Energy-Based Modeling and Tracking Control of Rigid Body Systems with Practical Multicopter Applications

Dissertation

zur Erlangung des Grades

des Doktors der Ingenieurwissenschaften
(Dr.-Ing.)

der Naturwissenschaftlich-Technischen Fakultät II
– Physik und Mechatronik –
der Universität des Saarlandes

von

Matthias Konz

Saarbrücken

2021

Kurzfassung

Der theoretische Teil dieser Arbeit befasst sich mit der Modellierung und Folgeregelung von Starrkörpersystemen. Ein Starrkörpersystem besteht aus miteinander verkoppelten starren Körpern, welche wiederum als Menge von speziell verkoppelten Massepunkten angesehen werden können.

Das erste Kapitel diskutiert allgemeine Systeme von verkoppelten Massepunkten. Ziel ist die Herleitung von Energien und Bewegungsgelichungen und deren Zusammenhang. Im Kontrast zum Großteil der Literator wird hier eine recht allgemeine Parametrierung in Positions- und Geschwindigkeitskoordinaten betrachtet.

Im zweiten Kapitel wird diese Theorie auf die spezielleren Fälle eines einzelnen freien Starrkörperes und allgemeiner Starrkörpersysteme angewendet. Als Resultat ergeben sich in kompakte Formulierungen für die Starrkörperenergien als auch für die Bewegungsgleichungen, welche in dieser Form wohl noch nicht in der Literatur etabliert sind.

Im dritten Kapitel wird ein Ansatz zur Folgeregelung von Starrköpersystemen durch statische Zustandsrückführung präsentiert. Zunächst wird ein, basierend auf den Erkentnissen der Modellierung, allgemeines, aber geometrisch passendes Wunschverhalten definiert. Das Regelgesetz ergibt sich aus der Minimierung der Differenz von realisierbarer Beschleunigung und Wunschbeschleunigung. Die Performance dieses Ansatzes wird an meheren Beispielen und Simulationsergebnissen diskutiert.

Zur praktischen Evaluation des Reglerkonzepts werden die am LSR entwickelten Tri- und Quadcopter verwendet. Die Performance des realisierten Regelkonzepts wird an meherern aerobatischen Beispielmaneuvern demonstriert.

Abstract

Many machines, vehicles and robots may be modeled as rigid body systems, i.e. a number of interconnected, undeformable bodies subject to inertia, gravity, and other forces. Energy-based methods for derivation of their equations of motion, like the Lagrange formalism, are standard in engineering education and well established in the dedicated literature. These algorithms commonly rely on the use of a minimal set of generalized coordinates. This is appropriate for many applications, e.g. machines containing only one-dimensional joints. For systems whose configuration space is nonlinear, e.g. mobile robots whose config. space contains the rigid body attitude, the use of minimal coordinates necessarily leads to singularities. From the point of view of differential geometry, this is a well known fact.

This work resolves this problem by the use of (possibly) redundant config. coordinates and (possibly) nonholonomic velocity coordinates. The second chapter reviews several established formalisms of analytical mechanics and states them in terms these more general coordinates. The third chapter applies these results to rigid body systems. Though inertia is the crucial part of the dynamics, this work also investigates dissipation and stiffness. Finally, it presents an algorithm for the derivation of global equations of motion of general rigid body systems.

The literature states computed-torque is a standard approach for tracking control of fully actuated mechanical systems. However, this recipe relies on minimal coordinates and consequently suffers from the problems mentioned above. There is no established "standard" approach for the control of underactuated systems.

This work presents three sightly different algorithms for for tracking control of general rigid body systems by means of static state feedback. These essentially minimize the distance between the actual realizable acceleration of the model and a desired acceleration computed from a stable prototype system. The prototype system shares the geometry and kinematics of the actual model, but may have different constitutive properties (inertia, damping, stiffness). The resulting control law can be computed globally and explicitly for any rigid body system. The resulting closed loop system (which may differ from the prototype in the underactuated case), is invariant to the chosen coordinates, i.e. its formulation is covariant. However, so far, there is no general proof of stability. The performance of the proposed approaches are discussed on several examples and simulation results.

The last chapter of this work discusses the experimental realization of the control approach to two small UAVs. The performance is demonstrated on tracking control for several aerobatic maneuvers.

Contents

1 Introduction **2**

2 Calculus with redundant coordinates **4**

2.0.1 Redundant configuration coordinates . . . 4
2.0.2 Minimal velocity coordinates . . . 5
2.0.3 Directional derivative and Hessian . . . 7
2.0.4 Commutation coefficients . . . 11
2.0.5 Linearization about a trajectory . . . 12
2.0.6 Calculus of variations . . . 13

3 Analytical mechanics of particle systems **15**

3.1 System under consideration . . . 15
3.2 Motivation: A problem of the textbook approach . . . 16
3.3 Geometry and kinematics . . . 17
3.3.1 Redundant configuration coordinates . . . 18
3.3.2 Minimal velocity coordinates . . . 18
3.3.3 Directional derivative and Hessian . . . 21
3.3.4 Commutation coefficients . . . 24
3.3.5 Linearization about a trajectory . . . 26
3.3.6 Calculus of variations . . . 27
3.4 Principles of mechanics . . . 28
3.4.1 Principle of constraint release . . . 28
3.4.2 Lagrange-d'Alembert's principle . . . 29
3.4.3 Gauß' principle . . . 29
3.4.4 Hamilton's principle . . . 30

3.5 Kinetics . . . 31
3.5.1 Inertia . . . 31
3.5.2 Dissipation . . . 33
3.5.3 Linear springs . . . 33
3.5.4 Gravitation . . . 35
3.5.5 External forces . . . 35
3.6 Additional constraints . . . 35
3.7 Equations of motion and total energy . . . 37
3.8 Summary . . . 38

4 Rigid body systems **39**
4.1 A single free rigid body . . . 40
4.1.1 Blah . . . 40
4.1.2 Coordinates . . . 41
4.1.3 Inertia . . . 42
4.1.4 Gravitation . . . 44
4.1.5 Stiffness . . . 44
4.1.6 Dissipation . . . 48
4.1.7 External forces. . . . 48
4.2 A more abstract view on the rigid body . . . 48
4.2.1 The special Euclidean group . . . 48
4.2.2 Inner product . . . 50
4.2.3 Rigid body energies . . . 52
4.3 Rigid body systems . . . 53
4.3.1 Configuration . . . 53
4.3.2 Velocity . . . 57
4.3.3 Inertia . . . 59
4.3.4 Gravitation . . . 60
4.3.5 Stiffness . . . 61
4.3.6 Dissipation . . . 61
4.4 Summary . . . 62

5 Tracking control of rigid body systems **63**
5.1 Approach 1: Inspired by particle distribution . . . 65
5.1.1 Particle system . . . 65
5.1.2 Free rigid body . . . 66
5.1.3 Rigid body systems . . . 68
5.2 Approach 2: Body based approach . . . 71
5.2.1 Free rigid body . . . 71
5.2.2 Rigid body systems . . . 72
5.3 Approach 3: Inspired by total energy . . . 73
5.3.1 Overall structure . . . 73
5.3.2 Special cases . . . 78
5.3.3 Free rigid body . . . 79
5.3.4 Rigid body systems . . . 80
5.4 Constant reference and Linearization . . . 82
5.5 Underactuated systems . . . 83
5.5.1 Control law through static optimization . . . 83
5.5.2 Matching condition . . . 84
5.5.3 Approximations . . . 85
5.5.4 Systems with input constraints . . . 86
5.6 Summary and recipe . . . 87
5.7 Examples of fully actuated systems . . . 89
5.7.1 Prismatic joint . . . 89
5.7.2 Revolute joint . . . 91
5.7.3 Rigid body orientation . . . 94
5.7.4 Planar rigid body . . . 96
5.7.5 Decoupling of translational and rotational motion . . . 100
5.7.6 SCARA robot . . . 103
5.7.7 Robot arm . . . 108
5.8 Examples of underactuated systems . . . 111
5.8.1 Two masses connected by a spring . . . 111
5.8.2 PVTOL . . . 114

5.8.3 Quadcopter 118
5.8.4 Bicopter 121

A Some math **125**

A.1 Linear Algebra 125
A.1.1 Matrix sets 125
A.1.2 Trace 125
A.1.3 Vee & wedge operatos 126
A.1.4 Inner product 126
A.1.5 Matrix decompositions 128
A.1.6 Moore-Penrose pseudoinverse 129
A.1.7 Orthogonal projectors 129
A.2 Differential geometry 131
A.2.1 Embedded manifold 131
A.2.2 Vectors 132
A.2.3 Riemannian geometry 135
A.2.4 The induced metric 136
A.2.5 Riemannian geometry and Lagrangian mechanics. 137
A.2.6 Relation to Euclidean geometry. 138
A.3 Transformation rules 138
A.3.1 Recap coordinate expressions 138
A.3.2 Change of configuration coordinates 139
A.3.3 Change of velocity coordinates / basis 140
A.4 Static optimization 143
A.4.1 Nonlinear programming 143
A.4.2 Quadratic programming 145
A.4.3 Blah 146
A.4.4 Attitude potential 147
A.5 Attitude parameterizations 150
A.5.1 Axis-angle representation 150
A.5.2 Unit quaternion 151

A.5.3 Euler angles 152
A.5.4 Minimal rotation matrix 153
A.6 Constraint stabilization 154
A.7 Alternative from for the body Jacobian 155

B On nonholonomic systems 157
B.1 Hamilton's principle 161
B.2 On error coordinates 163
B.3 A possible generalization of the rigid body energies 164

C Templates 171
C.1 Math fonts 171

Chapter 1

Introduction

This work deals with the modeling and control of rigid body systems. Modeling here means to mathematically capture the behavior of a system. For rigid body systems this is a set of ordinary differential equations with time being the independent variable. Control here means the derivation of an algorithm, the control law, which in combination with the model equations, yields a stable system. In particular we are interested in tracking control, i.e. a controlled system that converges to a given (possibly accelerated) reference trajectory, in contrast to stabilization of a constant configuration. The rigid body is an important concept of classical dynamics: From an engineering or physics point of view, it could be regarded as a solid body whose deformation is negligible compared to its overall motion. A rigid body system is the collection of one or more rigid bodies that interact with each other and or the surrounding space.

Rigid body systems are omnipresent in mechanical engineering and robotics. Thus they are standard content for undergraduate courses and there are plenty of textbooks on the subject (e.g. [Goldstein, 1951] or [Spong et al., 2006]). The common approach therein is the Lagrange formalism which relies on a parameterization of the system in terms of a minimal set of generalized coordinates. A popular control approach is the so-called computed torque method which can be regarded as a simple case of feedback linearization. Application of these methods only requires basic calculus and linear algebra. At least from a theoretical point of view, these methods are well suited for many robotics applications. Common examples are robotic manipulators which involve n one-dimensional joints. Their configuration space is $\mathbb{R}^n$. One the other hand, for e.g. mobile robots, whose configuration space is usually not isomorphic to $\mathbb{R}^n$, the use of minimal coordinates will inevitably lead to singularities in model and control law. The problem is brought to the point in [Roberson and Schwertassek, 1988, sec. 1.1.1]:

> *[The scientists of the Eighteenth Century] recognized that there was something about rotation [...] which somehow made the analysis of rotation a problem of higher order difficulty. We now know that the problem is in the mathematics, not the physics, but the problem is still with us.*

The problem in the mathematics arises from the use of minimal coordinates. From the point of view of differential geometry, this corresponds to a single chart on the configuration space, and it is evident that this parameterization can only be local if the configuration space is nonlinear, i.e. not isomorphic to the Euclidean space $\mathbb{R}^n$ of same dimension.

Acknowledging this problem, the more evolved literature, like e.g. [Marsden and Ratiu, 1998] or [Bullo and Lewis, 2004], utilizes differential geometry and Lie group theory. However, since this requires a much tougher mathematical background, these methods are beyond the reach of many engineers. Furthermore, for an actual implementation of a simulator or a control law, we do need coordinates.

There are generally two possibilities for a global parameterization (e.g. for the rotation of a rigid body):

1. several sets of minimal coordinates on overlapping charts (e.g. 4 sets of Euler-angles)

2. one set of redundant coordinates that embed the configuration space in the Euclidean space (e.g. rotation matrix or unit quaternion)

Popular methods like the Lagrange formalism or computed torque require minimal coordinates. One major goals of this work is to review and extend these methods to also work with redundant coordinates.

particles

Not only inertia but also gravitation, stiffness, damping

This work is organized as follows: The second chapter establishes redundant configuration coordinates and nonholonomic velocity coordinates from a purely mathematical point of view. We will derive the necessary tools, like differential or calculus of variations, for the next chapters.

The second chapter reviews some principles of mechanics for systems of particles. It derives a formulation for their equations of motion.

The forth chapter deals with modeling of rigid body systems. Firstly, it discusses the important example of a single, free rigid body. Secondly, it develops a formulation for derivation of equations of motion for general systems of interconnected rigid bodies.

The fifth chapter deals with tracking control of rigid body systems. It proposes three recipes and applies these to several examples.

The sixth chapter shows the practical application of the proposed control approach for the example of a fully actuated tricopter and an underactuated quadcopter. It discusses the experimental results in comparison to others from the contemporary literature.

The last chapter summarizes this work, points out open problems and possible future extensions.

Chapter 2

Calculus with redundant coordinates

The particle configuration space $\mathfrak{X} = \{\mathfrak{x} \in \mathbb{R}^{3\mathfrak{N}} \,|\, \mathfrak{c}(\mathfrak{x}) = \mathbf{0}\}$ can be seen as an embedded differentiable manifold with $\dim \mathfrak{X} = 3\mathfrak{N} - \operatorname{rank} \frac{\partial \mathfrak{c}}{\partial \mathfrak{x}} = n$. We call the manifold *nonlinear*, if it is *not homeomorphic* to $\mathbb{R}^n$, i.e. there exists no global, continuous function $\mathbb{R}^n \to \mathfrak{X}$ that has a continuous inverse (a global homeomorphism). The common way to tackle this in differential geometry is to use an atlas, a set of overlapping *local* charts. For the previous Example **??**, this was done four different charts in [Grafarend and Kühnel, 2011]. In this work we will not pursue this path.

2.0.1 Redundant configuration coordinates

Another way for a global parameterization of nonlinear configuration manifolds is motivated from the *Whitney embedding theorem* (see e.g. [Lee, 2003, Theo. 6.14]), that states: *Every smooth manifold of dimension n can be smoothly embedded in the Euclidean space $\mathbb{R}^{2n}$.* Note that $2n$ is a worst case bound, i.e. for a particular example a lower dimension for the embedding space might work and a higher dimension is permitted anyway. In the notation of this work, we use $\nu > 0$ generalized coordinates $\boldsymbol{x}(t) = [x^1(t), \ldots, x^\nu(t)]^\top \in \mathbb{R}^\nu$ that might be constrained by $c \geq 0$ smooth functions of the form $\boldsymbol{\phi}(\boldsymbol{x}) = [\phi^1(\boldsymbol{x}), \ldots, \phi^c(\boldsymbol{x})]^\top = \mathbf{0}$. For $c > 0$ these coordinates are not independent and are commonly called *redundant*. The set of mutually admissible coordinates is

$$\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^\nu \,|\, \boldsymbol{\phi}(\boldsymbol{x}) = \mathbf{0}\} \tag{2.1}$$

with $\dim \mathbb{X} = \nu - \operatorname{rank} \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} = n$. By the Whitney embedding theorem, for a suitable number of coordinates ν and suitable constraints $\boldsymbol{\phi}(\boldsymbol{x}) = \mathbf{0}$, this manifolds is homeomorphic to the particle configuration space $\mathbb{X} \cong \mathfrak{X}$. This means there is a global, invertible function $\mathfrak{x}(\boldsymbol{x})$ from the coordinates to the particle positions.

2.0.2 Minimal velocity coordinates

For the following it is crucial to note that a geometric constraint is equivalent to its derivative supplemented with a suitable initial condition

$$\phi^\kappa(\boldsymbol{x}) = 0 \tag{2.2a}$$

$$\Leftrightarrow \qquad \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x})\dot{x}^\alpha = 0, \qquad \phi^\kappa(\boldsymbol{x}_0) = 0 \tag{2.2b}$$

$$\Leftrightarrow \qquad \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x})\ddot{x}^\alpha + \frac{\partial^2 \phi^\kappa}{\partial x^\alpha \partial x^\beta}(\boldsymbol{x})\dot{x}^\beta\dot{x}^\alpha = 0, \qquad \phi^\kappa(\boldsymbol{x}_0) = 0,\ \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x}_0)\dot{x}^\alpha_0 = 0 \tag{2.2c}$$

$$\dots$$

where $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$. Even though (3.13a) might be nonlinear, its derivative (3.13b) is always *linear* in the velocities $\dot{\boldsymbol{x}}$. So here it is reasonable to choose *minimal velocity coordinates*: Let $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu \times n}$ be a matrix with the properties $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\boldsymbol{A} = \boldsymbol{0}$ and $\operatorname{rank} \boldsymbol{A} = n$. The first property of $\boldsymbol{A}(\boldsymbol{x})$ is that these columns of $\boldsymbol{A}(\boldsymbol{x})$ are orthogonal to the rows of $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}(\boldsymbol{x})$. The second property implies that the columns of $\boldsymbol{A}(\boldsymbol{x})$ are linearly independent. So the columns of $\boldsymbol{A}(\boldsymbol{x})$ can be interpreted as a *basis vectors* for the tangent space $\mathsf{T}_{\boldsymbol{x}}\mathbb{X}$. We can capture all allowed velocities $\dot{\boldsymbol{x}}(t)$ by the minimal velocity coordinates $\boldsymbol{\xi}(t) \in \mathbb{R}^n$ through

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi} \tag{2.3}$$

This kinematic relation (3.14) ensures that the time derivative (3.13b) of the geometric constraint is fulfilled, and consequently the geometric constraint only has to be imposed on the initial condition $\boldsymbol{\phi}(\boldsymbol{x}(t_0)) = \boldsymbol{0}$.

Existence for A? guaranteed for Lie groups: the Lie algebra at the identity can be translated around the manifold by the group operation

Construction of $\boldsymbol{A}$ by matrix inversion

Example 1. Consider a single particle constrained to a circle of radius ρ as illustrated in Figure 3.1.

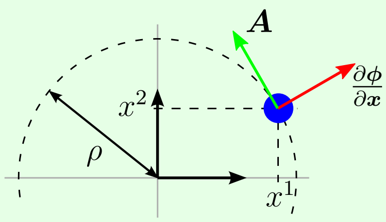


Figure 2.1: Particle on a circle

We use the its Cartesian position $[x^1, x^2]^\top \in \mathbb{R}^2$ constrained by $\phi = (x^1)^2 + (x^2)^2 - \rho^2 = 0$ as configuration coordinates. A reasonable choice for the kinematics matrix $\boldsymbol{A}$ is motivated from

$$\underbrace{\begin{bmatrix} 2x^1 & 2x^2 \end{bmatrix}}_{\frac{\partial \phi}{\partial \boldsymbol{x}}} \underbrace{\begin{bmatrix} -x^2 \\ x^1 \end{bmatrix}}_{\boldsymbol{A}} = 0 \tag{2.4}$$

Example 2. Consider again the system from Example ??. Instead of parameterizing the rotation matrix $\boldsymbol{R}$ by minimal coordinates we now take its 9 coefficients $\boldsymbol{x} = [R^{\mathsf{x}}_{\mathsf{x}}, R^{\mathsf{y}}_{\mathsf{x}}, R^{\mathsf{z}}_{\mathsf{x}}, R^{\mathsf{x}}_{\mathsf{y}}, R^{\mathsf{y}}_{\mathsf{y}}, R^{\mathsf{z}}_{\mathsf{y}}, R^{\mathsf{x}}_{\mathsf{z}}, R^{\mathsf{y}}_{\mathsf{z}}, R^{\mathsf{z}}_{\mathsf{z}}]^\top \in \mathbb{R}^9$ as configuration coordinates. The constraints $\boldsymbol{R}^\top \boldsymbol{R} = \mathbf{I}_3$ and $\det \boldsymbol{R} = 1$ read

$$\boldsymbol{\phi}(\boldsymbol{x}) = \begin{bmatrix} (R^{\mathsf{x}}_{\mathsf{x}})^2 + (R^{\mathsf{y}}_{\mathsf{x}})^2 + (R^{\mathsf{z}}_{\mathsf{x}})^2 - 1 \\ (R^{\mathsf{x}}_{\mathsf{y}})^2 + (R^{\mathsf{y}}_{\mathsf{y}})^2 + (R^{\mathsf{z}}_{\mathsf{y}})^2 - 1 \\ (R^{\mathsf{x}}_{\mathsf{z}})^2 + (R^{\mathsf{y}}_{\mathsf{z}})^2 + (R^{\mathsf{z}}_{\mathsf{z}})^2 - 1 \\ R^{\mathsf{x}}_{\mathsf{y}} R^{\mathsf{x}}_{\mathsf{z}} + R^{\mathsf{y}}_{\mathsf{y}} R^{\mathsf{y}}_{\mathsf{z}} + R^{\mathsf{z}}_{\mathsf{y}} R^{\mathsf{z}}_{\mathsf{z}} \\ R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{x}}_{\mathsf{z}} + R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{z}} + R^{\mathsf{z}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{z}} \\ R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{x}}_{\mathsf{y}} + R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{y}} + R^{\mathsf{z}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{y}} \\ R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{y}} R^{\mathsf{z}}_{\mathsf{z}} + R^{\mathsf{x}}_{\mathsf{y}} R^{\mathsf{y}}_{\mathsf{z}} R^{\mathsf{z}}_{\mathsf{x}} + R^{\mathsf{x}}_{\mathsf{z}} R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{y}} - R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{z}} R^{\mathsf{z}}_{\mathsf{y}} - R^{\mathsf{x}}_{\mathsf{y}} R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{z}} - R^{\mathsf{x}}_{\mathsf{z}} R^{\mathsf{y}}_{\mathsf{y}} R^{\mathsf{z}}_{\mathsf{x}} - 1 \end{bmatrix} = \mathbf{0}. \tag{2.5}$$

The 9 conditions $\boldsymbol{R}^\top \boldsymbol{R} = \mathbf{I}_3$ yields due to symmetry only 6 constraints and already imply $\det \boldsymbol{R} = \pm 1$. Since the determinant is a smooth function, the corresponding manifold must consist of two disjoint components, one with $\det \boldsymbol{R} = +1$ (proper rotations) and one with $\det \boldsymbol{R} = -1$ (rotations with reflection). So the additional constraint $\det \boldsymbol{R} = +1$ does not change the dimension of the configuration space. Formally this means $\operatorname{rank} \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} = 6$ and consequently $\dim \mathbb{X} = 9 - 6 = 3$. A kinematics matrix with $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} \boldsymbol{A} = \mathbf{0}$ and $\operatorname{rank} \boldsymbol{A} = 3$ is given by

$$\boldsymbol{A}(\boldsymbol{x}) = \begin{bmatrix} 0 & -R^{\mathsf{x}}_{\mathsf{z}} & R^{\mathsf{x}}_{\mathsf{y}} \\ 0 & -R^{\mathsf{y}}_{\mathsf{z}} & R^{\mathsf{y}}_{\mathsf{y}} \\ 0 & -R^{\mathsf{z}}_{\mathsf{z}} & R^{\mathsf{z}}_{\mathsf{y}} \\ R^{\mathsf{x}}_{\mathsf{z}} & 0 & -R^{\mathsf{x}}_{\mathsf{x}} \\ R^{\mathsf{y}}_{\mathsf{z}} & 0 & -R^{\mathsf{y}}_{\mathsf{x}} \\ R^{\mathsf{z}}_{\mathsf{z}} & 0 & -R^{\mathsf{z}}_{\mathsf{x}} \\ -R^{\mathsf{x}}_{\mathsf{y}} & R^{\mathsf{x}}_{\mathsf{x}} & 0 \\ -R^{\mathsf{y}}_{\mathsf{y}} & R^{\mathsf{y}}_{\mathsf{x}} & 0 \\ -R^{\mathsf{z}}_{\mathsf{y}} & R^{\mathsf{z}}_{\mathsf{x}} & 0 \end{bmatrix}. \tag{2.6}$$

The resulting kinematic equation $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}$ can be reordered to the matrix equation $\dot{\boldsymbol{R}} = \boldsymbol{R} \operatorname{wed}(\boldsymbol{\xi})$ by introducing the *wedge operator* defined as

$$\operatorname{wed} \begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} = \begin{bmatrix} 0 & -\xi^3 & \xi^2 \\ \xi^3 & 0 & -\xi^1 \\ -\xi^2 & \xi^1 & 0 \end{bmatrix}. \tag{2.7}$$

Some identities involving the pseudo-inverse. For any matrix $\boldsymbol{S} \in \mathbb{R}^{m \times n}$ there exists a unique *(Moore-Penrose) pseudoinverse* $\boldsymbol{S}^+ \in \mathbb{R}^{n \times m}$ determined by the following conditions [Penrose, 1955, Theo. 1]:

$$\boldsymbol{S}\boldsymbol{S}^{+}\boldsymbol{S} = \boldsymbol{S}, \tag{2.8a}$$

$$\boldsymbol{S}^{+}\boldsymbol{S}\boldsymbol{S}^{+} = \boldsymbol{S}^{+}, \tag{2.8b}$$

$$(\boldsymbol{S}\boldsymbol{S}^{+})^{\top} = \boldsymbol{S}\boldsymbol{S}^{+}, \tag{2.8c}$$

$$(\boldsymbol{S}^{+}\boldsymbol{S})^{\top} = \boldsymbol{S}^{+}\boldsymbol{S}. \tag{2.8d}$$

If the matrix $\boldsymbol{S}$ has linearly independent columns, its pseudoinverse is $\boldsymbol{S}^{+} = (\boldsymbol{S}^{\top}\boldsymbol{S})^{-1}\boldsymbol{S}^{\top}$. Similarly, if $\boldsymbol{S}$ has linearly independent rows, its pseudoinverse is $\boldsymbol{S}^{+} = \boldsymbol{S}^{\top}(\boldsymbol{S}\boldsymbol{S}^{\top})^{-1}$.

Define $\boldsymbol{Y}(\boldsymbol{x}) \in \mathbb{R}^{n\times\nu}$ as $\boldsymbol{Y} = \boldsymbol{A}^{+} = (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}$, i.e. the pseudoinverse of the kinematics matrix $\boldsymbol{A}$. Note that this implies $\boldsymbol{Y}\boldsymbol{A} = \mathbf{I}_n$, but $\boldsymbol{A}\boldsymbol{Y} \neq \mathbf{I}_\nu$. We also introduce the matrices $\boldsymbol{\Phi} = \frac{\partial \phi}{\partial \boldsymbol{x}}$ and $\boldsymbol{\Psi} = \boldsymbol{\Phi}^{+}$. With $\boldsymbol{\Phi}\boldsymbol{A} = \mathbf{0}$ and the Penrose conditions (3.19), we can show[1] that $\boldsymbol{\Psi}^{\top}\boldsymbol{A} = \mathbf{0}$ and $\boldsymbol{Y}^{\top}\boldsymbol{\Phi} = \mathbf{0}$. Furthermore, since $\operatorname{rank}\boldsymbol{\Psi} = \operatorname{rank}\boldsymbol{\Phi} = \nu - n$ the columns of $\boldsymbol{\Psi}(\boldsymbol{x})$ span the complementary space $(\mathsf{T}_{\boldsymbol{x}}\mathbb{X})^{\perp}$ though they might not be a basis since the columns might not be linearly independent.

The matrix $\boldsymbol{P} = \boldsymbol{A}\boldsymbol{Y}$ is an *orthogonal projector*, i.e. $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\boldsymbol{P}^{\top} = \boldsymbol{P}$ which result directly from the Penrose conditions (3.19). It is in fact the unique orthogonal projector from $\mathbb{R}^{\nu}$ to the tangent space $\mathsf{T}_{\boldsymbol{x}}\mathbb{X}$. The unique orthogonal projector to the complementary space $(\mathsf{T}_{\boldsymbol{x}}\mathbb{X})^{\perp}$ is $\boldsymbol{P}^{\perp} = \mathbf{I}_\nu - \boldsymbol{P}$. On the other hand, since $\boldsymbol{\Psi}$ spans the complementary space the complementary projector can also be expressed as $\boldsymbol{P}^{\perp} = \boldsymbol{\Psi}\boldsymbol{\Phi}$. This leads to the identity

$$\underbrace{\boldsymbol{A}\boldsymbol{Y}}_{\boldsymbol{P}} + \underbrace{\boldsymbol{\Psi}\boldsymbol{\Phi}}_{\boldsymbol{P}^{\perp}} = \mathbf{I}_\nu. \tag{2.9}$$

2.0.3 Directional derivative and Hessian

Consider a function $\mathcal{V} : \mathbb{R}^{\nu} \to \mathbb{R}$ and a curve $\boldsymbol{x} : \mathbb{R} \to \mathbb{X}$. Since $\mathbb{X} \subset \mathbb{R}^{\nu}$, their composition $\mathcal{V} \circ \boldsymbol{x} = f : \mathbb{R} \to \mathbb{R}$ is a scalar function and has the the Taylor expansion

$$\begin{aligned}\underbrace{\mathcal{V}(\boldsymbol{x}(t))}_{f(t)} = \underbrace{\mathcal{V}(\boldsymbol{x}(0))}_{f(0)} + t\,\underbrace{\frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\boldsymbol{x}(0))\dot{x}^{\alpha}(0)}_{\dot{f}(0)} \\ + \tfrac{1}{2}t^2\underbrace{\left(\frac{\partial^2 \mathcal{V}}{\partial x^{\beta}\partial x^{\alpha}}(\boldsymbol{x}(0))\dot{x}^{\alpha}(0)\dot{x}^{\beta}(0) + \frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\boldsymbol{x}(0))\ddot{x}^{\alpha}(0)\right)}_{\ddot{f}(0)} + \mathcal{O}(t^3).\end{aligned} \tag{2.10}$$

Now let the curve be parameterized by $\dot{\boldsymbol{x}}(t) = \boldsymbol{A}(\boldsymbol{x}(t))\boldsymbol{\xi}(t)$ and we use the shorthand notations $\bar{\boldsymbol{x}} = \boldsymbol{x}(0)$, $\bar{\boldsymbol{\xi}} = \boldsymbol{\xi}(0)$ and $\bar{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{x}(0))$ to write

$$\begin{aligned}\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + t\frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\bar{\boldsymbol{x}})\bar{A}^{\alpha}_{i}\bar{\xi}^{i} \\ + \tfrac{1}{2}t^2\Big(\frac{\partial^2 \mathcal{V}}{\partial x^{\beta}\partial x^{\alpha}}(\bar{\boldsymbol{x}})\bar{A}^{\alpha}_{i}\bar{A}^{\beta}_{j}\bar{\xi}^{i}\bar{\xi}^{j} + \frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\bar{\boldsymbol{x}})\Big(\frac{\partial A^{\alpha}_{i}}{\partial x^{\beta}}(\bar{\boldsymbol{x}})\bar{A}^{\beta}_{j}\bar{\xi}^{i}\bar{\xi}^{j} + \bar{A}^{\alpha}_{i}\dot{\bar{\xi}}^{i}\Big)\Big) + \mathcal{O}(t^3)\end{aligned} \tag{2.11}$$

[1] $\boldsymbol{\Psi}^{\top}\boldsymbol{A} = (\boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{\Psi})^{\top}\boldsymbol{A} = \boldsymbol{\Psi}^{\top}(\boldsymbol{\Psi}\boldsymbol{\Phi})^{\top}\boldsymbol{A} = \boldsymbol{\Psi}^{\top}\boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{A} = \mathbf{0}$

Introducing the notation

$$\partial_i = A_i^\alpha \frac{\partial}{\partial x^\alpha}, \ i = 1, \dots, n \tag{2.12}$$

for the derivative in the direction of the i-th basis vector, we can state the Taylor expansion as

$$\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + t\, \partial_i \mathcal{V}(\bar{\boldsymbol{x}}) \bar{\xi^i} + \tfrac{1}{2} t^2 \big(\partial_i \partial_j \mathcal{V}(\bar{\boldsymbol{x}}) \bar{\xi^i} \bar{\xi^j} + \partial_i \mathcal{V}(\bar{\boldsymbol{x}}) \dot{\bar{\xi^i}}\big) + \mathcal{O}(t^3). \tag{2.13}$$

There are two more things we can derive from this equation:

- If $\partial_i \mathcal{V}(\bar{\boldsymbol{x}}) = 0, i = 1, \dots, n$ then $\bar{\boldsymbol{x}}$ is called a *critical point* of $\mathcal{V}$. At a critical point the expansion (3.24) reduces to

$$\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + \tfrac{1}{2} t^2 \underbrace{(\partial_i \partial_j \mathcal{V})(\bar{\boldsymbol{x}})}_{\bar{H}_{ij}} \bar{\xi^i} \bar{\xi^j} + \mathcal{O}(t^3). \tag{2.14}$$

  This relation holds for any sufficiently smooth curve $t \mapsto \boldsymbol{x}(t)$ through $\bar{\boldsymbol{x}}$ and consequently for any velocity vector $\bar{\boldsymbol{\xi}}$ at the critical point. So if the matrix $\bar{\boldsymbol{H}}$ is positive (negative) definite, then $\bar{\boldsymbol{x}}$ is a local minimum (maximum) of $\mathcal{V}$.

- Assume the curve $t \mapsto \boldsymbol{x}(t)$ is a *geodesic*, i.e. $\dot{\xi^i} = -\mathit{\Gamma}^i_{jk} \xi^j \xi^k$ with the connection coefficients $\mathit{\Gamma}^i_{jk}$ that will be discussed later. Plugging this into (3.24) we find a coordinate form of the *Hessian tensor* $\boldsymbol{\nabla}^2 \mathcal{V}$ of the potential:

$$\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + t (\partial_i \mathcal{V})(\bar{\boldsymbol{x}}) \bar{\xi^i} + \tfrac{1}{2} t^2 \underbrace{\big(\partial_i \partial_j \mathcal{V} - \mathit{\Gamma}^k_{ij} \partial_k \mathcal{V}\big)}_{(\boldsymbol{\nabla}^2 \mathcal{V})_{ij}} (\bar{\boldsymbol{x}}) \bar{\xi^i} \bar{\xi^j} + \mathcal{O}(t^3). \tag{2.15}$$

  At a critical point $\bar{\boldsymbol{x}}$, the Hessian of the potential is independent of the connection coefficients $\mathit{\Gamma}^i_{jk}$ and consequently of the underlying metric. There it coincides with the matrix $\bar{\boldsymbol{H}}$ defined in (3.25).

Example 3. Consider the function

$$\mathcal{V}(\boldsymbol{R}) = \operatorname{tr}\big(\boldsymbol{\mathit{\Pi}}' (\mathrm{I}_3 - \boldsymbol{R})\big). \tag{2.16}$$

proposed in [Koditschek, 1989], where $\boldsymbol{\mathit{\Pi}}' = \boldsymbol{\mathit{\Pi}}'^\top \in \mathbb{R}^{3\times 3}$ and $\boldsymbol{R} \in \mathbb{SO}(3)$. It can be regarded as a distance function on the configuration space $\mathbb{X} = \mathbb{SO}(3)$ with six tunable parameters which makes it quite useful for control purposes, see e.g. [Bullo and Murray, 1999] or [Lee et al., 2010]. This function will appear several times in later chapters, so it will be discussed in detail here.

As in the previous example, we regard the coefficients of the rotation matrix $\boldsymbol{R}$ as redundant configuration coordinates. Using the basis from (3.17) we compute the differential and the Hessian at a critical point $\boldsymbol{R}_0$ as

$$\boldsymbol{\nabla} \mathcal{V}(\boldsymbol{R}_0) = \operatorname{vee2}(\boldsymbol{\mathit{\Pi}}' \boldsymbol{R}_0) \stackrel{!}{=} \mathbf{0}, \tag{2.17a}$$

$$\boldsymbol{\nabla}^2 \mathcal{V}(\boldsymbol{R}_0) = \operatorname{tr}(\boldsymbol{\mathit{\Pi}}' \boldsymbol{R}_0) \mathrm{I}_3 - \boldsymbol{\mathit{\Pi}}' \boldsymbol{R}_0. \tag{2.17b}$$

with the vee2 operator defined as

$$\mathrm{vee2}:\mathbb{R}^{3\times 3}\to\mathbb{R}^3,\ \begin{bmatrix} * & A_{12} & A_{13}\\ A_{21} & * & A_{23}\\ A_{31} & A_{32} & * \end{bmatrix}\mapsto\begin{bmatrix} A_{32}-A_{23}\\ A_{13}-A_{31}\\ A_{21}-A_{12}\end{bmatrix}. \tag{2.18}$$

So the condition for a critical point at $\boldsymbol{R}_0$ is that the product $\boldsymbol{\Pi}'\boldsymbol{R}_0$ is skew-symmetric. Obviously the function $\mathcal{V}$ has a critical point at

$$\boldsymbol{R}_{0,1}=\mathbf{I}_3\ :\qquad \mathcal{V}(\boldsymbol{R}_{0,1})=0,\qquad \boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,1})=\mathrm{tr}(\boldsymbol{\Pi}')\mathbf{I}_3-\boldsymbol{\Pi}'=:\boldsymbol{\Pi}.$$

It is a minimum if the matrix $\boldsymbol{\Pi}$ is positive definite.

There are more critical points: For their investigation it will be useful to consider the eigenvalue decomposition $\boldsymbol{\Pi}'=\boldsymbol{X}\boldsymbol{\Lambda}'\boldsymbol{X}^\top$ with $\boldsymbol{\Lambda}'=\mathrm{diag}(\lambda_1',\lambda_2',\lambda_3')$, $\boldsymbol{X}\in\mathbb{SO}(3)$. Note that the matrix $\boldsymbol{\Pi}$ has the same eigenvectors, i.e. $\boldsymbol{\Pi}=\boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^\top$ with $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)$ and the eigenvalues are related by

$$\boldsymbol{\Lambda}=\mathrm{tr}(\boldsymbol{\Lambda}')\mathbf{I}_3-\boldsymbol{\Lambda}'\qquad\Leftrightarrow\qquad\boldsymbol{\Lambda}'=\tfrac{1}{2}\,\mathrm{tr}(\boldsymbol{\Lambda})\mathbf{I}_3-\boldsymbol{\Lambda} \tag{2.19a}$$

$$\left.\begin{aligned}\lambda_1&=\lambda_2'+\lambda_3',\\ \lambda_2&=\lambda_3'+\lambda_1',\\ \lambda_3&=\lambda_1'+\lambda_2'\end{aligned}\right\}\qquad\Leftrightarrow\qquad\left\{\begin{aligned}\lambda_1'&=\tfrac{1}{2}(\lambda_2+\lambda_3-\lambda_1),\\ \lambda_2'&=\tfrac{1}{2}(\lambda_3+\lambda_1-\lambda_2),\\ \lambda_3'&=\tfrac{1}{2}(\lambda_1+\lambda_2-\lambda_3)\end{aligned}\right. \tag{2.19b}$$

Using the transformed orientation $\tilde{\boldsymbol{R}}_0=\boldsymbol{X}^\top\boldsymbol{R}_0\boldsymbol{X}$ we have

$$\boldsymbol{\nabla}\mathcal{V}(\boldsymbol{R}_0)=\boldsymbol{X}\,\mathrm{vee2}(\boldsymbol{\Lambda}'\tilde{\boldsymbol{R}}_0)\overset{!}{=}\mathbf{0}, \tag{2.20a}$$

$$\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_0)=\boldsymbol{X}\big(\,\mathrm{tr}(\boldsymbol{\Lambda}'\tilde{\boldsymbol{R}}_0)\mathbf{I}_3-\boldsymbol{\Lambda}'\tilde{\boldsymbol{R}}_0\big)\boldsymbol{X}^\top. \tag{2.20b}$$

For critical points in addition to $\boldsymbol{R}_{0,1}$ we need to distinguish the following cases:

- *distinct eigenvalues* $\lambda_i\neq\lambda_j\ \Leftrightarrow\ \lambda_i'\neq\lambda_j'$, $i\neq j$:

$$\begin{aligned}\tilde{\boldsymbol{R}}_{0,2}&=\mathrm{diag}(1,-1,-1)\ :\ \mathcal{V}(\boldsymbol{R}_{0,2})=2\lambda_1,\ \ \mathrm{eig}\big(\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,2})\big)=\{-\lambda_1,\lambda_3-\lambda_1,\lambda_2-\lambda_1\}\\ \tilde{\boldsymbol{R}}_{0,3}&=\mathrm{diag}(-1,1,-1)\ :\ \mathcal{V}(\boldsymbol{R}_{0,3})=2\lambda_2,\ \ \mathrm{eig}\big(\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,3})\big)=\{\lambda_3-\lambda_2,-\lambda_2,\lambda_1-\lambda_2\}\\ \tilde{\boldsymbol{R}}_{0,4}&=\mathrm{diag}(-1,-1,1)\ :\ \mathcal{V}(\boldsymbol{R}_{0,4})=2\lambda_3,\ \ \mathrm{eig}\big(\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,4})\big)=\{\lambda_2-\lambda_3,\lambda_1-\lambda_3,-\lambda_3\}\end{aligned}$$

- *double eigenvalue* $\lambda_1=\lambda_2\neq\lambda_3$: the function $\mathcal{V}$ is stationary at the point $\boldsymbol{R}_{0,4}$ and on the circular submanifold

$$\tilde{\boldsymbol{R}}_{0,5}=\begin{bmatrix}-c & s & 0\\ s & c & 0\\ 0 & 0 & -1\end{bmatrix},\ c^2+s^2=1\ :$$

$$\mathcal{V}(\boldsymbol{R}_{0,5})=2\lambda_1,\qquad\mathrm{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,5})\big)=\{-\lambda_1,\lambda_3-\lambda_1,0\}$$

  which includes the points $\boldsymbol{R}_{0,2}$ and $\boldsymbol{R}_{0,3}$. For the cases $\lambda_1=\lambda_3\neq\lambda_2$ and $\lambda_2=\lambda_3\neq\lambda_1$ we have analogous results.

- *triple eigenvalue* $\lambda_1 = \lambda_2 = \lambda_3 \Rightarrow \boldsymbol{\Lambda}' = \boldsymbol{\Pi}' = \frac{1}{2}\lambda_1 \mathbf{I}_3$, $\boldsymbol{X} = \mathbf{I}_3$: The function $\mathcal{V}$ is stationary on the spherical submanifold

$$\boldsymbol{R}_{0,6} = \begin{bmatrix} 1-2(a_2^2+a_3^2) & 2a_2a_1 & 2a_1a_3 \\ 2a_1a_2 & 1-2(a_1^2+a_3^2) & 2a_2a_3 \\ 2a_1a_3 & 2a_2a_3 & 1-2(a_1^2+a_2^2) \end{bmatrix}, \; a_1^2+a_2^2+a_3^2 = 1 \; :$$

$$\mathcal{V}(\boldsymbol{R}_{0,6}) = 2\lambda_1, \qquad \operatorname{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,6})\big) = \{-\lambda_1, 0, 0\}$$

which includes $\tilde{\boldsymbol{R}}_{0,2}$, $\tilde{\boldsymbol{R}}_{0,3}$ and $\tilde{\boldsymbol{R}}_{0,4}$ The matrix $\boldsymbol{R}_{0,6}$ is a 180° rotation about the unit axis $\boldsymbol{a} = [a_1, a_2, a_3]^\top$.

If one or more eigenvalues are zero we have even more critical points

- *one zero eigenvalue* $\lambda_3 = 0$: the function $\mathcal{V}$ is stationary on the circular submanifold

$$\tilde{\boldsymbol{R}}_{0,7} = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; c^2+s^2 = 1 \; :$$

$$\mathcal{V}(\boldsymbol{R}_{0,7}) = 0, \qquad \operatorname{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,5})\big) = \{\lambda_1, \lambda_2, 0\}$$

which includes the points $\boldsymbol{R}_{0,1} = \mathbf{I}_3$ and $\boldsymbol{R}_{0,4}$.

- *two zero eigenvalues* $\lambda_2 = \lambda_3 = 0$: the function $\mathcal{V}$ is stationary on the spherical submanifold

$$\tilde{\boldsymbol{R}}_{0,8} = \begin{bmatrix} q_\mathsf{w}^2 - q_\mathsf{y}^2 - q_\mathsf{z}^2 & -2q_\mathsf{w}q_\mathsf{z} & 2q_\mathsf{w}q_\mathsf{y} \\ 2q_\mathsf{w}q_\mathsf{z} & q_\mathsf{w}^2 + q_\mathsf{y}^2 - q_\mathsf{z}^2 & 2q_\mathsf{y}q_\mathsf{z} \\ -2q_\mathsf{w}q_\mathsf{y} & 2q_\mathsf{y}q_\mathsf{z} & q_\mathsf{w}^2 - q_\mathsf{y}^2 + q_\mathsf{z}^2 \end{bmatrix}, \; q_\mathsf{w}^2 + q_\mathsf{y}^2 + q_\mathsf{z}^2 = 1 \; :$$

$$\mathcal{V}(\boldsymbol{R}_{0,8}) = 0, \qquad \operatorname{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,8})\big) = \{\lambda_1, 0, 0\}$$

which includes the points $\boldsymbol{R}_{0,1} = \mathbf{I}_3$ and $\boldsymbol{R}_{0,3}$, $\boldsymbol{R}_{0,4}$.

- *three zero eigenvalues* $\lambda_3 = \lambda_2 = \lambda_3 = 0 \;\Rightarrow\; \boldsymbol{\Pi} = \boldsymbol{\Pi}' = \mathbf{0}$: the function $\mathcal{V}$ degenerates to a constant $\mathcal{V}(\boldsymbol{R}) = 0 \,\forall\, \boldsymbol{R} \in \mathbb{SO}(3)$.

Comparing the eigenvalues $\operatorname{eig}(\nabla^2\mathcal{V})$ of the Hessians at the critical points, it is clear that the function $\mathcal{V}$ always has exactly one minimum and exactly one maximum, i.e. all other critical points are saddles, independently of the signs of λ_1, λ_2 and λ_3. If one or more $\lambda_i = 0$ or $\lambda_i = \lambda_j$ then the minimum or maximum is taken not on a single point but on a one or two-dimensional submanifold.

The most useful case for the following is $\lambda_1, \lambda_2, \lambda_3 > 0$, i.e. $\boldsymbol{\Pi}$ is positive definite and $\boldsymbol{R} = \mathbf{I}_3$ is the global minimum of $\mathcal{V}$. It is also worth noting that for $\boldsymbol{\Pi} = \mathbf{I}_3$ the function $\mathcal{V}$ is related to the angle θ of the rotation matrix $\boldsymbol{R}$ by $\mathcal{V} = \frac{1}{2}\operatorname{tr}(\mathbf{I}_3 - \boldsymbol{R}) = 1 - \cos\theta$.

