010.

[148] N. J. Sonin. About determining maximal and minimal properties of plane curves. *Warsawskye Universitetskye Izvestiya*, 1-2:1–68, 1886.

[149] E. D. Sontag. *Mathematical control theory*, volume 6 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1990. Deterministic finite-dimensional systems.

[150] M. W. Spong, P. Corke, and R. Lozano. Nonlinear control of the inertia wheel pendulum. *Automatica*, 37:1845–1851, 1999.

[151] H. J. Sussmann. Geometry and optimal control. In *Mathematical control theory*, pages 140–198. Springer, New York, 1999.

[152] F. Takens. A global version of the inverse problem of the calculus of variations. *J. Differential Geom.*, 14(4):543–562 (1981), 1979.

[153] W. M. Tulczyjew. Les sous-variétés lagrangiennes et la dynamique hamiltonienne. *C. R. Acad. Sci. Paris Sér. A-B*, 283(1):Ai, A15–A18, 1976.

[154] W. M. Tulczyjew. Les sous-variétés lagrangiennes et la dynamique lagrangienne. *C. R. Acad. Sci. Paris Sér. A-B*, 283(8):Av, A675–A678, 1976.

[155] K. Uchino. Lagrangian calculus on Dirac manifolds. *J. Math. Soc. Japan*, 57(3):803–825, 2005.

[156] I. Vaisman. *Symplectic geometry and secondary characteristic classes*, volume 72 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1987.

[157] I. Vaisman. *Lectures on symplectic and Poisson geometry*. Textos de Matemática. Série B [Texts in Mathematics. Series B], 23. Universidade de Coimbra Departamento de Matemática, Coimbra, 2000. Notes edited by Fani Petalidou.

[158] A. van der Schaft. Port-Hamiltonian systems: an introductory survey. In *International Congress of Mathematicians. Vol. III*, pages 1339–1365. Eur. Math. Soc., Zürich, 2006.

[159] A. J. van der Schaft and B. M. Maschke. On the Hamiltonian formulation of nonholonomic mechanical systems. *Rep. Math. Phys.*, 34(2):225–233, 1994.

[160] R. Vitolo. Variational sequences. In *Handbook of global analysis*, pages 1115–1163, 1217. Elsevier Sci. B. V., Amsterdam, 2008.

[161] H. von Helmholtz. Ueber die physikalische bedeutung des prinicips der kleinsten wirkung. *Journal für die reine und angewandte Mathematik*, 100:137–166, 1887.

[162] A. Weinstein. Symplectic manifolds and their Lagrangian submanifolds. *Advances in Math.*, 6:329–346, 1971.

[163] A. Weinstein. Symplectic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 5(1):1–13, 1981.

[164] A. Weinstein. The local structure of Poisson manifolds. *J. Differential Geom.*, 18(3):523–557, 1983.

[165] A. Weinstein. Lagrangian mechanics and groupoids. In *Mechanics day (Waterloo, ON, 1992)*, volume 7 of *Fields Inst. Commun.*, pages 207–231. Amer. Math. Soc., Providence, RI, 1996.

[166] S. Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 2 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2003.

[167] K. Yano and S. Ishihara. *Tangent and cotangent bundles: differential geometry*. Marcel Dekker, Inc., New York, 1973. Pure and Applied Mathematics, No. 16.