2.0.4 Commutation coefficients

For a function $\mathcal{V} : \mathbb{R}^\nu \to \mathbb{R}$ we are used to the fact that partial derivatives commute, i.e. $\partial^2\mathcal{V}/\partial x^\alpha \partial x^\beta = \partial^2\mathcal{V}/\partial x^\beta \partial x^\alpha$. Unfortunately this is (in general) not the case for a directional derivatives like ∂_i defined in (3.23). Consequently we investigate the following commutation relation

$$
\begin{aligned}
\partial_i\partial_j\mathcal{V} - \partial_j\partial_i\mathcal{V} &= A_i^\alpha \frac{\partial}{\partial x^\alpha}\left(A_j^\beta \frac{\partial\mathcal{V}}{\partial x^\beta}\right) - A_j^\beta \frac{\partial}{\partial x^\beta}\left(A_i^\alpha \frac{\partial\mathcal{V}}{\partial x^\alpha}\right) \\
&= A_i^\alpha \frac{\partial A_j^\beta}{\partial x^\alpha}\frac{\partial\mathcal{V}}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\frac{\partial\mathcal{V}}{\partial x^\alpha} + A_i^\alpha A_j^\beta \underbrace{\left(\frac{\partial^2\mathcal{V}}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2\mathcal{V}}{\partial x^\beta \partial x^\alpha}\right)}_{=0} \\
&= \left(A_i^\beta \frac{\partial A_j^\alpha}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\right)\frac{\partial\mathcal{V}}{\partial x^\alpha}.
\end{aligned}
\tag{2.21}
$$

Now using the identity (3.20) with $\Phi_\alpha^\kappa = \frac{\partial\phi^\kappa}{\partial x^\alpha}$ and $\Phi_\alpha^\kappa A_i^\alpha = 0 \Rightarrow \Phi_\alpha^\kappa \frac{\partial A_i^\alpha}{\partial x^\beta} = -\frac{\partial \Phi_\alpha^\kappa}{\partial x^\beta} A_i^\alpha$ to shape this expressions a bit further

$$
\begin{aligned}
\partial_i\partial_j\mathcal{V} - \partial_j\partial_i\mathcal{V} &= \left(A_i^\beta \frac{\partial A_j^\alpha}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\right)\overbrace{(A_k^\sigma Y_\alpha^k + \Psi_\kappa^\sigma \Phi_\alpha^\kappa)}^{\delta_\alpha^\sigma}\frac{\partial\mathcal{V}}{\partial x^\sigma} \\
&= \left(A_i^\beta \frac{\partial A_j^\alpha}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\right) Y_\alpha^k A_k^\sigma \frac{\partial\mathcal{V}}{\partial x^\sigma} - \left(A_i^\beta A_j^\alpha \frac{\partial \Phi_\alpha^\kappa}{\partial x^\beta} - A_j^\beta A_i^\alpha \frac{\partial \Phi_\alpha^\kappa}{\partial x^\beta}\right)\Psi_\kappa^\sigma \frac{\partial\mathcal{V}}{\partial x^\sigma} \\
&= \underbrace{\left(A_i^\beta \frac{\partial A_j^\alpha}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\right) Y_\alpha^k}_{\gamma_{ij}^k} \underbrace{A_k^\sigma \frac{\partial\mathcal{V}}{\partial x^\sigma}}_{\partial_k\mathcal{V}} - A_j^\beta A_i^\alpha \underbrace{\left(\frac{\partial^2\phi^\kappa}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2\phi^\kappa}{\partial x^\beta \partial x^\alpha}\right)}_{=0}\Psi_\kappa^\sigma \frac{\partial\mathcal{V}}{\partial x^\sigma}.
\end{aligned}
\tag{2.22}
$$

Since this relation holds for any function $\mathcal{V}$ we can state it in operator form and introduce the *commutation coefficients* γ_{ij}^k as

$$
\partial_i\partial_j - \partial_j\partial_i = \gamma_{ij}^k \partial_k, \qquad \gamma_{ij}^k = \left(\partial_i A_j^\alpha - \partial_j A_i^\alpha\right)(A^+)_\alpha^k. \tag{2.23}
$$

It is interesting to note that the commutation coefficients are *invariant* to the choice of configuration coordinates $\boldsymbol{x}$, even though the coordinates appear explicitly in the definition: For a change of configuration coordinates $\boldsymbol{x} = f(\hat{\boldsymbol{x}})$ the commutation symbols transform like $\hat{\gamma}_{ij}^k(\hat{\boldsymbol{x}}) = \gamma_{ij}^k(f(\hat{\boldsymbol{x}}))$. This might be obvious from a geometric point of view, but the explicit calculation of the coordinate transformation is shown in see section A.3. It will even turn out that for most of our examples the coefficients will be constants.

The right hand side of (3.34) appears in the context of Lagrange's equation in [Boltzmann, 1902] and [Hamel, 1904] for the case of minimal configuration coordinates and consequently with square matrices $\boldsymbol{Y}$ and $\boldsymbol{A}$. In the contemporary literature on this context these quantities γ_{ij}^k are sometimes called the *Boltzmann three-index symbols* [Lurie, 2002, sec. 1.8] or *Hamel coefficients* [Bremer, 2008, p. 75]. The left hand side of (3.34) appears in the context of tensor algebra in [Misner et al., 1973, Box 8.4] where γ_{ij}^k are called the *commutation coefficients*. From the way γ_{ij}^k is defined here (3.34), this naming seems most fitting.

The commutation coefficients γ^i_{jk} vanish if the corresponding velocity coordinates ξ^i are *integrable*, i.e.

$$\exists\, \pi^i \,:\, \dot{\pi}^i = \xi^i = Y^i_\alpha \dot{x}^\alpha \qquad \Rightarrow \qquad Y^i_\alpha = \frac{\partial \pi^i}{\partial x^\beta}$$
$$\Rightarrow \quad \frac{\partial Y^i_\alpha}{\partial x^\alpha} = \frac{\partial^2 \pi^i}{\partial x^\beta \partial x^\alpha} = \frac{\partial Y^i_\beta}{\partial x^\alpha}, \qquad \Rightarrow \qquad \gamma^i_{jk} = 0 \tag{2.24}$$

which is not the case in general. Nevertheless the quantities π are commonly introduced as *nonholonomic coordinates* in [Boltzmann, 1902] (also called *quasi coordinates* in [Lurie, 2002, sec. 1.5]). Then we could write $\partial_i(\partial_j f) - \partial_j(\partial_i f) = \partial^2 f/\partial\pi^i\partial\pi^j - \partial^2 f/\partial\pi^j\partial\pi^i \neq 0$ what might lead to the conception that partial derivatives do not commute. The commutativity clearly holds, the issue is rather that the coordinates π do not exist. To avoid confusion of this kind we do not pick up this notation here.

Example 4. The commutation coefficients γ^k_{ij} associated with the kinematics matrix $\boldsymbol{A}$ from (3.17) are

$$\gamma^k_{ij} = \begin{cases} +1, & (i,j,k) \text{ even permutation of } (1,2,3) \\ -1, & (i,j,k) \text{ odd permutation of } (1,2,3) \\ \;\;\,0, & \text{else} \end{cases} . \tag{2.25}$$

This coincides with the three dimensional Levi-Civita symbol. It is related to the 3 dimensional *cross product* by $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3 \,:\, [\gamma^k_{ij} a^i b^j]_{k=1..3} = \boldsymbol{a} \times \boldsymbol{b}$.

2.0.5 Linearization about a trajectory

Let $\bar{\boldsymbol{x}} : [t_1, t_2] \to \mathbb{X}$ be a smooth curve with the velocity coordinates $\bar{\boldsymbol{\xi}} : [t_1, t_2] \to \mathbb{R}^n : t \mapsto \boldsymbol{A}^+(\bar{\boldsymbol{x}}(t))\dot{\bar{\boldsymbol{x}}}(t)$. For a small deviation $\boldsymbol{x} \approx \bar{\boldsymbol{x}}$ with $\boldsymbol{x} \in \mathbb{X}$ we may approximate the geometric constraint as

$$\boldsymbol{\phi}(\boldsymbol{x}) \approx \underbrace{\boldsymbol{\phi}(\bar{\boldsymbol{x}})}_{=\boldsymbol{0}} + \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) = \boldsymbol{0}. \tag{2.26}$$

Since this constraint is affine w.r.t. $\boldsymbol{x}$ it is reasonable to use a the basis $\boldsymbol{\varepsilon}(t) \in \mathbb{R}^n$ for the deviated configuration coordinates:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{A}(\bar{\boldsymbol{x}})\boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} = \boldsymbol{A}^+(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}), \tag{2.27}$$

For the velocity coordinates $\boldsymbol{\xi}$ of the deviated curve $\boldsymbol{x}$ we use again the first order approximation and $\boldsymbol{Y} = \boldsymbol{A}^+$:

$$\begin{aligned}
\xi^i &= Y^i_\alpha(\boldsymbol{x})\dot{x}^\alpha \\
&\approx Y^i_\alpha(\bar{\boldsymbol{x}} + \boldsymbol{A}(\bar{\boldsymbol{x}})\boldsymbol{\varepsilon})\tfrac{\mathrm{d}}{\mathrm{d}t}\big(\bar{x}^\alpha + A^\alpha_j(\bar{\boldsymbol{x}})\,\varepsilon^j\big) \\
&\approx Y^i_\alpha(\bar{\boldsymbol{x}})\big(\dot{x}^\alpha_{\mathsf{R}} + \tfrac{\partial A^\alpha_j}{\partial x^\beta}(\bar{\boldsymbol{x}})\,\dot{x}^\beta_{\mathsf{R}}\varepsilon^j + A^\alpha_j(\bar{\boldsymbol{x}})\,\dot{\varepsilon}^j\big) + \tfrac{\partial Y^i_\alpha}{\partial x^\beta}(\bar{\boldsymbol{x}})A^\beta_j(\bar{\boldsymbol{x}})\,\varepsilon^j\dot{x}^\alpha_{\mathsf{R}} \\
&= \bar{\xi}^i + \dot{\varepsilon}^i + \gamma^i_{kj}(\bar{\boldsymbol{x}})\bar{\xi}^k\varepsilon^j
\end{aligned} \tag{2.28}$$

Using these results we may formulate an approximation of a general smooth function f along the trajectory $t \mapsto \bar{\boldsymbol{x}}(t)$ as

$$
\begin{aligned}
f(\boldsymbol{x},\boldsymbol{\xi},\dot{\boldsymbol{\xi}}) &\approx f(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}}) + \frac{\partial f}{\partial x^\alpha}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})(x^\alpha - \bar{x}^\alpha) + \frac{\partial f}{\partial \xi^i}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})(\xi^i - \bar{\xi}^i) \\
&\quad + \frac{\partial f}{\partial \dot{\xi}^i}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})(\dot{\xi}^i - \dot{\bar{\xi}}^i) \\
&\approx f(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}}) + (\partial_i f)(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})\varepsilon^i + \frac{\partial f}{\partial \xi^i}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})(\dot{\varepsilon}^i + \gamma^i_{kj}(\bar{\boldsymbol{x}})\bar{\xi}^k\varepsilon^j) \\
&\quad + \frac{\partial f}{\partial \dot{\xi}^i}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})(\ddot{\varepsilon}^i + \gamma^i_{kj}(\bar{\boldsymbol{x}})\bar{\xi}^k\dot{\varepsilon}^j + \gamma^i_{kj}(\bar{\boldsymbol{x}})\dot{\bar{\xi}}^k\varepsilon^j + \partial_l\gamma^i_{kj}(\bar{\boldsymbol{x}})\bar{\xi}^l\bar{\xi}^k\varepsilon^j) \\
&= \bar{f} + \bar{F}^0_i\varepsilon^i + \bar{F}^1_i\dot{\varepsilon}^i + \bar{F}^2_i\ddot{\varepsilon}^i
\end{aligned}
\tag{2.29}
$$

where

$$
\begin{aligned}
\bar{f} &= f(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}}), \\
\bar{F}^0_i &= (\partial_i f)(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}}) + \frac{\partial f}{\partial \xi^j}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})\gamma^j_{ki}(\bar{\boldsymbol{x}})\bar{\xi}^k + \frac{\partial f}{\partial \dot{\xi}^j}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})\big(\gamma^j_{ki}(\bar{\boldsymbol{x}})\dot{\bar{\xi}}^k + \partial_l\gamma^j_{ki}(\bar{\boldsymbol{x}})\bar{\xi}^l\bar{\xi}^k\big), \\
\bar{F}^1_i &= \frac{\partial f}{\partial \xi^i}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}}) + \frac{\partial f}{\partial \dot{\xi}^j}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}})\gamma^j_{ki}(\bar{\boldsymbol{x}})\bar{\xi}^k, \\
\bar{F}^2_i &= \frac{\partial f}{\partial \dot{\xi}^i}(\bar{\boldsymbol{x}},\bar{\boldsymbol{\xi}},\dot{\bar{\boldsymbol{\xi}}}).
\end{aligned}
$$

Evidently, the expressions simplify significantly is the velocity coordinates are holonomic, i.e. $\gamma = 0$, or if the approximation is about a static point $\bar{\boldsymbol{x}} = \mathit{const.} \Rightarrow \boldsymbol{\xi} = \boldsymbol{0}$.

2.0.6 Calculus of variations

The calculus of variations is concerned with the extremals of functionals, i.e. functions of functions. For the particular context of classical mechanics we are interested in the curves $t \mapsto \boldsymbol{x}(t)$ for which the functional

$$
\mathcal{J}[\boldsymbol{x}] = \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{x}(t),\boldsymbol{\xi}(t),t)\,\mathrm{d}t \tag{2.30}
$$

for given boundary conditions $\boldsymbol{x}(t_1)$ and $\boldsymbol{x}(t_2)$ is *stationary*. The *Lagrangian* $\mathcal{L}$ is here a function of the configuration coordinates $\boldsymbol{x}$, its derivatives $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi}$ parameterized in the velocity coordinates $\boldsymbol{\xi}$ and may depend explicitly on the time t as well.

For the standard case, $\boldsymbol{x} = \boldsymbol{q}$ and $\boldsymbol{\xi} = \dot{\boldsymbol{q}}$, a derivation of may be found in e.g. [Lanczos, 1986, ch. II], [Arnold, 1989, sec. 12] or [Courant and Hilbert, 1924, chap. 4, §3]. For the present case we modify the well known derivation slightly: Suppose that $\boldsymbol{x} : [t_1, t_2] \mapsto \mathbb{X}$ is the solution to the variational problem. With the function $\boldsymbol{\chi}(t) \in \mathbb{R}^\nu$ and the parameter $\varepsilon \in \mathbb{R}$ we define a perturbation to it by

$$
\bar{\boldsymbol{x}} = \boldsymbol{x} + \varepsilon\boldsymbol{\chi}. \tag{2.31}
$$

We need $\bar{\boldsymbol{x}}(t) \in \mathbb{X}$ and consequently $\boldsymbol{\phi}(\bar{\boldsymbol{x}}) = \mathbf{0}$. Assuming ε to be sufficiently small, we may use the first order approximation analog to subsection 3.3.5: With the *variation coordinates* $\boldsymbol{h} : [t_1, t_2] \to \mathbb{R}^n$ we parameterize $\boldsymbol{\chi} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}$. Using the inverse kinematic relation $\boldsymbol{\xi} = \boldsymbol{Y}(\boldsymbol{x})\dot{\boldsymbol{x}}$ we can write the functional for the varied path as

$$\mathcal{J}[\bar{\boldsymbol{x}}] = \int_{t_1}^{t_2} \mathcal{L}\big(\boldsymbol{x} + \varepsilon \boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}, \boldsymbol{Y}(\boldsymbol{x} + \varepsilon \boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}) \tfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{x} + \varepsilon \boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}), t\big)\, \mathrm{d}t =: \mathcal{P}(\varepsilon) \tag{2.32}$$

Now if $\boldsymbol{x}(t)$ is indeed the solution to the variational problem, then $\mathcal{P}(\varepsilon)$ must have a minimum at $\mathcal{P}(0)$ and consequently $\partial\mathcal{P}/\partial\varepsilon(0) = 0$. Evaluation of this "ordinary" differentiation yields

$$\begin{aligned} 0 = \frac{\partial \mathcal{P}}{\partial \varepsilon}\Big|_{\varepsilon=0} &= \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial x^\alpha} A_i^\alpha h^i + \frac{\partial \mathcal{L}}{\partial \xi^i} \left(\frac{\partial Y_\alpha^i}{\partial x^\beta} A_j^\beta h^j \dot{x}^\alpha + Y_\alpha^i \frac{\partial A_j^\alpha}{\partial x^\beta} h^j \dot{x}^\beta + \dot{h}^i \right) \right) \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left(\partial_i \mathcal{L}\, h^i + \frac{\partial \mathcal{L}}{\partial \xi^i} \big(\gamma_{kj}^i h^j \xi^k + \dot{h}^i \big) \right) \mathrm{d}t \end{aligned} \tag{2.33}$$

where we have found again the commutation coefficients γ_{kj}^i previously derived in (3.34). Integrating by parts with the boundary conditions $\boldsymbol{h}(t_1) = \boldsymbol{h}(t_2) = \mathbf{0}$ gives

$$\int_{t_0}^{t_1} h^i \left(A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} - \gamma_{ij}^k \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} \right) \mathrm{d}t = 0. \tag{2.34}$$

Since the variation coordinates $h^i, i = 1, \ldots, n$ are independent by definition, the *fundamental lemma of the calculus of variations* (see e.g. [Arnold, 1989, p. 57] or [Courant and Hilbert, 1924, p. 166]) states that, for the integral to vanish, the terms in the brackets have to vanish, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} + \gamma_{ij}^k \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} - A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} = 0, \qquad i = 1, \ldots, n. \tag{2.35}$$

This, combined with the kinematic relation $\dot{x}^\alpha = A_i^\alpha \xi^i, \alpha = 1 \ldots \nu$, is the necessary condition for the functional (3.41) to be stationary.

For the special case $\boldsymbol{x}(t) = \boldsymbol{q}(t) \in \mathbb{R}^n$ and $\boldsymbol{\xi}(t) = \dot{\boldsymbol{q}}(t)$ we have $\boldsymbol{A} = \mathbf{I}_n$ and $\gamma = 0$. Then (3.46) coincides with the *Euler-Lagrange equation.*

Example 5. Consider the configuration coordinates $\boldsymbol{x} = [\boldsymbol{R}_\mathsf{x}^\top, \boldsymbol{R}_\mathsf{y}^\top, \boldsymbol{R}_\mathsf{z}^\top]^\top$ and the velocity coordinates $\boldsymbol{\xi} = \boldsymbol{\omega}$ related by $\dot{\boldsymbol{R}} = \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega})$. For the Lagrangian

$$\mathcal{L} = \tfrac{1}{2} \boldsymbol{\omega}^\top \boldsymbol{\Theta} \boldsymbol{\omega} + \mathrm{tr}(\boldsymbol{\Pi}'(\mathbf{I}_3 - \boldsymbol{R})) \tag{2.36}$$

and taking into account the results from Example 9 and 10 we obtain

$$\Big[\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} + \gamma_{ij}^k \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} - A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} \Big]_{i=1,2,3} = \boldsymbol{\Theta} \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \boldsymbol{\Theta} \boldsymbol{\omega} + \mathrm{vee2}(\boldsymbol{\Pi}' \boldsymbol{R}). \tag{2.37}$$

Chapter 3

Analytical mechanics of particle systems

Goal. Established approaches of analytical mechanics commonly rely on the parameterization of the system in terms of *minimal* generalized coordinates $\boldsymbol{q}$ and their derivatives $\dot{\boldsymbol{q}}$. In contrast to this we will allow *redundant* generalized coordinates $\boldsymbol{x}$, i.e. coordinates that may not be completely free, but may themselves be constrained by equations of the form $\boldsymbol{\phi}(\boldsymbol{x}) = \mathbf{0}$. Furthermore we will consider *velocity coordinates* $\boldsymbol{\xi}$ that parameterize the change $\dot{\boldsymbol{x}}$ of the configuration coordinates. In this chapter we like to review some basic concepts of analytical mechanics for systems of constrained particles with this more general parameterization. The resulting formulations might become more cumbersome, but with some examples we like to show that it is worth it.

3.1 System under consideration

Particle positions. For this chapter we consider systems of $\mathfrak{N}$ particles under geometric constraints: The *position* of a particle in a given reference frame at a given time t is $\mathfrak{r}_p(t) \in \mathbb{R}^3, p = 1, \dots, \mathfrak{N}$ and the collection of all particle positions is $\mathfrak{x} = [\mathfrak{r}_1^\top, \dots, \mathfrak{r}_\mathfrak{N}^\top]^\top \in \mathbb{R}^{3\mathfrak{N}}$. *Geometric constraints* on the particles are captured in $\mathfrak{H} \geq 0$ smooth functions of the form $\mathfrak{c}(\mathfrak{x}) = [\mathfrak{c}^1(\mathfrak{x}), \dots, \mathfrak{c}^\mathfrak{H}(\mathfrak{x})]^\top = \mathbf{0}$. The set of all mutually admissible particle positions

$$\mathfrak{X} = \{\mathfrak{x} \in \mathbb{R}^{3\mathfrak{N}} \,|\, \mathfrak{c}(\mathfrak{x}) = \mathbf{0}\} \tag{3.1}$$

is called the *configuration space*. We require $\frac{\partial \mathfrak{c}}{\partial \mathfrak{x}}(\mathfrak{x})$ to have a constant, though not necessarily full rank. The dimension of the configuration space is $\dim \mathfrak{X} = 3\mathfrak{N} - \operatorname{rank} \frac{\partial \mathfrak{c}}{\partial \mathfrak{x}} = n$.

Coordinates. Now let the admissible particle positions $\mathfrak{x} \in \mathfrak{X}$ be parameterized $\mathfrak{r}_p = \mathfrak{r}_p(\boldsymbol{x})$ by possibly redundant coordinates $\boldsymbol{x} \in \mathbb{X}$. This means $\boldsymbol{\phi}(\boldsymbol{x}) = \mathbf{0} \;\Rightarrow\; \mathfrak{c}(\mathfrak{x}(\boldsymbol{x})) = \mathbf{0}$ and consequently $\boldsymbol{x} \in \mathbb{X} \;\Rightarrow\; \mathfrak{x}(\boldsymbol{x}) \in \mathfrak{X}$.

With a suitable kinematics matrix $\boldsymbol{A}(\boldsymbol{x})$, as discussed in the previous section, we can parameterize the particle velocity as

$$\dot{\mathfrak{r}}_p = \frac{\partial \mathfrak{r}_p}{\partial \boldsymbol{x}} \boldsymbol{A}\boldsymbol{\xi} = \partial_i \mathfrak{r}_p \xi^i \tag{3.2}$$

with minimal velocity coordinates $\boldsymbol{\xi} \in \mathbb{R}^n$. Finally we may express the particle accelerations as

$$\ddot{\mathfrak{r}}_p = \partial_i \mathfrak{r}_p(\boldsymbol{x})\dot{\xi}^i + \partial_j \partial_i \mathfrak{r}_p(\boldsymbol{x})\xi^i \xi^j, \tag{3.3}$$

From the parameterization it is evident that

$$\partial_i \mathfrak{r}_p = \frac{\partial \dot{\mathfrak{r}}_p}{\partial \xi^i} = \frac{\partial \ddot{\mathfrak{r}}_p}{\partial \dot{\xi}^i}. \tag{3.4}$$

3.2 Motivation: A problem of the textbook approach

A very common recipe for the derivation of equations of motion for mechanical systems, which we call the *standard Lagrange formalism*, can be found in many textbooks (e.g. [Murray et al., 1994, sec. 4.2], [Landau and Lifshitz, 1960, §5], ...):

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}^i} - \frac{\partial \mathcal{L}}{\partial q^i} = Q_i, \quad i = 1, \ldots, n. \tag{3.5}$$

where the Lagrangian $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{T}(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \mathcal{V}(\boldsymbol{q})$ for mechanical systems is the difference between kinetic energy $\mathcal{T}$ and potential energy $\mathcal{V}$ formulated in *generalized coordinates* $\boldsymbol{q}(t) \in \mathbb{R}^n$ and $\boldsymbol{Q}(t) \in \mathbb{R}^n$ are external, generalized forces.

For the system described above we must express the particle positions $\mathfrak{r} = \mathfrak{r}(\boldsymbol{q})$ in terms of the generalized coordinates such that the geometric constraints are automatically fulfilled, i.e. $\mathfrak{c}(\mathfrak{r}(\boldsymbol{q})) = \boldsymbol{0}\,\forall\, \boldsymbol{q} \in \mathbb{R}^n$. The kinetic energy in terms of the generalized coordinates is[1]

$$\mathcal{T}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \sum\nolimits_p \tfrac{1}{2}\mathfrak{m}_p \|\dot{\mathfrak{r}}_p(\boldsymbol{q}, \dot{\boldsymbol{q}})\|^2 = \tfrac{1}{2} \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \tfrac{\partial \mathfrak{r}_p}{\partial q^i}(\boldsymbol{q}), \tfrac{\partial \mathfrak{r}_p}{\partial q^j}(\boldsymbol{q}) \rangle}_{M_{ij}(\boldsymbol{q})} \dot{q}^i \dot{q}^j \tag{3.6}$$

where we introduced the *system inertia matrix* $\boldsymbol{M}(\boldsymbol{q}) \in \mathbb{R}^{n \times n}$. Having this particular structure for the Lagrangian, we can evaluate (3.5) to

$$M_{ij}\ddot{q}^j + \underbrace{\tfrac{1}{2}\big(\tfrac{\partial M_{ij}}{\partial q^k} + \tfrac{\partial M_{ik}}{\partial q^j} - \tfrac{\partial M_{jk}}{\partial q^i}\big)}_{\Gamma_{ijk}} \dot{q}^k \dot{q}^j + \tfrac{\partial \mathcal{V}}{\partial q^i} = Q_i, \quad i = 1, \ldots, n. \tag{3.7}$$

where we dropped the explicit dependencies $M_{ij} = M_{ij}(\boldsymbol{q})$ for the sake of readability. The terms Γ_{ijk} are called the *Christoffel symbols* for the inertia matrix M_{ij}.

[1] Here and throughout this work we use the summation convention $a^i b_i \equiv \sum_{i=1}^n a^i b_i$.

Example 6. A common way to parameterize the system from Example **??** with minimal coordinates $\boldsymbol{q}$ uses *Euler angles*: Set the particle positions as

$$\mathfrak{r}_1(\boldsymbol{q}) = l_1 \boldsymbol{R}_{\mathsf{x}}(\boldsymbol{q}), \qquad \mathfrak{r}_2(\boldsymbol{q}) = l_2 \boldsymbol{R}_{\mathsf{y}}(\boldsymbol{q}), \qquad \mathfrak{r}_3(\boldsymbol{q}) = l_3 \boldsymbol{R}_{\mathsf{z}}(\boldsymbol{q}) \tag{3.8}$$

where $\boldsymbol{R}_{\mathsf{x}}, \boldsymbol{R}_{\mathsf{y}}, \boldsymbol{R}_{\mathsf{z}}$ are the columns of a rotation matrix $\boldsymbol{R}$. In the *roll-pitch-yaw* convention $\boldsymbol{q} = [\alpha, \beta, \varphi]^\top$ and with the short notation $\mathrm{c}_\varphi = \cos\varphi$, $\mathrm{s}_\varphi = \sin\varphi$ this is

$$\boldsymbol{R}(\boldsymbol{q}) = [\boldsymbol{R}_{\mathsf{x}}(\boldsymbol{q}), \boldsymbol{R}_{\mathsf{y}}(\boldsymbol{q}), \boldsymbol{R}_{\mathsf{z}}(\boldsymbol{q})] = \begin{bmatrix} \mathrm{c}_\varphi \mathrm{c}_\beta & \mathrm{c}_\varphi \mathrm{s}_\beta \mathrm{s}_\alpha - \mathrm{s}_\varphi \mathrm{c}_\alpha & \mathrm{c}_\varphi \mathrm{s}_\beta \mathrm{c}_\alpha + \mathrm{s}_\varphi \mathrm{s}_\alpha \\ \mathrm{s}_\varphi \mathrm{c}_\beta & \mathrm{s}_\varphi \mathrm{s}_\beta \mathrm{s}_\alpha + \mathrm{c}_\varphi \mathrm{c}_\alpha & \mathrm{s}_\varphi \mathrm{s}_\beta \mathrm{c}_\alpha - \mathrm{c}_\varphi \mathrm{s}_\alpha \\ -\mathrm{s}_\beta & \mathrm{c}_\beta \mathrm{s}_\alpha & \mathrm{c}_\beta \mathrm{c}_\alpha \end{bmatrix}. \tag{3.9}$$

The kinetic energy in terms of these coordinates is

$$\mathcal{T} = \tfrac{1}{2} \underbrace{[\dot\alpha, \dot\beta, \dot\varphi]}_{\dot{\boldsymbol{q}}^\top} \underbrace{\begin{bmatrix} \Theta_{\mathsf{x}} & 0 & -\Theta_{\mathsf{x}} \mathrm{s}_\beta \\ 0 & \Theta_{\mathsf{y}} \mathrm{c}_\alpha^2 + \Theta_{\mathsf{z}} \mathrm{s}_\alpha^2 & (\Theta_{\mathsf{y}} - \Theta_{\mathsf{z}}) \mathrm{c}_\alpha \mathrm{s}_\alpha \mathrm{c}_\beta \\ -\Theta_{\mathsf{x}} \mathrm{s}_\beta & (\Theta_{\mathsf{y}} - \Theta_{\mathsf{z}}) \mathrm{c}_\alpha \mathrm{s}_\alpha \mathrm{c}_\beta & \Theta_{\mathsf{x}} \mathrm{s}_\beta^2 + (\Theta_{\mathsf{y}} \mathrm{s}_\alpha^2 + \Theta_{\mathsf{z}} \mathrm{c}_\alpha^2) \mathrm{c}_\beta^2 \end{bmatrix}}_{M(\boldsymbol{q})} \underbrace{\begin{bmatrix} \dot\alpha \\ \dot\beta \\ \dot\varphi \end{bmatrix}}_{\dot{\boldsymbol{q}}} \tag{3.10}$$

where

$$\Theta_{\mathsf{x}} = \mathfrak{m}_2 l_2^2 + \mathfrak{m}_3 l_3^2, \qquad \Theta_{\mathsf{y}} = \mathfrak{m}_1 l_1^2 + \mathfrak{m}_3 l_3^2, \qquad \Theta_{\mathsf{z}} = \mathfrak{m}_1 l_1^2 + \mathfrak{m}_2 l_2^2. \tag{3.11}$$

The equations of motion result by plugging the inertia matrix into (3.7). However, even before the evaluation we can already see a mayor problem: The inertia matrix $\boldsymbol{M}$, with $\det \boldsymbol{M} = \Theta_{\mathsf{x}} \Theta_{\mathsf{y}} \Theta_{\mathsf{z}} \mathrm{c}_\beta^2$, is singular at $\mathrm{c}_\beta = 0$. This means that the equations of motion cannot be solved at this point.

It should be stressed that there is no physical reason for the singularity in Example 6, it is rather a consequence of an unsuitable parameterization of the system. The problem arises from the fact the the configuration space $\mathfrak{X}$ determined by the geometric constraints (**??**), though having $\dim \mathfrak{X} = 3$, is not homeomorphic to $\mathbb{R}^3$.

3.3 Geometry and kinematics

The particle configuration space $\mathfrak{X} = \{\mathfrak{x} \in \mathbb{R}^{3\mathfrak{N}} \,|\, \mathfrak{c}(\mathfrak{x}) = \mathbf{0}\}$ can be seen as an embedded differentiable manifold with $\dim \mathfrak{X} = 3\mathfrak{N} - \operatorname{rank} \frac{\partial \mathfrak{c}}{\partial \mathfrak{x}} = n$. We call the manifold *nonlinear*, if it is *not homeomorphic* to $\mathbb{R}^n$, i.e. there exists no global, continuous function $\mathbb{R}^n \to \mathfrak{X}$ that has a continuous inverse (a global homeomorphism). The common way to tackle this in differential geometry is to use an atlas, a set of overlapping *local* charts. For the previous Example **??**, this was done four different charts in [Grafarend and Kühnel, 2011]. In this work we will not pursue this path.

3.3.1 Redundant configuration coordinates

Another way for a global parameterization of nonlinear configuration manifolds is motivated from the *Whitney embedding theorem* (see e.g. [Lee, 2003, Theo. 6.14]), that states: *Every smooth manifold of dimension n can be smoothly embedded in the Euclidean space $\mathbb{R}^{2n}$.* Note that $2n$ is a worst case bound, i.e. for a particular example a lower dimension for the embedding space might work and a higher dimension is permitted anyway. In the notation of this work, we use $\nu > 0$ generalized coordinates $\boldsymbol{x}(t) = [x^1(t), \dots, x^\nu(t)]^\top \in \mathbb{R}^\nu$ that might be constrained by $c \geq 0$ smooth functions of the form $\boldsymbol{\phi}(\boldsymbol{x}) = [\phi^1(\boldsymbol{x}), \dots, \phi^c(\boldsymbol{x})]^\top = \boldsymbol{0}$. For $c > 0$ these coordinates are not independent and are commonly called *redundant*. The set of mutually admissible coordinates is

$$\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^\nu \,|\, \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}\} \tag{3.12}$$

with $\dim \mathbb{X} = \nu - \operatorname{rank} \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} = n$. By the Whitney embedding theorem, for a suitable number of coordinates ν and suitable constraints $\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}$, this manifolds is homeomorphic to the particle configuration space $\mathbb{X} \cong \mathfrak{X}$. This means there is a global, invertible function $\mathfrak{x}(\boldsymbol{x})$ from the coordinates to the particle positions.

3.3.2 Minimal velocity coordinates

For the following it is crucial to note that a geometric constraint is equivalent to its derivative supplemented with a suitable initial condition

$$\phi^\kappa(\boldsymbol{x}) = 0 \tag{3.13a}$$

$$\Leftrightarrow \qquad \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x})\dot{x}^\alpha = 0, \qquad \phi^\kappa(\boldsymbol{x}_0) = 0 \tag{3.13b}$$

$$\Leftrightarrow \qquad \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x})\ddot{x}^\alpha + \frac{\partial^2 \phi^\kappa}{\partial x^\alpha \partial x^\beta}(\boldsymbol{x})\dot{x}^\beta \dot{x}^\alpha = 0, \qquad \phi^\kappa(\boldsymbol{x}_0) = 0, \; \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x}_0)\dot{x}_0^\alpha = 0 \tag{3.13c}$$

$$\dots$$

where $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$. Even though (3.13a) might be nonlinear, its derivative (3.13b) is always *linear* in the velocities $\dot{\boldsymbol{x}}$. So here it is reasonable to choose *minimal velocity coordinates*: Let $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu \times n}$ be a matrix with the properties $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\boldsymbol{A} = \boldsymbol{0}$ and $\operatorname{rank} \boldsymbol{A} = n$. The first property of $\boldsymbol{A}(\boldsymbol{x})$ is that these columns of $\boldsymbol{A}(\boldsymbol{x})$ are orthogonal to the rows of $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}(\boldsymbol{x})$. The second property implies that the columns of $\boldsymbol{A}(\boldsymbol{x})$ are linearly independent. So the columns of $\boldsymbol{A}(\boldsymbol{x})$ can be interpreted as a *basis vectors* for the tangent space $\mathsf{T}_{\boldsymbol{x}}\mathbb{X}$. We can capture all allowed velocities $\dot{\boldsymbol{x}}(t)$ by the minimal velocity coordinates $\boldsymbol{\xi}(t) \in \mathbb{R}^n$ through

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi} \tag{3.14}$$

This kinematic relation (3.14) ensures that the time derivative (3.13b) of the geometric constraint is fulfilled, and consequently the geometric constraint only has to be imposed on the initial condition $\boldsymbol{\phi}(\boldsymbol{x}(t_0)) = \boldsymbol{0}$.

Existence for A? guaranteed for Lie groups: the Lie algebra at the identity can be translated around the manifold by the group operation

Construction of $\boldsymbol{A}$ by matrix inversion

Example 7. Consider a single particle constrained to a circle of radius ρ as illustrated in Figure 3.1.

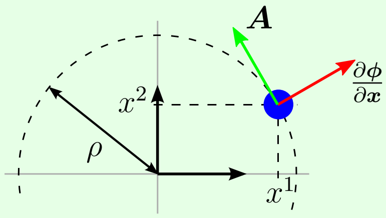


Figure 3.1: Particle on a circle

We use the its Cartesian position $[x^1, x^2]^\top \in \mathbb{R}^2$ constrained by $\phi = (x^1)^2 + (x^2)^2 - \rho^2 = 0$ as configuration coordinates. A reasonable choice for the kinematics matrix $\boldsymbol{A}$ is motivated from

$$\underbrace{\begin{bmatrix} 2x^1 & 2x^2 \end{bmatrix}}_{\frac{\partial \phi}{\partial \boldsymbol{x}}} \underbrace{\begin{bmatrix} -x^2 \\ x^1 \end{bmatrix}}_{\boldsymbol{A}} = 0 \tag{3.15}$$

Example 8. Consider again the system from Example **??**. Instead of parameterizing the rotation matrix $\boldsymbol{R}$ by minimal coordinates we now take its 9 coefficients $\boldsymbol{x} = [R^{\mathsf{x}}_{\mathsf{x}}, R^{\mathsf{y}}_{\mathsf{x}}, R^{\mathsf{z}}_{\mathsf{x}}, R^{\mathsf{x}}_{\mathsf{y}}, R^{\mathsf{y}}_{\mathsf{y}}, R^{\mathsf{z}}_{\mathsf{y}}, R^{\mathsf{x}}_{\mathsf{z}}, R^{\mathsf{y}}_{\mathsf{z}}, R^{\mathsf{z}}_{\mathsf{z}}]^\top \in \mathbb{R}^9$ as configuration coordinates. The constraints $\boldsymbol{R}^\top \boldsymbol{R} = \mathbf{I}_3$ and $\det \boldsymbol{R} = 1$ read

$$\boldsymbol{\phi}(\boldsymbol{x}) = \begin{bmatrix} (R^{\mathsf{x}}_{\mathsf{x}})^2 + (R^{\mathsf{y}}_{\mathsf{x}})^2 + (R^{\mathsf{z}}_{\mathsf{x}})^2 - 1 \\ (R^{\mathsf{x}}_{\mathsf{y}})^2 + (R^{\mathsf{y}}_{\mathsf{y}})^2 + (R^{\mathsf{z}}_{\mathsf{y}})^2 - 1 \\ (R^{\mathsf{x}}_{\mathsf{z}})^2 + (R^{\mathsf{y}}_{\mathsf{z}})^2 + (R^{\mathsf{z}}_{\mathsf{z}})^2 - 1 \\ R^{\mathsf{x}}_{\mathsf{y}} R^{\mathsf{x}}_{\mathsf{z}} + R^{\mathsf{y}}_{\mathsf{y}} R^{\mathsf{y}}_{\mathsf{z}} + R^{\mathsf{z}}_{\mathsf{y}} R^{\mathsf{z}}_{\mathsf{z}} \\ R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{x}}_{\mathsf{z}} + R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{z}} + R^{\mathsf{z}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{z}} \\ R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{x}}_{\mathsf{y}} + R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{y}} + R^{\mathsf{z}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{y}} \\ R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{y}} R^{\mathsf{z}}_{\mathsf{z}} + R^{\mathsf{x}}_{\mathsf{y}} R^{\mathsf{y}}_{\mathsf{z}} R^{\mathsf{z}}_{\mathsf{x}} + R^{\mathsf{x}}_{\mathsf{z}} R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{y}} - R^{\mathsf{x}}_{\mathsf{x}} R^{\mathsf{y}}_{\mathsf{z}} R^{\mathsf{z}}_{\mathsf{y}} - R^{\mathsf{x}}_{\mathsf{y}} R^{\mathsf{y}}_{\mathsf{x}} R^{\mathsf{z}}_{\mathsf{z}} - R^{\mathsf{x}}_{\mathsf{z}} R^{\mathsf{y}}_{\mathsf{y}} R^{\mathsf{z}}_{\mathsf{x}} - 1 \end{bmatrix} = \mathbf{0}. \tag{3.16}$$

The 9 conditions $\boldsymbol{R}^\top \boldsymbol{R} = \mathbf{I}_3$ yields due to symmetry only 6 constraints and already imply $\det \boldsymbol{R} = \pm 1$. Since the determinant is a smooth function, the corresponding manifold must consist of two disjoint components, one with $\det \boldsymbol{R} = +1$ (proper rotations) and one with $\det \boldsymbol{R} = -1$ (rotations with reflection). So the additional constraint $\det \boldsymbol{R} = +1$ does not change the dimension of the configuration space. Formally this means $\operatorname{rank} \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} = 6$ and consequently $\dim \mathbb{X} = 9 - 6 = 3$. A kinematics matrix with $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} \boldsymbol{A} = \mathbf{0}$ and $\operatorname{rank} \boldsymbol{A} = 3$

is given by

$$\boldsymbol{A}(\boldsymbol{x}) = \begin{bmatrix} 0 & -R_\mathsf{z}^\mathsf{x} & R_\mathsf{y}^\mathsf{x} \\ 0 & -R_\mathsf{z}^\mathsf{y} & R_\mathsf{y}^\mathsf{y} \\ 0 & -R_\mathsf{z}^\mathsf{z} & R_\mathsf{y}^\mathsf{z} \\ R_\mathsf{z}^\mathsf{x} & 0 & -R_\mathsf{x}^\mathsf{x} \\ R_\mathsf{z}^\mathsf{y} & 0 & -R_\mathsf{x}^\mathsf{y} \\ R_\mathsf{z}^\mathsf{z} & 0 & -R_\mathsf{x}^\mathsf{z} \\ -R_\mathsf{y}^\mathsf{x} & R_\mathsf{x}^\mathsf{x} & 0 \\ -R_\mathsf{y}^\mathsf{y} & R_\mathsf{x}^\mathsf{y} & 0 \\ -R_\mathsf{y}^\mathsf{z} & R_\mathsf{x}^\mathsf{z} & 0 \end{bmatrix}. \tag{3.17}$$

The resulting kinematic equation $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}$ can be reordered to the matrix equation $\dot{\boldsymbol{R}} = \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\xi})$ by introducing the *wedge operator* defined as

$$\mathrm{wed}\begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} = \begin{bmatrix} 0 & -\xi^3 & \xi^2 \\ \xi^3 & 0 & -\xi^1 \\ -\xi^2 & \xi^1 & 0 \end{bmatrix}. \tag{3.18}$$

Some identities involving the pseudo-inverse. For any matrix $\boldsymbol{S} \in \mathbb{R}^{m\times n}$ there exists a unique *(Moore-Penrose) pseudoinverse* $\boldsymbol{S}^+ \in \mathbb{R}^{n\times m}$ determined by the following conditions [Penrose, 1955, Theo. 1]:

$$\boldsymbol{S}\boldsymbol{S}^+\boldsymbol{S} = \boldsymbol{S}, \tag{3.19a}$$
$$\boldsymbol{S}^+\boldsymbol{S}\boldsymbol{S}^+ = \boldsymbol{S}^+, \tag{3.19b}$$
$$(\boldsymbol{S}\boldsymbol{S}^+)^\top = \boldsymbol{S}\boldsymbol{S}^+, \tag{3.19c}$$
$$(\boldsymbol{S}^+\boldsymbol{S})^\top = \boldsymbol{S}^+\boldsymbol{S}. \tag{3.19d}$$

If the matrix $\boldsymbol{S}$ has linearly independent columns, its pseudoinverse is $\boldsymbol{S}^+ = (\boldsymbol{S}^\top\boldsymbol{S})^{-1}\boldsymbol{S}^\top$. Similarly, if $\boldsymbol{S}$ has linearly independent rows, its pseudoinverse is $\boldsymbol{S}^+ = \boldsymbol{S}^\top(\boldsymbol{S}\boldsymbol{S}^\top)^{-1}$.

Define $\boldsymbol{Y}(\boldsymbol{x}) \in \mathbb{R}^{n\times\nu}$ as $\boldsymbol{Y} = \boldsymbol{A}^+ = (\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top$, i.e. the pseudoinverse of the kinematics matrix $\boldsymbol{A}$. Note that this implies $\boldsymbol{Y}\boldsymbol{A} = \mathbf{I}_n$, but $\boldsymbol{A}\boldsymbol{Y} \neq \mathbf{I}_\nu$. We also introduce the matrices $\boldsymbol{\Phi} = \frac{\partial\phi}{\partial\boldsymbol{x}}$ and $\boldsymbol{\Psi} = \boldsymbol{\Phi}^+$. With $\boldsymbol{\Phi}\boldsymbol{A} = \mathbf{0}$ and the Penrose conditions (3.19), we can show[2] that $\boldsymbol{\Psi}^\top\boldsymbol{A} = \mathbf{0}$ and $\boldsymbol{Y}^\top\boldsymbol{\Phi} = \mathbf{0}$. Furthermore, since $\operatorname{rank}\boldsymbol{\Psi} = \operatorname{rank}\boldsymbol{\Phi} = \nu - n$ the columns of $\boldsymbol{\Psi}(\boldsymbol{x})$ span the complementary space $(\mathsf{T}_{\boldsymbol{x}}\mathbb{X})^\perp$ though they might not be a basis since the columns might not be linearly independent.

The matrix $\boldsymbol{P} = \boldsymbol{A}\boldsymbol{Y}$ is an *orthogonal projector*, i.e. $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\boldsymbol{P}^\top = \boldsymbol{P}$ which result directly from the Penrose conditions (3.19). It is in fact the unique orthogonal projector from $\mathbb{R}^\nu$ to the tangent space $\mathsf{T}_{\boldsymbol{x}}\mathbb{X}$. The unique orthogonal projector to the complementary space $(\mathsf{T}_{\boldsymbol{x}}\mathbb{X})^\perp$ is $\boldsymbol{P}^\perp = \mathbf{I}_\nu - \boldsymbol{P}$. On the other hand, since $\boldsymbol{\Psi}$ spans the complementary space the complementary projector can also be expressed as $\boldsymbol{P}^\perp = \boldsymbol{\Psi}\boldsymbol{\Phi}$. This leads to the

[2] $\boldsymbol{\Psi}^\top\boldsymbol{A} = (\boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{\Psi})^\top\boldsymbol{A} = \boldsymbol{\Psi}^\top(\boldsymbol{\Psi}\boldsymbol{\Phi})^\top\boldsymbol{A} = \boldsymbol{\Psi}^\top\boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{A} = \mathbf{0}$

identity

$$\underbrace{\boldsymbol{A}\boldsymbol{Y}}_{\boldsymbol{P}} + \underbrace{\boldsymbol{\Psi}\boldsymbol{\Phi}}_{\boldsymbol{P}^{\perp}} = \mathbf{I}_{\nu}. \tag{3.20}$$

3.3.3 Directional derivative and Hessian

Consider a function $\mathcal{V} : \mathbb{R}^{\nu} \to \mathbb{R}$ and a curve $\boldsymbol{x} : \mathbb{R} \to \mathbb{X}$. Since $\mathbb{X} \subset \mathbb{R}^{\nu}$, their composition $\mathcal{V} \circ \boldsymbol{x} = f : \mathbb{R} \to \mathbb{R}$ is a scalar function and has the the Taylor expansion

$$\begin{aligned}\underbrace{\mathcal{V}(\boldsymbol{x}(t))}_{f(t)} = \underbrace{\mathcal{V}(\boldsymbol{x}(0))}_{f(0)} + t \underbrace{\frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\boldsymbol{x}(0))\dot{x}^{\alpha}(0)}_{\dot{f}(0)} \\ + \tfrac{1}{2}t^2 \underbrace{\Big(\frac{\partial^2 \mathcal{V}}{\partial x^{\beta}\partial x^{\alpha}}(\boldsymbol{x}(0))\dot{x}^{\alpha}(0)\dot{x}^{\beta}(0) + \frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\boldsymbol{x}(0))\ddot{x}^{\alpha}(0)\Big)}_{\ddot{f}(0)} + \mathcal{O}(t^3).\end{aligned} \tag{3.21}$$

Now let the curve be parameterized by $\dot{\boldsymbol{x}}(t) = \boldsymbol{A}(\boldsymbol{x}(t))\boldsymbol{\xi}(t)$ and we use the shorthand notations $\bar{\boldsymbol{x}} = \boldsymbol{x}(0)$, $\bar{\boldsymbol{\xi}} = \boldsymbol{\xi}(0)$ and $\bar{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{x}(0))$ to write

$$\begin{aligned}\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + t \frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\bar{\boldsymbol{x}})\bar{A}^{\alpha}_{i}\bar{\xi}^{i} \\ + \tfrac{1}{2}t^2\Big(\frac{\partial^2 \mathcal{V}}{\partial x^{\beta}\partial x^{\alpha}}(\bar{\boldsymbol{x}})\bar{A}^{\alpha}_{i}\bar{A}^{\beta}_{j}\bar{\xi}^{i}\bar{\xi}^{j} + \frac{\partial \mathcal{V}}{\partial x^{\alpha}}(\bar{\boldsymbol{x}})\Big(\frac{\partial A^{\alpha}_{i}}{\partial x^{\beta}}(\bar{\boldsymbol{x}})\bar{A}^{\beta}_{j}\bar{\xi}^{i}\bar{\xi}^{j} + \bar{A}^{\alpha}_{i}\dot{\bar{\xi}}^{i}\Big)\Big) + \mathcal{O}(t^3)\end{aligned} \tag{3.22}$$

Introducing the notation

$$\partial_i = A^{\alpha}_{i}\frac{\partial}{\partial x^{\alpha}}, \ i = 1, \ldots, n \tag{3.23}$$

for the derivative in the direction of the i-th basis vector, we can state the Taylor expansion as

$$\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + t\,\partial_i\mathcal{V}(\bar{\boldsymbol{x}})\bar{\xi}^{i} + \tfrac{1}{2}t^2\big(\partial_i\partial_j\mathcal{V}(\bar{\boldsymbol{x}})\bar{\xi}^{i}\bar{\xi}^{j} + \partial_i\mathcal{V}(\bar{\boldsymbol{x}})\dot{\bar{\xi}}^{i}\big) + \mathcal{O}(t^3). \tag{3.24}$$

There are two more things we can derive from this equation:

- If $\partial_i\mathcal{V}(\bar{\boldsymbol{x}}) = 0, i = 1, \ldots, n$ then $\bar{\boldsymbol{x}}$ is called a *critical point* of $\mathcal{V}$. At a critical point the expansion (3.24) reduces to

  $$\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + \tfrac{1}{2}t^2 \underbrace{(\partial_i\partial_j\mathcal{V})(\bar{\boldsymbol{x}})}_{\bar{H}_{ij}}\bar{\xi}^{i}\bar{\xi}^{j} + \mathcal{O}(t^3). \tag{3.25}$$

  This relation holds for any sufficiently smooth curve $t \mapsto \boldsymbol{x}(t)$ through $\bar{\boldsymbol{x}}$ and consequently for any velocity vector $\bar{\boldsymbol{\xi}}$ at the critical point. So if the matrix $\bar{\boldsymbol{H}}$ is positive (negative) definite, then $\bar{\boldsymbol{x}}$ is a local minimum (maximum) of $\mathcal{V}$.

- Assume the curve $t \mapsto \boldsymbol{x}(t)$ is a *geodesic*, i.e. $\dot{\xi}^i = -\Gamma^i_{jk}\xi^j\xi^k$ with the connection coefficients Γ^i_{jk} that will be discussed later. Plugging this into (3.24) we find a coordinate form of the *Hessian tensor* $\boldsymbol{\nabla}^2\mathcal{V}$ of the potential:
$$\mathcal{V}(\boldsymbol{x}(t)) = \mathcal{V}(\bar{\boldsymbol{x}}) + t(\partial_i\mathcal{V})(\bar{\boldsymbol{x}})\bar{\xi}^i + \tfrac{1}{2}t^2\underbrace{\big(\partial_i\partial_j\mathcal{V} - \Gamma^k_{ij}\partial_k\mathcal{V}\big)}_{(\boldsymbol{\nabla}^2\mathcal{V})_{ij}}(\bar{\boldsymbol{x}})\bar{\xi}^i\bar{\xi}^j + \mathcal{O}(t^3). \tag{3.26}$$
At a critical point $\bar{\boldsymbol{x}}$, the Hessian of the potential is independent of the connection coefficients Γ^i_{jk} and consequently of the underlying metric. There it coincides with the matrix $\bar{\boldsymbol{H}}$ defined in (3.25).

Example 9. Consider the function
$$\mathcal{V}(\boldsymbol{R}) = \operatorname{tr}\big(\boldsymbol{\Pi}'(\mathbf{I}_3 - \boldsymbol{R})\big). \tag{3.27}$$
proposed in [Koditschek, 1989], where $\boldsymbol{\Pi}' = \boldsymbol{\Pi}'^\top \in \mathbb{R}^{3\times3}$ and $\boldsymbol{R} \in \mathbb{SO}(3)$. It can be regarded as a distance function on the configuration space $\mathbb{X} = \mathbb{SO}(3)$ with six tunable parameters which makes it quite useful for control purposes, see e.g. [Bullo and Murray, 1999] or [Lee et al., 2010]. This function will appear several times in later chapters, so it will be discussed in detail here.

As in the previous example, we regard the coefficients of the rotation matrix $\boldsymbol{R}$ as redundant configuration coordinates. Using the basis from (3.17) we compute the differential and the Hessian at a critical point $\boldsymbol{R}_0$ as
$$\boldsymbol{\nabla}\mathcal{V}(\boldsymbol{R}_0) = \operatorname{vee2}(\boldsymbol{\Pi}'\boldsymbol{R}_0) \overset{!}{=} \mathbf{0}, \tag{3.28a}$$
$$\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_0) = \operatorname{tr}(\boldsymbol{\Pi}'\boldsymbol{R}_0)\mathbf{I}_3 - \boldsymbol{\Pi}'\boldsymbol{R}_0. \tag{3.28b}$$
with the vee2 operator defined as
$$\operatorname{vee2} : \mathbb{R}^{3\times3} \to \mathbb{R}^3, \begin{bmatrix} * & A_{12} & A_{13} \\ A_{21} & * & A_{23} \\ A_{31} & A_{32} & * \end{bmatrix} \mapsto \begin{bmatrix} A_{32} - A_{23} \\ A_{13} - A_{31} \\ A_{21} - A_{12} \end{bmatrix}. \tag{3.29}$$
So the condition for a critical point at $\boldsymbol{R}_0$ is that the product $\boldsymbol{\Pi}'\boldsymbol{R}_0$ is skew-symmetric. Obviously the function $\mathcal{V}$ has a critical point at
$$\boldsymbol{R}_{0,1} = \mathbf{I}_3 \ : \qquad \mathcal{V}(\boldsymbol{R}_{0,1}) = 0, \qquad \boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,1}) = \operatorname{tr}(\boldsymbol{\Pi}')\mathbf{I}_3 - \boldsymbol{\Pi}' =: \boldsymbol{\Pi}.$$
It is a minimum if the matrix $\boldsymbol{\Pi}$ is positive definite.

There are more critical points: For their investigation it will be useful to consider the eigenvalue decomposition $\boldsymbol{\Pi}' = \boldsymbol{X}\boldsymbol{\Lambda}'\boldsymbol{X}^\top$ with $\boldsymbol{\Lambda}' = \operatorname{diag}(\lambda_1', \lambda_2', \lambda_3')$, $\boldsymbol{X} \in \mathbb{SO}(3)$. Note that the matrix $\boldsymbol{\Pi}$ has the same eigenvectors, i.e. $\boldsymbol{\Pi} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^\top$ with $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ and the eigenvalues are related by
$$\boldsymbol{\Lambda} = \operatorname{tr}(\boldsymbol{\Lambda}')\mathbf{I}_3 - \boldsymbol{\Lambda}' \qquad \Leftrightarrow \qquad \boldsymbol{\Lambda}' = \tfrac{1}{2}\operatorname{tr}(\boldsymbol{\Lambda})\mathbf{I}_3 - \boldsymbol{\Lambda} \tag{3.30a}$$
$$\left.\begin{aligned} \lambda_1 &= \lambda_2' + \lambda_3', \\ \lambda_2 &= \lambda_3' + \lambda_1', \\ \lambda_3 &= \lambda_1' + \lambda_2' \end{aligned}\right\} \qquad \Leftrightarrow \qquad \left\{\begin{aligned} \lambda_1' &= \tfrac{1}{2}(\lambda_2 + \lambda_3 - \lambda_1), \\ \lambda_2' &= \tfrac{1}{2}(\lambda_3 + \lambda_1 - \lambda_2), \\ \lambda_3' &= \tfrac{1}{2}(\lambda_1 + \lambda_2 - \lambda_3) \end{aligned}\right. \tag{3.30b}$$

Using the transformed orientation $\tilde{\boldsymbol{R}}_0 = \boldsymbol{X}^\top \boldsymbol{R}_0 \boldsymbol{X}$ we have

$$\boldsymbol{\nabla}\mathcal{V}(\boldsymbol{R}_0) = \boldsymbol{X}\,\mathrm{vee2}(\boldsymbol{\Lambda}'\tilde{\boldsymbol{R}}_0) \stackrel{!}{=} \mathbf{0}, \tag{3.31a}$$

$$\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_0) = \boldsymbol{X}\big(\,\mathrm{tr}(\boldsymbol{\Lambda}'\tilde{\boldsymbol{R}}_0)\mathbf{I}_3 - \boldsymbol{\Lambda}'\tilde{\boldsymbol{R}}_0\big)\boldsymbol{X}^\top. \tag{3.31b}$$

For critical points in addition to $\boldsymbol{R}_{0,1}$ we need to distinguish the following cases:

- *distinct eigenvalues* $\lambda_i \neq \lambda_j \Leftrightarrow \lambda'_i \neq \lambda'_j,\ i \neq j$:

$$\begin{aligned}
\tilde{\boldsymbol{R}}_{0,2} &= \mathrm{diag}(1,-1,-1)\ :\ \ \mathcal{V}(\boldsymbol{R}_{0,2}) = 2\lambda_1,\ \ \mathrm{eig}\big(\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,2})\big) = \{-\lambda_1, \lambda_3-\lambda_1, \lambda_2-\lambda_1\}\\
\tilde{\boldsymbol{R}}_{0,3} &= \mathrm{diag}(-1,1,-1)\ :\ \ \mathcal{V}(\boldsymbol{R}_{0,3}) = 2\lambda_2,\ \ \mathrm{eig}\big(\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,3})\big) = \{\lambda_3-\lambda_2, -\lambda_2, \lambda_1-\lambda_2\}\\
\tilde{\boldsymbol{R}}_{0,4} &= \mathrm{diag}(-1,-1,1)\ :\ \ \mathcal{V}(\boldsymbol{R}_{0,4}) = 2\lambda_3,\ \ \mathrm{eig}\big(\boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_{0,4})\big) = \{\lambda_2-\lambda_3, \lambda_1-\lambda_3, -\lambda_3\}
\end{aligned}$$

- *double eigenvalue* $\lambda_1 = \lambda_2 \neq \lambda_3$: the function $\mathcal{V}$ is stationary at the point $\boldsymbol{R}_{0,4}$ and on the circular submanifold

$$\tilde{\boldsymbol{R}}_{0,5} = \begin{bmatrix} -c & s & 0\\ s & c & 0\\ 0 & 0 & -1 \end{bmatrix},\ c^2+s^2 = 1:$$

$$\mathcal{V}(\boldsymbol{R}_{0,5}) = 2\lambda_1, \qquad \mathrm{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,5})\big) = \{-\lambda_1, \lambda_3-\lambda_1, 0\}$$

  which includes the points $\boldsymbol{R}_{0,2}$ and $\boldsymbol{R}_{0,3}$. For the cases $\lambda_1 = \lambda_3 \neq \lambda_2$ and $\lambda_2 = \lambda_3 \neq \lambda_1$ we have analogous results.

- *triple eigenvalue* $\lambda_1 = \lambda_2 = \lambda_3 \Rightarrow \boldsymbol{\Lambda}' = \boldsymbol{\Pi}' = \frac{1}{2}\lambda_1\mathbf{I}_3,\ \boldsymbol{X} = \mathbf{I}_3$: The function $\mathcal{V}$ is stationary on the spherical submanifold

$$\boldsymbol{R}_{0,6} = \begin{bmatrix} 1-2(a_2^2+a_3^2) & 2a_2a_1 & 2a_1a_3\\ 2a_1a_2 & 1-2(a_1^2+a_3^2) & 2a_2a_3\\ 2a_1a_3 & 2a_2a_3 & 1-2(a_1^2+a_2^2) \end{bmatrix},\ a_1^2+a_2^2+a_3^2 = 1\ :$$

$$\mathcal{V}(\boldsymbol{R}_{0,6}) = 2\lambda_1, \qquad \mathrm{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,6})\big) = \{-\lambda_1, 0, 0\}$$

  which includes $\tilde{\boldsymbol{R}}_{0,2}$, $\tilde{\boldsymbol{R}}_{0,3}$ and $\tilde{\boldsymbol{R}}_{0,4}$ The matrix $\boldsymbol{R}_{0,6}$ is a 180° rotation about the unit axis $\boldsymbol{a} = [a_1, a_2, a_3]^\top$.

If one or more eigenvalues are zero we have even more critical points

- *one zero eigenvalue* $\lambda_3 = 0$: the function $\mathcal{V}$ is stationary on the circular submanifold

$$\tilde{\boldsymbol{R}}_{0,7} = \begin{bmatrix} c & -s & 0\\ s & c & 0\\ 0 & 0 & 1 \end{bmatrix},\ c^2+s^2 = 1:$$

$$\mathcal{V}(\boldsymbol{R}_{0,7}) = 0, \qquad \mathrm{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,5})\big) = \{\lambda_1, \lambda_2, 0\}$$

  which includes the points $\boldsymbol{R}_{0,1} = \mathbf{I}_3$ and $\boldsymbol{R}_{0,4}$.

- *two zero eigenvalues* $\lambda_2 = \lambda_3 = 0$: the function $\mathcal{V}$ is stationary on the spherical submanifold

$$\tilde{\boldsymbol{R}}_{0,8} = \begin{bmatrix} q_\mathsf{w}^2 - q_\mathsf{y}^2 - q_\mathsf{z}^2 & -2q_\mathsf{w}q_\mathsf{z} & 2q_\mathsf{w}q_\mathsf{y} \\ 2q_\mathsf{w}q_\mathsf{z} & q_\mathsf{w}^2 + q_\mathsf{y}^2 - q_\mathsf{z}^2 & 2q_\mathsf{y}q_\mathsf{z} \\ -2q_\mathsf{w}q_\mathsf{y} & 2q_\mathsf{y}q_\mathsf{z} & q_\mathsf{w}^2 - q_\mathsf{y}^2 + q_\mathsf{z}^2 \end{bmatrix}, \; q_\mathsf{w}^2 + q_\mathsf{y}^2 + q_\mathsf{z}^2 = 1 :$$
$$\mathcal{V}(\boldsymbol{R}_{0,8}) = 0, \qquad \operatorname{eig}\big(\nabla^2\mathcal{V}(\boldsymbol{R}_{0,8})\big) = \{\lambda_1, 0, 0\}$$

  which includes the points $\boldsymbol{R}_{0,1} = \mathbf{I}_3$ and $\boldsymbol{R}_{0,3}$, $\boldsymbol{R}_{0,4}$.

- *three zero eigenvalues* $\lambda_3 = \lambda_2 = \lambda_3 = 0 \;\Rightarrow\; \boldsymbol{\mathit{\Pi}} = \boldsymbol{\mathit{\Pi}}' = \mathbf{0}$: the function $\mathcal{V}$ degenerates to a constant $\mathcal{V}(\boldsymbol{R}) = 0 \,\forall\, \boldsymbol{R} \in \mathbb{SO}(3)$.

Comparing the eigenvalues $\operatorname{eig}(\nabla^2\mathcal{V})$ of the Hessians at the critical points, it is clear that the function $\mathcal{V}$ always has exactly one minimum and exactly one maximum, i.e. all other critical points are saddles, independently of the signs of λ_1, λ_2 and λ_3. If one or more $\lambda_i = 0$ or $\lambda_i = \lambda_j$ then the minimum or maximum is taken not on a single point but on a one or two-dimensional submanifold.

The most useful case for the following is $\lambda_1, \lambda_2, \lambda_3 > 0$, i.e. $\boldsymbol{\mathit{\Pi}}$ is positive definite and $\boldsymbol{R} = \mathbf{I}_3$ is the global minimum of $\mathcal{V}$. It is also worth noting that for $\boldsymbol{\mathit{\Pi}} = \mathbf{I}_3$ the function $\mathcal{V}$ is related to the angle θ of the rotation matrix $\boldsymbol{R}$ by $\mathcal{V} = \frac{1}{2}\operatorname{tr}(\mathbf{I}_3 - \boldsymbol{R}) = 1 - \cos\theta$.

3.3.4 Commutation coefficients

For a function $\mathcal{V} : \mathbb{R}^\nu \to \mathbb{R}$ we are used to the fact that partial derivatives commute, i.e. $\partial^2\mathcal{V}/\partial x^\alpha \partial x^\beta = \partial^2\mathcal{V}/\partial x^\beta \partial x^\alpha$. Unfortunately this is (in general) not the case for a directional derivatives like ∂_i defined in (3.23). Consequently we investigate the following commutation relation

$$\begin{aligned} \partial_i\partial_j\mathcal{V} - \partial_j\partial_i\mathcal{V} &= A_i^\alpha \frac{\partial}{\partial x^\alpha}\left(A_j^\beta \frac{\partial\mathcal{V}}{\partial x^\beta}\right) - A_j^\beta \frac{\partial}{\partial x^\beta}\left(A_i^\alpha \frac{\partial\mathcal{V}}{\partial x^\alpha}\right) \\ &= A_i^\alpha \frac{\partial A_j^\beta}{\partial x^\alpha}\frac{\partial\mathcal{V}}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\frac{\partial\mathcal{V}}{\partial x^\alpha} + A_i^\alpha A_j^\beta \underbrace{\left(\frac{\partial^2\mathcal{V}}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2\mathcal{V}}{\partial x^\beta \partial x^\alpha}\right)}_{=0} \\ &= \left(A_i^\beta \frac{\partial A_j^\alpha}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\right)\frac{\partial\mathcal{V}}{\partial x^\alpha}. \end{aligned} \tag{3.32}$$

Now using the identity (3.20) with $\varPhi^\kappa_\alpha = \frac{\partial \phi^\kappa}{\partial x^\alpha}$ and $\varPhi^\kappa_\alpha A^\alpha_i = 0 \Rightarrow \varPhi^\kappa_\alpha \frac{\partial A^\alpha_i}{\partial x^\beta} = -\frac{\partial \varPhi^\kappa_\alpha}{\partial x^\beta} A^\alpha_i$ to shape this expressions a bit further

$$
\begin{aligned}
\partial_i \partial_j \mathcal{V} - \partial_j \partial_i \mathcal{V} &= \left(A^\beta_i \frac{\partial A^\alpha_j}{\partial x^\beta} - A^\beta_j \frac{\partial A^\alpha_i}{\partial x^\beta} \right) \overbrace{\left(A^\sigma_k Y^k_\alpha + \varPsi^\sigma_\kappa \varPhi^\kappa_\alpha \right)}^{\delta^\sigma_\alpha} \frac{\partial \mathcal{V}}{\partial x^\sigma} \\
&= \left(A^\beta_i \frac{\partial A^\alpha_j}{\partial x^\beta} - A^\beta_j \frac{\partial A^\alpha_i}{\partial x^\beta} \right) Y^k_\alpha A^\sigma_k \frac{\partial \mathcal{V}}{\partial x^\sigma} - \left(A^\beta_i A^\alpha_j \frac{\partial \varPhi^\kappa_\alpha}{\partial x^\beta} - A^\beta_j A^\alpha_i \frac{\partial \varPhi^\kappa_\alpha}{\partial x^\beta} \right) \varPsi^\sigma_\kappa \frac{\partial \mathcal{V}}{\partial x^\sigma} \\
&= \underbrace{\left(A^\beta_i \frac{\partial A^\alpha_j}{\partial x^\beta} - A^\beta_j \frac{\partial A^\alpha_i}{\partial x^\beta} \right) Y^k_\alpha}_{\gamma^k_{ij}} \underbrace{A^\sigma_k \frac{\partial \mathcal{V}}{\partial x^\sigma}}_{\partial_k \mathcal{V}} - A^\beta_j A^\alpha_i \underbrace{\left(\frac{\partial^2 \phi^\kappa}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2 \phi^\kappa}{\partial x^\beta \partial x^\alpha} \right)}_{=0} \varPsi^\sigma_\kappa \frac{\partial \mathcal{V}}{\partial x^\sigma}.
\end{aligned}
\tag{3.33}
$$

Since this relation holds for any function $\mathcal{V}$ we can state it in operator form and introduce the *commutation coefficients* γ^k_{ij} as

$$
\partial_i \partial_j - \partial_j \partial_i = \gamma^k_{ij} \partial_k, \qquad \gamma^k_{ij} = \left(\partial_i A^\alpha_j - \partial_j A^\alpha_i \right) (A^+)^k_\alpha. \tag{3.34}
$$

It is interesting to note that the commutation coefficients are *invariant* to the choice of configuration coordinates $\boldsymbol{x}$, even though the coordinates appear explicitly in the definition: For a change of configuration coordinates $\boldsymbol{x} = f(\hat{\boldsymbol{x}})$ the commutation symbols transform like $\hat{\gamma}^k_{ij}(\hat{\boldsymbol{x}}) = \gamma^k_{ij}(f(\hat{\boldsymbol{x}}))$. This might be obvious from a geometric point of view, but the explicit calculation of the coordinate transformation is shown in see section A.3. It will even turn out that for most of our examples the coefficients will be constants.

The right hand side of (3.34) appears in the context of Lagrange's equation in [Boltzmann, 1902] and [Hamel, 1904] for the case of minimal configuration coordinates and consequently with square matrices $\boldsymbol{Y}$ and $\boldsymbol{A}$. In the contemporary literature on this context these quantities γ^k_{ij} are sometimes called the *Boltzmann three-index symbols* [Lurie, 2002, sec. 1.8] or *Hamel coefficients* [Bremer, 2008, p. 75]. The left hand side of (3.34) appears in the context of tensor algebra in [Misner et al., 1973, Box 8.4] where γ^k_{ij} are called the *commutation coefficients*. From the way γ^k_{ij} is defined here (3.34), this naming seems most fitting.

The commutation coefficients γ^i_{jk} vanish if the corresponding velocity coordinates ξ^i are *integrable*, i.e.

$$
\begin{aligned}
\exists\ \pi^i : \dot{\pi}^i = \xi^i = Y^i_\alpha \dot{x}^\alpha \qquad &\Rightarrow \qquad Y^i_\alpha = \frac{\partial \pi^i}{\partial x^\beta} \\
&\Rightarrow \qquad \frac{\partial Y^i_\alpha}{\partial x^\alpha} = \frac{\partial^2 \pi^i}{\partial x^\beta \partial x^\alpha} = \frac{\partial Y^i_\beta}{\partial x^\alpha}, \qquad \Rightarrow \qquad \gamma^i_{jk} = 0
\end{aligned}
\tag{3.35}
$$

which is not the case in general. Nevertheless the quantities π are commonly introduced as *nonholonomic coordinates* in [Boltzmann, 1902] (also called *quasi coordinates* in [Lurie, 2002, sec. 1.5]). Then we could write $\partial_i(\partial_j f) - \partial_j(\partial_i f) = \partial^2 f / \partial \pi^i \partial \pi^j - \partial^2 f / \partial \pi^j \partial \pi^i \neq 0$ what might lead to the conception that partial derivatives do not commute. The commutativity clearly holds, the issue is rather that the coordinates π do not exist. To avoid confusion of this kind we do not pick up this notation here.

Example 10. The commutation coefficients γ_{ij}^k associated with the kinematics matrix $\boldsymbol{A}$ from (3.17) are

$$\gamma_{ij}^k = \begin{cases} +1, & (i,j,k) \text{ even permutation of } (1,2,3) \\ -1, & (i,j,k) \text{ odd permutation of } (1,2,3) \\ 0, & \text{else} \end{cases} . \tag{3.36}$$

This coincides with the three dimensional Levi-Civita symbol. It is related to the 3 dimensional *cross product* by $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3 \, : \, [\gamma_{ij}^k a^i b^j]_{k=1..3} = \boldsymbol{a} \times \boldsymbol{b}$.

3.3.5 Linearization about a trajectory

Let $\bar{\boldsymbol{x}} : [t_1, t_2] \to \mathbb{X}$ be a smooth curve with the velocity coordinates $\bar{\boldsymbol{\xi}} : [t_1, t_2] \to \mathbb{R}^n : t \mapsto \boldsymbol{A}^+(\bar{\boldsymbol{x}}(t))\dot{\bar{\boldsymbol{x}}}(t)$. For a small deviation $\boldsymbol{x} \approx \bar{\boldsymbol{x}}$ with $\boldsymbol{x} \in \mathbb{X}$ we may approximate the geometric constraint as

$$\boldsymbol{\phi}(\boldsymbol{x}) \approx \underbrace{\boldsymbol{\phi}(\bar{\boldsymbol{x}})}_{=\boldsymbol{0}} + \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) = \boldsymbol{0}. \tag{3.37}$$

Since this constraint is affine w.r.t. $\boldsymbol{x}$ it is reasonable to use a the basis $\boldsymbol{\varepsilon}(t) \in \mathbb{R}^n$ for the deviated configuration coordinates:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{A}(\bar{\boldsymbol{x}})\boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} = \boldsymbol{A}^+(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}), \tag{3.38}$$

For the velocity coordinates $\boldsymbol{\xi}$ of the deviated curve $\boldsymbol{x}$ we use again the first order approximation and $\boldsymbol{Y} = \boldsymbol{A}^+$:

$$\begin{aligned} \xi^i &= Y_\alpha^i(\boldsymbol{x})\dot{x}^\alpha \\ &\approx Y_\alpha^i(\bar{\boldsymbol{x}} + \boldsymbol{A}(\bar{\boldsymbol{x}})\boldsymbol{\varepsilon})\tfrac{\mathrm{d}}{\mathrm{d}t}\big(\bar{x}^\alpha + A_j^\alpha(\bar{\boldsymbol{x}})\,\varepsilon^j\big) \\ &\approx Y_\alpha^i(\bar{\boldsymbol{x}})\big(\dot{x}_{\mathsf{R}}^\alpha + \tfrac{\partial A_j^\alpha}{\partial x^\beta}(\bar{\boldsymbol{x}})\,\dot{x}_{\mathsf{R}}^\beta \varepsilon^j + A_j^\alpha(\bar{\boldsymbol{x}})\,\dot{\varepsilon}^j\big) + \tfrac{\partial Y_\alpha^i}{\partial x^\beta}(\bar{\boldsymbol{x}})A_j^\beta(\bar{\boldsymbol{x}})\,\varepsilon^j \dot{x}_{\mathsf{R}}^\alpha \\ &= \bar{\xi}^i + \dot{\varepsilon}^i + \gamma_{kj}^i(\bar{\boldsymbol{x}})\bar{\xi}^k\varepsilon^j \end{aligned} \tag{3.39}$$

Using these results we may formulate an approximation of a general smooth function f along the trajectory $t \mapsto \bar{\boldsymbol{x}}(t)$ as

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{\xi}, \dot{\boldsymbol{\xi}}) &\approx f(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}}) + \frac{\partial f}{\partial x^\alpha}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})(x^\alpha - \bar{x}^\alpha) + \frac{\partial f}{\partial \xi^i}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})(\xi^i - \bar{\xi}^i) \\ &\quad + \frac{\partial f}{\partial \dot{\xi}^i}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})(\dot{\xi}^i - \dot{\bar{\xi}}^i) \\ &\approx f(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}}) + (\partial_i f)(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})\varepsilon^i + \frac{\partial f}{\partial \xi^i}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})(\dot{\varepsilon}^i + \gamma_{kj}^i(\bar{\boldsymbol{x}})\bar{\xi}^k\varepsilon^j) \\ &\quad + \frac{\partial f}{\partial \dot{\xi}^i}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})(\ddot{\varepsilon}^i + \gamma_{kj}^i(\bar{\boldsymbol{x}})\bar{\xi}^k\dot{\varepsilon}^j + \gamma_{kj}^i(\bar{\boldsymbol{x}})\dot{\bar{\xi}}^k\varepsilon^j + \partial_l\gamma_{kj}^i(\bar{\boldsymbol{x}})\bar{\xi}^l\bar{\xi}^k\varepsilon^j) \\ &= \bar{f} + \bar{F}_i^0\varepsilon^i + \bar{F}_i^1\dot{\varepsilon}^i + \bar{F}_i^2\ddot{\varepsilon}^i \end{aligned} \tag{3.40}$$

where

$$\begin{aligned}
\bar{f} &= f(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}}),\\
\bar{F}_i^0 &= (\partial_i f)(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}}) + \frac{\partial f}{\partial \xi^j}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})\gamma^j_{ki}(\bar{\boldsymbol{x}})\bar{\xi}^k + \frac{\partial f}{\partial \dot{\xi}^j}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})\big(\gamma^j_{ki}(\bar{\boldsymbol{x}})\dot{\bar{\xi}}^k + \partial_l \gamma^j_{ki}(\bar{\boldsymbol{x}})\bar{\xi}^l\bar{\xi}^k\big),\\
\bar{F}_i^1 &= \frac{\partial f}{\partial \xi^i}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}}) + \frac{\partial f}{\partial \dot{\xi}^j}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}})\gamma^j_{ki}(\bar{\boldsymbol{x}})\bar{\xi}^k,\\
\bar{F}_i^2 &= \frac{\partial f}{\partial \dot{\xi}^i}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\xi}}, \dot{\bar{\boldsymbol{\xi}}}).
\end{aligned}$$

Evidently, the expressions simplify significantly is the velocity coordinates are holonomic, i.e. $\gamma = 0$, or if the approximation is about a static point $\bar{\boldsymbol{x}} = const. \Rightarrow \boldsymbol{\xi} = \boldsymbol{0}$.

3.3.6 Calculus of variations

The calculus of variations is concerned with the extremals of functionals, i.e. functions of functions. For the particular context of classical mechanics we are interested in the curves $t \mapsto \boldsymbol{x}(t)$ for which the functional

$$\mathcal{J}[\boldsymbol{x}] = \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{x}(t), \boldsymbol{\xi}(t), t)\,\mathrm{d}t \tag{3.41}$$

for given boundary conditions $\boldsymbol{x}(t_1)$ and $\boldsymbol{x}(t_2)$ is *stationary.* The *Lagrangian* $\mathcal{L}$ is here a function of the configuration coordinates $\boldsymbol{x}$, its derivatives $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi}$ parameterized in the velocity coordinates $\boldsymbol{\xi}$ and may depend explicitly on the time t as well.

For the standard case, $\boldsymbol{x} = \boldsymbol{q}$ and $\boldsymbol{\xi} = \dot{\boldsymbol{q}}$, a derivation of may be found in e.g. [Lanczos, 1986, ch. II], [Arnold, 1989, sec. 12] or [Courant and Hilbert, 1924, chap. 4, §3]. For the present case we modify the well known derivation slightly: Suppose that $\boldsymbol{x} : [t_1, t_2] \mapsto \mathbb{X}$ is the solution to the variational problem. With the function $\boldsymbol{\chi}(t) \in \mathbb{R}^\nu$ and the parameter $\varepsilon \in \mathbb{R}$ we define a perturbation to it by

$$\bar{\boldsymbol{x}} = \boldsymbol{x} + \varepsilon\boldsymbol{\chi}. \tag{3.42}$$

We need $\bar{\boldsymbol{x}}(t) \in \mathbb{X}$ and consequently $\boldsymbol{\phi}(\bar{\boldsymbol{x}}) = \boldsymbol{0}$. Assuming ε to be sufficiently small, we may use the first order approximation analog to subsection 3.3.5: With the *variation coordinates* $\boldsymbol{h} : [t_1, t_2] \to \mathbb{R}^n$ we parameterize $\boldsymbol{\chi} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}$. Using the inverse kinematic relation $\boldsymbol{\xi} = \boldsymbol{Y}(\boldsymbol{x})\dot{\boldsymbol{x}}$ we can write the functional for the varied path as

$$\mathcal{J}[\bar{\boldsymbol{x}}] = \int_{t_1}^{t_2} \mathcal{L}\big(\boldsymbol{x} + \varepsilon\boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}, \boldsymbol{Y}(\boldsymbol{x} + \varepsilon\boldsymbol{A}(\boldsymbol{x})\boldsymbol{h})\tfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{x} + \varepsilon\boldsymbol{A}(\boldsymbol{x})\boldsymbol{h}), t\big)\,\mathrm{d}t =: \mathcal{P}(\varepsilon) \tag{3.43}$$

Now if $\boldsymbol{x}(t)$ is indeed the solution to the variational problem, then $\mathcal{P}(\varepsilon)$ must have a minimum at $\mathcal{P}(0)$ and consequently $\partial\mathcal{P}/\partial\varepsilon(0) = 0$. Evaluation of this "ordinary" differentiation yields

$$\begin{aligned}
0 = \frac{\partial \mathcal{P}}{\partial \varepsilon}\Big|_{\varepsilon=0} &= \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial x^\alpha} A^\alpha_i h^i + \frac{\partial \mathcal{L}}{\partial \xi^i}\left(\frac{\partial Y^i_\alpha}{\partial x^\beta} A^\beta_j h^j \dot{x}^\alpha + Y^i_\alpha \frac{\partial A^\alpha_j}{\partial x^\beta} h^j \dot{x}^\beta + \dot{h}^i\right)\right)\mathrm{d}t\\
&= \int_{t_1}^{t_2} \left(\partial_i \mathcal{L}\, h^i + \frac{\partial \mathcal{L}}{\partial \xi^i}\big(\gamma^i_{kj} h^j \xi^k + \dot{h}^i\big)\right)\mathrm{d}t
\end{aligned} \tag{3.44}$$

where we have found again the commutation coefficients γ^i_{kj} previously derived in (3.34). Integrating by parts with the boundary conditions $\boldsymbol{h}(t_1) = \boldsymbol{h}(t_2) = \mathbf{0}$ gives

$$\int_{t_0}^{t_1} h^i \left(A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} - \gamma^k_{ij} \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} \right) \mathrm{d}t = 0. \tag{3.45}$$

Since the variation coordinates $h^i, i = 1, \ldots, n$ are independent by definition, the *fundamental lemma of the calculus of variations* (see e.g. [Arnold, 1989, p. 57] or [Courant and Hilbert, 1924, p. 166]) states that, for the integral to vanish, the terms in the brackets have to vanish, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} + \gamma^k_{ij} \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} - A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} = 0, \qquad i = 1, \ldots, n. \tag{3.46}$$

This, combined with the kinematic relation $\dot{x}^\alpha = A_i^\alpha \xi^i, \alpha = 1 \ldots \nu$, is the necessary condition for the functional (3.41) to be stationary.

For the special case $\boldsymbol{x}(t) = \boldsymbol{q}(t) \in \mathbb{R}^n$ and $\boldsymbol{\xi}(t) = \dot{\boldsymbol{q}}(t)$ we have $\boldsymbol{A} = \mathbf{I}_n$ and $\gamma = 0$. Then (3.46) coincides with the *Euler-Lagrange equation*.

Example 11. Consider the configuration coordinates $\boldsymbol{x} = [\boldsymbol{R}_\mathsf{x}^\top, \boldsymbol{R}_\mathsf{y}^\top, \boldsymbol{R}_\mathsf{z}^\top]^\top$ and the velocity coordinates $\boldsymbol{\xi} = \boldsymbol{\omega}$ related by $\dot{\boldsymbol{R}} = \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega})$. For the Lagrangian

$$\mathcal{L} = \tfrac{1}{2} \boldsymbol{\omega}^\top \boldsymbol{\Theta} \boldsymbol{\omega} + \mathrm{tr}(\boldsymbol{\Pi}'(\mathbf{I}_3 - \boldsymbol{R})) \tag{3.47}$$

and taking into account the results from Example 9 and 10 we obtain

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} + \gamma^k_{ij} \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} - A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} \right]_{i=1,2,3} = \boldsymbol{\Theta} \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \boldsymbol{\Theta} \boldsymbol{\omega} + \mathrm{vee2}(\boldsymbol{\Pi}' \boldsymbol{R}). \tag{3.48}$$

3.4 Principles of mechanics

Consider the system from section 3.1 and let each particle have an associated mass $\mathfrak{m}_p \in \mathbb{R}^+$ and applied force $\mathfrak{F}_p^\mathsf{A}(t) \in \mathbb{R}^3$. Newton's second law states that for a system of *free* particles (here $\mathfrak{c} = \varnothing$) the equations of motion are given by

$$\mathfrak{m}_p \ddot{\mathfrak{r}}_p = \mathfrak{F}_p^\mathsf{A}, \quad p = 1, \ldots, \mathfrak{N}. \tag{3.49}$$

3.4.1 Principle of constraint release

The *principle of constraint release* (see e.g. [Lurie, 2002, sec. 6.1], [Hamel, 1949, sec. 32]) states that the motion of system of geometrically constrained particles is governed by

$$\mathfrak{c}(\mathfrak{r}) = \mathbf{0}, \quad \mathfrak{m}_p \ddot{\mathfrak{r}}_p = \mathfrak{F}_p^\mathsf{A} + \lambda_\kappa \frac{\partial \mathfrak{c}^\kappa}{\partial \mathfrak{r}_p}, \quad p = 1, \ldots, \mathfrak{N}. \tag{3.50}$$

where $\boldsymbol{\lambda}(t) \in \mathbb{R}^{\mathfrak{H}}$ are commonly called *Lagrange multipliers*. This is called Lagrange's equation of the first kind.

Formulation of the particle accelerations $\ddot{\mathfrak{r}}_p$ in terms of the coordinates $\boldsymbol{x}$ and $\boldsymbol{\xi}$ and summing up the projections of (3.50) on $\partial_i \mathfrak{r}_p$ yields

$$\underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^k \xi^j \rangle}_{f_i^{\mathrm{M}}} = \underbrace{\sum\nolimits_p \langle \partial_i \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{A}} \rangle}_{f_i^{\mathrm{A}}} + \underbrace{\sum\nolimits_p \langle \partial_i \mathfrak{r}_p, \lambda_\kappa \frac{\partial \mathfrak{c}^\kappa}{\partial \mathfrak{r}_p} \rangle}_{0}, \quad i = 1, \ldots, n. \tag{3.51}$$

The marked term vanishes, since $\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{c}^\kappa = \sum_p \frac{\partial \mathfrak{c}^\kappa}{\partial \mathfrak{r}_p} \partial_i \mathfrak{r}_p \xi^i = 0$ holds for any $\boldsymbol{\xi}$. These are n equations linear in the n coefficients of $\dot{\boldsymbol{\xi}}$. For the following we call $\boldsymbol{f}^{\mathrm{M}}$ the *generalized inertia force* and $\boldsymbol{f}^{\mathrm{A}}$ the *generalized applied force*.

3.4.2 Lagrange-d'Alembert's principle

The *Lagrange-d'Alembert principle* for a system of geometrically constrained particles states (e.g. [Goldstein, 1951, sec. 1.4] or [Lurie, 2002, sec. 6.3]):

$$\sum\nolimits_{p=1}^{\mathfrak{N}} \langle \delta \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{A}} - \mathfrak{m}_p \ddot{\mathfrak{r}}_p \rangle = 0. \tag{3.52}$$

The *virtual displacements* $\delta \mathfrak{r}_p$ are tangents to possible motions: For particle positions constrained by $\mathfrak{c}(\mathfrak{x}) = \mathbf{0}$ the displacements have to fulfill $\frac{\partial \mathfrak{c}}{\partial \mathfrak{x}} \delta \mathfrak{x} = \mathbf{0}$.

Now let the particle positions $\mathfrak{x} \in \mathfrak{X}$ be parameterized $\mathfrak{x} = \mathfrak{x}(\boldsymbol{x})$ by possibly redundant coordinates $\boldsymbol{x} \in \mathbb{X}$ such that $\boldsymbol{x} \in \mathbb{X} \Rightarrow \mathfrak{x}(\boldsymbol{x}) \in \mathfrak{X}$ or equivalently $\boldsymbol{\phi}(\boldsymbol{x}) = \mathbf{0} \Rightarrow \mathfrak{c}(\mathfrak{x}(\boldsymbol{x})) = \mathbf{0}$. With a suitable kinematics matrix $\boldsymbol{A}(\boldsymbol{x})$, as discussed in the previous section, we can parameterize the particle velocity as $\dot{\mathfrak{r}}_p = \frac{\partial \mathfrak{r}_p}{\partial \boldsymbol{x}} \boldsymbol{A} \boldsymbol{\xi} = \partial_i \mathfrak{r}_p \xi^i$ with minimal velocity coordinates $\boldsymbol{\xi} \in \mathbb{R}^n$. In the same way we parameterize the possible virtual displacements as $\delta \mathfrak{r}_p = \partial_i \mathfrak{r}_p h^i$ with the minimal *displacement coordinates* $\boldsymbol{h} \in \mathbb{R}^n$. Since the Lagrange-d'Alembert principle (3.52) has to hold for any displacement and the displacement coordinates $\boldsymbol{h}$ are independent, we can conclude

$$\sum\nolimits_p \langle \partial_i \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{A}} - \mathfrak{m}_p \ddot{\mathfrak{r}}_p \rangle = 0, \quad i = 1, \ldots, n. \tag{3.53}$$

3.4.3 Gauß' principle

The *Gauß' principle of least constraint* was originally described in [Gauß, 1829], in words rather than in equations. Maybe due to this one finds somewhat different mathematical formulations in more contemporary sources, e.g. [Hamel, 1949, sec. VII.8], [Lanczos, 1986, sec. IV.8], [Bremer, 2008, sec. 2.2]. For an extensive treatment of Gauß' principle and historical background see [Papastavridis, 2002, §6.6]. A differential geometric treatment of Gauß' principle and extension the general Lagrangian systems can be found in [Lewis, 1996].

For this work we use the formulation from [Päsler, 1968, sec. 7] of Gauß principle:

$$\begin{aligned} \min_{\ddot{\mathfrak{x}} \in \mathbb{R}^{3\mathfrak{N}}} \quad & \mathcal{G} = \tfrac{1}{2} \sum\nolimits_{p=1}^{\mathfrak{N}} \mathfrak{m}_p \|\ddot{\mathfrak{r}}_p - \ddot{\mathfrak{r}}_p^{\mathsf{f}}\|^2 \\ \text{s. t.} \quad & \ddot{\mathfrak{c}}(\mathfrak{x}, \dot{\mathfrak{x}}, \ddot{\mathfrak{x}}) = \mathbf{0} \end{aligned} \tag{3.54}$$

where $\mathfrak{x} = [\mathfrak{r}_1^\top, \ldots, \mathfrak{r}_{\mathfrak{N}}^\top]^\top$ and $\ddot{\mathfrak{r}}_p^{\mathsf{f}}$ are the particle accelerations of the unconstrained system and we call $\mathcal{G}$ the Gaussian constraint. Its crucial to note that the constraint equations $\ddot{\mathfrak{c}}(\mathfrak{x}, \dot{\mathfrak{x}}, \ddot{\mathfrak{x}}) = \mathbf{0}$ are *linear* in the accelerations $\ddot{\mathfrak{x}}$. Consequently, as stressed in [Gauß, 1829], the principle (3.54) can be regarded as a (static) quadratic optimization problem with linear constraints.

The unconstrained accelerations according to Newton's second law (3.49) are $\ddot{\mathfrak{r}}_p^{\mathsf{f}} = \mathfrak{F}_p^{\mathsf{A}}/\mathfrak{m}_p$. Expressing the particle accelerations as $\ddot{\mathfrak{r}}_p = \partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^k \xi^j$ in terms of coordinates, transforms (3.54) to an *unconstrained* problem

$$\min_{\dot{\boldsymbol{\xi}} \in \mathbb{R}^n} \mathcal{G} = \tfrac{1}{2} \sum\nolimits_p \mathfrak{m}_p \|\partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^k \xi^j - \tfrac{\mathfrak{F}_p^{\mathsf{A}}}{\mathfrak{m}_p}\|^2 \tag{3.55}$$

which has the necessary condition

$$\frac{\partial \mathcal{G}}{\partial \dot{\xi}^i} = \sum\nolimits_p \langle \partial_i \mathfrak{r}_p, \mathfrak{m}_p(\partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^k \xi^j) - \mathfrak{F}_p^{\mathsf{A}} \rangle = 0, \quad i = 1, \ldots, n. \tag{3.56}$$

This is obviously again the same result previously obtained in (3.51).

3.4.4 Hamilton's principle

Another principle can be stated as [Lurie, 2002, sec. 12.2] or [Szabó, 1956, sec. I.3]:

$$\int_{t_1}^{t_2} \big(\delta \mathcal{T} - \delta' \mathcal{W}\big) \mathrm{d}t = 0, \qquad \mathcal{T} = \tfrac{1}{2} \sum\nolimits_p \mathfrak{m}_p \|\dot{\mathfrak{r}}_p\|^2, \ \delta' \mathcal{W} = \langle \delta \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{A}} \rangle \tag{3.57}$$

where $\mathcal{T}$ is the kinetic energy and $\delta' \mathcal{W}$ is the virtual work of the applied forces. Using the result (3.46) from the calculus of variations we obtain

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{T}}{\partial \xi^i} + \gamma_{ij}^k \xi^j \frac{\partial \mathcal{T}}{\partial \xi^k} - \partial_i \mathcal{T}}_{-f_i^{\mathrm{M}}} = \underbrace{\langle \partial_i \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{A}} \rangle}_{f_i^{\mathrm{A}}}, \quad i = 1, \ldots, n. \tag{3.58}$$

Evaluation and some rearrangement of the left hand term $\boldsymbol{f}^{\mathrm{M}}$ shows that it indeed is the generalized inertia force previously defined in (3.51):

$$\begin{aligned} -f_i^{\mathrm{M}} &= \sum\nolimits_p \mathfrak{m}_p \Big(\frac{\mathrm{d}}{\mathrm{d}t} \langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \xi^j \rangle + \gamma_{ij}^k \xi^j \langle \partial_k \mathfrak{r}_p, \partial_l \mathfrak{r}_p \xi^l \rangle - \langle \partial_i \partial_j \mathfrak{r}_p \xi^j, \partial_k \mathfrak{r}_p \xi^k \rangle \Big) \\ &= \sum\nolimits_p \mathfrak{m}_p \Big(\langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^j \xi^k \rangle + \langle \partial_l \mathfrak{r}_p \xi^l, \underbrace{\big(\partial_j \partial_i \mathfrak{r}_p - \partial_i \partial_j \mathfrak{r}_p + \gamma_{ij}^k \partial_k \mathfrak{r}_p\big)}_{0} \xi^j \rangle \Big). \end{aligned} \tag{3.59}$$

The principle of least action is closely related to (3.57), but there is a crucial difference: The virtual work $\delta' \mathcal{W}$ is, in general, not a variation of a whatever function, so we cannot pull the variation out of the integral and (3.57) implies, in general, no variational statement. If the applied forces are conservative, then the virtual work $\delta'\mathcal{W} = \delta\mathcal{V}$ can be stated as the variation of a potential $\mathcal{V}$ and we can define the "action" as $\mathcal{S} = \int_{t_1}^{t_2}(\mathcal{T} - \mathcal{V})\mathrm{d}t$. For non-autonomous or dissipative systems, which are considered in this text, one should not use the naming "of least action" since the "action" is simply not defined. See [Lurie, 2002, sec. 12.2] for an extensive discussion of this subject.

3.5 Kinetics

Kinetics is the the part of mechanics that relates motion to its causes, i.e. forces.In the previous section we derived the *kinetic equation* (3.51) as a balance of generalized inertia force $f_i^{\mathrm{M}} = \sum_p \langle \partial_i \mathfrak{r}_p, -\mathfrak{m}_p \ddot{\mathfrak{r}}_p \rangle$ and the generalized applied force $f_i^{\mathrm{A}} = \sum_p \langle \partial_i \mathfrak{r}_p, \boldsymbol{\mathfrak{F}}_p^{\mathsf{A}} \rangle$. The applied force $\boldsymbol{\mathfrak{F}}_p^{\mathsf{A}}$ sums all forces on the particle that are not due to inertia or constraints. For the following we split it as $\boldsymbol{\mathfrak{F}}_p^{\mathsf{A}} = \boldsymbol{\mathfrak{F}}_p^{\mathsf{D}} + \boldsymbol{\mathfrak{F}}_p^{\mathsf{K}} + \boldsymbol{\mathfrak{F}}_p^{\mathsf{G}} + \boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}$ into dissipative force $\boldsymbol{\mathfrak{F}}_p^{\mathsf{D}}$, force originating from springs $\boldsymbol{\mathfrak{F}}_p^{\mathsf{K}}$, gravitational force $\boldsymbol{\mathfrak{F}}_p^{\mathsf{G}}$, and $\boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}$ collects everything else, e.g. control forces.

3.5.1 Inertia

Expressing the particle accelerations $\ddot{\mathfrak{r}}_p$ in terms of the velocity coordinates $\boldsymbol{\xi}$ and rearranging some terms we find

$$\begin{aligned} -f_i^{\mathrm{M}} &= \sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \overbrace{\partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^k \xi^j}^{\ddot{\mathfrak{r}}_p} \rangle \\ &= \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \rangle}_{M_{ij}} \dot{\xi}^j + \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_k \partial_j \mathfrak{r}_p \rangle}_{\Gamma_{ijk}} \xi^k \xi^j \end{aligned} \tag{3.60}$$

where we introduced the *system inertia matrix* $\boldsymbol{M}(\boldsymbol{x}) \in \mathbb{R}^{n \times n}$ and the *connection coefficients* $\boldsymbol{\Gamma}(\boldsymbol{x}) \in \mathbb{R}^{n \times n \times n}$. These are related by the following identities

$$\partial_k M_{ij} = \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_k \partial_j \mathfrak{r}_p \rangle}_{\Gamma_{ijk}} + \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_k \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \rangle}_{\Gamma_{jik}} \tag{3.61a}$$

$$\gamma_{ij}^s M_{sk} = \sum\nolimits_p \mathfrak{m}_p \langle \gamma_{ij}^s \partial_s \mathfrak{r}_p, \partial_k \mathfrak{r}_p \rangle = \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_i \partial_j \mathfrak{r}_p, \partial_k \mathfrak{r}_p \rangle}_{\Gamma_{kji}} - \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \partial_j \partial_i \mathfrak{r}_p, \partial_k \mathfrak{r}_p \rangle}_{\Gamma_{kij}} \tag{3.61b}$$

Plugging these together while permuting the indices, we find

$$\begin{aligned}
\varGamma_{ijk} &= \partial_k M_{ij} - \varGamma_{jik} \\
&= \partial_k M_{ij} + \gamma^s_{ik} M_{sj} - \varGamma_{jki} \\
&= \partial_k M_{ij} + \gamma^s_{ik} M_{sj} - \partial_i M_{kj} + \varGamma_{kji} \\
&= \partial_k M_{ij} + \gamma^s_{ik} M_{sj} - \partial_i M_{kj} + \gamma^s_{ij} M_{sk} + \varGamma_{kij} \\
&= \partial_k M_{ij} + \gamma^s_{ik} M_{sj} - \partial_i M_{kj} + \gamma^s_{ij} M_{sk} + \partial_j M_{ik} - \varGamma_{ikj} \\
&= \partial_k M_{ij} + \gamma^s_{ik} M_{sj} - \partial_i M_{kj} + \gamma^s_{ij} M_{sk} + \partial_j M_{ik} - \gamma^s_{kj} M_{si} - \varGamma_{ijk} \qquad (3.62)
\end{aligned}$$

$$\Leftrightarrow \quad \varGamma_{ijk} = \tfrac{1}{2}\big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \gamma^s_{ij} M_{sk} + \gamma^s_{ik} M_{sj} - \gamma^s_{jk} M_{si}\big). \qquad (3.63)$$

This means the connection coefficients are determined by the inertia matrix $\boldsymbol{M}$ and the geometric matrix $\boldsymbol{A}$ which also determines the commutation coefficients $\boldsymbol{\gamma}$.

Lagrange formulation. With some reordering and using the relations (3.34) and (3.4) we can write the inertial force as

$$\begin{aligned}
-f^{\mathrm{M}}_i &= \sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathbf{r}_p, \ddot{\mathbf{r}}_p \rangle \\
&= \sum\nolimits_p \mathfrak{m}_p \Big(\frac{\mathrm{d}}{\mathrm{d}t} \langle \partial_i \mathbf{r}_p, \dot{\mathbf{r}}_p \rangle - \langle \frac{\mathrm{d}}{\mathrm{d}t}\big(\partial_i \mathbf{r}_p\big), \dot{\mathbf{r}}_p \rangle \Big) \\
&= \sum\nolimits_p \mathfrak{m}_p \Big(\frac{\mathrm{d}}{\mathrm{d}t} \langle \partial_i \mathbf{r}_p, \dot{\mathbf{r}}_p \rangle + \langle \gamma^k_{ij} \xi^j \partial_k \mathbf{r}_p, \dot{\mathbf{r}}_p \rangle - \langle \partial_i \dot{\mathbf{r}}_p, \dot{\mathbf{r}}_p \rangle \Big) \\
&= \sum\nolimits_p \mathfrak{m}_p \Big(\frac{\mathrm{d}}{\mathrm{d}t} \langle \frac{\partial \dot{\mathbf{r}}_p}{\partial \xi^i}, \dot{\mathbf{r}}_p \rangle + \gamma^k_{ij} \xi^j \langle \frac{\partial \dot{\mathbf{r}}_p}{\partial \xi^k}, \dot{\mathbf{r}}_p \rangle - \langle \partial_i \dot{\mathbf{r}}_p, \dot{\mathbf{r}}_p \rangle \Big) \\
&= \Big(\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \xi^i} + \gamma^k_{ij} \xi^j \frac{\partial}{\partial \xi^k} - \partial_i \Big) \underbrace{\Big(\tfrac{1}{2} \sum\nolimits_p \mathfrak{m}_p \|\dot{\mathbf{r}}_p\|^2 \Big)}_{\mathcal{T}}, \qquad i = 1, \dots, n \qquad (3.64)
\end{aligned}$$

with the kinetic energy $\mathcal{T}$. The derivation is essentially what was done in (3.59) backwards. The formulation (3.64) is a slight generalization of the "classical" Lagrange formulation (3.5) allowing redundant configuration coordinates $\boldsymbol{x}$ and velocity coordinates $\boldsymbol{\xi}$. It is very similar to the formulations proposed in [Boltzmann, 1902] and [Hamel, 1904] up to the definition of the directional derivative ∂_i and the commutation coefficients γ^k_{ij} which therein restrict to *minimal* configuration coordinates.

Gibbs-Appell formulation. Yet another formulation can be obtained using (3.4):

$$\begin{aligned}
-f^{\mathrm{M}}_i &= \sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathbf{r}_p, \ddot{\mathbf{r}}_p \rangle \\
&= \sum\nolimits_p \mathfrak{m}_p \langle \frac{\partial \ddot{\mathbf{r}}_p}{\partial \dot{\xi}^i}, \ddot{\mathbf{r}}_p \rangle \\
&= \frac{\partial}{\partial \dot{\xi}^i} \underbrace{\Big(\tfrac{1}{2} \sum\nolimits_p \mathfrak{m}_p \|\ddot{\mathbf{r}}_p\|^2 \Big)}_{\mathcal{S}} \qquad i = 1, \dots, n \qquad (3.65)
\end{aligned}$$

with the acceleration energy $\mathcal{S}$. This formulation (3.65) was first proposed by [Gibbs, 1879] and by [Appell, 1900] with a focus on nonholonomic systems.

3.5.2 Dissipation

[Goldstein, 1951, p. 24] "Frictional forces of this type may be derived in terms of a function, known as Rayleigh's dissipation function"

[Papastavridis, 2002, p. 519] same as above

For dissipative forces we pick up the concept from [Rayleigh, 1877, §81]: "Suppose that each particle of the system is retarded by forces proportional to its component velocities". So for each particle we assume a damping force

$$\mathfrak{F}_p^{\mathsf{D}} = -\mathfrak{d}_p \dot{\mathfrak{r}}_p, \qquad p = 1, \dots, \mathfrak{N} \tag{3.66}$$

with the damping parameters $\mathfrak{d}_p \in \mathbb{R}^+$. For a parameterized system (**??**) we have the *generalized dissipation force*:

$$f_i^{\mathsf{D}} = \sum\nolimits_p \langle \partial_i \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{D}} \rangle = -\underbrace{\sum\nolimits_p \mathfrak{d}_p \langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \rangle}_{D_{ij}} \xi^j \tag{3.67}$$

where we introduced the *system dissipation matrix* $\boldsymbol{D}(\boldsymbol{x}) \in \mathbb{SYM}_0^+(n)$. Furthermore, these quantities can be derived from the so-called *Rayleigh dissipation function* $\mathcal{R} : \mathsf{T}\mathbb{X} \to \mathbb{R}$ using the rules (3.4):

$$\mathcal{R} = \sum\nolimits_p \tfrac{1}{2} \mathfrak{d}_p \|\dot{\mathfrak{r}}_p\|^2 = \tfrac{1}{2} D_{ij} \xi^i \xi^j, \qquad f_i^{\mathsf{D}} = -\frac{\partial \mathcal{R}}{\partial \xi^i}, \qquad D_{ij} = \frac{\partial^2 \mathcal{R}}{\partial \xi^i \partial \xi^j}. \tag{3.68}$$

Note that the non-negativity of the damping parameters $\mathfrak{d}_p \geq 0, p = 1, \dots, \mathfrak{N}$ implies non-negativity of the dissipation function $\mathcal{R} \geq 0$ and positive semi-definiteness of the dissipation matrix $\boldsymbol{D} \geq 0$.

General dissipation

$$\mathcal{R} = \tfrac{1}{2} \sum_{p,q=1}^{\mathfrak{N}} \mathfrak{d}_{pq} \|\dot{\mathfrak{r}}_p - \dot{\mathfrak{r}}_q\|^2 \tag{3.69}$$

$$= \tfrac{1}{2} \underbrace{\Big(\sum_{p,q=1}^{\mathfrak{N}} \mathfrak{d}_{pq} \langle \partial_i \mathfrak{r}_p - \partial_i \mathfrak{r}_q, \partial_j \mathfrak{r}_p - \partial_j \mathfrak{r}_q \rangle \Big)}_{D_{ij}} \xi^j \xi^i \tag{3.70}$$

3.5.3 Linear springs

Assume that the particle with index p is connected to another particle with index q by a spring. In the simplest case the resulting spring force obeys Hooke's law [Hooke, 1678]: The force on particle p connected by a spring, with the spring constant $\mathfrak{k}_{pq} \in \mathbb{R}^+$, to particle q is

$$\mathfrak{F}_{pq}^{\mathsf{K}} = \mathfrak{k}_{pq} (\mathfrak{r}_q - \mathfrak{r}_p). \tag{3.71}$$

Naturally, the opposing force $\mathfrak{F}_{qp}^{\mathsf{K}} = -\mathfrak{F}_{pq}^{\mathsf{K}}$ acts on particle q. For a particle system we can assume that each particle may be connected to each other particle and set $\mathfrak{k}_{pq} = 0$ if there is no spring. Then the overall force on each particle is

$$\mathfrak{F}_p^{\mathsf{K}} = \sum_{q=1,q\neq p}^{\mathfrak{N}} \mathfrak{k}_{pq}(\mathfrak{r}_q - \mathfrak{r}_p), \qquad \mathfrak{k}_{pq} = \mathfrak{k}_{qp}. \tag{3.72}$$

For a parameterized system (**??**) we have the *generalized stiffness force*:

$$\begin{aligned} f_i^{\mathrm{K}} &= \sum_{p=1}^{\mathfrak{N}} \langle \partial_i \mathfrak{r}_p, \mathfrak{F}_p^{\mathsf{K}} \rangle = \sum_{p=1}^{\mathfrak{N}} \sum_{q=1,q\neq p}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i \mathfrak{r}_p, \mathfrak{r}_q - \mathfrak{r}_p \rangle \\ &= \tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i \mathfrak{r}_p, \mathfrak{r}_q - \mathfrak{r}_p \rangle + \tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{qp} \langle \partial_i \mathfrak{r}_q, \mathfrak{r}_p - \mathfrak{r}_q \rangle \\ &= -\tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i (\mathfrak{r}_q - \mathfrak{r}_p), \mathfrak{r}_q - \mathfrak{r}_p \rangle \\ &= -\partial_i \underbrace{\left(\tfrac{1}{4} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \|\mathfrak{r}_q - \mathfrak{r}_p\|^2 \right)}_{\mathcal{V}^{\mathrm{K}}} = -\partial_i \underbrace{\left(\tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=p+1}^{\mathfrak{N}} \mathfrak{k}_{pq} \|\mathfrak{r}_q - \mathfrak{r}_p\|^2 \right)}_{\mathcal{V}^{\mathrm{K}}} \end{aligned} \tag{3.73}$$

Since it can be derived from a potential $\mathcal{V}^{\mathrm{K}} : \mathbb{X} \to \mathbb{R}$, it is called a *conservative force.* Note that the non-negativity of the spring constants $\mathfrak{k}_{pq} \geq 0, p, q = 1, \ldots, \mathfrak{N}$ implies the non-negativity of the potential $\mathcal{V}^{\mathrm{K}} \geq 0$.

potential energy

$$\mathcal{V}^{\mathrm{K}} = \tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=p+1}^{\mathfrak{N}} \mathfrak{k}_{pq} \|\mathfrak{r}_p - \mathfrak{r}_q\|^2, \qquad f_i^{\mathrm{K}} = \partial_i \mathcal{V}^{\mathrm{K}} \tag{3.74}$$

This identity might not be so obvious so we display the computation explicitly: First note that due to $\mathfrak{k}_{pq} = \mathfrak{k}_{qp}$ and the obvious fact that the term for $q = p$ vanishes anyway, the potential can be expressed as

$$\mathcal{V}^{\mathrm{K}} = \tfrac{1}{4} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \|\mathfrak{r}_p - \mathfrak{r}_q\|^2 \tag{3.75}$$

with this the directional derivatives are

$$\begin{aligned} \partial_i \mathcal{V}^{\mathrm{K}} &= \tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i \mathfrak{r}_p - \partial_i \mathfrak{r}_q, \mathfrak{r}_p - \mathfrak{r}_q \rangle \\ &= \tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i \mathfrak{r}_p, \mathfrak{r}_p - \mathfrak{r}_q \rangle + \tfrac{1}{2} \sum_{p=1}^{\mathfrak{N}} \sum_{q=1}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i \mathfrak{r}_q, \mathfrak{r}_q - \mathfrak{r}_p \rangle \\ &= \sum_{p=1}^{\mathfrak{N}} \sum_{q=1,q\neq p}^{\mathfrak{N}} \mathfrak{k}_{pq} \langle \partial_i \mathfrak{r}_p, \mathfrak{r}_p - \mathfrak{r}_q \rangle \end{aligned} \tag{3.76}$$

3.5.4 Gravitation

For far most engineering applications we are dealing with systems that move close to the surface of the earth and where Galilei's gravitation principle holds, see [Galilei, 1638, Day 3]. In a contemporary formulation it states that a particle with mass $\mathfrak{m}_p$ is subject to the gravitational force

$$\boldsymbol{\mathfrak{F}}_p^{\mathsf{G}} = \mathfrak{m}_p \boldsymbol{a}_{\mathsf{G}} \tag{3.77}$$

where $\boldsymbol{a}_{\mathsf{G}}$ are the coefficients of the gravitational acceleration of the earth w.r.t. the chosen inertial frame. Commonly the inertial frame is chosen such that the $\boldsymbol{e}_{\mathsf{z}}$ axis is opposing gravity and we have $\boldsymbol{a}_{\mathsf{G}} = [0, 0, -g]^\top$ with the *gravity of earth* $g = 9.8\,\frac{\mathrm{m}}{\mathrm{s}^2}$.

The resulting generalized force for a system of particles is

$$f_i^{\mathrm{G}} = \sum\nolimits_p \langle \partial_i \mathfrak{r}_p, \boldsymbol{\mathfrak{F}}_p^{\mathsf{G}} \rangle = \sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \boldsymbol{a}_{\mathsf{G}} \rangle = \partial_i \underbrace{\sum\nolimits_p \mathfrak{m}_p \langle \mathfrak{r}_p, \boldsymbol{a}_{\mathsf{G}} \rangle}_{\mathcal{V}^{\mathrm{G}}} . \tag{3.78}$$

Since it can be derived from a potential $\mathcal{V}^{\mathrm{G}} : \mathbb{X} \to \mathbb{R}$, it is called a *conservative force.*

3.5.5 External forces

Let external forces $\boldsymbol{\mathfrak{F}}_p^{\mathsf{E}} = \boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u})$

$$\boldsymbol{f}^{\mathrm{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u}) = \sum\nolimits_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u}) \rangle \tag{3.79}$$

3.6 Additional constraints

In addition to geometric constraints $\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}$ that have been incorporated by the chosen coordinates $\boldsymbol{x}$ and $\boldsymbol{\xi}$ we like to incorporate additional constraints of various forms:

- geometric constraints:

$$\begin{aligned} & \boldsymbol{\psi}(\boldsymbol{x}) = \boldsymbol{0} \\ \Leftrightarrow \quad & \underbrace{\partial_i \psi^\kappa(\boldsymbol{x})}_{Z_i^\kappa(\boldsymbol{x})} \dot{\xi}^i = \underbrace{-\partial_j \partial_i \psi^\kappa(\boldsymbol{x}) \xi^i \xi^j}_{b^\kappa(\boldsymbol{x}, \boldsymbol{\xi})}, \quad \psi^\kappa(\boldsymbol{x}_0) = 0,\ \partial_i \psi^\kappa(\boldsymbol{x}_0) \xi_0^i = 0 \end{aligned} \tag{3.80a}$$

- linear kinematic constraints (possibly nonholonomic):

$$\begin{aligned} & \boldsymbol{N}(\boldsymbol{x}) \dot{\boldsymbol{x}} = \underbrace{\boldsymbol{N}(\boldsymbol{x}) \boldsymbol{A}(\boldsymbol{x})}_{\boldsymbol{Z}(\boldsymbol{x})} \boldsymbol{\xi} = \boldsymbol{0} \\ \Leftrightarrow \quad & Z_i^\kappa(\boldsymbol{x}) \dot{\xi}^i = \underbrace{-\partial_j Z_i^\kappa(\boldsymbol{x}) \xi^i \xi^j}_{b^\kappa(\boldsymbol{x}, \boldsymbol{\xi})},\ Z_i^\kappa(\boldsymbol{x}_0) \xi_0^\kappa = 0 \end{aligned} \tag{3.80b}$$

- general kinematic constraints

$$\begin{aligned} &\boldsymbol{\eta}(\boldsymbol{x},\boldsymbol{\xi},t) = \mathbf{0} \\ \Leftrightarrow \quad &\underbrace{\tfrac{\partial \eta^\kappa}{\partial \xi^i}(\boldsymbol{x},\boldsymbol{\xi},t)}_{Z^\kappa_i(\boldsymbol{x},\boldsymbol{\xi},t)}\dot{\xi}^i = \underbrace{-\partial_i \eta^\kappa(\boldsymbol{x},\boldsymbol{\xi},t)\xi^i - \tfrac{\partial \eta^\kappa}{\partial t}(\boldsymbol{x},\boldsymbol{\xi},t)}_{b^\kappa(\boldsymbol{x},\boldsymbol{\xi},t)}, \quad \eta^\kappa(\boldsymbol{x}_0,\boldsymbol{\xi}_0,t_0) = 0 \end{aligned} \tag{3.80c}$$

- linear[3] acceleration constraints.

$$\boldsymbol{Z}(\boldsymbol{x},\boldsymbol{\xi},t)\dot{\boldsymbol{\xi}} = \boldsymbol{b}(\boldsymbol{x},\boldsymbol{\xi},t) \tag{3.80d}$$

All these constraints can be formulated as *linear acceleration constraints* $\boldsymbol{Z}\dot{\boldsymbol{\xi}} = \boldsymbol{b}$ possibly supplemented by suitable conditions on the initial coordinates $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$ and $\boldsymbol{\xi}_0 = \boldsymbol{\xi}(t_0)$. Gauß suggested in [Gauß, 1829] also the application to inequality constraints which will not be discussed here. For a contemporary discussion and applications of this see [Pfeiffer and Glocker, 1996, sec. 6.1].

In terms of the chosen coordinates $\boldsymbol{x}$ and $\boldsymbol{\xi}$ the Gaussian constraint $\mathcal{G}$ can be written as

$$\begin{aligned} \mathcal{G} &= \tfrac{1}{2}\sum_p \mathfrak{m}_p \| \overbrace{\partial_j \mathfrak{r}_p \dot{\xi}^j + \partial_k \partial_j \mathfrak{r}_p \xi^k \xi^j}^{\ddot{\mathfrak{r}}_p} - \tfrac{\mathfrak{F}^{\mathrm{A}}_p}{\mathfrak{m}_p}\|^2 \\ &= \underbrace{\tfrac{1}{2}\sum_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p\rangle}_{M_{ij}} \dot{\xi}^i \dot{\xi}^j + \underbrace{\sum_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_k \partial_j \mathfrak{r}_p\rangle \xi^j \xi^k}_{c_i} \dot{\xi}^i + \underbrace{\sum_p \langle \partial_i \mathfrak{r}_p, \mathfrak{F}^{\mathrm{A}}_p\rangle}_{f^{\mathrm{A}}_i} \dot{\xi}^i + \mathcal{G}_0 \\ &= \tfrac{1}{2}\dot{\boldsymbol{\xi}}^\top \boldsymbol{M}\dot{\boldsymbol{\xi}} + \dot{\boldsymbol{\xi}}^\top(\boldsymbol{c} - \boldsymbol{f}^{\mathrm{A}}) + \mathcal{G}_0 \end{aligned} \tag{3.81}$$

where $\mathcal{G}_0$ collects the terms independent of the acceleration $\dot{\boldsymbol{\xi}}$. Overall we can state the Gauß principle with the additional constraints $\boldsymbol{Z}\dot{\boldsymbol{\xi}} = \boldsymbol{b}$ originating from (3.80) as

$$\begin{aligned} \min_{\dot{\boldsymbol{\xi}} \in \mathbb{R}^n} \quad & \mathcal{G} = \tfrac{1}{2}\dot{\boldsymbol{\xi}}^\top \boldsymbol{M}\dot{\boldsymbol{\xi}} + \dot{\boldsymbol{\xi}}^\top(\boldsymbol{c} - \boldsymbol{f}^{\mathrm{A}}) + \mathcal{G}_0 \\ \text{s. t.} \quad & \boldsymbol{Z}\dot{\boldsymbol{\xi}} = \boldsymbol{b} \end{aligned} \tag{3.82}$$

For whatever reason we do not want to eliminate the constraints by a change of coordinates but use the concept of the *Lagrange multipliers* (e.g. [Luenberger and Ye, 2015, ch. 14]): By defining the auxiliary function $\bar{\mathcal{G}} = \mathcal{G} - \boldsymbol{\lambda}^\top(\boldsymbol{Z}\dot{\boldsymbol{\xi}} - \boldsymbol{b})$ the solution to (3.82) can be stated as

$$\begin{bmatrix} \frac{\partial \bar{\mathcal{G}}}{\partial \dot{\boldsymbol{\xi}}} \\ \frac{\partial \bar{\mathcal{G}}}{\partial \boldsymbol{\lambda}} \end{bmatrix} = \mathbf{0} \quad \Leftrightarrow \quad \begin{bmatrix} \boldsymbol{M} & -\boldsymbol{Z}^\top \\ \boldsymbol{Z} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{\xi}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}^{\mathrm{A}} - \boldsymbol{c} \\ \boldsymbol{b} \end{bmatrix}. \tag{3.83}$$

The additional quantities $\boldsymbol{\lambda}$ are called the *Lagrange multipliers* and can be interpreted as reaction forces.

The possibility of handling such a variety of constraints might put Gauß' principle in a superior position compared to other principles as pointed out in [Hamel, 1949, p. 525].

[3]Nonlinear acceleration constraints could be handled as well, but with a more sophisticated solution than (3.83)

The Lagrange-d'Alembert principle handles linear kinematic constraints $\boldsymbol{N}\dot{\boldsymbol{x}} = \boldsymbol{0}$ by requiring $\boldsymbol{N}\delta\boldsymbol{x} = \boldsymbol{0}$. However, there are no real world examples of nonlinear nonholonomic constraints [Hamel, 1949, p. 499], [Neimark and Fufaev, 1972, ch. IV], [Roberson and Schwertassek, 1988, p. 96] and none of acceleration constraints [Hamel, 1949, p. 505 & 525], so one should be careful with this context. Note for example that for nonlinear kinematic and acceleration constraints the reaction forces $\boldsymbol{\lambda}$ enter the balance of energy.

Example?

3.7 Equations of motion and total energy

force balance

$$\underbrace{M_{ij}(\boldsymbol{x})\dot{\xi}^j + \Gamma_{ijk}(\boldsymbol{x})\xi^k\xi^j}_{f_i^{\mathrm{M}}} + \underbrace{D_{ij}(\boldsymbol{x})\xi^j}_{f_i^{\mathrm{D}}} + \underbrace{\partial_i\mathcal{V}(\boldsymbol{x})}_{f_i^{\mathrm{V}}} = f_i^{\mathrm{E}}(\boldsymbol{x},\boldsymbol{u}), \qquad i = 1,\dots,n. \tag{3.84}$$

For the change of kinetic energy $\mathcal{T}(\boldsymbol{x},\boldsymbol{\xi}) = M_{ij}(\boldsymbol{x})\xi^i\xi^j$ we have

$$\dot{\mathcal{T}} = M_{ij}\xi^i\dot{\xi}^j + \tfrac{1}{2}\partial_k M_{ij}\xi^i\xi^j\xi^k = \xi^i f_i^{\mathrm{M}} + \underbrace{\left(\tfrac{1}{2}\partial_k M_{ij} - \Gamma_{ijk}\right)\xi^i\xi^j}_{0}\xi^k \tag{3.85}$$

The marked term vanishes due to (3.61a). Furthermore we have similar relations for the potential energy $\mathcal{V}$, the dissipation function $\mathcal{R}$ and we define the external power $\mathcal{P}^{\mathrm{E}}$ as

$$\dot{\mathcal{V}} = \xi^i f_i^{\mathrm{V}}, \qquad 2\mathcal{R} = \xi^i f_i^{\mathrm{D}}, \qquad \mathcal{P}^{\mathrm{E}} = \xi^i f_i^{\mathrm{E}}. \tag{3.86}$$

Defining the *total energy* $\mathcal{W}$ as the sum of kinetic and potential energy and taking into account the kinetic equation $\boldsymbol{f}^{\mathrm{M}} + \boldsymbol{f}^{\mathrm{D}} + \boldsymbol{f}^{\mathrm{V}} = \boldsymbol{f}^{\mathrm{E}}$ we find

$$\mathcal{W} = \mathcal{T} + \mathcal{V}, \qquad \dot{\mathcal{W}} = \mathcal{P}^{\mathrm{E}} - 2\mathcal{R}. \tag{3.87}$$

3.8 Summary

The equations of motion of a system of constrained particles, parameterized by the (possibly redundant) configuration coordinates $\boldsymbol{x}(t) \in \mathbb{R}^\nu$, $\boldsymbol{\phi}(\boldsymbol{x}) = \mathbf{0}$ and minimal velocity coordinates $\boldsymbol{\xi}(t) \in \mathbb{R}^n$, may be formulated as

$$\dot{x}^\alpha = A_i^\alpha(\boldsymbol{x})\xi^i, \qquad \alpha = 1, \dots, \nu \tag{3.88a}$$

$$\underbrace{M_{ij}(\boldsymbol{x})\dot{\xi}^j + \Gamma_{ijk}(\boldsymbol{x})\xi^k\xi^j}_{f_i^{\mathrm{M}}} + \underbrace{D_{ij}(\boldsymbol{x})\xi^j}_{f_i^{\mathrm{D}}} + \underbrace{\partial_i \mathcal{V}(\boldsymbol{x})}_{f_i^{\mathrm{V}}} = f_i^{\mathrm{E}}(\boldsymbol{x}, \boldsymbol{u}), \qquad i = 1, \dots, n. \tag{3.88b}$$

The kinematics matrix $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu \times n}$ must fulfill $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\boldsymbol{A} = \mathbf{0}$ and $\operatorname{rank} \boldsymbol{A} = n$. The directional derivative ∂_i and the commutation coefficients $\boldsymbol{\gamma}(\boldsymbol{x}) \in \mathbb{R}^{n\times n\times n}$ for a given kinematics matrix $\boldsymbol{A}$ are

$$\partial_i = A_i^\alpha \frac{\partial}{\partial x^\alpha}, \qquad \gamma_{ij}^k = \big(\partial_i A_j^\alpha - \partial_j A_i^\alpha\big)(A^+)_\alpha^k. \tag{3.89}$$

The connection coefficients $\boldsymbol{\Gamma}(\boldsymbol{x}) \in \mathbb{R}^{n\times n\times n}$ for a given inertia matrix $\boldsymbol{M}(\boldsymbol{x}) \in \mathbb{R}^{n\times n}$ are

$$\Gamma_{ijk} = \tfrac{1}{2}\big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \gamma_{ij}^s M_{sk} + \gamma_{ik}^s M_{sj} - \gamma_{jk}^s M_{si}\big). \tag{3.90}$$

The inertia matrix for a system of particles with positions $\mathfrak{r}_p(\boldsymbol{x}) \in \mathbb{R}^3$ and masses $\mathfrak{m}_p$ is

$$M_{ij} = \sum\nolimits_p \mathfrak{m}_p \langle \partial_i \mathfrak{r}_p, \partial_j \mathfrak{r}_p \rangle \tag{3.91}$$

For the corresponding inertia force $\boldsymbol{f}^{\mathrm{M}}$ we also have the Lagrange formulation

$$f_i^{\mathrm{M}} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{T}}{\partial \xi^i} + \gamma_{ij}^k \xi^j \frac{\partial \mathcal{T}}{\partial \xi^k} - \partial_i \mathcal{T}, \qquad \mathcal{T} = \tfrac{1}{2}\sum\nolimits_p \mathfrak{m}_p \|\dot{\mathfrak{r}}_p\|^2 = \tfrac{1}{2} M_{ij}\xi^j\xi^i \tag{3.92}$$

and the Gibbs-Appell formulation

$$f_i^{\mathrm{M}} = \frac{\partial \mathcal{S}}{\partial \dot{\xi}^i}, \qquad \mathcal{S} = \tfrac{1}{2}\sum\nolimits_p \mathfrak{m}_p \|\ddot{\mathfrak{r}}_p\|^2 = \tfrac{1}{2} M_{ij}\dot{\xi}^j\dot{\xi}^i + \Gamma_{ijk}\xi^k\xi^j\dot{\xi}^i + \mathcal{S}_0 \tag{3.93}$$

dissipation, potential, external forces

special case with minimal config coord

Chapter 4

Rigid body systems

A *rigid body* can be regarded as a special case of a particle system whose particles have constant distance to each other. Several rigid bodies that may be constrained to each other and/or to the surrounding space constitute a *rigid body system*. This chapter applies the results from the previous section to them. As the configuration space of a free rigid body is nonlinear, the use of redundant coordinates is very appropriate.

Goal. The goal for this section is to motivate a recipe or formalism for the derivation of the equations of motion for rigid body systems. In contrast to established formalisms, the approach here should be most flexible w.r.t. the parameterization of the configuration and velocity. Furthermore, it should not only apply to the inertia of a system but also to damping and stiffness, as motivated in the previous chapter. Overall, this should give deeper insight into the structure of rigid body systems which will be exploited in the next chapter on control of rigid body systems.

[Goldstein, 1951, chap. 4]: A rigid body is defined as a system of mass points subject top the holonomic constraints thet the distances between all pairs of points remain constant throughout the motion.
[Landau and Lifshitz, 1960, §31]: A rigid body may be defined in mechanics as a system of particles such that the distances between the particles do not vary.
[Boltzmann, 1897, §44]: ...starrer Körper, d.h. für ein System materieller Punkte, welche so verbunden sind, dass sich während ihrer ganzen Bewegung ihre relative Lage nicht ändern kann.
[Arnold, 1989, §28]: A rigid body is a system of point masses, constrained by holonomic relations expressed by the fact that the distance between points is constant.
[Bremer, 2008, sec. 4.1]: A body can be defined as a number N of particles or mass points with N going to infinity.
[Hamel, 1949, sec. 23]: nicht wirklich hilfreich
[Roberson and Schwertassek, 1988, sec. 3.1.1]: definition by frames
[Kane and Levinson, 1985] no direct definition
[Shabana, 2005] no direct definition
[Abraham and Marsden, 1978, p. 229] no direct definition, but "The constrained system par excellence is the rigid body"

4.1 A single free rigid body

Established textbooks on physics (e.g. [Landau and Lifshitz, 1960, §31] or [Boltzmann, 1897, §44]) define a *rigid body* as a a system of $\mathfrak{N}$ particles such that the distances $d_{pq} = \|\mathfrak{r}_p - \mathfrak{r}_q\|$ between their positions $\mathfrak{r}_p$ are constant. For a system with $\mathfrak{N} > 3$ particles that do not lie in a plane, this would still allow a mirroring of the body about an arbitrary plane. Since this does change the shape of the body, additional constraints are necessary: For all particles we require $\langle(\mathfrak{r}_q - \mathfrak{r}_p) \times (\mathfrak{r}_s - \mathfrak{r}_p), \mathfrak{r}_t - \mathfrak{r}_p\rangle = c_{pqst} = const.$. The magnitude $|c_{pqst}|$ is already fixed by the corresponding particle distances d_{pq}, but the sign of c_{pqst} is positive (negative) if the vectors $\mathfrak{r}_q - \mathfrak{r}_p$, $\mathfrak{r}_s - \mathfrak{r}_p$ and $\mathfrak{r}_t - \mathfrak{r}_p$ form a right (left) handed basis of $\mathbb{R}^3$ and $c_{pqst} = 0$ if the four particles lie on a plane. For the following we require that for at least one quadruple of particles $c_{pqst} \neq 0$, i.e. not all particles lie in a common plane.

Based on the discussion above the configuration space for a rigid body as a system of constrained particles can be written as

$$\mathfrak{X} = \left\{ \begin{bmatrix} \mathfrak{r}_1 \\ \vdots \\ \mathfrak{r}_{\mathfrak{N}} \end{bmatrix} \in \mathbb{R}^{3\mathfrak{N}} \;\middle|\; \begin{array}{c} \|\mathfrak{r}_p - \mathfrak{r}_q\| = d_{pq}, \\ \langle(\mathfrak{r}_q - \mathfrak{r}_p) \times (\mathfrak{r}_s - \mathfrak{r}_p), \mathfrak{r}_t - \mathfrak{r}_p\rangle = c_{pqst}, \\ p, q, s, t = 1, \ldots, \mathfrak{N} \end{array} \right\} \tag{4.1}$$

Evidently, for large $\mathfrak{N}$ the number of constraints surpasses the count $3\mathfrak{N}$ of particle coordinates, so the geometric constraints cannot be independent.

4.1.1 Blah

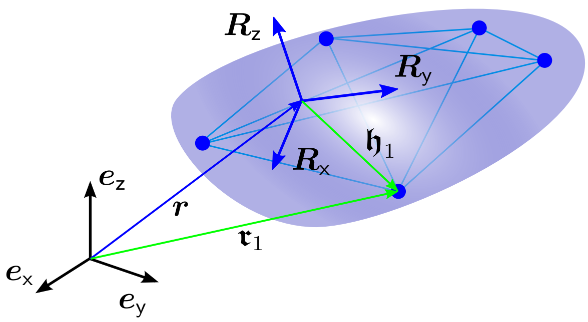


Figure 4.1: RigidBodyIllustration

For a given set of particle positions $(\mathfrak{r}_1, \ldots, \mathfrak{r}_{\mathfrak{N}}) \in \mathfrak{X}$ and the body configuration $(\boldsymbol{r}, \boldsymbol{R}) \in \mathbb{R}^3 \times \mathbb{SO}(3)$ we can compute the quantities

$$\mathfrak{h}_p = \boldsymbol{R}^\top(\mathfrak{r}_p - \boldsymbol{r}), \qquad p = 1, \ldots, \mathfrak{N}. \tag{4.2}$$

There is a mapping $h : \mathbb{X} = \mathbb{R}^3 \times \mathbb{SO}(3) \to \mathfrak{X}$ defined as

$$\begin{bmatrix} \mathfrak{r}_1 \\ \vdots \\ \mathfrak{r}_{\mathfrak{N}} \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_1 \\ \vdots \\ \boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_{\mathfrak{N}} \end{bmatrix}}_{h} \tag{4.3}$$

For the inverse function we use the result from subsection 4.1.5:

$$\mathcal{H} = \sum\nolimits_p \|\boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p - \mathfrak{r}_p\|^2 \geq 0 \tag{4.4}$$

has a minimum at

$$(\boldsymbol{r}, \boldsymbol{R}) = (\boldsymbol{p} - \bar{\boldsymbol{Y}}\bar{\boldsymbol{X}}^\top \boldsymbol{h}, \bar{\boldsymbol{Y}}\bar{\boldsymbol{X}}^\top) \tag{4.5}$$

where

$$k = \sum\nolimits_p \mathfrak{k}_p, \qquad \boldsymbol{h} = k^{-1}\sum\nolimits_p \mathfrak{k}_p\mathfrak{h}_p, \qquad \boldsymbol{p} = k^{-1}\sum\nolimits_p \mathfrak{k}_p\mathfrak{p}_p, \qquad \boldsymbol{P} = \sum\nolimits_p \mathfrak{k}_p\mathfrak{h}_p\mathfrak{p}_p^\top. \tag{4.6}$$

4.1.2 Coordinates

Configuration coordinates. Position $\boldsymbol{r}(t) \in \mathbb{R}^3$ and orientation $\boldsymbol{R} \in \mathbb{SO}(3)$.

The key argument here is

$$\mathfrak{r}_p = \boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p, \quad p = 1, \ldots, \mathfrak{N} \tag{4.7}$$

where $\boldsymbol{r}(t) \in \mathbb{R}^3$, $\boldsymbol{R}(t) \in \mathbb{SO}(3)$ and the *constant* particle position $\mathfrak{h}_p \in \mathbb{R}^3$ w.r.t. the body fixed frame.

$$\boldsymbol{x} = [\boldsymbol{r}^\top, \boldsymbol{R}_\mathsf{x}^\top, \boldsymbol{R}_\mathsf{y}^\top, \boldsymbol{R}_\mathsf{z}^\top]^\top \;\in\; \mathbb{X} \cong \mathbb{R}^3 \times \mathbb{SO}(3) \cong \mathfrak{X}. \tag{4.8}$$

$$d_{pq} = \|\mathfrak{r}_p - \mathfrak{r}_q\| = \|\boldsymbol{R}(\mathfrak{h}_p - \mathfrak{h}_q)\| \tag{4.9}$$

$$c_{pqst} = \langle(\boldsymbol{R}(\mathfrak{h}_q - \mathfrak{h}_p)) \times (\boldsymbol{R}(\mathfrak{h}_s - \mathfrak{h}_p)), \boldsymbol{R}(\mathfrak{h}_t - \mathfrak{h}_p)\rangle, \tag{4.10}$$

Velocity coordinates.

$$\underbrace{\begin{bmatrix} \dot{\boldsymbol{r}} \\ \dot{\boldsymbol{R}}_\mathsf{x} \\ \dot{\boldsymbol{R}}_\mathsf{y} \\ \dot{\boldsymbol{R}}_\mathsf{z} \end{bmatrix}}_{\dot{\boldsymbol{x}}} = \underbrace{\begin{bmatrix} \boldsymbol{R}_\mathsf{x} & \boldsymbol{R}_\mathsf{y} & \boldsymbol{R}_\mathsf{z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\boldsymbol{R}_\mathsf{z} & \boldsymbol{R}_\mathsf{y} \\ 0 & 0 & 0 & \boldsymbol{R}_\mathsf{z} & 0 & -\boldsymbol{R}_\mathsf{x} \\ 0 & 0 & 0 & -\boldsymbol{R}_\mathsf{y} & \boldsymbol{R}_\mathsf{x} & 0 \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} v_\mathsf{x} \\ v_\mathsf{y} \\ v_\mathsf{z} \\ \omega_\mathsf{x} \\ \omega_\mathsf{y} \\ \omega_\mathsf{z} \end{bmatrix}}_{\boldsymbol{\xi}}. \tag{4.11}$$

This can be written in the probably more familiar form

$$\dot{\boldsymbol{r}} = \boldsymbol{R}\boldsymbol{v}, \qquad \dot{\boldsymbol{R}} = \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega}). \tag{4.12}$$

With this we can express the particle velocities $\dot{\mathfrak{r}}_p$ and accelerations $\ddot{\mathfrak{r}}_p$ in terms of the configuration $(\boldsymbol{r}, \boldsymbol{R})$ and velocity coordinates $(\boldsymbol{v}, \boldsymbol{\omega})$ as

$$\mathfrak{r}_p = \boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p, \tag{4.13a}$$

$$\dot{\mathfrak{r}}_p = \boldsymbol{R}\big(\boldsymbol{v} - \mathrm{wed}(\mathfrak{h}_p)\boldsymbol{\omega}\big), \tag{4.13b}$$

$$\ddot{\mathfrak{r}}_p = \boldsymbol{R}\big(\dot{\boldsymbol{v}} - \mathrm{wed}(\mathfrak{h}_p)\dot{\boldsymbol{\omega}} + \mathrm{wed}(\boldsymbol{\omega})(\boldsymbol{v} - \mathrm{wed}(\mathfrak{h}_p)\boldsymbol{\omega})\big), \quad p = 1, \ldots, \mathfrak{N} \tag{4.13c}$$

and

$$\nabla \mathfrak{r}_p = \frac{\partial \dot{\mathfrak{r}}_p}{\partial \boldsymbol{\xi}} = \frac{\partial \ddot{\mathfrak{r}}_p}{\partial \dot{\boldsymbol{\xi}}} = \big[\boldsymbol{R},\ -\boldsymbol{R}\,\mathrm{wed}(\mathfrak{h}_p)\big] \tag{4.14}$$

Commutation coefficients. For the next steps we will also need the commutation coefficients γ for the chosen coordinates. To indicate the special case for the rigid body we use the sans font $\boldsymbol{\gamma}$ here. Plugging the kinematic matrix $\boldsymbol{A}$ from (4.11) into the definition (3.34) yields

$$\gamma^1_{26} = \gamma^1_{53} = \gamma^2_{34} = \gamma^2_{61} = \gamma^3_{15} = \gamma^3_{42} = \gamma^4_{56} = \gamma^5_{64} = \gamma^6_{45} = 1\ , \tag{4.15}$$

and the remaining coefficients vanish. It will be useful to define the matrix

$$\left[\gamma^k_{ij}\xi^j\right]^{k=1\ldots6}_{i=1\ldots6} = \begin{bmatrix} \mathrm{wed}(\boldsymbol{\omega}) & 0 \\ \mathrm{wed}(\boldsymbol{v}) & \mathrm{wed}(\boldsymbol{\omega}) \end{bmatrix} = -\mathrm{ad}^\top_{\boldsymbol{\xi}} \tag{4.16}$$

whose naming will be discussed later.

4.1.3 Inertia

Kinetic energy. Expressing the particle velocities $\dot{\mathfrak{r}}_p$ in terms of the chosen coordinates (4.13b) the kinetic energy $\mathcal{T}$ of a free rigid body is

$$\begin{aligned}
\mathcal{T} &= \tfrac{1}{2}\sum\nolimits_p \mathfrak{m}_p \| \overbrace{\boldsymbol{R}\big(\boldsymbol{v} - \mathrm{wed}(\mathfrak{h}_p)\boldsymbol{\omega}\big)}^{\dot{\mathfrak{r}}_p} \|^2 \\
&= \tfrac{1}{2}\underbrace{\sum\nolimits_p \mathfrak{m}_p}_{m} \|\boldsymbol{v}\|^2 - \boldsymbol{v}^\top \underbrace{\sum\nolimits_p \mathfrak{m}_p\,\mathrm{wed}(\mathfrak{h}_p)}_{m\,\mathrm{wed}(\boldsymbol{s})}\,\boldsymbol{\omega} + \tfrac{1}{2}\boldsymbol{\omega}^\top \underbrace{\sum\nolimits_p \mathfrak{m}_p\,\mathrm{wed}(\mathfrak{h}_p)^\top\,\mathrm{wed}(\mathfrak{h}_p)}_{\boldsymbol{\Theta}}\,\boldsymbol{\omega} \\
&= \tfrac{1}{2}\underbrace{\begin{bmatrix}\boldsymbol{v}^\top & \boldsymbol{\omega}^\top\end{bmatrix}}_{\boldsymbol{\xi}^\top} \underbrace{\begin{bmatrix} m\mathbf{I}_3 & m\,\mathrm{wed}(\boldsymbol{s})^\top \\ m\,\mathrm{wed}(\boldsymbol{s}) & \boldsymbol{\Theta}\end{bmatrix}}_{\mathsf{M}} \underbrace{\begin{bmatrix}\boldsymbol{v} \\ \boldsymbol{\omega}\end{bmatrix}}_{\boldsymbol{\xi}} .
\end{aligned} \tag{4.17}$$

Here we have substituted some well established inertia parameters: the total mass m, the center of mass $\boldsymbol{s} = m^{-1}\sum_p \mathfrak{m}_p\mathfrak{h}_p$ and the moment of inertia $\boldsymbol{\Theta} = \boldsymbol{\Theta}^\top$. Assuming that the particle masses are positive $\mathfrak{m}_p > 0, p = 1, \ldots, \mathfrak{N}$ implies that the total mass is positive $m > 0$. Furthermore, if the rigid body has at least three particles that do not lie on a line, the inertia matrix is positive definite $\boldsymbol{\Theta} > 0$. It is important to notice that the inertia matrix M for the chosen coordinates is *constant*[1].

[1] One reason for the choice of $\boldsymbol{v}$ as velocity coordinates it the fact that the inertia matrix M is *constant*. If we choose instead $\dot{\boldsymbol{r}}$ as velocity coordinates we have

$$\mathcal{T} = \tfrac{1}{2}[\dot{\boldsymbol{r}}^\top, \boldsymbol{\omega}^\top] \begin{bmatrix} m\mathbf{I}_3 & m\boldsymbol{R}\,\mathrm{wed}(\boldsymbol{s})^\top \\ m\,\mathrm{wed}(\boldsymbol{s})\boldsymbol{R}^\top & \boldsymbol{\Theta}\end{bmatrix} \begin{bmatrix}\dot{\boldsymbol{r}} \\ \boldsymbol{\omega}\end{bmatrix}.$$

Obviously the body inertia matrix depends on the orientation $\boldsymbol{R}$ of the body and is not constant unless the reference position $\boldsymbol{r}$ coincides with the *center of mass*, i.e. $\boldsymbol{s} = 0$. Actually many textbooks, e.g. [Murray et al., 1994, p. 167] or [Shabana, 2005, p. 153], restrict to this case for their expressions of the kinetic energy. In the next section on rigid body systems we will see that it can be quite useful to use *geometrically* meaningful body fixed frames rather than restricting to the center of mass.

Connection coefficients. As the inertia matrix M in (4.17) is constant, the corresponding connection coefficients Γ_{ijk} only consist of the terms with the commutation coefficients γ:

$$\Gamma_{ijk} = \tfrac{1}{2}\big(\gamma^s_{ij}\mathsf{M}_{sk} + \gamma^s_{ik}\mathsf{M}_{sj} - \gamma^s_{jk}\mathsf{M}_{si}\big) = -\Gamma_{jik}. \tag{4.18}$$

Taking into account this skew symmetry and the commutation coefficients γ given in (4.15), the non-zero connection coefficients are

$$\begin{aligned}
\Gamma_{324} = \Gamma_{135} = \Gamma_{216} &= m, && (4.19a)\\
\Gamma_{254} = \Gamma_{364} = \Gamma_{515} = \Gamma_{616} &= m s_{\mathsf{x}}, && (4.19b)\\
\Gamma_{424} = \Gamma_{145} = \Gamma_{365} = \Gamma_{626} &= m s_{\mathsf{y}}, && (4.19c)\\
\Gamma_{434} = \Gamma_{535} = \Gamma_{146} = \Gamma_{256} &= m s_{\mathsf{z}}, && (4.19d)\\
\Gamma_{654} &= \Theta'_{\mathsf{xx}} = \tfrac{1}{2}\big(\Theta_{\mathsf{yy}} + \Theta_{\mathsf{zz}} - \Theta_{\mathsf{xx}}\big), && (4.19e)\\
\Gamma_{465} &= \Theta'_{\mathsf{yy}} = \tfrac{1}{2}\big(\Theta_{\mathsf{xx}} + \Theta_{\mathsf{zz}} - \Theta_{\mathsf{yy}}\big), && (4.19f)\\
\Gamma_{546} &= \Theta'_{\mathsf{zz}} = \tfrac{1}{2}\big(\Theta_{\mathsf{xx}} + \Theta_{\mathsf{yy}} - \Theta_{\mathsf{zz}}\big), && (4.19g)\\
\Gamma_{464} = \Gamma_{655} &= \Theta'_{\mathsf{xy}} = -\Theta_{\mathsf{xy}}, && (4.19h)\\
\Gamma_{544} = \Gamma_{656} &= \Theta'_{\mathsf{xz}} = -\Theta_{\mathsf{xz}}, && (4.19i)\\
\Gamma_{545} = \Gamma_{466} &= \Theta'_{\mathsf{yz}} = -\Theta_{\mathsf{yz}}. && (4.19j)
\end{aligned}$$

The parameters $\boldsymbol{\Theta}' = \frac{1}{2}\operatorname{tr}(\boldsymbol{\Theta})\mathbf{I}_3 - \boldsymbol{\Theta}$ will appear several times in the following.

Acceleration energy. Expressing the particle accelerations $\ddot{\mathfrak{r}}_p$ in terms of the coordinates (4.13c) we find the acceleration energy $\mathcal{S}$ for the free rigid body as[2]

$$\begin{aligned}
\mathcal{S} = \tfrac{1}{2}\sum\nolimits_p \mathfrak{m}_p \| &\overbrace{\boldsymbol{R}\big(\dot{\boldsymbol{v}} - \operatorname{wed}(\mathfrak{h}_p)\dot{\boldsymbol{\omega}} + \operatorname{wed}(\boldsymbol{\omega})(\boldsymbol{v} - \operatorname{wed}(\mathfrak{h}_p)\boldsymbol{\omega})\big)}^{\ddot{\mathfrak{r}}_p} \|^2 \\
= \tfrac{1}{2}&\underbrace{\sum\nolimits_p \mathfrak{m}_p}_{m} \|\dot{\boldsymbol{v}}\|^2 - \dot{\boldsymbol{v}}^\top \underbrace{\sum\nolimits_p \mathfrak{m}_p \operatorname{wed}(\mathfrak{h}_p)}_{m \operatorname{wed}(\boldsymbol{s})} \dot{\boldsymbol{\omega}} + \tfrac{1}{2}\dot{\boldsymbol{\omega}}^\top \underbrace{\sum\nolimits_p \mathfrak{m}_p \operatorname{wed}(\mathfrak{h}_p)^\top \operatorname{wed}(\mathfrak{h}_p)}_{\boldsymbol{\Theta}} \dot{\boldsymbol{\omega}} \\
&+ \dot{\boldsymbol{v}}^\top \operatorname{wed}(\boldsymbol{\omega})\Big(\underbrace{\sum\nolimits_p \mathfrak{m}_p}_{m} \boldsymbol{v} - \underbrace{\sum\nolimits_p \mathfrak{m}_p \operatorname{wed}(\mathfrak{h}_p)}_{m \operatorname{wed}(\boldsymbol{s})} \boldsymbol{\omega}\Big) \\
&+ \dot{\boldsymbol{\omega}}^\top \Big(\underbrace{\sum\nolimits_p \mathfrak{m}_p \operatorname{wed}(\mathfrak{h}_p)}_{m \operatorname{wed}(\boldsymbol{s})} \operatorname{wed}(\boldsymbol{\omega})\boldsymbol{v} + \operatorname{wed}(\boldsymbol{\omega}) \underbrace{\sum\nolimits_p \mathfrak{m}_p \operatorname{wed}(\mathfrak{h}_p)^\top \operatorname{wed}(\mathfrak{h}_p)}_{\boldsymbol{\Theta}} \boldsymbol{\omega}\Big) \\
&+ \tfrac{1}{2}\underbrace{\sum\nolimits_p \mathfrak{m}_p}_{m} \|\operatorname{wed}(\boldsymbol{\omega})\boldsymbol{v}\|^2 + \boldsymbol{v}^\top \operatorname{wed}(\boldsymbol{\omega})^2 \underbrace{\sum\nolimits_p \mathfrak{m}_p \operatorname{wed}(\mathfrak{h}_p)}_{m \operatorname{wed}(\boldsymbol{s})} \boldsymbol{\omega} \\
&+ \tfrac{1}{2}\operatorname{tr}\Big(\underbrace{\sum\nolimits_p \mathfrak{m}_p \mathfrak{h}_p \mathfrak{h}_p^\top}_{\boldsymbol{\Theta}'} \operatorname{wed}(\boldsymbol{\omega})^4\Big)
\end{aligned} \tag{4.20}$$

[2]using the Jacobi identity $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$: $\operatorname{wed}(\boldsymbol{a})\operatorname{wed}(\boldsymbol{b})\boldsymbol{c} + \operatorname{wed}(\boldsymbol{b})\operatorname{wed}(\boldsymbol{c})\boldsymbol{a} + \operatorname{wed}(\boldsymbol{c})\operatorname{wed}(\boldsymbol{a})\boldsymbol{b} = \boldsymbol{0}$

As for the kinetic energy $\mathcal{T}$, the acceleration energy $\mathcal{S}$ for a free rigid body can be formulated using the inertia parameters $(m, \boldsymbol{s}, \boldsymbol{\Theta})$ as

$$
\begin{aligned}
\mathcal{S} = \tfrac{1}{2} \underbrace{\begin{bmatrix} \dot{\boldsymbol{v}}^\top & \dot{\boldsymbol{\omega}}^\top \end{bmatrix}}_{\dot{\boldsymbol{\xi}}^\top} & \underbrace{\begin{bmatrix} m\mathbf{I}_3 & m\,\mathrm{wed}(\boldsymbol{s})^\top \\ m\,\mathrm{wed}(\boldsymbol{s}) & \boldsymbol{\Theta} \end{bmatrix}}_{\mathsf{M}} \underbrace{\begin{bmatrix} \dot{\boldsymbol{v}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}} \\
& + \underbrace{\begin{bmatrix} \dot{\boldsymbol{v}}^\top & \dot{\boldsymbol{\omega}}^\top \end{bmatrix}}_{\dot{\boldsymbol{\xi}}^\top} \underbrace{\begin{bmatrix} \mathrm{wed}(\boldsymbol{\omega}) & \mathbf{0} \\ \mathrm{wed}(\boldsymbol{v}) & \mathrm{wed}(\boldsymbol{\omega}) \end{bmatrix}}_{-\mathrm{ad}_{\boldsymbol{\xi}}^\top} \underbrace{\begin{bmatrix} m\mathbf{I}_3 & m\,\mathrm{wed}(\boldsymbol{s})^\top \\ m\,\mathrm{wed}(\boldsymbol{s}) & \boldsymbol{\Theta} \end{bmatrix}}_{\mathsf{M}} \underbrace{\begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{bmatrix}}_{\boldsymbol{\xi}} \\
& + \underbrace{\tfrac{1}{2} m \|\,\mathrm{wed}(\boldsymbol{\omega})\boldsymbol{v}\|^2 + \boldsymbol{v}^\top \mathrm{wed}(\boldsymbol{\omega})^2 m\,\mathrm{wed}(\boldsymbol{s})\boldsymbol{\omega} + \tfrac{1}{2}\,\mathrm{tr}\left(\boldsymbol{\Theta}'\,\mathrm{wed}(\boldsymbol{\omega})^4\right)}_{\mathcal{S}_0}. \qquad (4.21)
\end{aligned}
$$

Note that $\mathcal{S}_0$ is independent of the generalized acceleration $\dot{\boldsymbol{\xi}}$ and consequently does not contribute to the generalized inertia force $\boldsymbol{f}^{\mathsf{M}} = \partial\mathcal{S}/\partial\dot{\boldsymbol{\xi}}$.

Inertia force. Finally, the inertia force of the free rigid body $\boldsymbol{f}^{\mathsf{M}}$, derived from any of the of the previous equivalent formulations, reads

$$\boldsymbol{f}^{\mathsf{M}} = \mathsf{M}\dot{\boldsymbol{\xi}} - \mathrm{ad}_{\boldsymbol{\xi}}^\top \mathsf{M}\boldsymbol{\xi}. \qquad (4.22)$$

4.1.4 Gravitation

The potential energy $\mathcal{V}^{\mathrm{G}}$ of a rigid body due to a gravitational acceleration $\boldsymbol{a}_{\mathsf{G}}$ according to (3.78) is

$$\mathcal{V}^{\mathrm{G}} = \sum_p \langle \boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p, \mathfrak{m}_p \boldsymbol{a}_{\mathsf{G}} \rangle = \langle \underbrace{\sum_p \mathfrak{m}_p}_{m} \boldsymbol{r} + \boldsymbol{R} \underbrace{\sum_p \mathfrak{m}_p \mathfrak{h}_p}_{m\boldsymbol{s}}, \boldsymbol{a}_{\mathsf{G}} \rangle, \qquad (4.23)$$

and the resulting generalized force is

$$\boldsymbol{f}^{\mathsf{G}} = \boldsymbol{\nabla}\mathcal{V}^{\mathrm{G}} = \begin{bmatrix} \boldsymbol{R}^\top m \boldsymbol{a}_{\mathsf{G}} \\ \mathrm{wed}(\boldsymbol{s})\boldsymbol{R}^\top m \boldsymbol{a}_{\mathsf{G}} \end{bmatrix}. \qquad (4.24)$$

4.1.5 Stiffness

Assume that every particle of the rigid body is connected to a position $\mathfrak{p}_p \in \mathbb{R}^3$ by a linear spring with stiffness $\mathfrak{k}_p \in \mathbb{R}^{(+)}$, see Figure 4.2. The resulting potential energy in terms of the rigid body coordinates $\boldsymbol{x} \cong (\boldsymbol{r}, \boldsymbol{R})$ is

$$\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) = \tfrac{1}{2} \sum_p \mathfrak{k}_p \|\boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p - \mathfrak{p}_p\|^2. \qquad (4.25)$$

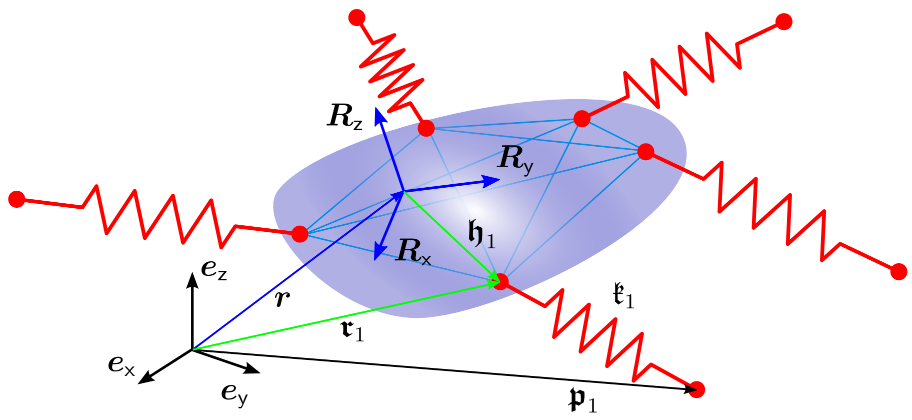


Figure 4.2: Rigid body stiffness

Some identities. In the following we use some basic identities

$$\begin{aligned}
\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n &: \boldsymbol{a}^\top \boldsymbol{b} = \mathrm{tr}(\boldsymbol{a}\boldsymbol{b}^\top) && (4.26a)\\
\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3 &: \mathrm{wed}(\boldsymbol{a})\boldsymbol{b} = 2\,\mathrm{vee}(\boldsymbol{b}\boldsymbol{a}^\top) && (4.26b)\\
&\quad \mathrm{wed}(\boldsymbol{a})\,\mathrm{wed}(\boldsymbol{b}) = \boldsymbol{b}\boldsymbol{a}^\top - (\boldsymbol{b}^\top\boldsymbol{a})\mathbf{I}_3 && (4.26c)\\
\boldsymbol{A} \in \mathbb{R}^{3\times 3}, \boldsymbol{b} \in \mathbb{R}^3 &: \mathrm{tr}(\boldsymbol{A}\,\mathrm{wed}(\boldsymbol{b})) = -2\boldsymbol{b}^\top\,\mathrm{vee}(\boldsymbol{A}) && (4.26d)\\
&\quad 2\,\mathrm{vee}(\mathrm{wed}(\boldsymbol{b})\boldsymbol{A}) = (\mathrm{tr}(\boldsymbol{A})\mathbf{I}_3 - \boldsymbol{A})\boldsymbol{b} && (4.26e)
\end{aligned}$$

Each may be checked by direct computation.

Stiffness parameters. Rearrange (4.25) to

$$\begin{aligned}
\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) &= \tfrac{1}{2}\sum\nolimits_p \mathfrak{k}_p \|\boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p - \mathfrak{p}_p\|^2 \\
&= \tfrac{1}{2}\sum\nolimits_p \mathfrak{k}_p \big(\|\boldsymbol{r}\|^2 + \underbrace{\|\boldsymbol{R}\mathfrak{h}_p\|^2}_{const.} + \underbrace{\|\mathfrak{p}_p\|^2}_{const.} + 2\langle \boldsymbol{r}, \boldsymbol{R}\mathfrak{h}_p\rangle - 2\langle \boldsymbol{r}, \mathfrak{p}_p\rangle - 2\langle \boldsymbol{R}\mathfrak{h}_p, \mathfrak{p}_p\rangle\big) \\
&= \sum\nolimits_p \mathfrak{k}_p \big(\tfrac{1}{2}\|\boldsymbol{r}\|^2 + \langle \boldsymbol{r}, \boldsymbol{R}\mathfrak{h}_p\rangle - \langle \boldsymbol{r}, \mathfrak{p}_p\rangle - \mathrm{tr}(\boldsymbol{R}\mathfrak{h}_p\mathfrak{p}_p^\top)\big) + \underbrace{\tfrac{1}{2}\sum\nolimits_p \mathfrak{k}_p\big(\|\mathfrak{h}_p\|^2 + \|\mathfrak{p}_p\|^2\big)}_{\mathcal{V}_c^{\mathrm{K}} = const.} \\
&= \tfrac{1}{2}k\|\boldsymbol{r}\|^2 + k\langle \boldsymbol{r}, \boldsymbol{R}\boldsymbol{h}\rangle - k\langle \boldsymbol{r}, \boldsymbol{p}\rangle - \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}) + \mathcal{V}_c^{\mathrm{K}} \qquad (4.27)
\end{aligned}$$

with substitution of the constant parameters

$$k = \sum\nolimits_p \mathfrak{k}_p, \quad \boldsymbol{h} = k^{-1}\sum\nolimits_p \mathfrak{k}_p\mathfrak{h}_p, \quad \boldsymbol{p} = k^{-1}\sum\nolimits_p \mathfrak{k}_p\mathfrak{p}_p, \quad \boldsymbol{P} = \sum\nolimits_p \mathfrak{k}_p\mathfrak{h}_p\mathfrak{p}_p^\top. \qquad (4.28)$$

Since there is no assumption on the particle and spring constellation, any constellation may be captured by the $1+3+3+9+1 = 17$ parameters within $(k, \boldsymbol{h}, \boldsymbol{p}, \boldsymbol{P}, \mathcal{V}_c^{\mathrm{K}})$.

Critical points. The time derivatives of the potential may be written as

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}^{\mathrm{K}} &= k\langle \boldsymbol{r}, \boldsymbol{R}\boldsymbol{v}\rangle + k\langle \boldsymbol{R}\boldsymbol{v}, \boldsymbol{R}\boldsymbol{h}\rangle + k\langle \boldsymbol{r}, \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega})\boldsymbol{h}\rangle - k\langle \boldsymbol{R}\boldsymbol{v}, \boldsymbol{p}\rangle - \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega}))\\ &= \boldsymbol{\xi}^{\top}\underbrace{\begin{bmatrix} k(\boldsymbol{R}^{\top}(\boldsymbol{r}-\boldsymbol{p})+\boldsymbol{h}) \\ k\,\mathrm{wed}(\boldsymbol{h})\boldsymbol{R}^{\top}\boldsymbol{r} + 2\,\mathrm{vee}(\boldsymbol{P}\boldsymbol{R})\end{bmatrix}}_{\boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}}}\end{aligned} \tag{4.29}$$

$$\tfrac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{V}^{\mathrm{K}} = \dot{\boldsymbol{\xi}}^{\top}\boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}} + \boldsymbol{\xi}^{\top}\underbrace{\begin{bmatrix} k\mathbf{I}_3 & k\,\mathrm{wed}\big(\boldsymbol{R}^{\top}(\boldsymbol{r}-\boldsymbol{p})\big)^{\top} \\ k\,\mathrm{wed}(\boldsymbol{h}) & k\,\mathrm{wed}(\boldsymbol{h})\,\mathrm{wed}(\boldsymbol{R}^{\top}\boldsymbol{r}) + \mathrm{tr}(\boldsymbol{P}\boldsymbol{R})\mathbf{I}_3 - (\boldsymbol{P}\boldsymbol{R})^{\top}\end{bmatrix}}_{\boldsymbol{\nabla}^2\mathcal{V}^{\mathrm{K}}}\boldsymbol{\xi} \tag{4.30}$$

We are interested in configurations $\boldsymbol{x}_{\mathsf{R}} \cong (\boldsymbol{r}_{\mathsf{R}}, \boldsymbol{R}_{\mathsf{R}})$ at which the potential is stationary $\boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}_{\mathsf{R}}) = \mathbf{0}$. From the upper part of (4.29) we get the condition

$$\boldsymbol{r}_{\mathsf{R}} = \boldsymbol{p} - \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h}. \tag{4.31}$$

Plugging this into the lower part of (4.29) we obtain

$$k\,\mathrm{wed}(\boldsymbol{h})\boldsymbol{R}_{\mathsf{R}}^{\top}(\boldsymbol{p} - \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h}) + 2\,\mathrm{vee}(\boldsymbol{P}\boldsymbol{R}_{\mathsf{R}}) = 2\,\mathrm{vee}(\underbrace{(\boldsymbol{P} - k\boldsymbol{h}\boldsymbol{p}^{\top})}_{\boldsymbol{P}_s}\boldsymbol{R}_{\mathsf{R}}) = \mathbf{0}. \tag{4.32}$$

So we need $\boldsymbol{R}_{\mathsf{R}} \in \mathbb{SO}(3)$ such that $\boldsymbol{P}_s\boldsymbol{R}_{\mathsf{R}}$ is symmetric. Let $\boldsymbol{P}_s = \boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{Y}^{\top}$ with $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{O}(3)$ and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$, $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$ be a singular value decomposition and define

$$\bar{\boldsymbol{X}} = \boldsymbol{X}\,\mathrm{diag}(1, 1, \det\boldsymbol{X}) \in \mathbb{SO}(3), \tag{4.33a}$$
$$\bar{\boldsymbol{Y}} = \boldsymbol{Y}\,\mathrm{diag}(1, 1, \det\boldsymbol{Y}) \in \mathbb{SO}(3), \tag{4.33b}$$
$$\bar{\boldsymbol{\Sigma}} = \mathrm{diag}(\sigma_1, \sigma_2, \det\boldsymbol{X}\,\det\boldsymbol{Y}\,\sigma_3). \tag{4.33c}$$

With this we may write $\boldsymbol{P}_s\boldsymbol{R}_{\mathsf{R}} = \bar{\boldsymbol{X}}\bar{\boldsymbol{\Sigma}}\bar{\boldsymbol{X}}^{\top}\bar{\boldsymbol{X}}\bar{\boldsymbol{Y}}^{\top}\boldsymbol{R}_{\mathsf{R}}$ which is clearly symmetric for $\boldsymbol{R}_{\mathsf{R}} = \bar{\boldsymbol{Y}}\bar{\boldsymbol{X}}^{\top}$. The Hessian at this critical point is

$$\begin{aligned}\boldsymbol{\nabla}^2\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}_{\mathsf{R}}) &= \begin{bmatrix} k\mathbf{I}_3 & k\,\mathrm{wed}(\boldsymbol{h})^{\top} \\ k\,\mathrm{wed}(\boldsymbol{h}) & \mathrm{tr}(\boldsymbol{P}_s\boldsymbol{R}_{\mathsf{R}})\mathbf{I}_3 - \boldsymbol{P}_s\boldsymbol{R}_{\mathsf{R}} + k\,\mathrm{wed}(\boldsymbol{h})\,\mathrm{wed}(\boldsymbol{h})^{\top}\end{bmatrix}\\ &= \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \\ \mathrm{wed}(\boldsymbol{h}) & \mathbf{I}_3\end{bmatrix}\begin{bmatrix} k\mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Pi}_{\mathsf{s}}\end{bmatrix}\begin{bmatrix} \mathbf{I}_3 & \mathrm{wed}(\boldsymbol{h})^{\top} \\ \mathbf{0} & \mathbf{I}_3\end{bmatrix}\end{aligned} \tag{4.34}$$

with $\boldsymbol{\Pi}_{\mathsf{s}} = \mathrm{tr}(\boldsymbol{P}_s\boldsymbol{R}_{\mathsf{R}})\mathbf{I}_3 - \boldsymbol{P}_s\boldsymbol{R}_{\mathsf{R}} = \bar{\boldsymbol{X}}(\mathrm{tr}(\bar{\boldsymbol{\Sigma}})\mathbf{I}_3 - \bar{\boldsymbol{\Sigma}})\bar{\boldsymbol{X}}^{\top} \in \mathbb{SYM}_0^+(3)$. The positive semi-definiteness of $\mathrm{tr}(\bar{\boldsymbol{\Sigma}})\mathbf{I}_3 - \bar{\boldsymbol{\Sigma}}$ is a consequence of σ_3 being the smallest singular value.

The rotational part of $\mathcal{V}^{\mathrm{K}}$ is analyzed thoroughly in subsection A.4.4 which also discusses saddle points and maxima. From the conclusions there we may deduct: $\boldsymbol{x}_{\mathsf{R}} \cong (\boldsymbol{r}_{\mathsf{R}}, \boldsymbol{R}_{\mathsf{R}}) = (\boldsymbol{p} - \bar{\boldsymbol{Y}}\bar{\boldsymbol{X}}^{\top}\boldsymbol{h}, \bar{\boldsymbol{Y}}\bar{\boldsymbol{X}}^{\top})$ is a minimum of $\mathcal{V}^{\mathrm{K}}$. It is the strict, global minimum, if and only if, $\boldsymbol{\Pi}_{\mathsf{s}}$ is positive definite.

Stiffness parameters cont'd. We may express the parameters $\boldsymbol{p}$ and $\boldsymbol{P}$ in terms of the minimum configuration $(\boldsymbol{r}_{\mathsf{R}}, \boldsymbol{R}_{\mathsf{R}})$ and the moment of stiffness matrix $\boldsymbol{\Pi}_{\mathsf{s}}$ as

$$\boldsymbol{p} = \boldsymbol{r}_{\mathsf{R}} + \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h} \tag{4.35}$$
$$\boldsymbol{P} = \boldsymbol{P}_s + k\boldsymbol{h}\boldsymbol{p}^{\top} = (\tfrac{1}{2}\,\mathrm{tr}(\boldsymbol{\Pi}_{\mathsf{s}})\mathbf{I}_3 - \boldsymbol{\Pi}_{\mathsf{s}})\boldsymbol{R}_{\mathsf{R}}^{\top} + k\boldsymbol{h}(\boldsymbol{r}_{\mathsf{R}} + \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h})^{\top} \tag{4.36}$$

Plugging this into (4.27) yields

$$\begin{aligned}\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) &= \tfrac{1}{2}k\|\boldsymbol{r}\|^2 + k\langle \boldsymbol{r}, \boldsymbol{R}\boldsymbol{h}\rangle - k\langle \boldsymbol{r}, \boldsymbol{r}_{\mathsf{R}} + \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h}\rangle \\ &\qquad - \operatorname{tr}\big(((\tfrac{1}{2}\operatorname{tr}(\boldsymbol{\Pi}_{\mathsf{s}})\mathbf{I}_3 - \boldsymbol{\Pi}_{\mathsf{s}})\boldsymbol{R}_{\mathsf{R}}^{\top} + k\boldsymbol{h}(\boldsymbol{r}_{\mathsf{R}} + \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h})^{\top})\boldsymbol{R}\big) + \mathcal{V}_c^{\mathrm{K}} \\ &= \tfrac{1}{2}k\|\boldsymbol{r}\|^2 - k\langle \boldsymbol{r}, \boldsymbol{r}_{\mathsf{R}}\rangle + k\langle \boldsymbol{r}, \boldsymbol{R}\boldsymbol{h}\rangle - k\langle \boldsymbol{r}, \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h}\rangle - k\langle \boldsymbol{r}_{\mathsf{R}}, \boldsymbol{R}\boldsymbol{h}\rangle \\ &\qquad - \operatorname{tr}\big(\underbrace{(\tfrac{1}{2}\operatorname{tr}(\boldsymbol{\Pi}_{\mathsf{s}})\mathbf{I}_3 - \boldsymbol{\Pi}_{\mathsf{s}} + k\boldsymbol{h}\boldsymbol{h}^{\top})}_{\boldsymbol{\Pi}'}\boldsymbol{R}_{\mathsf{R}}^{\top}\boldsymbol{R}\big) + \mathcal{V}_c^{\mathrm{K}} \\ &= \tfrac{1}{2}k\|\boldsymbol{r} - \boldsymbol{r}_{\mathsf{R}}\|^2 + k\langle \boldsymbol{r} - \boldsymbol{r}_{\mathsf{R}}, (\boldsymbol{R} - \boldsymbol{R}_{\mathsf{R}})\boldsymbol{h}\rangle + \operatorname{tr}\big(\boldsymbol{\Pi}'(\mathbf{I}_3 - \boldsymbol{R}_{\mathsf{R}}^{\top}\boldsymbol{R})\big) \\ &\qquad + \underbrace{\mathcal{V}_c^{\mathrm{K}} - \tfrac{1}{2}k\|\boldsymbol{r}_{\mathsf{R}}\|^2 - k\langle \boldsymbol{r}_{\mathsf{R}}, \boldsymbol{R}_{\mathsf{R}}\boldsymbol{h}\rangle - \operatorname{tr}(\boldsymbol{\Pi}')}_{\mathcal{V}_0^{\mathrm{K}} = const.}\end{aligned} \tag{4.37}$$

The potential $\mathcal{V}_0^{\mathrm{K}}$ at the minimum is

$$\mathcal{V}_0^{\mathrm{K}} = \mathcal{V}^{\mathrm{K}}(\boldsymbol{x}_{\mathsf{R}}) = \tfrac{1}{2}\sum\nolimits_p \mathfrak{k}_p\|\boldsymbol{r}_{\mathsf{R}} + \boldsymbol{R}_{\mathsf{R}}\mathfrak{h}_p - \mathfrak{p}_p\|^2 \geq 0. \tag{4.38}$$

The differential may be written as

$$\boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) = \begin{bmatrix} k\boldsymbol{R}^{\top}(\boldsymbol{r}-\boldsymbol{r}_{\mathsf{R}}) + (\mathbf{I}_3-\boldsymbol{R}^{\top}\boldsymbol{R}_{\mathsf{R}})k\boldsymbol{h} \\ k\operatorname{wed}(\boldsymbol{h})\boldsymbol{R}^{\top}(\boldsymbol{r}-\boldsymbol{r}_{\mathsf{R}}) + \operatorname{vee2}\big(\boldsymbol{\Pi}'\boldsymbol{R}_{\mathsf{R}}^{\top}\boldsymbol{R}\big)\end{bmatrix} \tag{4.39}$$

and the Hessian at the minimum is

$$\boldsymbol{\nabla}^2\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}_{\mathsf{R}}) = \begin{bmatrix} k\mathbf{I}_3 & k\operatorname{wed}(\boldsymbol{h})^{\top} \\ k\operatorname{wed}(\boldsymbol{h}) & \boldsymbol{\Pi}\end{bmatrix} \geq 0 \tag{4.40}$$

where $\boldsymbol{\Pi} = \operatorname{tr}(\boldsymbol{\Pi}')\mathbf{I}_3 - \boldsymbol{\Pi}' = \boldsymbol{\Pi}_{\mathsf{s}} + k\operatorname{wed}(\boldsymbol{h})\operatorname{wed}(\boldsymbol{h})^{\top} \in \mathbb{SYM}_0^+(3)$.

Conclusion. The conclusion of this subsection is that any constellation of linear springs attached to a rigid body may be captured by the potential $\mathcal{V}^{\mathrm{K}}$ from (4.37) and the resulting force $\boldsymbol{f}^{\mathrm{K}} = \boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}}$ from (4.39). It is parameterized by 6 parameters within $(\boldsymbol{r}_{\mathsf{R}}, \boldsymbol{R}_{\mathsf{R}}) \in \mathbb{R}^3 \times \mathbb{SO}(3)$ which describe the minimum, the $1+3+6 = 10$ parameters within the total stiffness $k \in \mathbb{R}_0^+$, the center of stiffness $\boldsymbol{h} \in \mathbb{R}^3$ and the moment of stiffness $\boldsymbol{\Pi} \in \mathbb{SYM}_0^+(3)$ and the offset $\mathcal{V}_0^{\mathrm{K}}$: $\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) \geq \mathcal{V}_0^{\mathrm{K}} \geq 0$. The naming is due to the obvious analogy to the inertia parameters from the previous section. The minimum is strict and global if, and only if, $k > 0$ and $\boldsymbol{\Pi}_{\mathsf{s}} > 0$.

4.1.6 Dissipation

Since the dissipation function $\mathcal{R}$ has the structure we already saw for the kinetic energy (4.17) we can immediately conclude

$$
\begin{aligned}
\mathcal{R} &= \tfrac{1}{2}\sum\nolimits_p \mathfrak{d}_p \|\dot{\mathfrak{r}}_p\|^2 \\
&= \tfrac{1}{2}\underbrace{\sum\nolimits_p \mathfrak{d}_p}_{d} \|\boldsymbol{v}\|^2 - \boldsymbol{v}^\top \underbrace{\sum\nolimits_p \mathfrak{d}_p \,\mathrm{wed}(\mathfrak{h}_p)}_{d\,\mathrm{wed}(\boldsymbol{l})} \boldsymbol{\omega} + \tfrac{1}{2}\boldsymbol{\omega}^\top \underbrace{\sum\nolimits_p \mathfrak{d}_p \,\mathrm{wed}(\mathfrak{h}_p)^\top \mathrm{wed}(\mathfrak{h}_p)}_{\boldsymbol{\Upsilon}} \boldsymbol{\omega} \\
&= \tfrac{1}{2}\underbrace{\begin{bmatrix}\boldsymbol{v}^\top & \boldsymbol{\omega}^\top\end{bmatrix}}_{\boldsymbol{\xi}^\top} \underbrace{\begin{bmatrix} d\mathbf{I}_3 & d\,\mathrm{wed}(\boldsymbol{l})^\top \\ d\,\mathrm{wed}(\boldsymbol{l}) & \boldsymbol{\Upsilon}\end{bmatrix}}_{\mathsf{D}} \underbrace{\begin{bmatrix}\boldsymbol{v} \\ \boldsymbol{\omega}\end{bmatrix}}_{\boldsymbol{\xi}}
\end{aligned}
\tag{4.41}
$$

The resulting generalized force is

$$
\boldsymbol{f}^{\mathsf{D}} = \frac{\partial \mathcal{R}}{\partial \boldsymbol{\xi}} = \mathsf{D}\boldsymbol{\xi} \tag{4.42}
$$

4.1.7 External forces.

Let external forces $\boldsymbol{\mathfrak{F}}_p^{\mathsf{E}} = \boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u})$

$$
\boldsymbol{f}^{\mathsf{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u}) = \Big[\sum\nolimits_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u})\rangle\Big]_{i=1,\dots,6} = \sum\nolimits_p \begin{bmatrix}\boldsymbol{R} \\ \mathrm{wed}(\mathfrak{h}_p)\boldsymbol{R}\end{bmatrix} \boldsymbol{\mathfrak{F}}_p^{\mathsf{E}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{u}) \tag{4.43}
$$

4.2 A more abstract view on the rigid body

Real, quadratic matrix sets

symmetric	$\mathbb{SYM}(m) = \{\boldsymbol{A} \in \mathbb{R}^{m\times m} \,\|\, \boldsymbol{A} = \boldsymbol{A}^\top\}$
symmetric, pos. def.	$\mathbb{SYM}^+(m) = \{\boldsymbol{A} \in \mathbb{SYM}(m) \,\|\, \boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} > 0 \,\forall\, \boldsymbol{x} \in \mathbb{R}^m \backslash \{\boldsymbol{0}\}\}$
symmetric, pos. semi-def.	$\mathbb{SYM}_0^+(m) = \{\boldsymbol{A} \in \mathbb{SYM}(m) \,\|\, \boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} \geq 0 \,\forall\, \boldsymbol{x} \in \mathbb{R}^m \backslash \{\boldsymbol{0}\}\}$

4.2.1 The special Euclidean group

Matrix Lie groups

$$
\text{special orthogonal} \qquad \mathbb{SO}(m) = \{\boldsymbol{R} \in \mathbb{R}^{m\times m} \mid \boldsymbol{R}^\top\boldsymbol{R} = \mathbf{I}_m,\ \det \boldsymbol{R} = +1\} \tag{4.45a}
$$

$$
\text{special Euclidean} \qquad \mathbb{SE}(m) = \left\{ \begin{bmatrix}\boldsymbol{R} & \boldsymbol{r} \\ \boldsymbol{0} & 1\end{bmatrix} \;\middle|\; \boldsymbol{r} \in \mathbb{R}^m, \boldsymbol{R} \in \mathbb{SO}(m) \right\} \tag{4.45b}
$$

The group operation is the matrix multiplication, the identity element is the identity matrix $\mathbf{I}_m$ resp. $\mathbf{I}_{m+1}$ and the inverse element is the inverse matrix. The associated Lie algebras are the vector spaces

$$\mathfrak{so}(m) = \{\boldsymbol{\Omega} \in \mathbb{R}^{m\times m} \,|\, \boldsymbol{\Omega}^\top = -\boldsymbol{\Omega}\} \tag{4.46a}$$

$$\mathfrak{se}(m) = \left\{ \begin{bmatrix} \boldsymbol{\Omega} & \boldsymbol{v} \\ \mathbf{0} & 0 \end{bmatrix} \,\middle|\, \boldsymbol{v} \in \mathbb{R}^m, \boldsymbol{\Omega} \in \mathfrak{so}(m) \right\} \tag{4.46b}$$

with the bracket operation $[\boldsymbol{X}, \boldsymbol{Y}] = \boldsymbol{XY} - \boldsymbol{YX}$ for $\boldsymbol{X}, \boldsymbol{Y} \in \mathfrak{so}(m)$ resp. $\mathfrak{se}(m)$, i.e. the commutator of the matrix multiplication.

Minimal coordinates. For the following we restrict to the $m = 3$ case. Introducing the *wedge* operator, we can parameterize the $n = 3$ resp. $n = 6$ dimensional Lie algebras (4.46) by a minimal set of coordinates

$$\mathrm{wed} : \mathbb{R}^3 \to \mathfrak{so}(3) \, : \, \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \mapsto \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}, \quad \mathfrak{so}(3) = \{\mathrm{wed}\, \boldsymbol{\omega} \,|\, \boldsymbol{\omega} \in \mathbb{R}^3\}, \tag{4.47a}$$

$$\mathrm{wed} : \mathbb{R}^6 \to \mathfrak{se}(3) \, : \, \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{bmatrix} \mapsto \begin{bmatrix} \mathrm{wed}\, \boldsymbol{\omega} & \boldsymbol{v} \\ \mathbf{0} & 0 \end{bmatrix}, \qquad \mathfrak{se}(3) = \{\mathrm{wed}\, \boldsymbol{\xi} \,|\, \boldsymbol{\xi} \in \mathbb{R}^6\}. \tag{4.47b}$$

The inverse operator is denoted by $\mathrm{vee}(\cdot)$.

Take the derivative of the geometric constraint for $\boldsymbol{R}(t) \in \mathbb{SO}(3)$ yields

$$\boldsymbol{R}^\top \boldsymbol{R} = \mathbf{I}_3 \quad \Rightarrow \quad \underbrace{\boldsymbol{R}^\top \dot{\boldsymbol{R}}}_{\boldsymbol{\Omega}} + \underbrace{\dot{\boldsymbol{R}}^\top \boldsymbol{R}}_{\boldsymbol{\Omega}^\top} = \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{\Omega} = -\boldsymbol{\Omega}^\top = \mathrm{wed}(\boldsymbol{\omega}) \tag{4.48}$$

With this we can give a kinematic relation that ensures that a group element remains in the group and that parameterizes its change by minimal velocity coordinates

$$\dot{\boldsymbol{R}} = \boldsymbol{R}\, \mathrm{wed}(\boldsymbol{\omega}), \qquad \boldsymbol{R}(t) \in \mathbb{SO}(3),\ \boldsymbol{\omega}(t) \in \mathbb{R}^3, \tag{4.49a}$$

$$\dot{\mathbf{G}} = \mathbf{G}\, \mathrm{wed}(\boldsymbol{\xi}), \qquad \mathbf{G}(t) \in \mathbb{SE}(3),\ \boldsymbol{\xi}(t) \in \mathbb{R}^6. \tag{4.49b}$$

Adjoint representations. It can be useful to represent the elements of a Lie Group $\mathbb{G}$ by their linear transformation of the associated Lie algebra. For $\boldsymbol{A} \in \mathbb{G}$ and $\boldsymbol{X}, \boldsymbol{Y} \in \mathfrak{g}$ define the *adjoint representation* as [Hall, 2015, Def. 3.32]:

$$\mathrm{Ad}_{\boldsymbol{A}} \boldsymbol{X} = \boldsymbol{AXA}^{-1}, \quad \mathrm{ad}_{\boldsymbol{X}} \boldsymbol{Y} = \boldsymbol{XY} - \boldsymbol{YX}. \tag{4.50}$$

Since we introduced minimal coordinates in (4.47) we can adjust these representations to act on the minimal coordinates of the Lie algebra: For $\boldsymbol{A} \in \mathbb{G}$ and $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^n$, where $n = \dim \mathbb{G} = \dim \mathfrak{g}$, define

$$\mathrm{Ad}_{\boldsymbol{A}} \boldsymbol{\xi} = \mathrm{vee}\left(\boldsymbol{A}\, \mathrm{wed}\, \boldsymbol{\xi}\, \boldsymbol{A}^{-1}\right), \qquad \mathrm{ad}_{\boldsymbol{\xi}} \boldsymbol{\eta} = \mathrm{vee}\left(\mathrm{wed}\, \boldsymbol{\xi}\, \mathrm{wed}\, \boldsymbol{\eta} - \mathrm{wed}\, \boldsymbol{\xi}\, \mathrm{wed}\, \boldsymbol{\eta}\right). \tag{4.51}$$

The adjoint representations for the particular Lie groups with $\boldsymbol{R} \in \mathbb{SO}(3)$, $\boldsymbol{\omega} \in \mathbb{R}^3$ and $\mathbf{G} = \left[\begin{smallmatrix} \boldsymbol{R} & \boldsymbol{r} \\ \mathbf{0} & 1 \end{smallmatrix}\right] \in \mathbb{SE}(3)$, $\boldsymbol{\xi} = \left[\begin{smallmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{smallmatrix}\right] \in \mathbb{R}^6$ are

$$\mathrm{Ad}_{\boldsymbol{R}} = \boldsymbol{R}, \qquad \mathrm{ad}_{\boldsymbol{\omega}} = \mathrm{wed}\, \boldsymbol{\omega} \tag{4.52a}$$

$$\mathrm{Ad}_{\mathbf{G}} = \begin{bmatrix} \boldsymbol{R} & \mathrm{wed}(\boldsymbol{r})\, \boldsymbol{R} \\ \mathbf{0} & \boldsymbol{R} \end{bmatrix}, \qquad \mathrm{ad}_{\boldsymbol{\xi}} = \begin{bmatrix} \mathrm{wed}\, \boldsymbol{\omega} & \mathrm{wed}\, \boldsymbol{v} \\ \mathbf{0} & \mathrm{wed}\, \boldsymbol{\omega} \end{bmatrix} \tag{4.52b}$$

Furthermore we have the identities

$$\mathrm{Ad}_{\mathbf{G}^{-1}} = \mathrm{Ad}_{\mathbf{G}}^{-1}, \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{Ad}_{\mathbf{G}}\big) = \mathrm{Ad}_{\mathbf{G}}\,\mathrm{ad}_{\boldsymbol{\xi}} \tag{4.53}$$

Further operators. Define the Vee operator through

$$\mathrm{tr}\left(\mathrm{wed}\,\boldsymbol{\xi}\,\boldsymbol{A}(\mathrm{wed}\,\boldsymbol{\eta})^\top\right) = \boldsymbol{\eta}^\top(\mathrm{Vee}\,\boldsymbol{A})\boldsymbol{\xi}. \tag{4.54}$$

The particular important cases for this work are

$$\mathrm{Vee}: \mathbb{R}^{3\times3} \to \mathbb{R}^{3\times3} \;:\; \boldsymbol{A} \mapsto \mathrm{tr}(\boldsymbol{A})\mathbf{I}_3 - \boldsymbol{A}, \tag{4.55a}$$

$$\mathrm{Vee}: \mathbb{R}^{4\times4} \to \mathbb{R}^{6\times6} \;:\; \begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c}^\top & d \end{bmatrix} \mapsto \begin{bmatrix} d\mathbf{I}_3 & (\mathrm{wed}\,\boldsymbol{b})^\top \\ \mathrm{wed}\,\boldsymbol{c} & \mathrm{Vee}\,\boldsymbol{A} \end{bmatrix}, \tag{4.55b}$$

The inverse operator is denoted by Wed. The Vee operator maps the coefficients of the map (4.54) with the arguments $\mathrm{wed}\,\boldsymbol{\xi}, \mathrm{wed}\,\boldsymbol{\eta} \in \mathfrak{g}$ to the coefficients of the corresponding map with the minimal coordinates $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^{\dim \mathfrak{g}}$. The naming is due to this similarity to the previously (4.47) defined vee operator.

Define the vee2 operator through

$$\mathrm{tr}\left(\boldsymbol{A}(\mathrm{wed}\,\boldsymbol{\xi})^\top\right) = \boldsymbol{\xi}^\top\,\mathrm{vee2}(\boldsymbol{A}). \tag{4.56}$$

The particular important cases for this work are

$$\mathrm{vee2}: \mathbb{R}^{3\times3} \to \mathbb{R}^{3} \;:\; \begin{bmatrix} * & A_{12} & A_{13} \\ A_{21} & * & A_{23} \\ A_{31} & A_{32} & * \end{bmatrix} \mapsto \begin{bmatrix} A_{32} - A_{23} \\ A_{13} - A_{31} \\ A_{21} - A_{12} \end{bmatrix}, \tag{4.57a}$$

$$\mathrm{vee2}: \mathbb{R}^{4\times4} \to \mathbb{R}^{6} \;:\; \begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ * & * \end{bmatrix} \mapsto \begin{bmatrix} \boldsymbol{b} \\ \mathrm{vee2}\,\boldsymbol{A} \end{bmatrix}. \tag{4.57b}$$

Note that for $\boldsymbol{\Omega} \in \mathfrak{so}(3) \subset \mathbb{R}^{3\times3}$ we have $\mathrm{vee2}(\boldsymbol{\Omega}) = 2\,\mathrm{vee}(\boldsymbol{\Omega})$, thus giving the motivation for the name.

Combining the definitions (4.54) and (4.56) also yields

$$\mathrm{vee2}\left(\mathrm{wed}\,\boldsymbol{\xi}\,\boldsymbol{A}\right) = \mathrm{Vee}\,\boldsymbol{A}\,\boldsymbol{\xi} \tag{4.58}$$

further identities

$$\mathrm{Ad}_{\mathbf{G}}^\top\,\mathrm{Vee}\,\boldsymbol{A}\,\mathrm{Ad}_{\mathbf{G}} = \mathrm{Vee}\left(\mathbf{G}^{-1}\boldsymbol{A}(\mathbf{G}^{-1})^\top\right), \tag{4.59a}$$

$$\mathrm{Ad}_{\mathbf{G}}^\top\,\mathrm{vee2}\,\boldsymbol{A} = \mathrm{vee2}\left(\mathbf{G}^\top\boldsymbol{A}(\mathbf{G}^{-1})^\top\right) \tag{4.59b}$$

$$\mathrm{vee2}(\mathbf{G}\,\mathrm{wed}(\boldsymbol{\xi})\boldsymbol{A}) = \mathrm{Vee}(\mathbf{G}\boldsymbol{A})\mathrm{Ad}_{\mathbf{G}}\boldsymbol{\xi} \tag{4.59c}$$

4.2.2 Inner product

For matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times m}$ and a symmetric, positive definite matrix $\boldsymbol{K} \in \mathbb{SYM}^+(n)$, define an *inner product* as

$$\langle \boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{K}} = \mathrm{tr}(\boldsymbol{A}^\top\boldsymbol{K}\boldsymbol{B}). \tag{4.60}$$

with the following properties $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{n\times m}$, $\lambda \in \mathbb{R}$:

$$\text{Symmetry} \qquad \langle \boldsymbol{B}, \boldsymbol{A}\rangle_{\boldsymbol{K}} = \langle \boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{K}} \tag{4.61a}$$

$$\text{Linearity} \qquad \langle \lambda\boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{K}} = \lambda\langle \boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{K}} = \langle \boldsymbol{A}, \lambda\boldsymbol{B}\rangle_{\boldsymbol{K}}, \tag{4.61b}$$

$$\langle \boldsymbol{A}+\boldsymbol{C}, \boldsymbol{B}\rangle_{\boldsymbol{K}} = \langle \boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{K}} + \langle \boldsymbol{C}, \boldsymbol{B}\rangle_{\boldsymbol{K}} \tag{4.61c}$$

$$\text{Pos. definit.} \qquad \langle \boldsymbol{A}, \boldsymbol{A}\rangle_{\boldsymbol{K}} \geq 0, \quad \langle \boldsymbol{A}, \boldsymbol{A}\rangle_{\boldsymbol{K}} = 0 \;\Leftrightarrow\; \boldsymbol{A} = \boldsymbol{0}. \tag{4.61d}$$

Let $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ be the columns of $\boldsymbol{A}$ and $\boldsymbol{B}$, then

$$\langle \boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{K}} = \operatorname{tr}\left(\begin{bmatrix}\boldsymbol{a}_1^\top\\ \vdots \\ \boldsymbol{a}_n^\top\end{bmatrix}\boldsymbol{K}\begin{bmatrix}\boldsymbol{b}_1 \cdots \boldsymbol{b}_n\end{bmatrix}\right) = \operatorname{tr}\begin{bmatrix}\boldsymbol{a}_1^\top\boldsymbol{K}\boldsymbol{b}_1 & \cdots & \boldsymbol{a}_1^\top\boldsymbol{K}\boldsymbol{b}_n\\ \vdots & \ddots & \vdots\\ \boldsymbol{a}_n^\top\boldsymbol{K}\boldsymbol{b}_1 & \cdots & \boldsymbol{a}_n^\top\boldsymbol{K}\boldsymbol{b}_n\end{bmatrix} = \sum_{i=1}^{n} \boldsymbol{a}_i^\top\boldsymbol{K}\boldsymbol{b}_i. \tag{4.62}$$

With this the preceding properties should be clear. Setting $\boldsymbol{K} = \mathbf{I}_n$ in the definition (A.8) is called the *Frobenius inner product* in [Horn and Johnson, 1985, sec. 5.2] or *Hilbert-Schmidt inner product* in [Hall, 2015, sec. A.6]. Furthermore, for $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times 1}$ it coincides with the common *dot product*.

The induced norm and metric are

$$\|\boldsymbol{A}\|_{\boldsymbol{K}} = \sqrt{\langle \boldsymbol{A}, \boldsymbol{A}\rangle_{\boldsymbol{K}}}, \qquad d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{B}) = \|\boldsymbol{A} - \boldsymbol{B}\|_{\boldsymbol{K}}. \tag{4.63}$$

Translation of particular arguments. For $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times n}$ and $\boldsymbol{R} \in \mathbb{SO}(n)$ we have the properties

$$\langle (\boldsymbol{R}\mathbf{X}_1)^\top, (\boldsymbol{R}\mathbf{X}_2)^\top\rangle_{\boldsymbol{K}} = \langle \mathbf{X}_1^\top, \mathbf{X}_2^\top\rangle_{\boldsymbol{K}}, \tag{4.64a}$$

$$\langle (\mathbf{X}_1\boldsymbol{R})^\top, (\mathbf{X}_2\boldsymbol{R})^\top\rangle_{\boldsymbol{K}} = \langle \mathbf{X}_1^\top, \mathbf{X}_2^\top\rangle_{\boldsymbol{R}\boldsymbol{K}\boldsymbol{R}^\top} \tag{4.64b}$$

For $\boldsymbol{\Xi}_1 = \left[\begin{smallmatrix}\mathbf{X}_1 & \mathbf{x}_1\\ \mathbf{0} & 0\end{smallmatrix}\right]$, $\boldsymbol{\Xi}_2 = \left[\begin{smallmatrix}\mathbf{X}_2 & \mathbf{x}_2\\ \mathbf{0} & 0\end{smallmatrix}\right]$ with $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times n}$ and $\mathbf{G} \in \mathbb{SE}(n)$ we have

$$\langle (\mathbf{G}\boldsymbol{\Xi}_1)^\top, (\mathbf{G}\boldsymbol{\Xi}_2)^\top\rangle_{\boldsymbol{K}} = \langle \boldsymbol{\Xi}_1^\top, \boldsymbol{\Xi}_2^\top\rangle_{\boldsymbol{K}}, \tag{4.65a}$$

$$\langle (\boldsymbol{\Xi}_1\mathbf{G})^\top, (\boldsymbol{\Xi}_2\mathbf{G})^\top\rangle_{\boldsymbol{K}} = \langle \boldsymbol{\Xi}_1^\top, \boldsymbol{\Xi}_2^\top\rangle_{\mathbf{G}\boldsymbol{K}\mathbf{G}^\top} \tag{4.65b}$$

Note that this case includes in particular $\boldsymbol{\Xi}_1, \boldsymbol{\Xi}_2 \in \mathfrak{se}(n)$.

Derivative. This might be useful for the following

$$\mathcal{W} = \tfrac{1}{2}\|(\operatorname{wed}(\boldsymbol{\xi}) + \mathbf{X})^\top\|_{\boldsymbol{K}}^2 = \tfrac{1}{2}\boldsymbol{\xi}^\top\boldsymbol{K}\boldsymbol{\xi} + \boldsymbol{\xi}^\top \operatorname{vee2}(\mathbf{X}\boldsymbol{K}) + \tfrac{1}{2}\operatorname{tr}(\mathbf{X}\boldsymbol{K}\mathbf{X}^\top) \tag{4.66}$$

$$\frac{\partial\mathcal{W}}{\partial\boldsymbol{\xi}} = \operatorname{Vee}(\boldsymbol{K})\boldsymbol{\xi} + \operatorname{vee2}(\mathbf{X}\boldsymbol{K}) = \operatorname{vee2}((\operatorname{wed}(\boldsymbol{\xi}) + \mathbf{X})\boldsymbol{K}) \tag{4.67}$$

$$\frac{\partial^2\mathcal{W}}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}} = \operatorname{Vee}(\boldsymbol{K}) \tag{4.68}$$

4.2.3 Rigid body energies

For the variables $\mathbf{x} \in \mathbb{R}^3$, $\mathbf{X} \in \mathbb{R}^{3\times 3}$ and the parameters $\mathfrak{k}_p \in \mathbb{R}$, $\mathfrak{h}_p \in \mathbb{R}^3$, $p = 1 \dots \mathfrak{N}$, consider the following calculation[3]

$$\begin{aligned}
\sum\nolimits_p \mathfrak{k}_p \|\mathbf{x} + \mathbf{X}\mathfrak{h}_p\|^2 &= \sum\nolimits_p \mathfrak{k}_p \Big(\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathfrak{h}_p \\ 1 \end{bmatrix} \Big)^\top \Big(\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathfrak{h}_p \\ 1 \end{bmatrix} \Big) \\
&= \sum\nolimits_p \mathfrak{k}_p \operatorname{tr} \Big(\Big(\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathfrak{h}_p \\ 1 \end{bmatrix} \Big) \Big(\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathfrak{h}_p \\ 1 \end{bmatrix} \Big)^\top \Big) \\
&= \operatorname{tr} \Big(\underbrace{\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{0} & 0 \end{bmatrix}}_{\mathbf{\Xi}} \underbrace{\sum\nolimits_p \mathfrak{k}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix}}_{\mathbf{K}'} \underbrace{\begin{bmatrix} \mathbf{X}^\top & \mathbf{0} \\ \mathbf{x}^\top & 0 \end{bmatrix}}_{\mathbf{\Xi}^\top} \Big) = \|\mathbf{\Xi}^\top\|^2_{\mathbf{K}'}.
\end{aligned} \tag{4.69}$$

Recalling the definitions of the kinetic energy $\mathcal{T}$ of a free rigid body (4.17), acceleration energy $\mathcal{S}$ in (4.20), dissipation function $\mathcal{R}$ in (4.41) and the potential energy $\mathcal{V}$ due to linear springs (??), we find that they all have the structure of (4.69), i.e.

$$\mathcal{T} = \tfrac{1}{2} \sum\nolimits_p \mathfrak{m}_p \| \underbrace{\dot{\boldsymbol{r}} + \dot{\boldsymbol{R}}\mathfrak{h}_p}_{\dot{\mathfrak{r}}_p} \|^2 = \tfrac{1}{2} \|\dot{\mathbf{G}}^\top\|^2_{\mathbf{M}'} \tag{4.70a}$$

$$\mathcal{S} = \tfrac{1}{2} \sum\nolimits_p \mathfrak{m}_p \| \underbrace{\ddot{\boldsymbol{r}} + \ddot{\boldsymbol{R}}\mathfrak{h}_p}_{\ddot{\mathfrak{r}}_p} \|^2 = \tfrac{1}{2} \|\ddot{\mathbf{G}}^\top\|^2_{\mathbf{M}'} \tag{4.70b}$$

$$\mathcal{R} = \tfrac{1}{2} \sum\nolimits_p \mathfrak{d}_p \| \underbrace{\dot{\boldsymbol{r}} + \dot{\boldsymbol{R}}\mathfrak{h}_p}_{\dot{\mathfrak{r}}_p} \|^2 = \tfrac{1}{2} \|\dot{\mathbf{G}}^\top\|^2_{\mathbf{D}'} \tag{4.70c}$$

$$\mathcal{V} = \tfrac{1}{2} \sum\nolimits_p \mathfrak{k}_p \| \underbrace{\boldsymbol{r} - \boldsymbol{r}_{\mathsf{R}} + (\boldsymbol{R} - \boldsymbol{R}_{\mathsf{R}})\mathfrak{h}_p)}_{\mathfrak{r}_p - \mathfrak{r}_{p\mathsf{R}}} \|^2 = \tfrac{1}{2} \|(\mathbf{G} - \mathbf{G}_{\mathsf{R}})^\top\|^2_{\mathbf{K}'} \tag{4.70d}$$

where

$$\mathbf{M}' = \sum\nolimits_p \mathfrak{m}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Theta}' & m\boldsymbol{s} \\ m\boldsymbol{s}^\top & m \end{bmatrix} \tag{4.71a}$$

$$\mathbf{D}' = \sum\nolimits_p \mathfrak{d}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Upsilon}' & d\boldsymbol{l} \\ d\boldsymbol{l}^\top & d \end{bmatrix} \tag{4.71b}$$

$$\mathbf{K}' = \sum\nolimits_p \mathfrak{k}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Pi}' & k\boldsymbol{h} \\ k\boldsymbol{h}^\top & k \end{bmatrix} \tag{4.71c}$$

The corresponding forces

$$\boldsymbol{f}^{\mathrm{K}} = \boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}} = \operatorname{vee2}\big((\mathbf{I}_4 - \mathbf{G}^{-1}\mathbf{G}_{\mathsf{R}})\mathbf{K}'\big) \tag{4.72a}$$

$$\boldsymbol{f}^{\mathrm{D}} = \frac{\partial \mathcal{R}}{\partial \boldsymbol{\xi}} = \operatorname{vee2}\big(\operatorname{wed}(\boldsymbol{\xi})\mathbf{D}'\big) \tag{4.72b}$$

$$\boldsymbol{f}^{\mathrm{M}} = \frac{\partial \mathcal{S}}{\partial \dot{\boldsymbol{\xi}}} = \operatorname{vee2}\big((\operatorname{wed}(\dot{\boldsymbol{\xi}}) + \operatorname{wed}(\boldsymbol{\xi})^2)\mathbf{M}'\big) \tag{4.72c}$$

note that $\boldsymbol{c} = \operatorname{Vee}(\operatorname{wed}(\boldsymbol{\xi})\mathbf{M}')\boldsymbol{\xi} = -\operatorname{ad}^\top_{\boldsymbol{\xi}} \operatorname{Vee}(\mathbf{M}')\boldsymbol{\xi}$.

[3] using the identity $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$: $\boldsymbol{a}^\top\boldsymbol{b} = \operatorname{tr}(\boldsymbol{a}\boldsymbol{b}^\top)$

Gravitation. Though the potential energy $\mathcal{V}^{\mathrm{G}}$ due to gravitation has a different form than the energies above, it can be written in the form

$$\mathcal{V}^{\mathrm{G}} = \langle \mathsf{G}^\top, \mathrm{wed}(\bar{\boldsymbol{\alpha}}_{\mathsf{G}})^\top \rangle_{\mathsf{M}'}, \qquad \bar{\boldsymbol{\alpha}}_{\mathsf{G}}^\top = [\boldsymbol{a}_{\mathsf{G}}^\top, \mathbf{0}_{1\times 3}]. \tag{4.73}$$

Its differential is

$$\boldsymbol{f}^{\mathsf{G}} = \boldsymbol{\nabla}\mathcal{V}^{\mathrm{G}} = \mathsf{M}\,\mathrm{Ad}_{\mathsf{G}}^{-1}\bar{\boldsymbol{\alpha}}_{\mathsf{G}}. \tag{4.74}$$

Its crucial that this only holds with this specific argument $\bar{\boldsymbol{\alpha}}_{\mathsf{G}}$, i.e. the fact that there is no "angular gravitational acceleration".

Change of body frame. Assume ${}_a\mathsf{M}'$ is the inertia matrix associated with the body frame ${}_a\mathsf{G}$. The inertia matrix ${}_b\mathsf{M}'$ associated with the body frame ${}_b\mathsf{G} = {}_a\mathsf{G}{}^a_b\mathsf{G}$, with ${}^a_b\mathsf{G} = const.$ can be computed using the translation rule from ??:

$$\mathcal{T} = W_{\mathrm{RB}}\big({}_a\mathsf{M}'; {}_a\dot{\mathsf{G}}\big) = W_{\mathrm{RB}}\big({}_a\mathsf{M}'; {}_b\dot{\mathsf{G}}{}^a_b\mathsf{G}^{-1}\big) = W_{\mathrm{RB}}\big(\underbrace{{}^a_b\mathsf{G}^{-1}{}_a\mathsf{M}'({}^a_b\mathsf{G}^{-1})^\top}_{{}_b\mathsf{M}'}; {}_b\dot{\mathsf{G}}\big) \tag{4.75}$$

So

$$\begin{aligned}
\underbrace{\begin{bmatrix} {}_b\boldsymbol{\Theta}' & {}_bm\,{}_b\boldsymbol{s} \\ {}_bm\,{}_b\boldsymbol{s}^\top & {}_bm \end{bmatrix}}_{{}_b\mathsf{M}'} &= \underbrace{\begin{bmatrix} {}^a_b\boldsymbol{R}^\top & -{}^a_b\boldsymbol{R}^\top{}^a_b\boldsymbol{r} \\ 0 & 1 \end{bmatrix}}_{{}^a_b\mathsf{G}^{-1}} \underbrace{\begin{bmatrix} {}_a\boldsymbol{\Theta}' & {}_am\,{}_a\boldsymbol{s} \\ {}_am\,{}_a\boldsymbol{s}^\top & {}_am \end{bmatrix}}_{{}_a\mathsf{M}'} \underbrace{\begin{bmatrix} {}^a_b\boldsymbol{R} & 0 \\ -{}^a_b\boldsymbol{r}^\top{}^a_b\boldsymbol{R} & 1 \end{bmatrix}}_{{}^a_b\mathsf{G}^{-\top}} \\
&= \begin{bmatrix} {}^a_b\boldsymbol{R}^\top({}_a\boldsymbol{\Theta}' + {}_am({}_a\boldsymbol{s} - {}^a_b\boldsymbol{r})({}_a\boldsymbol{s} - {}^a_b\boldsymbol{r})^\top - {}_am\,{}_a\boldsymbol{s}\,{}_a\boldsymbol{s}^\top){}^a_b\boldsymbol{R} & {}^a_b\boldsymbol{R}^\top{}_am({}_a\boldsymbol{s} - {}^a_b\boldsymbol{r}) \\ {}_am({}_a\boldsymbol{s} - {}^a_b\boldsymbol{r})^\top{}^a_b\boldsymbol{R} & {}_am \end{bmatrix}
\end{aligned} \tag{4.76}$$

4.3 Rigid body systems

A rigid body system consists of $N \geq 1$ rigid bodies which may be constrained to each other and/or to the surrounding space. As before, we restrict to geometric constraints. The derivation of equations of motion for such systems is treated in e.g. [Roberson and Schwertassek, 1988], [Murray et al., 1994], [Kane and Levinson, 1985])

The goal of this section is to present an algorithm for the computation of the equations of motion for rigid body systems that allows for a rather flexible parameterization.

4.3.1 Configuration

As motivated for the single rigid body, let there be a body fixed frame for each body of the system as illustrated in Figure 4.3. The components of the position of the b-th body w.r.t. the inertial frame are ${}^0_b\boldsymbol{r} \in \mathbb{R}^3$ and the components of its attitude are ${}^0_b\boldsymbol{R} = [{}^0_b\boldsymbol{R}_{\mathsf{x}}, {}^0_b\boldsymbol{R}_{\mathsf{y}}, {}^0_b\boldsymbol{R}_{\mathsf{z}}] \in \mathbb{SO}(3)$. The configuration can also be expressed w.r.t. any other body: ${}^a_b\boldsymbol{r}$ is the position of the b-th frame w.r.t. the frame of the a-th body and analog of the attitude ${}^a_b\boldsymbol{R}$. The left side indices are used for readability but also to emphasize their

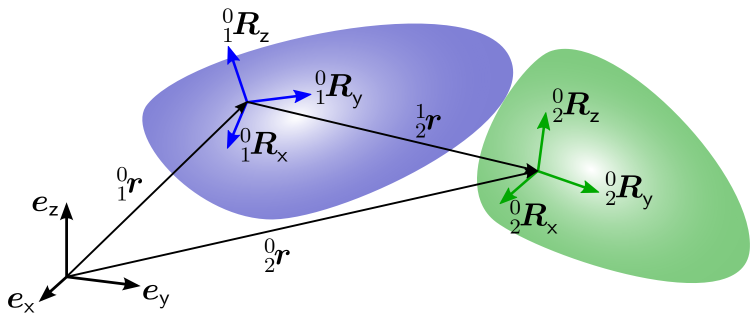


Figure 4.3: inertial frame and body fixed frames

different nature compared to the right side indices. The sum convention does not apply to these indices.

For the positions and attitudes we have the following relations

$$
\begin{aligned}
{}^{a}_{c}\boldsymbol{r} &= {}^{a}_{b}\boldsymbol{r} + {}^{a}_{b}\boldsymbol{R}\,{}^{b}_{c}\boldsymbol{r}, & {}^{a}_{c}\boldsymbol{R} &= {}^{a}_{b}\boldsymbol{R}\,{}^{b}_{c}\boldsymbol{R}, & & \qquad (4.77a)\\
{}^{b}_{a}\boldsymbol{r} &= -{}^{a}_{b}\boldsymbol{R}^{\top}\,{}^{a}_{b}\boldsymbol{r}, & {}^{b}_{a}\boldsymbol{R} &= {}^{a}_{b}\boldsymbol{R}^{\top}, & & \qquad (4.77b)\\
{}^{a}_{a}\boldsymbol{r} &= 0, & {}^{a}_{a}\boldsymbol{R} &= \mathbf{I}_3, & a,b,c &= 0,\dots,N. \qquad (4.77c)
\end{aligned}
$$

As already motivated in the last section, it will be convenient to merge position ${}^{0}_{b}\boldsymbol{r} \in \mathbb{R}^3$ and rotation matrix ${}^{0}_{b}\boldsymbol{R} \in \mathbb{SO}(3)$ into the (rigid body) configuration matrix

$$
{}^{a}_{b}\mathbf{G} = \begin{bmatrix} {}^{a}_{b}\boldsymbol{R} & {}^{a}_{b}\boldsymbol{r} \\ 0 & 1 \end{bmatrix} \in \mathbb{SE}(3). \qquad (4.78)
$$

Then (4.77) is equivalent to

$$
{}^{a}_{c}\mathbf{G} = {}^{a}_{b}\mathbf{G}\,{}^{b}_{c}\mathbf{G}, \qquad {}^{b}_{a}\mathbf{G} = {}^{a}_{b}\mathbf{G}^{-1}, \qquad {}^{a}_{a}\mathbf{G} = \mathbf{I}_4, \qquad a,b,c = 0,\dots,N. \qquad (4.79)
$$

For a system of N body fixed frames and a inertial frame there are $(N+1)^2$ transformations, but due to the rules (4.79), only N can be independent. So a RBS can have at most $6N$ degrees of freedom which is the case if there are no constraints (like joints) between the bodies. Constraints of a joint between body a and b can be captured inside the corresponding transformation ${}^{a}_{b}\mathbf{G}$. We will discuss this in the following example.

Example 12. Tricopter with suspended load: configuration. Consider the Tricopter with a suspended load as shown in Figure 4.4. The top part of the figure shows the body fixed frames which are attached to geometrically meaningful points. The numbering of the bodies is rather arbitrary.

The Tricopter flies freely in space, i.e. there are no constraints between the inertial frame and any body of the system. So we chose to describe the configuration of the central body w.r.t. the inertial frame as

$$
{}^{0}_{1}\mathbf{G} = \begin{bmatrix} {}^{0}_{1}R^{\mathsf{x}}_{\mathsf{x}} & {}^{0}_{1}R^{\mathsf{x}}_{\mathsf{y}} & {}^{0}_{1}R^{\mathsf{x}}_{\mathsf{z}} & {}^{0}_{1}r^{\mathsf{x}} \\ {}^{0}_{1}R^{\mathsf{y}}_{\mathsf{x}} & {}^{0}_{1}R^{\mathsf{y}}_{\mathsf{y}} & {}^{0}_{1}R^{\mathsf{y}}_{\mathsf{z}} & {}^{0}_{1}r^{\mathsf{y}} \\ {}^{0}_{1}R^{\mathsf{z}}_{\mathsf{x}} & {}^{0}_{1}R^{\mathsf{z}}_{\mathsf{y}} & {}^{0}_{1}R^{\mathsf{z}}_{\mathsf{z}} & {}^{0}_{1}r^{\mathsf{z}} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \qquad (4.80a)
$$

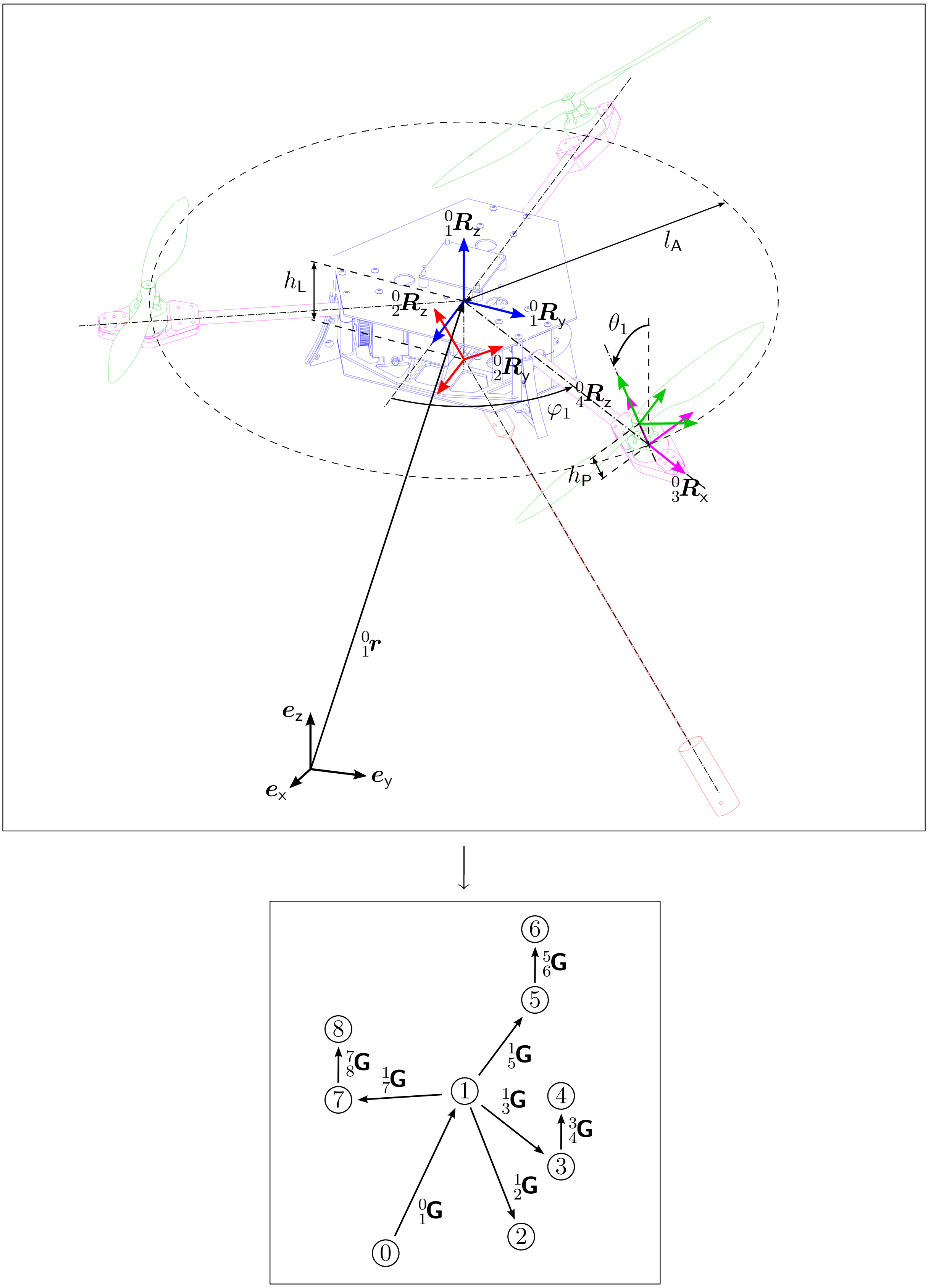


Figure 4.4: Frames attached to the Tricopter bodies (top) and the configuration graph (bottom)

The suspended load is a rigid body that is attached by a spherical joint to the central body. The body fixed frame of the load was placed in the center of this spherical joint. As a consequence the position of the load (the position of its body fixed frame not the position of its center of mass) w.r.t. the central body is constant. This is reflected by the configuration

$$ {}^{1}_{2}\mathsf{G} = \begin{bmatrix} {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{x}} & {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{y}} & {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{z}} & 0 \\ {}^{1}_{2}R^{\mathsf{y}}_{\mathsf{x}} & {}^{1}_{2}R^{\mathsf{y}}_{\mathsf{y}} & {}^{1}_{2}R^{\mathsf{y}}_{\mathsf{z}} & 0 \\ {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{x}} & {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{z}} & {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{z}} & h_{\mathsf{L}} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.80b} $$

The three arms are connected to the central body each by revolute joints with tilt angles θ_k, $k = 1, 2, 3$. The joint axis lie in the plane spanned by ${}^{0}_{1}\boldsymbol{R}_{\mathsf{x}}$ and ${}^{0}_{1}\boldsymbol{R}_{\mathsf{y}}$ and their angles to ${}^{0}_{1}\boldsymbol{R}_{\mathsf{x}}$ are $\varphi_1 = \frac{\pi}{3}, \varphi_2 = \pi, \varphi_3 = -\frac{\pi}{3}$. The body fixed axes are placed such that ${}_{2k+1}^{\;\;0}\boldsymbol{R}_{\mathsf{x}}$ coincide with the tilt axis and ${}_{2k+1}^{\;\;0}\boldsymbol{R}_{\mathsf{z}}$, $k = 1, 2, 3$ coincide with the propeller spinning axis. The configuration of the k-th arm w.r.t. the central body is

$$ {}_{2k+1}^{\;\;1}\mathsf{G} = \begin{bmatrix} \cos\varphi_k & -\sin\varphi_k\cos\theta_k & \sin\varphi_k\sin\theta_k & l_{\mathsf{A}}\cos\varphi_k \\ \sin\varphi_k & \cos\varphi_k\cos\theta_k & -\cos\varphi_k\sin\theta_k & l_{\mathsf{A}}\sin\varphi_k \\ 0 & \sin\theta_k & \cos\theta_k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad k = 1, \dots, 3. \tag{4.80c} $$

The propellers are connected by revolute joints to the arms. The body fixed frame is attached to the geometric center of the propeller (which will be an important point for the aerodynamics). The configuration w.r.t. the corresponding arm is

$$ {}_{2k+2}^{2k+1}\mathsf{G} = \begin{bmatrix} \mathrm{c}_k & -\mathrm{s}_k & 0 & 0 \\ \mathrm{s}_k & \mathrm{c}_k & 0 & 0 \\ 0 & 0 & 1 & h_{\mathsf{P}} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad k = 1, \dots, 3. \tag{4.80d} $$

The set of configurations $\mathcal{G}_0 = \{{}^{0}_{1}\mathsf{G}, {}^{1}_{2}\mathsf{G}, {}^{1}_{3}\mathsf{G}, {}^{3}_{4}\mathsf{G}, {}^{1}_{5}\mathsf{G}, {}^{5}_{6}\mathsf{G}, {}^{1}_{7}\mathsf{G}, {}^{7}_{8}\mathsf{G}\}$ form a directed graph as shown at the bottom of Figure 4.4. With them and the rules from (4.79) we can compute the configuration ${}^{a}_{b}\mathsf{G}$ of any body w.r.t. any other body or the inertial frame.

The configurations can be seen as functions ${}^{a}_{b}\mathsf{G}(\boldsymbol{x})$ of the system coordinates

$$ \boldsymbol{x} = \left[{}^{0}_{1}r^{\mathsf{x}}, {}^{0}_{1}r^{\mathsf{y}}, {}^{0}_{1}r^{\mathsf{z}}, {}^{0}_{1}R^{\mathsf{x}}_{\mathsf{x}}, \dots, {}^{0}_{1}R^{\mathsf{z}}_{\mathsf{z}}, {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{x}}, \dots, {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{z}}, \theta_1, \theta_2, \theta_3, \mathrm{c}_1, \mathrm{s}_1, \mathrm{c}_2, \mathrm{s}_2, \mathrm{c}_3, \mathrm{s}_3\right]^{\top} \in \mathbb{R}^{30} \tag{4.81} $$

and the constant parameters $h_{\mathsf{L}}, l_{\mathsf{A}}, \varphi_1, \varphi_2, \varphi_3, h_{\mathsf{P}}$. From the rules (4.79) emerge the geometric constraints

$$ \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0} \qquad \cong \qquad \begin{cases} {}^{0}_{1}\boldsymbol{R}^{\top}\,{}^{0}_{1}\boldsymbol{R} = \mathbf{I}_3, \ \det {}^{0}_{1}\boldsymbol{R} = +1, \\ {}^{1}_{2}\boldsymbol{R}^{\top}\,{}^{1}_{2}\boldsymbol{R} = \mathbf{I}_3, \ \det {}^{1}_{2}\boldsymbol{R} = +1, \\ (\mathrm{c}_k)^2 + (\mathrm{s}_k)^2 = 1, \ k = 1, 2, 3 \end{cases} \tag{4.82} $$

The configuration space of the rigid body system is

$$ \mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^{30} \,|\, \boldsymbol{\phi}(\boldsymbol{x}) = 0\} \ \cong \ \mathbb{SE}(3) \times \mathbb{SO}(3) \times \mathbb{R}^3 \times (\mathbb{S}^1)^3. \tag{4.83} $$

This example was mainly chosen as the Tricopter will be discussed in the following chapters. However, it is also an example of a system that is complex enough that one probably does not want to derive the equations of motion by hand without a formalism. It also covers the most common manifolds encountered in rigid body mechanics. Even though the revolute joints for the propeller tilt and the propeller spinning axes both imply a $\mathbb{S}^1$ manifold, the local parameterization by the angle $\theta_k, k = 1, 2, 3$ is chosen. This has a practical motivation: The tilt mechanism also twists the cables to the propeller motor and so $\theta_k = 0$ and $\theta_k = 2\pi$ are really different situations in practice. On the other hand it should also show that the following algorithm handles minimal coordinates just as fine.

A generalization of the example states: A rigid body system can be parameterized by a set of ν (possibly redundant) coordinates $\boldsymbol{x}$ which again parameterize a set of configurations ${}^a_b\mathbf{G}(\boldsymbol{x})$ which form a *connected* graph. The property connected is essential: it ensures that, with the rules (4.79), all remaining configurations of the graph can be computed i.e. the corresponding *complete* graph. Loops in the graph and the property ${}^a_b\mathbf{G} \in \mathbb{SE}(3)$ may imply geometric constraints.

The use of graph theory in the context of algorithms for rigid body systems is quite common, see e.g. [Roberson and Schwertassek, 1988, sec. 8.2] or [Wittenburg, 2008, sec. 5.3]. However, we will not go any deeper into this. All we need for the following is that any configuration ${}^a_b\mathbf{G}(\boldsymbol{x}), a, b = 0 \dots N$ can be expressed in terms of the configuration coordinates $\boldsymbol{x}$.

4.3.2 Velocity

In (4.11) we motivated particular velocity coordinates $\boldsymbol{\xi} = [\boldsymbol{v}^\top, \boldsymbol{\omega}^\top]^\top$ for the free rigid body, which did lead to a convenient mathematical expressions. In the context of rigid body systems we may associate a *body velocity* ${}^a_b\boldsymbol{\xi} = [{}^a_b\boldsymbol{v}^\top, {}^a_b\boldsymbol{\omega}^\top]^\top$ with any configuration ${}^a_b\mathbf{G}, a, b = 0, \dots, N$ defined by

$$ {}^a_b\boldsymbol{\xi} = \operatorname{vee}\left({}^b_a\mathbf{G}\,{}^a_b\dot{\mathbf{G}}\right), \quad a, b = 0, \dots, N. \tag{4.84}$$

From the rules (4.79) for the configurations we can conclude similar rules for their velocities: For the composition ${}^a_c\mathbf{G} = {}^a_b\mathbf{G}\,{}^b_c\mathbf{G}$ we get

$$ {}^a_c\boldsymbol{\xi} = \operatorname{vee}\left({}^c_b\mathbf{G}\,{}^b_a\mathbf{G}({}^a_b\dot{\mathbf{G}}\,{}^b_c\mathbf{G} + {}^a_b\mathbf{G}\,{}^b_c\dot{\mathbf{G}})\right) = \operatorname{vee}\left({}^c_b\mathbf{G}\operatorname{wed}({}^a_b\boldsymbol{\xi})\,{}^b_c\mathbf{G}\right) + {}^b_c\boldsymbol{\xi} = \operatorname{Ad}_{{}^c_b\mathbf{G}}\,{}^a_b\boldsymbol{\xi} + {}^b_c\boldsymbol{\xi} \tag{4.85a}$$

Differentiation of the rule for the inverse yields

$$ \tfrac{\mathrm{d}}{\mathrm{d}t}\left({}^a_b\mathbf{G}\,{}^b_a\mathbf{G}\right) = {}^a_b\dot{\mathbf{G}}\,{}^b_a\mathbf{G} + {}^a_b\mathbf{G}\,{}^b_a\dot{\mathbf{G}} = {}^a_b\mathbf{G}\operatorname{wed}({}^a_b\boldsymbol{\xi})\,{}^b_a\mathbf{G} + \operatorname{wed}({}^b_a\boldsymbol{\xi}) = \mathbf{0} \quad \Leftrightarrow \quad {}^b_a\boldsymbol{\xi} = -\operatorname{Ad}_{{}^a_b\mathbf{G}}\,{}^a_b\boldsymbol{\xi} \tag{4.85b}$$

and obviously

$$ {}^a_a\boldsymbol{\xi} = \mathbf{0}. \tag{4.85c}$$

System velocity and body Jacobians. Based on their definition (4.84), the body velocities ${}^{a}_{b}\boldsymbol{\xi}$ can be seen as a function of the system coordinates $\boldsymbol{x}$ and their derivatives $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}$. Crucially the velocity is linear in $\dot{\boldsymbol{x}}$ and consequently linear in the system velocity $\boldsymbol{\xi}$ and we can write

$$ {}^{a}_{b}\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{\xi}) = {}^{a}_{b}\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}, \qquad {}^{a}_{b}\boldsymbol{J}(\boldsymbol{x}) = \frac{\partial\, {}^{a}_{b}\boldsymbol{\xi}}{\partial \boldsymbol{\xi}}(\boldsymbol{x}) = \frac{\partial}{\partial \dot{\boldsymbol{x}}} \operatorname{vee}\Big({}^{b}_{a}\mathbf{G}(\boldsymbol{x})\frac{\mathrm{d}}{\mathrm{d}t}\big({}^{a}_{b}\mathbf{G}(\boldsymbol{x})\big)\Big)\boldsymbol{A}. \tag{4.86} $$

The matrix ${}^{a}_{b}\boldsymbol{J}(\boldsymbol{x}) \in \mathbb{R}^{6\times n}$ that maps the system velocity $\boldsymbol{\xi}$ to the body velocity ${}^{a}_{b}\boldsymbol{\xi}$ is commonly called the *body Jacobian*. An alternative formula for the body Jacobian, which might give additional geometric insight, is given in (A.186). The following rules emerge directly from (4.85):

$$ {}^{a}_{c}\boldsymbol{J} = \mathrm{Ad}_{{}^{c}_{b}\mathbf{G}}\, {}^{a}_{b}\boldsymbol{J} + {}^{b}_{c}\boldsymbol{J}, \qquad {}^{b}_{a}\boldsymbol{J} = -\mathrm{Ad}_{{}^{a}_{b}\mathbf{G}}\, {}^{a}_{b}\boldsymbol{J}, \qquad {}^{a}_{a}\boldsymbol{J} = \mathbf{0}. \tag{4.87} $$

Example 13. Tricopter with suspended load: kinematics. For the tricopter with load from Example 12 we chose the following velocity coordinates: The components of the body velocity ${}^{0}_{1}\boldsymbol{\xi}$ of the central body w.r.t. the inertial frame, the components of the angular velocity ${}^{0}_{2}\boldsymbol{\omega}$ of the load w.r.t. the inertial frame, the angular velocities $\dot{\theta}_k, k = 1, 2, 3$ of the arm tilt mechanism and the angular velocities $\varpi_k, k = 1, 2, 3$ of the propellers w.r.t. the arms. These velocity coordinates $\boldsymbol{\xi} = [{}^{0}_{1}\boldsymbol{\xi}^\top, {}^{0}_{2}\boldsymbol{\omega}^\top, \dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3, \varpi_1, \varpi_2, \varpi_3]^\top$ are related to the configuration coordinates $\boldsymbol{\xi}$ by the kinematic equation

$$ \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi} \qquad \cong \qquad \begin{cases} {}^{0}_{1}\dot{\mathbf{G}} = {}^{0}_{1}\mathbf{G}\,\operatorname{wed}({}^{0}_{1}\boldsymbol{\xi}), \\ {}^{1}_{2}\dot{\boldsymbol{R}} = {}^{1}_{2}\boldsymbol{R}\,\operatorname{wed}({}^{0}_{2}\boldsymbol{\omega}) - \operatorname{wed}({}^{0}_{1}\boldsymbol{\omega})\,{}^{1}_{2}\boldsymbol{R}, \\ \dot{\theta}_k = \dot{\theta}_k,\;\; k = 1,2,3 \\ \dot{\mathrm{c}}_k = -\mathrm{s}_k\varpi_k,\;\; k = 1,2,3 \\ \dot{\mathrm{s}}_k = \mathrm{c}_k\varpi_k,\;\; k = 1,2,3 \end{cases} . \tag{4.88} $$

The relative velocity ${}^{1}_{2}\boldsymbol{\omega} = \operatorname{vee}({}^{1}_{2}\boldsymbol{R}^\top {}^{1}_{2}\dot{\boldsymbol{R}})$ of the load would be another possible and probably more obvious choice. The absolute velocity ${}^{0}_{2}\boldsymbol{\omega}$ is mainly chosen to demonstrate the flexibility of the presented approach but the use of absolute velocities also leads to less cumbersome terms in the system inertia matrix.

The body velocities associated with the configuration matrices from (4.80) are

$$ {}^{0}_{1}\boldsymbol{\xi} = \begin{bmatrix} {}^{0}_{1}v^{\mathsf{x}} \\ {}^{0}_{1}v^{\mathsf{y}} \\ {}^{0}_{1}v^{\mathsf{z}} \\ {}^{0}_{1}\omega^{\mathsf{x}} \\ {}^{0}_{1}\omega^{\mathsf{y}} \\ {}^{0}_{1}\omega^{\mathsf{z}} \end{bmatrix}, \qquad {}^{1}_{2}\boldsymbol{\xi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ {}^{0}_{2}\omega^{\mathsf{x}} - {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{x}}\,{}^{0}_{2}\omega^{\mathsf{x}} - {}^{1}_{2}R^{\mathsf{y}}_{\mathsf{x}}\,{}^{0}_{2}\omega^{\mathsf{y}} - {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{x}}\,{}^{0}_{2}\omega^{\mathsf{z}} \\ {}^{0}_{2}\omega^{\mathsf{y}} - {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{y}}\,{}^{0}_{2}\omega^{\mathsf{x}} - {}^{1}_{2}R^{\mathsf{y}}_{\mathsf{y}}\,{}^{0}_{2}\omega^{\mathsf{y}} - {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{y}}\,{}^{0}_{2}\omega^{\mathsf{z}} \\ {}^{0}_{2}\omega^{\mathsf{z}} - {}^{1}_{2}R^{\mathsf{x}}_{\mathsf{z}}\,{}^{0}_{2}\omega^{\mathsf{x}} - {}^{1}_{2}R^{\mathsf{y}}_{\mathsf{z}}\,{}^{0}_{2}\omega^{\mathsf{y}} - {}^{1}_{2}R^{\mathsf{z}}_{\mathsf{z}}\,{}^{0}_{2}\omega^{\mathsf{z}} \end{bmatrix}, \tag{4.89a} $$

$$ {}^{1}_{2k+1}\boldsymbol{\xi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dot{\theta}_k \\ 0 \\ 0 \end{bmatrix}, \qquad {}^{2k+1}_{2k+2}\boldsymbol{\xi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \varpi_k \end{bmatrix}, \quad k = 1, 2, 3. \tag{4.89b} $$

From this it should clear how the corresponding body Jacobians look like, e.g. ${}^{0}_{1}\boldsymbol{J} = [\mathbf{I}_6\ 0]$.

For the formulation of the kinetic energy in the following subsection we will need the body Jacobians ${}^{0}_{b}\boldsymbol{J}, b = 1, \dots, N$. From the graph structure and the rules (4.87) we can compute them iteratively as

$$
\begin{aligned}
{}^{0}_{2}\boldsymbol{J} &= \mathrm{Ad}_{{}^{2}_{1}\mathbf{G}}\, {}^{0}_{1}\boldsymbol{J} + {}^{1}_{2}\boldsymbol{J}, && (4.90a)\\
{}^{0}_{2k+1}\boldsymbol{J} &= \mathrm{Ad}_{{}^{2k+1}_{1}\mathbf{G}}\, {}^{0}_{1}\boldsymbol{J} + {}^{1}_{2k+1}\boldsymbol{J}, && (4.90b)\\
{}^{0}_{2k+2}\boldsymbol{J} &= \mathrm{Ad}_{{}^{2k+2}_{2k+1}\mathbf{G}}\, {}^{0}_{2k+1}\boldsymbol{J} + {}^{2k+1}_{2k+2}\boldsymbol{J}. && (4.90c)
\end{aligned}
$$

These terms are significantly more cumbersome, so they are not displayed explicitly.

4.3.3 Inertia

Kinetic energy and inertia matrix. The kinetic energy $\mathcal{T}$ of a rigid body system is simply the sum of the kinetic energies of its bodies. Recalling the kinetic energy (4.17) for a single free rigid body, and formulation of the absolute body velocities ${}^{0}_{b}\boldsymbol{\xi}$ in terms of the chosen configuration coordinates $\boldsymbol{x}$ and velocity coordinates $\boldsymbol{\xi}$ according to (4.86) yields

$$
\mathcal{T}(\boldsymbol{x}, \boldsymbol{\xi}) = \sum\nolimits_b \tfrac{1}{2}\big({}^{0}_{b}\boldsymbol{\xi}(\boldsymbol{x}, \boldsymbol{\xi})\big)^{\top} {}^{0}_{b}\mathbf{M}\, {}^{0}_{b}\boldsymbol{\xi}(\boldsymbol{x}, \boldsymbol{\xi}) = \tfrac{1}{2}\boldsymbol{\xi}^{\top} \underbrace{\sum\nolimits_b \big({}^{0}_{b}\boldsymbol{J}(\boldsymbol{x})\big)^{\top} {}^{0}_{b}\mathbf{M}\, {}^{0}_{b}\boldsymbol{J}(\boldsymbol{x})}_{\boldsymbol{M}(\boldsymbol{x})} \boldsymbol{\xi} \tag{4.91}
$$

with the system inertia matrix $\boldsymbol{M}(\boldsymbol{x}) \in \mathbb{R}^{n\times n}$ computed from the body Jacobians ${}^{0}_{a}\boldsymbol{J}(\boldsymbol{x})$ and the body inertia matrices ${}^{0}_{b}\mathbf{M}$ defined in (4.17).

Connection coefficients. With a rather cumbersome computation (see (A.190)), it can be shown that the connection coefficients $\mathit{\Gamma}_{ijk}$ associated to the system inertia matrix $\boldsymbol{M}$ from (4.91) can be expressed in terms of the body Jacobians ${}^{0}_{b}\boldsymbol{J}$, the body inertia matrices ${}^{0}_{b}\mathbf{M}$ and the body connection coefficients ${}^{0}_{b}\Gamma_{pqr}$ from (4.19) or the body commutation coefficients Υ^{h}_{pq} from (4.15) as

$$
\mathit{\Gamma}_{ijk} = \sum\nolimits_b {}^{0}_{b}J^{p}_{i}\big({}^{0}_{b}\mathsf{M}_{pq}\partial_k {}^{0}_{b}J^{q}_{j} + \underbrace{\tfrac{1}{2}\big(\Upsilon^{h}_{pq}\,{}^{0}_{b}\mathsf{M}_{hr} + \Upsilon^{h}_{pr}\,{}^{0}_{b}\mathsf{M}_{hq} - \Upsilon^{h}_{qr}\,{}^{0}_{b}\mathsf{M}_{hp}\big)}_{{}^{0}_{b}\Gamma_{pqr}}\, {}^{0}_{b}J^{q}_{j}\,{}^{0}_{b}J^{r}_{k}\big) \tag{4.92}
$$

With this we can state the gyroscopic terms as

$$
c_i = \mathit{\Gamma}_{ijk}\xi^j\xi^k = \sum\nolimits_b {}^{0}_{b}J^{p}_{i}\big({}^{0}_{b}\mathsf{M}_{pq}\partial_k {}^{0}_{b}J^{q}_{j} + {}^{0}_{b}J^{q}_{j}\Upsilon^{h}_{pq}\,{}^{0}_{b}\mathsf{M}_{hr}\,{}^{0}_{b}J^{r}_{k}\big)\xi^j\xi^k \tag{4.93}
$$

Acceleration energy. Completely analog as above, but based on the acceleration energy (4.20) of a single free rigid body, we express the acceleration energy $\mathcal{S}$ for a rigid body system as

$$\begin{aligned}\mathcal{S}(\boldsymbol{x},\boldsymbol{\xi},\dot{\boldsymbol{\xi}}) = \tfrac{1}{2}\dot{\boldsymbol{\xi}}^\top &\underbrace{\textstyle\sum_b \big({}^0_b\boldsymbol{J}(\boldsymbol{x})\big)^\top {}^0_b\mathbf{M}\, {}^0_b\boldsymbol{J}(\boldsymbol{x})}_{\boldsymbol{M}(\boldsymbol{x})}\dot{\boldsymbol{\xi}} \\ &+ \dot{\boldsymbol{\xi}}^\top \underbrace{\textstyle\sum_b \big({}^0_b\boldsymbol{J}(\boldsymbol{x})\big)^\top \big({}^0_b\mathbf{M}\, {}^0_b\dot{\boldsymbol{J}}(\boldsymbol{x},\boldsymbol{\xi}) - \mathrm{ad}^\top_{{}^0_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}}\, {}^0_b\mathbf{M}\, {}^0_b\boldsymbol{J}(\boldsymbol{x})\big)\boldsymbol{\xi}}_{\boldsymbol{c}(\boldsymbol{x},\boldsymbol{\xi})} \\ &+ \underbrace{\textstyle\sum_b \tfrac{1}{2}\big\|\big(\,\mathrm{wed}\,\big({}^0_b\dot{\boldsymbol{J}}(\boldsymbol{x},\boldsymbol{\xi})\boldsymbol{\xi}\big) + \mathrm{wed}\,\big({}^0_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}\big)^2\big)^\top\big\|^2_{\mathrm{Wed}({}^0_b\mathbf{M})}}_{\mathcal{S}_0(\boldsymbol{x},\boldsymbol{\xi})}. \end{aligned} \tag{4.94}$$

For the explicit form of $\mathcal{S}_0$, the formulation (4.70b) is useful.

Inertia force. The generalized inertia force $\boldsymbol{f}^{\mathrm{M}}$ for a rigid body system can be computed from

- $f_i^{\mathrm{M}} = M_{ij}\dot{\xi}^j + \mathit{\Gamma}_{ijk}\xi^k\xi^j$ using the inertia matrix M_{ij} from (4.91) and the connection coefficients $\mathit{\Gamma}_{ijk}$ from (4.92).
- $f_i^{\mathrm{M}} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{T}}{\partial \xi^i} + \gamma^k_{ij}\xi^j\frac{\partial \mathcal{T}}{\partial \xi^k} - \partial_i\mathcal{T}$ with the kinetic energy $\mathcal{T}$ from (4.91).
- $f_i^{\mathrm{M}} = \frac{\partial \mathcal{S}}{\partial \dot{\xi}^i}$ using the acceleration energy $\mathcal{S}$ from (4.94).

which yield

$$\boldsymbol{f}^{\mathrm{M}}(\boldsymbol{x},\boldsymbol{\xi},\dot{\boldsymbol{\xi}}) = \boldsymbol{M}(\boldsymbol{x})\dot{\boldsymbol{\xi}} + \boldsymbol{c}(\boldsymbol{x},\boldsymbol{\xi}) \tag{4.95}$$

with the system inertia matrix $\boldsymbol{M}$ from (4.91) and the gyroscopic terms $\boldsymbol{c}$ from (4.94) or (4.93).

Similar results are called *the projection equation* in [Bremer, 2008, sec. 4.2.5] and *the Kane equations* [Kane and Levinson, 1985, chap. 6]. There is some controversy (starting in [Desloge, 1987]) about the naming, since the equations result rather directly (as shown above) from the Gibbs-Appell formulation. See [Lesser, 1992] or [Papastavridis, 2002, p. 714] for an overview.

In contrast to the derivations in the sources above, the formulation (4.95) poses no restrictions on the body fixed frames and allows redundant configuration coordinates.

4.3.4 Gravitation

The potential energy of gravitation of a rigid body system is the sum of the potentials of the individual bodies (4.23). Using the formulation from (4.73) this is

$$\begin{aligned}\mathcal{V}^{\mathrm{G}}(\boldsymbol{x}) &= \textstyle\sum_b {}^0_b m\,\langle {}^0_b\boldsymbol{r}(\boldsymbol{x}) + {}^0_b\boldsymbol{R}(\boldsymbol{x})\,{}^0_b\boldsymbol{s}, \boldsymbol{a}_{\mathsf{G}}\rangle \\ &= \textstyle\sum_b \langle ({}^0_b\mathbf{G}(\boldsymbol{x}))^\top, \mathrm{wed}(\bar{\boldsymbol{\alpha}}_{\mathsf{G}})^\top\rangle_{{}^0_b\mathbf{M}'}, \qquad \bar{\boldsymbol{\alpha}}^\top_{\mathsf{G}} = [\boldsymbol{a}^\top_{\mathsf{G}}, \mathbf{0}_{1\times 3}]. \end{aligned} \tag{4.96}$$

The resulting generalized force on the system may be written as

$$\boldsymbol{f}^{\mathrm{G}}(\boldsymbol{x}) = \boldsymbol{\nabla}\mathcal{V}^{\mathrm{G}}(\boldsymbol{x}) = \textstyle\sum_b ({}^0_b\boldsymbol{J}(\boldsymbol{x}))^\top\, {}^0_b\mathbf{M}\,\mathrm{Ad}^{-1}_{{}^0_b\mathbf{G}(\boldsymbol{x})}\bar{\boldsymbol{\alpha}}_{\mathsf{G}}. \tag{4.97}$$

4.3.5 Stiffness

Potential energy

$$\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) = \sum_{a,b} \tfrac{1}{2}\|({}^{a}_{b}\mathsf{G}(\boldsymbol{x}) - {}^{a}_{b}\mathsf{G}_{\mathsf{R}})^{\top}\|^{2}_{{}^{a}_{b}\mathsf{K}'} \tag{4.98}$$

stiffness force

$$\boldsymbol{f}^{\mathrm{K}}(\boldsymbol{x}) = \boldsymbol{\nabla}\mathcal{V}^{\mathrm{K}}(\boldsymbol{x}) = \sum_{a,b}({}^{a}_{b}\boldsymbol{J}(\boldsymbol{x}))^{\top}\underbrace{\mathrm{vee2}\left((\mathbf{I}_4 - ({}^{a}_{b}\mathsf{G}(\boldsymbol{x}))^{-1}\,{}^{a}_{b}\mathsf{G}_{\mathsf{R}})\,{}^{a}_{b}\mathsf{K}'\right)}_{{}^{a}_{b}\boldsymbol{f}^{\mathsf{K}}(\boldsymbol{x})} \tag{4.99}$$

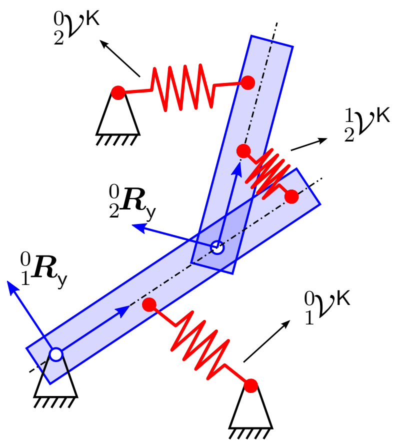


Figure 4.5: MultibodyPotentialIllustration

4.3.6 Dissipation

Dissipation function

$$\mathcal{R}(\boldsymbol{x},\boldsymbol{\xi}) = \sum_{a,b}\tfrac{1}{2}\|({}^{a}_{b}\dot{\mathsf{G}}(\boldsymbol{x},\boldsymbol{\xi}))^{\top}\|^{2}_{{}^{a}_{b}\mathsf{D}'} = \tfrac{1}{2}\boldsymbol{\xi}^{\top}\underbrace{\sum_{a,b}({}^{a}_{b}\boldsymbol{J}(\boldsymbol{x}))^{\top}\,{}^{a}_{b}\mathsf{D}\,{}^{a}_{b}\boldsymbol{J}(\boldsymbol{x})}_{\boldsymbol{D}(\boldsymbol{x})}\boldsymbol{\xi} \tag{4.100}$$

dissipative force

$$\boldsymbol{f}^{\mathrm{D}}(\boldsymbol{x},\boldsymbol{\xi}) = \frac{\partial\mathcal{R}}{\partial\boldsymbol{\xi}}(\boldsymbol{x},\boldsymbol{\xi}) = \boldsymbol{D}(\boldsymbol{x})\boldsymbol{\xi} \tag{4.101}$$

4.4 Summary

There is an algorithm:

$$\left(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{A}, {}^{a}_{b}\mathbf{G}, {}^{0}_{b}\mathbf{M}', {}^{a}_{b}\mathbf{D}', {}^{a}_{b}\mathbf{K}', {}^{a}_{b}\mathbf{G}_{\mathsf{R}}, \boldsymbol{a}_{\mathsf{G}}\right) \;\mapsto\; \left(\boldsymbol{M}, \boldsymbol{c}, \boldsymbol{f}^{\mathrm{D}}, \boldsymbol{f}^{\mathrm{K}}, \boldsymbol{f}^{\mathrm{G}}\right) \tag{4.102}$$

which is

$$ {}^{a}_{b}\boldsymbol{J}(\boldsymbol{x}) = \frac{\partial}{\partial \dot{\boldsymbol{x}}} \operatorname{vee}\left({}^{b}_{a}\mathbf{G}(\boldsymbol{x}) \frac{\mathrm{d}}{\mathrm{d}t}\left({}^{a}_{b}\mathbf{G}(\boldsymbol{x})\right)\right)\boldsymbol{A}(\boldsymbol{x}) \tag{4.103a}$$

fill the graph

$$ {}^{a}_{c}\mathbf{G} = {}^{a}_{b}\mathbf{G}\,{}^{b}_{c}\mathbf{G}, \qquad {}^{b}_{a}\mathbf{G} = {}^{a}_{b}\mathbf{G}^{-1}, \tag{4.103b}$$

$$ {}^{a}_{c}\boldsymbol{J} = \mathrm{Ad}_{{}^{c}_{b}\mathbf{G}}\,{}^{a}_{b}\boldsymbol{J} + {}^{b}_{c}\boldsymbol{J}, \qquad {}^{b}_{a}\boldsymbol{J} = -\mathrm{Ad}_{{}^{a}_{b}\mathbf{G}}\,{}^{a}_{b}\boldsymbol{J}, \tag{4.103c}$$

$$ {}^{a}_{c}\dot{\boldsymbol{J}} = \mathrm{Ad}_{{}^{c}_{b}\mathbf{G}}\left({}^{a}_{b}\dot{\boldsymbol{J}} + \mathrm{ad}_{{}^{c}_{b}\boldsymbol{J}\boldsymbol{\xi}}\,{}^{a}_{b}\boldsymbol{J}\right) + {}^{b}_{c}\dot{\boldsymbol{J}}, \qquad {}^{b}_{a}\dot{\boldsymbol{J}} = -\mathrm{Ad}_{{}^{a}_{b}\mathbf{G}}\left({}^{a}_{b}\dot{\boldsymbol{J}} + \mathrm{ad}_{{}^{a}_{b}\boldsymbol{J}\boldsymbol{\xi}}\,{}^{a}_{b}\boldsymbol{J}\right), \tag{4.103d}$$

system assembly

$$\boldsymbol{M} = \sum\nolimits_{b} {}^{0}_{b}\boldsymbol{J}^{\top}\, \mathrm{Vee}({}^{0}_{b}\mathbf{M}')\, {}^{0}_{b}\boldsymbol{J}, \tag{4.103e}$$

$$\boldsymbol{c} = \sum\nolimits_{b} {}^{0}_{b}\boldsymbol{J}^{\top}\, \mathrm{vee2}\left(\left(\mathrm{wed}({}^{0}_{b}\dot{\boldsymbol{J}}\boldsymbol{\xi}) + \mathrm{wed}({}^{0}_{b}\boldsymbol{J}\boldsymbol{\xi})^{2}\right) {}^{0}_{b}\mathbf{M}'\right) \tag{4.103f}$$

$$\boldsymbol{D} = \sum\nolimits_{a,b} {}^{a}_{b}\boldsymbol{J}^{\top}\, \mathrm{Vee}({}^{0}_{b}\mathbf{D}')\, {}^{a}_{b}\boldsymbol{J} \tag{4.103g}$$

$$\boldsymbol{f}^{\mathrm{K}} = \sum\nolimits_{a,b} {}^{a}_{b}\boldsymbol{J}^{\top}\, \mathrm{vee2}\left((\mathbf{I}_4 - {}^{b}_{a}\mathbf{G}\,{}^{a}_{b}\mathbf{G}_{\mathsf{R}})\,{}^{a}_{b}\mathbf{K}'\right) \tag{4.103h}$$

$$\boldsymbol{f}^{\mathrm{G}} = \sum\nolimits_{b} {}^{0}_{b}\boldsymbol{J}^{\top}\, \mathrm{Vee}({}^{0}_{b}\mathbf{M}')\,\mathrm{Ad}_{{}^{b}_{0}\mathbf{G}}\,\bar{\boldsymbol{\alpha}}_{\mathsf{G}}, \qquad \bar{\boldsymbol{\alpha}}_{\mathsf{G}} = [\boldsymbol{a}_{\mathsf{G}}^{\top}, \mathbf{0}_{1\times 3}]^{\top} \tag{4.103i}$$

$$\mathcal{T} = \tfrac{1}{2}\|\boldsymbol{\xi}\|^{2}_{\boldsymbol{M}} \tag{4.103j}$$

$$\mathcal{R} = \tfrac{1}{2}\|\boldsymbol{\xi}\|^{2}_{\boldsymbol{D}} \tag{4.103k}$$

$$\mathcal{V}^{\mathrm{K}} = \sum\nolimits_{a,b} \tfrac{1}{2}\|{}^{a}_{b}\mathbf{G}^{\top} - {}^{a}_{b}\mathbf{G}_{\mathsf{R}}^{\top}\|^{2}_{{}^{a}_{b}\mathbf{K}'} \tag{4.103l}$$

$$\mathcal{V}^{\mathrm{G}} = \sum\nolimits_{b} \langle {}^{0}_{b}\mathbf{G}^{\top}, \mathrm{wed}(\bar{\boldsymbol{\alpha}}_{\mathsf{G}})\rangle_{{}^{0}_{b}\mathbf{M}'}, \tag{4.103m}$$

Chapter 5

Tracking control of rigid body systems

This chapter motivates and discusses several approaches for a model based design of a tracking controller for a rigid body system by static feedback.

System model. The previous chapter discussed the equations of motion of rigid body systems: For chosen configuration coordinates $\boldsymbol{x}(t) \in \mathbb{X}$, velocity coordinates $\boldsymbol{\xi}(t) \in \mathbb{R}^n$ and the control inputs $\boldsymbol{u}(t) \in \mathbb{R}^p$ these have the form

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}, \qquad \overbrace{\boldsymbol{M}(\boldsymbol{x})\dot{\boldsymbol{\xi}} + \boldsymbol{c}(\boldsymbol{x},\boldsymbol{\xi})}^{\boldsymbol{f}^{\mathrm{M}}(\boldsymbol{x},\boldsymbol{\xi},\dot{\boldsymbol{\xi}})} + \overbrace{\boldsymbol{D}(\boldsymbol{x})\boldsymbol{\xi}}^{\boldsymbol{f}^{\mathrm{D}}(\boldsymbol{x},\boldsymbol{\xi})} + \overbrace{\nabla\mathcal{V}(\boldsymbol{x})}^{\boldsymbol{f}^{\mathrm{K}}(\boldsymbol{x})} = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u}. \tag{5.1}$$

The forces $\boldsymbol{f}^{\mathrm{M}}, \boldsymbol{f}^{\mathrm{D}}, \boldsymbol{f}^{\mathrm{K}}$ may be computed from the rigid body configurations ${}^{a}_{b}\mathbf{G}(\boldsymbol{x})$ and the constitutive parameters ${}^{0}_{b}\mathbf{M}$, ${}^{a}_{b}\mathbf{D}$, ${}^{a}_{b}\mathbf{K}$. This structure will be the main inspiration for the design of the controlled system.

However, mathematically, the control approach does not rely on the model having this structure. We may assume any model of the form

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}, \qquad \boldsymbol{M}(\boldsymbol{x})\dot{\boldsymbol{\xi}} + \boldsymbol{b}(\boldsymbol{x},\boldsymbol{\xi}) = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \tag{5.2}$$

where $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu\times n}$ and $\boldsymbol{B}(\boldsymbol{x}) \in \mathbb{R}^{n\times p}$ are full rank, the system inertia matrix $\boldsymbol{M}(\boldsymbol{x}) \in \mathbb{SYM}^+(n)$ is symmetric, positive definite, and $\boldsymbol{b}(\boldsymbol{x},\boldsymbol{\xi}) \in \mathbb{R}^n$ collects the remaining terms of the kinetic equation. The system is called *fully-actuated* for $p = n$ and *underactuated* for $0 < p < n$. Firstly we will restrict to fully-actuated systems and later try to expand the approach to underactuated systems.

Reference trajectory and tracking controller. Let there be a *reference trajectory* $t \mapsto \boldsymbol{x}_{\mathsf{R}}(t)$ which is compatible with the model (5.2): It must be feasible, i.e. $\boldsymbol{x}_{\mathsf{R}}(t) \in \mathbb{X}$, and sufficiently smooth, so we can define the reference velocity $\boldsymbol{\xi}_{\mathsf{R}} = \boldsymbol{A}^{+}(\boldsymbol{x}_{\mathsf{R}})\dot{\boldsymbol{x}}_{\mathsf{R}}$ and acceleration $\dot{\boldsymbol{\xi}}_{\mathsf{R}}$. For the underactuated case we also require the kinetic equation $\boldsymbol{M}(\boldsymbol{x}_{\mathsf{R}})\dot{\boldsymbol{\xi}}_{\mathsf{R}} + \boldsymbol{b}(\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}}) = \boldsymbol{B}(\boldsymbol{x}_{\mathsf{R}})\boldsymbol{u}_{\mathsf{R}}$ to have a solution for $\boldsymbol{u}_{\mathsf{R}}$.

The design task for a *tracking controller* is: Find a function $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}},\dot{\boldsymbol{\xi}}_{\mathsf{R}})$ (the controller) such that $t \mapsto \boldsymbol{x}_{\mathsf{R}}(t)$ is a stable and attractive trajectory of the closed loop which is the combination of model (5.2) and controller.

State of the art. This is a pretty general task and may be tackled by various standard approaches from control theory, see e.g. [Spong et al., 2006, chap. 7-10] for some overview. For fully actuated systems a popular approach is *computed torque*, see e.g. [Murray et al., 1994, sec. 4.5.2], also called *inverse dynamics* in [Spong et al., 2006, sec. 8.3]. It can be regarded as a particularly simple case of feedback linearization utilizing the fact that any set of minimal generalized coordinates $\boldsymbol{q}(t) \in \mathbb{R}^n$ is a *flat output* of a fully actuated mechanical system [Martin et al., 1997, sec. 7.1].

For underactuated systems there is no standard textbook approach. Examples for the flatness-based approach can be found in e.g. [Rathinam and Murray, 1998], [Murray et al., 1995] or [Martin et al., 1997, sec. 7.1]. General Lyapunov designs can be found in [Olfati-Saber, 2001] and a approach called *controlled Lagrangians* is proposed in [Bloch et al., 2000].

Outline for this chapter. With the computed torque method one might consider the topic to be solved for fully actuated systems. However, for system whose configuration space is not isomorph to $\mathbb{R}^n$ it is only local due to the requirement of minimal coordinates $\boldsymbol{q}(t) \in \mathbb{R}^n$. Furthermore, applying linear dynamics in these coordinates may result in intrinsic singularities for the closed loop. Recalling the satellite example from section 3.2, it should be clear that linear dynamics for the Euler angles would probably not be a good choice, also see [Konz and Rudolph, 2016] for further examples.

If linear dynamics are not a general choice, what *is* a good choice for the closed loop dynamics? This chapter motivates three approaches for designing a closed loop for fully actuated systems which rely on the underlying rigid body structure. In addition, the result will be extended to underactuated systems. Finally the approach will be applied to several example systems including the tricopter (fully-actuated), the quadcopter (underactuated but flat) and the bicopter (underactuated and probably not flat).

5.1 Approach 1: Inspired by particle distribution

5.1.1 Particle system

The basic idea. Consider a system of *free* particles with the equations of motion $\mathfrak{m}_p\ddot{\mathfrak{r}}_p = \mathfrak{F}_p^{\mathsf{A}}, p = 1,\ldots,\mathfrak{N}$ and the control inputs $\mathfrak{F}_p^{\mathsf{A}}$. We want the system to track a smooth reference trajectory $t \mapsto (\mathfrak{r}_{1\mathsf{R}},\ldots,\mathfrak{r}_{\mathfrak{N}\mathsf{R}})(t)$. Probably the simplest solution is the control law $\mathfrak{F}_p^{\mathsf{A}} = \mathfrak{m}_p\ddot{\mathfrak{r}}_{p\mathsf{R}} - \bar{\mathfrak{d}}_p\dot{\mathfrak{r}}_{p\mathrm{E}} - \bar{\mathfrak{k}}_p\mathfrak{r}_{p\mathrm{E}}$ with the position error $\mathfrak{r}_{p\mathrm{E}} = \mathfrak{r}_p - \mathfrak{r}_{p\mathsf{R}}$ and the design parameters $\bar{\mathfrak{k}}_p, \bar{\mathfrak{d}}_p \in \mathbb{R} > 0$. The resulting closed loop is

$$\mathfrak{m}_p\ddot{\mathfrak{r}}_{p\mathrm{E}} + \bar{\mathfrak{d}}_p\dot{\mathfrak{r}}_{p\mathrm{E}} + \bar{\mathfrak{k}}_p\mathfrak{r}_{p\mathrm{E}} = \mathbf{0}, \qquad p = 1\ldots\mathfrak{N}. \tag{5.3}$$

It is clearly exponentially stable and the *desired stiffness* $\bar{\mathfrak{k}}_p$ and *desired damping* $\bar{\mathfrak{d}}_p$ are intuitive tuning parameters.

For a system of particles with geometric constraints $\mathfrak{c}(\mathfrak{r}_1,\ldots,\mathfrak{r}_{\mathfrak{N}}) = \mathbf{0}$ we cannot achieve (5.3) in general. As the next best thing we can get as close as possible by formulation of the following constrained optimization problem

$$\begin{array}{ll} \underset{\ddot{\mathfrak{r}}\in\mathbb{R}^{3\mathfrak{N}}}{\text{minimize}} & \bar{\mathcal{G}} = \frac{1}{2}\sum_p \frac{1}{\bar{\mathfrak{m}}_p}\|\bar{\mathfrak{m}}_p\ddot{\mathfrak{r}}_{p\mathrm{E}} + \bar{\mathfrak{d}}_p\dot{\mathfrak{r}}_{p\mathrm{E}} + \bar{\mathfrak{k}}_p\mathfrak{r}_{p\mathrm{E}}\|^2 \\ \text{subject to} & \mathfrak{c}(\mathfrak{r}_1,\ldots,\mathfrak{r}_{\mathfrak{N}}) = \mathbf{0} \end{array}. \tag{5.4}$$

Note that we also replaced the particle masses $\mathfrak{m}_p$ by *desired masses* $\bar{\mathfrak{m}}_p$ as a additional design parameters. This will turn out crucial for control of underactuated systems.

The controlled system. The constrained problem (5.4) can be transformed to an unconstrained one by formulating the particle accelerations $\ddot{\mathfrak{r}}_p = \ddot{\mathfrak{r}}_p(\boldsymbol{x},\boldsymbol{\xi},\dot{\boldsymbol{\xi}})$, $p = 1\ldots\mathfrak{N}$ in terms of minimal acceleration coordinates $\dot{\boldsymbol{\xi}}$. Analogous, let the reference particle positions be formulated in terms of the reference coordinates $\boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}, \dot{\boldsymbol{\xi}}_{\mathsf{R}}$, i.e. $\mathfrak{r}_{p\mathsf{R}} = \mathfrak{r}_p(\boldsymbol{x}_{\mathsf{R}})$, $\dot{\mathfrak{r}}_{p\mathsf{R}} = \dot{\mathfrak{r}}_p(\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}})$, $\ddot{\mathfrak{r}}_{p\mathsf{R}} = \ddot{\mathfrak{r}}_p(\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}},\dot{\boldsymbol{\xi}}_{\mathsf{R}})$, and the position error $\mathfrak{r}_{p\mathrm{E}} = \mathfrak{r}_{p\mathrm{E}}(\boldsymbol{x},\boldsymbol{x}_{\mathsf{R}}) = \mathfrak{r}_p(\boldsymbol{x}) - \mathfrak{r}_p(\boldsymbol{x}_{\mathsf{R}})$, etc.. With this, the solution of (5.4) can be computed from

$$\begin{aligned} \frac{\partial\bar{\mathcal{G}}}{\partial\dot{\xi}^i} &= \sum_p \langle\bar{\mathfrak{m}}_p\ddot{\mathfrak{r}}_{p\mathrm{E}} + \bar{\mathfrak{d}}_p\dot{\mathfrak{r}}_{p\mathrm{E}} + \bar{\mathfrak{k}}_p\mathfrak{r}_{p\mathrm{E}}, \frac{\partial\ddot{\mathfrak{r}}_p}{\partial\dot{\xi}^i}\rangle \\ &= \sum_p \bar{\mathfrak{m}}_p\langle\ddot{\mathfrak{r}}_{p\mathrm{E}}, \frac{\partial\ddot{\mathfrak{r}}_p}{\partial\dot{\xi}^i}\rangle + \sum_p \bar{\mathfrak{d}}_p\langle\dot{\mathfrak{r}}_{p\mathrm{E}}, \frac{\partial\dot{\mathfrak{r}}_p}{\partial\xi^i}\rangle + \sum_p \bar{\mathfrak{k}}_p\langle\mathfrak{r}_{p\mathrm{E}}, \partial_i\mathfrak{r}_p\rangle \\ &= \underbrace{\frac{\partial}{\partial\dot{\xi}^i}\underbrace{\sum_p \tfrac{1}{2}\bar{\mathfrak{m}}_p\|\ddot{\mathfrak{r}}_{p\mathrm{E}}\|^2}_{\bar{\mathcal{S}}}}_{\bar{f}_i^{\mathsf{M}}} + \underbrace{\frac{\partial}{\partial\xi^i}\underbrace{\sum_p \tfrac{1}{2}\bar{\mathfrak{d}}_p\|\dot{\mathfrak{r}}_{p\mathrm{E}}\|^2}_{\bar{\mathcal{R}}}}_{\bar{f}_i^{\mathsf{D}}} + \underbrace{\partial_i\underbrace{\sum_p \tfrac{1}{2}\bar{\mathfrak{k}}_p\|\mathfrak{r}_{p\mathrm{E}}\|^2}_{\bar{\mathcal{V}}}}_{\bar{f}_i^{\mathsf{K}}} = 0, \quad i = 1,\ldots,n. \end{aligned} \tag{5.5}$$

Here we introduced formulations for the *controlled acceleration energy* $\bar{\mathcal{S}}(\boldsymbol{x},\boldsymbol{\xi},\dot{\boldsymbol{\xi}},\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}},\dot{\boldsymbol{\xi}}_{\mathsf{R}})$, the *controlled dissipation function* $\bar{\mathcal{R}}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}})$ and the *controlled potential energy* $\bar{\mathcal{V}}(\boldsymbol{x},\boldsymbol{x}_{\mathsf{R}})$. It is worth noting that the inertia force $\bar{\boldsymbol{f}}^{\mathsf{M}}$ could also be derived from the *controlled kinetic energy* $\bar{\mathcal{T}}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{x}_{\mathsf{R}},\boldsymbol{\xi}_{\mathsf{R}})$ as

$$\bar{f}_i^{\mathsf{M}} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{\mathcal{T}}}{\partial\xi^i} + \gamma_{ij}^k\xi^j\frac{\partial\bar{\mathcal{T}}}{\partial\xi^k} - \partial_i\bar{\mathcal{T}}, \qquad \bar{\mathcal{T}} = \tfrac{1}{2}\sum_p \bar{\mathfrak{m}}_p\|\dot{\mathfrak{r}}_{p\mathrm{E}}\|^2. \tag{5.6}$$

Also note that all the defined "energies" are symmetric in the sense that $\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = \bar{\mathcal{V}}(\boldsymbol{x}_\mathsf{R}, \boldsymbol{x})$, etc..

The corresponding forces expressed more explicitly are

$$\begin{aligned}
\bar{f}_i^{\mathsf{M}} = \frac{\partial \bar{\mathcal{S}}}{\partial \dot{\xi}^i} &= \underbrace{\sum\nolimits_p \bar{\mathfrak{m}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \partial_j \mathfrak{r}_p(\boldsymbol{x}) \rangle}_{\bar{M}_{ij}(\boldsymbol{x})} \dot{\xi}^j + \underbrace{\sum\nolimits_p \bar{\mathfrak{m}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \partial_k \partial_j \mathfrak{r}_p(\boldsymbol{x}) \rangle}_{\bar{\Gamma}_{ijk}(\boldsymbol{x})} \xi^j \xi^k \\
&\quad - \underbrace{\sum\nolimits_p \bar{\mathfrak{m}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \partial_j \mathfrak{r}_p(\boldsymbol{x}_\mathsf{R}) \rangle}_{\bar{M}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})} \dot{\xi}_\mathsf{R}^j - \underbrace{\sum\nolimits_p \bar{\mathfrak{m}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \partial_k \partial_j \mathfrak{r}_p(\boldsymbol{x}_\mathsf{R}) \rangle}_{\bar{\Gamma}'_{ijk}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})} \xi_\mathsf{R}^j \xi_\mathsf{R}^k, \quad (5.7a) \\
\bar{f}_i^{\mathsf{D}} = \frac{\partial \bar{\mathcal{R}}}{\partial \xi^i} &= \underbrace{\sum\nolimits_p \bar{\mathfrak{d}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \partial_j \mathfrak{r}_p(\boldsymbol{x}) \rangle}_{\bar{D}_{ij}(\boldsymbol{x})} \xi^j - \underbrace{\sum\nolimits_p \bar{\mathfrak{d}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \partial_j \mathfrak{r}_p(\boldsymbol{x}_\mathsf{R}) \rangle}_{\bar{D}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})} \xi_\mathsf{R}^j, \quad (5.7b) \\
\bar{f}_i^{\mathsf{K}} = \partial_i \bar{\mathcal{V}} &= \sum\nolimits_p \bar{\mathfrak{k}}_p \langle \partial_i \mathfrak{r}_p(\boldsymbol{x}), \mathfrak{r}_p(\boldsymbol{x}) - \mathfrak{r}_p(\boldsymbol{x}_\mathsf{R}) \rangle. \quad (5.7c)
\end{aligned}$$

So we can rewrite (5.5) as

$$\begin{aligned}
&\bar{M}_{ij}(\boldsymbol{x})\dot{\xi}^j - \bar{M}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\dot{\xi}_\mathsf{R}^j + \bar{\Gamma}_{ijk}(\boldsymbol{x})\xi^j\xi^k - \bar{\Gamma}'_{ijk}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^j\xi_\mathsf{R}^k \\
&\qquad + \bar{D}_{ij}(\boldsymbol{x})\xi^j - \bar{D}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^j + \partial_i \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = 0, \quad i = 1, \ldots, n. \quad (5.8)
\end{aligned}$$

Total energy. Having a definition for a kinetic energy $\bar{\mathcal{T}}$ and a potential energy $\bar{\mathcal{V}}$ it is worth investigating the total energy $\bar{\mathcal{W}} = \bar{\mathcal{T}} + \bar{\mathcal{V}}$ and its change along the solutions of closed loop (5.8). Using the substitutions defined in (5.7) and $\boldsymbol{\Psi}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) \in \mathbb{R}^{n \times n}$ defined through $\bar{M}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = \Psi_i^s(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\bar{M}_{sj}(\boldsymbol{x})$ we have

$$\bar{\mathcal{W}} = \overbrace{\tfrac{1}{2}\bar{M}_{ij}(\boldsymbol{x})\xi^i\xi^j - \bar{M}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\xi^i\xi_\mathsf{R}^j + \tfrac{1}{2}\bar{M}_{ij}(\boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^i\xi_\mathsf{R}^j}^{\bar{\mathcal{T}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_\mathsf{R}, \boldsymbol{\xi}_\mathsf{R})} + \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) \quad (5.9)$$

$$\begin{aligned}
\dot{\bar{\mathcal{W}}} &= \xi^i \big(\bar{M}_{ij}(\boldsymbol{x})\dot{\xi}^j + \bar{\Gamma}_{ijk}(\boldsymbol{x})\xi^j\xi^k - \bar{M}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\dot{\xi}_\mathsf{R}^j - \bar{\Gamma}'_{ijk}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^j\xi_\mathsf{R}^k + \partial_i \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) \big) \\
&\quad + \xi_\mathsf{R}^i \big(\bar{M}_{ij}(\boldsymbol{x}_\mathsf{R})\dot{\xi}_\mathsf{R}^j + \bar{\Gamma}_{ijk}(\boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^j\xi_\mathsf{R}^k - \bar{M}'_{ij}(\boldsymbol{x}_\mathsf{R}, \boldsymbol{x})\dot{\xi}^j - \bar{\Gamma}'_{ijk}(\boldsymbol{x}_\mathsf{R}, \boldsymbol{x})\xi^j\xi^k + \partial_i \bar{\mathcal{V}}(\boldsymbol{x}_\mathsf{R}, \boldsymbol{x}) \big) \\
&\overset{(5.8)}{=} -(\xi^i - \Psi_s^i(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^s)\big(\bar{D}_{ij}(\boldsymbol{x})\xi^j - \bar{D}'_{ij}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\xi_\mathsf{R}^j \big) \\
&\quad + \xi_\mathsf{R}^i \big(\partial_i \bar{\mathcal{V}}(\boldsymbol{x}_\mathsf{R}, \boldsymbol{x}) + \Psi_i^s(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\partial_s \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) + \big(\bar{M}_{ij}(\boldsymbol{x}_\mathsf{R}) - \Psi_i^s(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\bar{M}'_{sj}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) \big)\dot{\xi}_\mathsf{R}^j \\
&\quad + \big(\bar{\Gamma}_{ijk}(\boldsymbol{x}_\mathsf{R}) - \Psi_i^s(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\bar{\Gamma}'_{sjk}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) \big)\xi_\mathsf{R}^j\xi_\mathsf{R}^k - \big(\bar{\Gamma}'_{ijk}(\boldsymbol{x}_\mathsf{R}, \boldsymbol{x}) - \Psi_i^s(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\bar{\Gamma}_{sjk}(\boldsymbol{x}) \big)\xi^j\xi^k \big)
\end{aligned} \quad (5.10)$$

Obviously the total energy $\bar{\mathcal{W}}$ is not a Lyapunov function for a general reference trajectory. It is, however, for the very special case of $\boldsymbol{\xi}_\mathsf{R} = \mathbf{0}$, i.e. proves stability for a constant reference configuration $\boldsymbol{x}_\mathsf{R} = const.$.

5.1.2 Free rigid body

Consider the free rigid body discussed in section 4.1 as a special case of a particle system. As motivated there we use the position $\boldsymbol{r}(t) \in \mathbb{R}^3$ and orientation matrix $\boldsymbol{R}(t) \in \mathbb{SO}(3)$

merged into the configuration matrix $\mathbf{G}(t) = \left[\begin{smallmatrix} \boldsymbol{R}(t) & \boldsymbol{r}(t) \\ \mathbf{0} & 1 \end{smallmatrix}\right] \in \mathbb{SE}(3)$ as configuration coordinates. Expressing the particle positions as $\mathfrak{r}_p = \boldsymbol{r} + \boldsymbol{R}\mathfrak{h}_p$ and applying the same calculations as in subsection 4.2.3 we can write the energies from (5.5) as

$$\bar{\mathcal{V}} = \sum\nolimits_p \tfrac{1}{2}\bar{\mathfrak{k}}_p \|\mathfrak{r}_p - \mathfrak{r}_{p\mathsf{R}}\|^2 = \tfrac{1}{2}\|\big(\mathbf{G} - \mathbf{G}_\mathsf{R}\big)^\top\|^2_{\bar{\mathbf{K}}'} \tag{5.11a}$$

$$\bar{\mathcal{R}} = \sum\nolimits_p \tfrac{1}{2}\bar{\mathfrak{d}}_p \|\dot{\mathfrak{r}}_p - \dot{\mathfrak{r}}_{p\mathsf{R}}\|^2 = \tfrac{1}{2}\|\big(\dot{\mathbf{G}} - \dot{\mathbf{G}}_\mathsf{R}\big)^\top\|^2_{\bar{\mathbf{D}}'} \tag{5.11b}$$

$$\bar{\mathcal{S}} = \sum\nolimits_p \tfrac{1}{2}\bar{\mathfrak{m}}_p \|\ddot{\mathfrak{r}}_p - \ddot{\mathfrak{r}}_{p\mathsf{R}}\|^2 = \tfrac{1}{2}\|\big(\ddot{\mathbf{G}} - \ddot{\mathbf{G}}_\mathsf{R}\big)^\top\|^2_{\bar{\mathbf{M}}'} \tag{5.11c}$$

$$\bar{\mathcal{T}} = \sum\nolimits_p \tfrac{1}{2}\bar{\mathfrak{m}}_p \|\dot{\mathfrak{r}}_p - \dot{\mathfrak{r}}_{p\mathsf{R}}\|^2 = \tfrac{1}{2}\|\big(\dot{\mathbf{G}} - \dot{\mathbf{G}}_\mathsf{R}\big)^\top\|^2_{\bar{\mathbf{M}}'} \tag{5.11d}$$

where

$$\bar{\mathbf{K}}' = \sum\nolimits_p \bar{\mathfrak{k}}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\Pi}}' & \bar{k}\bar{\boldsymbol{h}} \\ \bar{k}\bar{\boldsymbol{h}}^\top & \bar{k} \end{bmatrix}, \tag{5.12a}$$

$$\bar{\mathbf{D}}' = \sum\nolimits_p \bar{\mathfrak{d}}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\Upsilon}}' & \bar{d}\bar{\boldsymbol{l}} \\ \bar{d}\bar{\boldsymbol{l}}^\top & \bar{d} \end{bmatrix}, \tag{5.12b}$$

$$\bar{\mathbf{M}}' = \sum\nolimits_p \bar{\mathfrak{m}}_p \begin{bmatrix} \mathfrak{h}_p\mathfrak{h}_p^\top & \mathfrak{h}_p^\top \\ \mathfrak{h}_p & 1 \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\Theta}}' & \bar{m}\bar{\boldsymbol{s}} \\ \bar{m}\bar{\boldsymbol{s}}^\top & \bar{m} \end{bmatrix}. \tag{5.12c}$$

As before we can interpret the entries of the *desired inertia matrix* $\bar{\mathbf{M}} = \mathrm{Vee}(\bar{\mathbf{M}}')$ as the *desired total mass* $\bar{m}$, *desired center of mass* $\bar{\boldsymbol{s}}$ and the *desired moment of inertia* $\bar{\boldsymbol{\Theta}} = \mathrm{Vee}(\bar{\boldsymbol{\Theta}}')$. The analog holds for the entries of the *desired damping matrix* $\bar{\mathbf{D}} = \mathrm{Vee}(\bar{\mathbf{D}}')$ and the *desired stiffness matrix* $\bar{\mathbf{K}} = \mathrm{Vee}(\bar{\mathbf{K}}')$.

Introduce the translational $\boldsymbol{v}(t) \in \mathbb{R}^3$ and angular velocity $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ merged into $\boldsymbol{\xi} = [\boldsymbol{v}^\top, \boldsymbol{\omega}^\top]^\top = \mathrm{vee}(\mathbf{G}^{-1}\dot{\mathbf{G}})$ as velocity coordinates. Furthermore, we introduce the *configuration error* $\mathbf{G}_\mathrm{E} = \mathbf{G}_\mathsf{R}^{-1}\mathbf{G}$ and exploit the invariance (A.15a) of the norm to left translation of the argument to express the desired energies (5.11) as

$$\bar{\mathcal{V}} = \tfrac{1}{2}\|\big(\mathbf{I}_4 - \mathbf{G}_\mathrm{E}^{-1}\big)^\top\|^2_{\bar{\mathbf{K}}'} \tag{5.13a}$$

$$\bar{\mathcal{R}} = \tfrac{1}{2}\|\big(\mathrm{wed}(\boldsymbol{\xi}) - \mathbf{G}_\mathrm{E}^{-1}\,\mathrm{wed}(\boldsymbol{\xi}_\mathsf{R})\big)^\top\|^2_{\bar{\mathbf{D}}'} \tag{5.13b}$$

$$\bar{\mathcal{S}} = \tfrac{1}{2}\|\big(\mathrm{wed}(\dot{\boldsymbol{\xi}}) + \mathrm{wed}(\boldsymbol{\xi})^2 - \mathbf{G}_\mathrm{E}^{-1}\big(\mathrm{wed}(\dot{\boldsymbol{\xi}}_\mathsf{R}) + \mathrm{wed}(\boldsymbol{\xi}_\mathsf{R})^2\big)\big)^\top\|^2_{\bar{\mathbf{M}}'} \tag{5.13c}$$

$$\bar{\mathcal{T}} = \tfrac{1}{2}\|\big(\mathrm{wed}(\boldsymbol{\xi}) - \mathbf{G}_\mathrm{E}^{-1}\,\mathrm{wed}(\boldsymbol{\xi}_\mathsf{R})\big)^\top\|^2_{\bar{\mathbf{M}}'} \tag{5.13d}$$

Using the vee2-operator defined in (4.56) the resulting forces can be expressed as

$$\bar{\boldsymbol{f}}^\mathsf{K} = \boldsymbol{\nabla}\bar{\mathcal{V}} = \mathrm{vee2}\left((\mathbf{I}_4 - \mathbf{G}_\mathrm{E}^{-1})\bar{\mathbf{K}}'\right), \tag{5.14a}$$

$$\bar{\boldsymbol{f}}^\mathsf{D} = \frac{\partial\bar{\mathcal{R}}}{\partial\boldsymbol{\xi}} = \mathrm{vee2}\left((\mathrm{wed}(\boldsymbol{\xi}) - \mathbf{G}_\mathrm{E}^{-1}\,\mathrm{wed}(\boldsymbol{\xi}_\mathsf{R}))\bar{\mathbf{D}}'\right), \tag{5.14b}$$

$$\bar{\boldsymbol{f}}^\mathsf{M} = \frac{\partial\bar{\mathcal{S}}}{\partial\dot{\boldsymbol{\xi}}} = \mathrm{vee2}\left(\big(\mathrm{wed}(\dot{\boldsymbol{\xi}}) + \mathrm{wed}(\boldsymbol{\xi})^2 - \mathbf{G}_\mathrm{E}^{-1}\big(\mathrm{wed}(\dot{\boldsymbol{\xi}}_\mathsf{R}) + \mathrm{wed}(\boldsymbol{\xi}_\mathsf{R})^2\big)\big)\bar{\mathbf{M}}'\right) \tag{5.14c}$$

or more explicitly

$$\bar{\boldsymbol{f}}^{\mathsf{K}} = \begin{bmatrix} \bar{k}\boldsymbol{R}_{\mathrm{E}}^{\top}(\boldsymbol{r}_{\mathrm{E}} + (\boldsymbol{R}_{\mathrm{E}} - \mathbf{I}_3)\bar{\boldsymbol{h}}) \\ \bar{k}\,\mathrm{wed}(\bar{\boldsymbol{h}})\boldsymbol{R}_{\mathrm{E}}^{\top}\boldsymbol{r}_{\mathrm{E}} + 2\,\mathrm{vee}\left(\mathrm{Vee}(\bar{\boldsymbol{\Pi}})(\boldsymbol{R}_{\mathrm{E}} - \mathbf{I}_3)\right) \end{bmatrix} \tag{5.15a}$$

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \underbrace{\begin{bmatrix} \bar{d}\mathbf{I}_3 & \bar{d}\,\mathrm{wed}(\bar{\boldsymbol{l}})^{\top} \\ \bar{d}\,\mathrm{wed}(\bar{\boldsymbol{l}}) & \bar{\boldsymbol{\Upsilon}} \end{bmatrix}}_{\bar{\mathsf{D}}} \underbrace{\begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{bmatrix}}_{\boldsymbol{\xi}} - \begin{bmatrix} \bar{d}\boldsymbol{R}_{\mathrm{E}}^{\top} & \boldsymbol{R}_{\mathrm{E}}^{\top}\bar{d}\,\mathrm{wed}(\bar{\boldsymbol{l}})^{\top} \\ \bar{d}\,\mathrm{wed}(\bar{\boldsymbol{l}})\boldsymbol{R}_{\mathrm{E}}^{\top} & \mathrm{Wed}(\boldsymbol{R}_{\mathrm{E}}^{\top}\,\mathrm{Vee}(\bar{\boldsymbol{\Upsilon}}))\boldsymbol{R}_{\mathrm{E}}^{\top} \end{bmatrix} \underbrace{\begin{bmatrix} \boldsymbol{v}_{\mathsf{R}} \\ \boldsymbol{\omega}_{\mathsf{R}} \end{bmatrix}}_{\boldsymbol{\xi}_{\mathsf{R}}} \tag{5.15b}$$

$$\begin{aligned}\bar{\boldsymbol{f}}^{\mathsf{M}} = {} & \underbrace{\begin{bmatrix} \bar{m}\mathbf{I}_3 & \bar{m}\,\mathrm{wed}(\bar{\boldsymbol{s}})^{\top} \\ \bar{m}\,\mathrm{wed}(\bar{\boldsymbol{s}}) & \bar{\boldsymbol{\Theta}} \end{bmatrix}}_{\bar{\mathsf{M}}} \underbrace{\begin{bmatrix} \dot{\boldsymbol{v}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}} + \begin{bmatrix} \bar{m}\,\mathrm{wed}(\boldsymbol{\omega}) & -\bar{m}\,\mathrm{wed}(\boldsymbol{\omega})\,\mathrm{wed}(\bar{\boldsymbol{s}}) \\ \bar{m}\,\mathrm{wed}(\bar{\boldsymbol{s}})\,\mathrm{wed}(\boldsymbol{\omega}) & \mathrm{wed}(\mathrm{Vee}(\bar{\boldsymbol{\Theta}})\boldsymbol{\omega}) \end{bmatrix} \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{\omega} \end{bmatrix} \\ & - \begin{bmatrix} \bar{m}\boldsymbol{R}_{\mathrm{E}}^{\top} & \boldsymbol{R}_{\mathrm{E}}^{\top}\bar{m}\,\mathrm{wed}(\bar{\boldsymbol{s}})^{\top} \\ \bar{m}\,\mathrm{wed}(\bar{\boldsymbol{s}})\boldsymbol{R}_{\mathrm{E}}^{\top} & \mathrm{Wed}(\boldsymbol{R}_{\mathrm{E}}^{\top}\,\mathrm{Vee}(\bar{\boldsymbol{\Theta}}))\boldsymbol{R}_{\mathrm{E}}^{\top} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{v}}_{\mathsf{R}} \\ \dot{\boldsymbol{\omega}}_{\mathsf{R}} \end{bmatrix} \\ & - \begin{bmatrix} \bar{m}\boldsymbol{R}_{\mathrm{E}}^{\top}\,\mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}}) & -\bar{m}\boldsymbol{R}_{\mathrm{E}}^{\top}\,\mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}})\,\mathrm{wed}(\bar{\boldsymbol{s}}) \\ \bar{m}\,\mathrm{wed}(\bar{\boldsymbol{s}})\boldsymbol{R}_{\mathrm{E}}^{\top}\,\mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}}) & \mathrm{Wed}(\boldsymbol{R}_{\mathrm{E}}^{\top}\,\mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}})\,\mathrm{Vee}(\bar{\boldsymbol{\Theta}}))\boldsymbol{R}_{\mathrm{E}}^{\top} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_{\mathsf{R}} \\ \boldsymbol{\omega}_{\mathsf{R}} \end{bmatrix}\end{aligned} \tag{5.15c}$$

The closed loop kinetic equation $\bar{\boldsymbol{f}}^{\mathsf{M}} + \bar{\boldsymbol{f}}^{\mathsf{D}} + \bar{\boldsymbol{f}}^{\mathsf{K}} = \mathbf{0}$ contains 30 tuning parameters within the matrices $\bar{\mathsf{M}}'$, $\bar{\mathsf{D}}'$, $\bar{\mathsf{K}}' \in \mathbb{SYM}^{+}(4)$. The characteristic polynomial of the first order approximation of the system about any constant configuration $\boldsymbol{x}_{\mathsf{R}} = const.$ is $\det(\bar{\mathsf{M}}\lambda^2 + \bar{\mathsf{D}}\lambda + \bar{\mathsf{K}})$ where $\bar{\mathsf{M}} = \mathrm{Vee}(\bar{\mathsf{M}}')$, etc..

blah

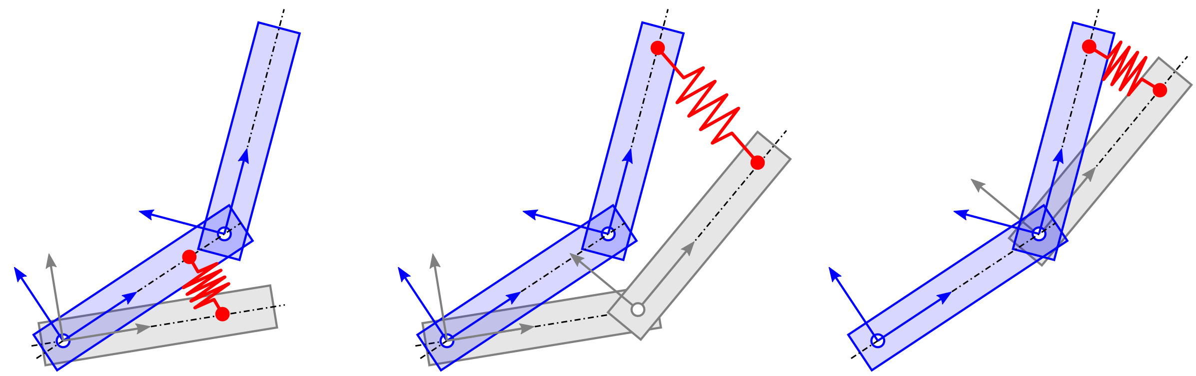
In contrast to this, the characteristic polynomial resulting from the computed torque method (see subsection 5.3.2) for this system is $\det(\mathbf{I}_6\lambda^2 + \boldsymbol{K}_1\lambda + \boldsymbol{K}_2)$, which has $n(n+1) = 42$ tuning parameters within the matrixes $\boldsymbol{K}_1, \boldsymbol{K}_2 \in \mathbb{SYM}^{+}(6)$.

blah

A possible generalization of the rigid body potential $\bar{\mathcal{V}}$ which works with all $\frac{1}{2}n(n+1) = 21$ tuning parameters in a matrix $\bar{\mathsf{K}} \in \mathbb{SYM}^{+}(6)$ is given in section B.3.

5.1.3 Rigid body systems

Let the particles belong to a system of N rigid bodies with the body configurations ${}^{a}_{b}\mathsf{G}$ as discussed in section 4.3. The potential energy from (5.5) may be formulated as $\bar{\mathcal{V}} = \sum_{b=1}^{N} \frac{1}{2} \| ({}^{0}_{b}\mathsf{G} - {}^{0}_{b}\mathsf{G}_{\mathsf{R}})^{\top} \|^2_{{}^{0}_{b}\bar{\mathsf{K}}'}$ with a body stiffness matrix ${}^{0}_{b}\bar{\mathsf{K}}'$ resulting from (5.12) for each body. This potential only captures stiffness w.r.t. the absolute configurations ${}^{0}_{b}\mathsf{G}$. Depending on the control objective it may be equally reasonable to consider a stiffness associated with the relative configurations ${}^{a}_{b}\mathsf{G}$ as illustrated in Figure 5.1. Considering the same argument for damping and inertia, we propose the following energies for the



${}^0_1\bar{\mathcal{V}}$ ${}^0_2\bar{\mathcal{V}}$ ${}^1_2\bar{\mathcal{V}}$

Figure 5.1: Different parts of the potential $\bar{\mathcal{V}}$ for a double pendulum

control of a rigid body system:

$$\bar{\mathcal{V}} = \sum\nolimits_{a,b=0}^{N} \overbrace{\tfrac{1}{2}\big\|\big({}^a_b\mathbf{G} - {}^a_b\mathbf{G}_\mathsf{R}\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{K}}'}}^{{}^a_b\bar{\mathcal{V}}}, \qquad {}^a_b\bar{\mathbf{K}}' \in \mathbb{SYM}_0^+(4) \tag{5.16a}$$

$$\bar{\mathcal{R}} = \sum\nolimits_{a,b=0}^{N} \tfrac{1}{2}\big\|\big({}^a_b\dot{\mathbf{G}} - {}^a_b\dot{\mathbf{G}}_\mathsf{R}\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{D}}'}, \qquad {}^a_b\bar{\mathbf{D}}' \in \mathbb{SYM}_0^+(4) \tag{5.16b}$$

$$\bar{\mathcal{S}} = \sum\nolimits_{a,b=0}^{N} \tfrac{1}{2}\big\|\big({}^a_b\ddot{\mathbf{G}} - {}^a_b\ddot{\mathbf{G}}_\mathsf{R}\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{M}}'}, \qquad {}^a_b\bar{\mathbf{M}}' \in \mathbb{SYM}_0^+(4) \tag{5.16c}$$

$$\bar{\mathcal{T}} = \sum\nolimits_{a,b=0}^{N} \tfrac{1}{2}\big\|\big({}^a_b\dot{\mathbf{G}} - {}^a_b\dot{\mathbf{G}}_\mathsf{R}\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{M}}'}. \tag{5.16d}$$

Note that ${}^a_b\bar{\mathbf{K}}' = {}^b_a\bar{\mathbf{K}}'$ implies ${}^a_b\bar{\mathcal{V}} = {}^b_a\bar{\mathcal{V}}$ and ${}^a_a\bar{\mathcal{V}} = 0$ since ${}^a_a\mathbf{G} = {}^a_a\mathbf{G}_\mathsf{R} = \mathbf{I}_4$ and analog for damping and inertia.

Let the body configurations ${}^a_b\mathbf{G}(\boldsymbol{x})$ and the body velocities ${}^a_b\boldsymbol{\xi} = {}^a_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}$ be formulated in terms of suitable system coordinates $\boldsymbol{x}$ and $\boldsymbol{\xi}$, as discussed in section 4.3. With the shorthand notations ${}^a_b\mathbf{G}_\mathrm{E} = {}^a_b\mathbf{G}_\mathrm{E}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = {}^a_b\mathbf{G}^{-1}(\boldsymbol{x}_\mathsf{R}){}^a_b\mathbf{G}(\boldsymbol{x})$ and ${}^a_b\boldsymbol{J} = {}^a_b\boldsymbol{J}(\boldsymbol{x})$, ${}^a_b\boldsymbol{J}_\mathsf{R} = {}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R})$ we can express (5.16) as

$$\bar{\mathcal{V}} = \sum\nolimits_{a,b} \tfrac{1}{2}\big\|\big(\mathbf{I}_4 - {}^a_b\mathbf{G}_\mathrm{E}^{-1}\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{K}}'}, \tag{5.17a}$$

$$\bar{\mathcal{R}} = \sum\nolimits_{a,b} \tfrac{1}{2}\big\|\big(\operatorname{wed}({}^a_b\boldsymbol{J}\boldsymbol{\xi}) - {}^a_b\mathbf{G}_\mathrm{E}^{-1}\operatorname{wed}({}^a_b\boldsymbol{J}_\mathsf{R}\boldsymbol{\xi}_\mathsf{R})\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{D}}'}, \tag{5.17b}$$

$$\begin{aligned}\bar{\mathcal{S}} = \sum\nolimits_{a,b} \tfrac{1}{2}\big\|\big(&\operatorname{wed}({}^a_b\boldsymbol{J}\dot{\boldsymbol{\xi}} + {}^a_b\dot{\boldsymbol{J}}\boldsymbol{\xi}) + \operatorname{wed}({}^a_b\boldsymbol{J}\boldsymbol{\xi})^2 \\ &- {}^a_b\mathbf{G}_\mathrm{E}^{-1}\big(\operatorname{wed}({}^a_b\boldsymbol{J}_\mathsf{R}\dot{\boldsymbol{\xi}}_\mathsf{R} + {}^a_b\dot{\boldsymbol{J}}_\mathsf{R}\boldsymbol{\xi}_\mathsf{R}) + \operatorname{wed}({}^a_b\boldsymbol{J}_\mathsf{R}\boldsymbol{\xi}_\mathsf{R})^2\big)\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{M}}'}\end{aligned} \tag{5.17c}$$

$$\bar{\mathcal{T}} = \sum\nolimits_{a,b} \tfrac{1}{2}\big\|\big(\operatorname{wed}({}^a_b\boldsymbol{J}\boldsymbol{\xi}) - {}^a_b\mathbf{G}_\mathrm{E}^{-1}\operatorname{wed}({}^a_b\boldsymbol{J}_\mathsf{R}\boldsymbol{\xi}_\mathsf{R})\big)^\top\big\|^2_{{}^a_b\bar{\mathbf{M}}'}. \tag{5.17d}$$

Plugging this into the original definition of the closed loop (5.7) we find:

The desired closed loop system for the particle based approach is given by

$$\bar{\boldsymbol{f}}^\mathsf{M} + \bar{\boldsymbol{f}}^\mathsf{D} + \bar{\boldsymbol{f}}^\mathsf{K} = \mathbf{0} \tag{5.18a}$$

where

$$\bar{\boldsymbol{f}}^{\mathsf{K}} = \boldsymbol{\nabla}\bar{\mathcal{V}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\left((\mathbf{I}_4 - {}^a_b\mathbf{G}_{\mathrm{E}}^{-1}){}^a_b\bar{\mathbf{K}}'\right) \tag{5.18b}$$

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \frac{\partial\bar{\mathcal{R}}}{\partial\boldsymbol{\xi}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\left((\operatorname{wed}({}^a_b\boldsymbol{J}\boldsymbol{\xi}) - {}^a_b\mathbf{G}_{\mathrm{E}}^{-1}\operatorname{wed}({}^a_b\boldsymbol{J}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}})){}^a_b\bar{\mathbf{D}}'\right) \tag{5.18c}$$

$$\begin{aligned}\bar{\boldsymbol{f}}^{\mathsf{M}} = \frac{\partial\bar{\mathcal{S}}}{\partial\dot{\boldsymbol{\xi}}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\Big(\Big(&\operatorname{wed}({}^a_b\boldsymbol{J}\dot{\boldsymbol{\xi}} + {}^a_b\dot{\boldsymbol{J}}\boldsymbol{\xi}) + \operatorname{wed}({}^a_b\boldsymbol{J}\boldsymbol{\xi})^2 \\ &- {}^a_b\mathbf{G}_{\mathrm{E}}^{-1}\Big(\operatorname{wed}({}^a_b\boldsymbol{J}_{\mathsf{R}}\dot{\boldsymbol{\xi}}_{\mathsf{R}} + {}^a_b\dot{\boldsymbol{J}}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}}) + \operatorname{wed}({}^a_b\boldsymbol{J}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}})^2\Big)\Big){}^a_b\bar{\mathbf{M}}'\Big)\end{aligned} \tag{5.18d}$$

The system inertia matrix $\bar{\boldsymbol{M}}$ can be recovered from the first term in (5.18d):

$$\sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\left(\operatorname{wed}({}^a_b\boldsymbol{J}\dot{\boldsymbol{\xi}}){}^a_b\bar{\mathbf{M}}'\right) = \underbrace{\sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{Wed}({}^a_b\bar{\mathbf{M}}'){}^a_b\boldsymbol{J}}_{\bar{\boldsymbol{M}}}\,\dot{\boldsymbol{\xi}} = \frac{\partial^2\bar{\mathcal{S}}}{\partial\dot{\boldsymbol{\xi}}\partial\dot{\boldsymbol{\xi}}} = \frac{\partial^2\bar{\mathcal{T}}}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}} \tag{5.19}$$

Though the body inertia matrices ${}^a_b\bar{\mathbf{M}}' \in \mathbb{SYM}_0^+(4)$ from (5.16c) are only required to be positive semi-definite, the resulting system inertia matrix $\bar{\boldsymbol{M}}(\boldsymbol{x}) \in \mathbb{SYM}^+(n)$ is required to be positive definite for the closed loop to be solvable.

5.2 Approach 2: Body based approach

The previous approach for the design of a closed loop has a vivid interpretation for the control energies and parameters. However, the total energy $\mathcal{W}$ does not, in general, serve as a Lyapunov function. In this section we attempt to modify the energies to account for this.

5.2.1 Free rigid body

Using the configuration error $\mathbf{G}_{\mathrm{E}} = \mathbf{G}_{\mathsf{R}}^{-1}\mathbf{G}$ and its velocity $\boldsymbol{\xi}_{\mathrm{E}} = \mathbf{G}_{\mathrm{E}}^{-1}\dot{\mathbf{G}}_{\mathrm{E}} = \boldsymbol{\xi} - \mathrm{Ad}_{\mathbf{G}_{\mathrm{E}}}^{-1}\boldsymbol{\xi}_{\mathsf{R}}$ we modify the energies from section 5.1 to

$$
\begin{aligned}
\bar{\mathcal{V}} &= \tfrac{1}{2}\|(\mathbf{G}_{\mathrm{E}} - \mathbf{I}_4)^\top\|_{\bar{\mathbf{K}}'}^2, && \bar{\mathbf{K}}' \in \mathbb{SYM}^+(4), && (5.20a)\\
\bar{\mathcal{R}} &= \tfrac{1}{2}\|\dot{\mathbf{G}}_{\mathrm{E}}^\top\|_{\bar{\mathbf{D}}'}^2 = \tfrac{1}{2}\|\operatorname{wed}(\boldsymbol{\xi}_{\mathrm{E}})^\top\|_{\bar{\mathbf{D}}'}^2 = \tfrac{1}{2}\boldsymbol{\xi}_{\mathrm{E}}^\top\bar{\mathbf{D}}\boldsymbol{\xi}_{\mathrm{E}}, && \bar{\mathbf{D}}' \in \mathbb{SYM}^+(4), && (5.20b)\\
\bar{\mathcal{S}} &= \tfrac{1}{2}\|\ddot{\mathbf{G}}_{\mathrm{E}}^\top\|_{\bar{\mathbf{M}}'}^2 = \tfrac{1}{2}\|(\operatorname{wed}(\dot{\boldsymbol{\xi}}_{\mathrm{E}}) + \operatorname{wed}(\boldsymbol{\xi}_{\mathrm{E}})^2)^\top\|_{\bar{\mathbf{M}}'}^2, && \bar{\mathbf{M}}' \in \mathbb{SYM}^+(4), && (5.20c)\\
\bar{\mathcal{T}} &= \tfrac{1}{2}\|\dot{\mathbf{G}}_{\mathrm{E}}^\top\|_{\bar{\mathbf{M}}'}^2 = \tfrac{1}{2}\|\operatorname{wed}(\boldsymbol{\xi}_{\mathrm{E}})^\top\|_{\bar{\mathbf{M}}'}^2 = \tfrac{1}{2}\boldsymbol{\xi}_{\mathrm{E}}^\top\bar{\mathbf{M}}\boldsymbol{\xi}_{\mathrm{E}} && && (5.20d)
\end{aligned}
$$

with usual substitution $\bar{\mathbf{M}} = \mathrm{Vee}(\bar{\mathbf{M}}')$. The closed loop forces are again defined as the derivatives of their corresponding energies. Using $\partial\boldsymbol{\xi}_{\mathrm{E}}/\partial\boldsymbol{\xi} = \mathbf{I}_6$ we have

$$
\begin{aligned}
\bar{\boldsymbol{f}}^{\mathsf{K}} &= \boldsymbol{\nabla}\bar{\mathcal{V}} = \operatorname{vee2}\big((\mathbf{I}_4 - \mathbf{G}_{\mathrm{E}}^{-1})\bar{\mathbf{K}}'\big) && (5.21a)\\
\bar{\boldsymbol{f}}^{\mathsf{D}} &= \frac{\partial\bar{\mathcal{R}}}{\partial\boldsymbol{\xi}} = \operatorname{vee2}\big(\operatorname{wed}(\boldsymbol{\xi}_{\mathrm{E}})\bar{\mathbf{D}}'\big) = \bar{\mathbf{D}}\boldsymbol{\xi}_{\mathrm{E}} && (5.21b)\\
\bar{\boldsymbol{f}}^{\mathsf{M}} &= \frac{\partial\bar{\mathcal{S}}}{\partial\dot{\boldsymbol{\xi}}} = \operatorname{vee2}\big(\big(\operatorname{wed}(\dot{\boldsymbol{\xi}}_{\mathrm{E}}) + \operatorname{wed}(\boldsymbol{\xi}_{\mathrm{E}})^2\big)\bar{\mathbf{M}}'\big) = \bar{\mathbf{M}}\dot{\boldsymbol{\xi}}_{\mathrm{E}} - \mathrm{ad}_{\boldsymbol{\xi}_{\mathrm{E}}}^\top\bar{\mathbf{M}}\boldsymbol{\xi}_{\mathrm{E}} && (5.21c)
\end{aligned}
$$

A crucial result of this approach is that the resulting closed loop equations can be written as *autonomous* equations for the configuration $\mathbf{G}_{\mathrm{E}}$ and velocity error $\boldsymbol{\xi}_{\mathrm{E}}$ as

$$
\dot{\mathbf{G}}_{\mathrm{E}} = \mathbf{G}\operatorname{wed}(\boldsymbol{\xi}_{\mathrm{E}}), \qquad \bar{\mathbf{M}}\dot{\boldsymbol{\xi}}_{\mathrm{E}} - \mathrm{ad}_{\boldsymbol{\xi}_{\mathrm{E}}}^\top\bar{\mathbf{M}}\boldsymbol{\xi}_{\mathrm{E}} + \bar{\mathbf{D}}\boldsymbol{\xi}_{\mathrm{E}} + \operatorname{vee2}\big((\mathbf{I}_4 - \mathbf{G}_{\mathrm{E}}^{-1})\bar{\mathbf{K}}'\big) = \mathbf{0}. \tag{5.22}
$$

A quite similar (though restricted to $\boldsymbol{s} = \boldsymbol{l} = \boldsymbol{h} = \mathbf{0}$) closed loop for the free rigid body is proposed in [Koditschek, 1989], though motivated from a Lie group point of view.

Total energy. The change of the total energy $\bar{\mathcal{W}} = \bar{\mathcal{T}} + \bar{\mathcal{V}}$ along the solutions of the closed loop (5.22) is

$$
\dot{\bar{\mathcal{W}}} = \boldsymbol{\xi}_{\mathrm{E}}^\top\big(\bar{\mathbf{M}}\dot{\boldsymbol{\xi}}_{\mathrm{E}} + \operatorname{vee2}\big((\mathbf{I}_4 - \mathbf{G}_{\mathrm{E}}^{-1})\bar{\mathbf{K}}'\big)\big) = \underbrace{\boldsymbol{\xi}_{\mathrm{E}}^\top\mathrm{ad}_{\boldsymbol{\xi}_{\mathrm{E}}}^\top}_{=0}\bar{\mathbf{M}}\boldsymbol{\xi}_{\mathrm{E}} - \boldsymbol{\xi}_{\mathrm{E}}^\top\bar{\mathbf{D}}\boldsymbol{\xi}_{\mathrm{E}} = -2\bar{\mathcal{R}}. \tag{5.23}
$$

Note that $\bar{\mathbf{K}}', \bar{\mathbf{D}}', \bar{\mathbf{M}}' \in \mathbb{SYM}^+(4)$ imply the positive definiteness of the total energy $\bar{\mathcal{W}}$ and the dissipation function $\bar{\mathcal{R}}$. Using this with the techniques from [Koditschek, 1989], one can show that "almost" all solutions of (5.22) converge to $\mathbf{G}_{\mathrm{E}} = \mathbf{I}_4$ and $\boldsymbol{\xi}_{\mathrm{E}} = \mathbf{0}$. The remaining solutions are the constant ($\boldsymbol{\xi}_{\mathrm{E}} = \mathbf{0}$) configurations $\mathbf{G}_{\mathrm{E}} \neq \mathbf{I}_4$ which are critical points of the potential $\bar{\mathcal{V}}$, see subsection 4.1.5. Roughly speaking, the configuration in which the body is 180° rotated to its reference.

5.2.2 Rigid body systems

For a rigid body system, let the body configurations ${}^a_b\mathbf{G} = {}^a_b\mathbf{G}(\boldsymbol{x})$ and the body velocities ${}^a_b\boldsymbol{\xi} = {}^a_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}$ be parameterized by the configuration $\boldsymbol{x}$ and velocity coordinates $\boldsymbol{\xi}$. So the body configuration errors ${}^a_b\mathbf{G}_\mathrm{E}$ and body velocity errors ${}^a_b\boldsymbol{\xi}_\mathrm{E}$ may be expressed as

$$
\begin{aligned}
{}^a_b\mathbf{G}_\mathrm{E}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) &= {}^a_b\mathbf{G}^{-1}(\boldsymbol{x}_\mathsf{R})\,{}^a_b\mathbf{G}(\boldsymbol{x}), && (5.24a)\\
{}^a_b\boldsymbol{\xi}_\mathrm{E}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_\mathsf{R}, \boldsymbol{\xi}_\mathsf{R}) &= {}^a_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi} - \mathrm{Ad}^{-1}_{{}^a_b\mathbf{G}_\mathrm{E}(\boldsymbol{x},\boldsymbol{x}_\mathsf{R})}\,{}^a_b\boldsymbol{J}_\mathsf{R}(\boldsymbol{x}_\mathsf{R})\boldsymbol{\xi}_\mathsf{R}. && (5.24b)
\end{aligned}
$$

As done in subsection 5.1.3, the system energies are simply the sum over the energies associated with the absolute and relative body configurations:

$$
\begin{aligned}
\bar{\mathcal{V}} &= \sum\nolimits_{a,b} \tfrac{1}{2}\|({}^a_b\mathbf{G}_\mathrm{E} - \mathbf{I}_4)^\top\|^2_{{}^a_b\bar{\mathbf{K}}'}, && (5.25a)\\
\bar{\mathcal{R}} &= \sum\nolimits_{a,b} \tfrac{1}{2}\|\operatorname{wed}({}^a_b\boldsymbol{\xi}_\mathrm{E})^\top\|^2_{{}^a_b\bar{\mathbf{D}}'}, && (5.25b)\\
\bar{\mathcal{S}} &= \sum\nolimits_{a,b} \tfrac{1}{2}\|(\operatorname{wed}({}^a_b\dot{\boldsymbol{\xi}}_\mathrm{E}) + \operatorname{wed}({}^a_b\boldsymbol{\xi}_\mathrm{E})^2)^\top\|^2_{{}^a_b\bar{\mathbf{M}}'}, && (5.25c)\\
\bar{\mathcal{T}} &= \sum\nolimits_{a,b} \tfrac{1}{2}\|\operatorname{wed}({}^a_b\boldsymbol{\xi}_\mathrm{E})^\top\|^2_{{}^a_b\bar{\mathbf{M}}'}. && (5.25d)
\end{aligned}
$$

Overall, the desired controlled system for the body based approach takes the form

$$
\bar{\boldsymbol{f}}^\mathsf{M} + \bar{\boldsymbol{f}}^\mathsf{D} + \bar{\boldsymbol{f}}^\mathsf{K} = \mathbf{0} \qquad (5.26a)
$$

where

$$
\begin{aligned}
\bar{\boldsymbol{f}}^\mathsf{K} &= \boldsymbol{\nabla}\bar{\mathcal{V}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\big((\mathbf{I}_4 - {}^a_b\mathbf{G}_\mathrm{E}^{-1})\,{}^a_b\bar{\mathbf{K}}'\big), && (5.26b)\\
\bar{\boldsymbol{f}}^\mathsf{D} &= \frac{\partial\bar{\mathcal{R}}}{\partial\boldsymbol{\xi}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top\,{}^a_b\bar{\mathbf{D}}\,{}^a_b\boldsymbol{\xi}_\mathrm{E} && (5.26c)\\
\bar{\boldsymbol{f}}^\mathsf{M} &= \frac{\partial\bar{\mathcal{S}}}{\partial\dot{\boldsymbol{\xi}}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top\big({}^a_b\bar{\mathbf{M}}\,{}^a_b\dot{\boldsymbol{\xi}}_\mathrm{E} - \operatorname{ad}^\top_{{}^a_b\boldsymbol{\xi}_\mathrm{E}}\,{}^a_b\bar{\mathbf{M}}\,{}^a_b\boldsymbol{\xi}_\mathrm{E}\big) && (5.26d)
\end{aligned}
$$

The change of the total energy $\bar{\mathcal{W}} = \bar{\mathcal{T}} + \bar{\mathcal{V}}$ along the solutions of (5.26) does *not* take a similar form to (5.23). Consequently there is no simple conclusion about stability.

5.3 Approach 3: Inspired by total energy

The previous approaches for the design of a closed loop also motived a total energy $\bar{\mathcal{W}}$. Unfortunately, it did, in general, not turn out to be useful for stability analysis. In this section we like to motivate yet another approach for the design of a closed loop dynamics for the tracking problem which is based on the total energy as Lyapunov function.

5.3.1 Overall structure

Total energy. Initially we drop the rigid body structure of the system and only consider the coordinates $\boldsymbol{x}, \boldsymbol{\xi}$ and their kinematic relation $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}$. Define the "total error energy" as

$$\bar{\mathcal{W}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) = \underbrace{\tfrac{1}{2}\| \overbrace{\boldsymbol{\xi} - \boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})\boldsymbol{\xi}_{\mathsf{R}}}^{\boldsymbol{\xi}_{\mathrm{E}}} \|^2_{\bar{\boldsymbol{M}}(\boldsymbol{x})}}_{\bar{\mathcal{T}}} + \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) \tag{5.27}$$

with the positive definite error potential $\bar{\mathcal{V}} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}^+$, $\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) = 0 \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{x}_{\mathsf{R}}$ and the positive definite inertia matrix $\bar{\boldsymbol{M}}(\boldsymbol{x}) \in \mathbb{SYM}^+(n)$. So far the transport map $\boldsymbol{Q} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}^{n \times n}$ may be any regular matrix with $\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}) = \mathbf{I}_n$. Combination of these requirements yields the positive definiteness of the total energy, i.e.

$$\bar{\mathcal{W}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) \geq 0, \quad \bar{\mathcal{W}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) = 0 \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{x}_{\mathsf{R}},\, \boldsymbol{\xi} = \boldsymbol{\xi}_{\mathsf{R}}. \tag{5.28}$$

Change of total energy. The time derivative of the total energy is

$$\begin{aligned} \dot{\bar{\mathcal{W}}} &= \boldsymbol{\xi}_{\mathrm{E}}^\top \big(\bar{\boldsymbol{M}}\dot{\boldsymbol{\xi}}_{\mathrm{E}} + \tfrac{1}{2}\dot{\bar{\boldsymbol{M}}}\boldsymbol{\xi}_{\mathrm{E}}\big) + \boldsymbol{\xi}^\top \boldsymbol{\nabla}\bar{\mathcal{V}} + \boldsymbol{\xi}_{\mathsf{R}}^\top \boldsymbol{\nabla}_{\mathsf{R}}\bar{\mathcal{V}}. \\ &= \boldsymbol{\xi}_{\mathrm{E}}^\top \big(\bar{\boldsymbol{M}}\dot{\boldsymbol{\xi}}_{\mathrm{E}} + \tfrac{1}{2}\dot{\bar{\boldsymbol{M}}}\boldsymbol{\xi}_{\mathrm{E}} + \boldsymbol{\nabla}\bar{\mathcal{V}}\big) + \boldsymbol{\xi}_{\mathsf{R}}^\top \big(\boldsymbol{\nabla}_{\mathsf{R}}\bar{\mathcal{V}} + \boldsymbol{Q}^\top \boldsymbol{\nabla}\bar{\mathcal{V}}\big). \end{aligned} \tag{5.29}$$

where $\boldsymbol{\nabla}_{\mathsf{R}} = \boldsymbol{A}^\top(\boldsymbol{x}_{\mathsf{R}})\frac{\partial}{\partial \boldsymbol{x}_{\mathsf{R}}}$. The second term vanishes if we require the transport map $\boldsymbol{Q}$ to fulfill

$$\boldsymbol{\nabla}_{\mathsf{R}}\bar{\mathcal{V}} = -\boldsymbol{Q}^\top \boldsymbol{\nabla}\bar{\mathcal{V}}. \tag{5.30}$$

The first term is non-positive if we set the closed loop kinetics as

$$\bar{\boldsymbol{M}}\dot{\boldsymbol{\xi}}_{\mathrm{E}} + (\tfrac{1}{2}\dot{\bar{\boldsymbol{M}}} + \bar{\boldsymbol{S}})\boldsymbol{\xi}_{\mathrm{E}} + \bar{\boldsymbol{D}}\boldsymbol{\xi}_{\mathrm{E}} + \boldsymbol{\nabla}\bar{\mathcal{V}} = \mathbf{0}. \tag{5.31}$$

with the positive definite damping matrix $\bar{\boldsymbol{D}} \in \mathbb{SYM}_0^+(n)$ and a skew symmetric matrix $\bar{\boldsymbol{S}} = -\bar{\boldsymbol{S}}^\top \in \mathbb{R}^{n \times n}$. Plugging the closed loop kinetics (5.31) and the requirement on the transport map (5.30) into the change of energy (5.29) yields

$$\dot{\bar{\mathcal{W}}} = -\boldsymbol{\xi}_{\mathrm{E}}^\top \bar{\boldsymbol{D}}\boldsymbol{\xi}_{\mathrm{E}} = -\|\boldsymbol{\xi}_{\mathrm{E}}\|^2_{\bar{\boldsymbol{D}}} = -2\bar{\mathcal{R}}. \tag{5.32}$$

Covariance of the closed loop. The skew symmetric matrix $\bar{\boldsymbol{S}}$ cancels out in the balance of energy, so is of no interest for tuning purposes. Instead it is used to ensure that the closed loop (5.31) is *covariant*, i.e. its definition is unchanged under a change of coordinates. While the stiffness force $\bar{f}_i^{\mathsf{K}} = \partial_i \bar{\mathcal{V}}$ and the dissipative force $\bar{f}_i^{\mathsf{D}} = \bar{D}_{ij}\xi_{\mathrm{E}}^j$ are tensors, the inertia force $\bar{f}_i^{\mathsf{M}} = \bar{M}_{ij}\dot{\xi}_{\mathrm{E}}^j + \frac{1}{2}\partial_k \bar{M}_{ij}\xi^k\xi_{\mathrm{E}}^j + \bar{S}_{ij}\xi_{\mathrm{E}}^j$ is not. A universal way is to derive a transformation law and put it as an additional requirement for the closed loop. The rather lengthy compututation is given in section A.3. It shoves that there are many co-vectors $\bar{\boldsymbol{f}}^{\mathsf{M}}$ such that $\dot{\bar{\mathcal{T}}} = \xi_{\mathrm{E}}^i \bar{f}_i^{\mathsf{M}}$.

However, a unique choice can be done by recognizing the kinetic energy $\bar{\mathcal{T}}$ as a Riemannian metric, respectively the inerita matrix $\bar{\boldsymbol{M}}$ as its coefficients, and using the *Levi-Civita connection*. The derivation requires more notation of differential geometry and is done in section A.3. The result takes the familiar form

$$\dot{\mathcal{T}} = \xi_{\mathrm{E}}^i \underbrace{\left(M_{ij}\dot{\xi}_{\mathrm{E}}^j + \Gamma_{ijk}\xi^k\xi_{\mathrm{E}}^j\right)}_{\bar{f}_i^{\mathsf{M}}} \tag{5.33}$$

For the coefficients $\bar{f}_i^{\mathsf{M}}$ to form a tensor we can derive the transformation rule (under a change of coordinates $\xi^i = W^i_{\hat{i}}(\boldsymbol{x})\hat{\xi}^{\hat{i}}$ see section A.3)

$$
\begin{aligned}
\hat{f}^{\mathrm{M}}_{\hat{i}} &= \hat{M}_{\hat{i}\hat{j}}\dot{\hat{\xi}}^{\hat{j}}_{\mathrm{E}} + \tfrac{1}{2}\partial_{\hat{k}}\hat{M}_{\hat{i}\hat{j}}\hat{\xi}^{\hat{k}}\hat{\xi}^{\hat{j}}_{\mathrm{E}} + \hat{S}_{\hat{i}\hat{j}}\hat{\xi}^{\hat{j}}_{\mathrm{E}} \\
&= W^i_{\hat{i}}M_{ij}W^j_{\hat{j}}\big(Z^{\hat{j}}_k\dot{\xi}^k_{\mathrm{E}} + \partial_l Z^{\hat{j}}_k\xi^l\xi^k_{\mathrm{E}}\big) \\
&+ \tfrac{1}{2}\big(W^i_{\hat{i}}W^j_{\hat{j}}W^k_{\hat{k}}\partial_k M_{ij} + (W^j_{\hat{j}}\partial_{\hat{k}}W^i_{\hat{i}} + W^i_{\hat{i}}\partial_{\hat{k}}W^j_{\hat{j}})M_{ij}\big)Z^{\hat{k}}_l\xi^l Z^{\hat{j}}_s\xi^s_{\mathrm{E}} \\
&+ \hat{S}_{\hat{i}\hat{j}}Z^{\hat{j}}_j\xi^j_{\mathrm{E}} \\
&= W^i_{\hat{i}}M_{ij}\dot{\xi}^j_{\mathrm{E}} + W^i_{\hat{i}}M_{ij}W^j_{\hat{j}}\partial_l Z^{\hat{j}}_k\xi^l\xi^k_{\mathrm{E}} \\
&+ \tfrac{1}{2}W^i_{\hat{i}}\partial_k M_{ij}\xi^k\xi^j_{\mathrm{E}} + \tfrac{1}{2}W^i_{\hat{i}}Z^{\hat{k}}_i\partial_l W^k_{\hat{k}}M_{kj}\xi^l\xi^j_{\mathrm{E}} + \tfrac{1}{2}W^i_{\hat{i}}\partial_k W^j_{\hat{j}}M_{ij}\xi^k Z^{\hat{j}}_s\xi^s_{\mathrm{E}} \\
&+ W^i_{\hat{i}}Z^{\hat{k}}_i\hat{S}_{\hat{k}\hat{j}}Z^{\hat{j}}_j\xi^j_{\mathrm{E}} \\
&= W^i_{\hat{i}}\Big(M_{ij}\dot{\xi}^j_{\mathrm{E}} + \tfrac{1}{2}\partial_k M_{ij}\xi^k\xi^j_{\mathrm{E}} \\
&\quad + \underbrace{\big(Z^{\hat{k}}_i\hat{S}_{\hat{k}\hat{j}}Z^{\hat{j}}_j + \big(\boldsymbol{M}[il]W^l_{\hat{j}}\partial_k Z^{\hat{j}}_j + \tfrac{1}{2}Z^{\hat{k}}_i\partial_k W^l_{\hat{k}}\boldsymbol{M}[lj] + \tfrac{1}{2}\partial_k W^s_{\hat{j}}M_{is}Z^{\hat{j}}_j\big)\xi^k\big)}_{S_{ij}}\xi^j_{\mathrm{E}}\Big)
\end{aligned}
\tag{5.34}
$$

So

$$\hat{\bar{S}}'_{\hat{i}\hat{j}} = W^i_{\hat{i}}W^j_{\hat{j}}\bar{S}'_{ij} + \tfrac{1}{2}\bar{M}_{ij}\big(W^i_{\hat{i}}\partial_k W^j_{\hat{j}} - W^j_{\hat{j}}\partial_k W^i_{\hat{i}}\big)\xi^k \tag{5.35}$$

$$\hat{\bar{S}}_{\hat{i}\hat{j}} = \bar{S}_{ij}W^i_{\hat{i}}W^j_{\hat{j}} + \tfrac{1}{2}\bar{M}_{ij}\big(W^i_{\hat{i}}\partial_{\hat{k}}W^j_{\hat{j}} - W^j_{\hat{j}}\partial_{\hat{k}}W^i_{\hat{i}}\big)\hat{\xi}^{\hat{k}}, \qquad \hat{i},\hat{j} = 1,\ldots,n. \tag{5.36}$$

Recall the coefficients $\bar{\Gamma}_{ijk}$ of the Levi-Civita connection for the chosen metric $\bar{\boldsymbol{M}}$:

$$\bar{\Gamma}_{ijk} = \tfrac{1}{2}\big(\partial_k\bar{M}_{ij} + \partial_j\bar{M}_{ik} - \partial_i\bar{M}_{jk} + \gamma^s_{ij}\bar{M}_{sk} + \gamma^s_{ik}\bar{M}_{sj} - \gamma^s_{jk}\bar{M}_{si}\big), \quad i,j,k = 1,\ldots,n \tag{5.37}$$

and their transformation rule

$$\hat{\bar{\Gamma}}_{\hat{i}\hat{j}\hat{k}} = \bar{\Gamma}_{ijk}W^i_{\hat{i}}W^j_{\hat{j}}W^k_{\hat{k}} + \bar{M}_{ij}W^i_{\hat{i}}\partial_{\hat{k}}W^j_{\hat{j}}, \qquad \hat{i},\hat{j},\hat{k} = 1,\ldots,n. \tag{5.38}$$

With this one definition for $\bar{S}_{ij}$ that obeys (5.48) is

$$\bar{S}_{ij} = \tfrac{1}{2}\big(\bar{\Gamma}_{ijk} - \bar{\Gamma}_{jik}\big)\xi^k, \qquad i,j = 1,\ldots,n. \tag{5.39}$$

Plugging this into (5.47) we finally obtain a covariant and energy conserving inertia force as

$$f^{\mathrm{M}}_i = \bar{M}_{ij}\dot{\xi}^j_{\mathrm{E}} + \bar{\Gamma}_{ijk}\xi^k\xi^j_{\mathrm{E}}, \qquad i = 1,\ldots,n. \tag{5.40}$$

Note that $\bar{\boldsymbol{f}}^{\mathsf{M}}$ from this definition can not be derived by Lagrange's formulation from the kinetic energy $\bar{\mathcal{T}}$. It is however completely determined by given metric coefficients $\bar{\boldsymbol{M}}$.

The definition of the inertia force $\bar{\boldsymbol{f}}^{\mathsf{M}}$ is not unique: For any antisymmetric *tensor* $S'_{ij} = -S'_{ji}$ the inertia force $\bar{f}^{\mathsf{M}}_i + S'_{ij}\xi^j_{\mathrm{E}}$ is covariant and energy conserving as well.

Define the closed loop kinematics as

$$\underbrace{\bar{M}_{ij}\dot{\xi}^j_{\mathrm{E}} + \bar{\Gamma}_{ijk}\xi^k\xi^j_{\mathrm{E}}}_{\bar{f}^{\mathsf{M}}_i} + \underbrace{\bar{D}_{ij}\xi^j_{\mathrm{E}}}_{\bar{f}^{\mathsf{D}}_i} + \underbrace{\partial_i\bar{\mathcal{V}}}_{\bar{f}^{\mathsf{K}}_i} = 0, \qquad i = 1,\ldots,n. \tag{5.41}$$

Error potential and velocity. The first thing we need is a quantification of how far the system configuration $\boldsymbol{x}(t) \in \mathbb{X}$ is away from its desired configuration $\boldsymbol{x}_{\mathsf{R}}(t) \in \mathbb{X}$: For this consider the *error potential* as a smooth, positive definite function $\bar{\mathcal{V}} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$, i.e.

$$\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) \geq 0, \qquad \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) = 0 \ \Leftrightarrow \ \boldsymbol{x} = \boldsymbol{x}_{\mathsf{R}}. \tag{5.42}$$

For the following we require the existence of a *transport map* as introduced in [Bullo and Murray, 1999]. That is a smooth function $\boldsymbol{Q} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}^{n\times n}$ such that

$$\underbrace{A^\alpha_i(\boldsymbol{x}_{\mathsf{R}})\frac{\partial\bar{\mathcal{V}}}{\partial x^\alpha_{\mathsf{R}}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})}_{\partial^{\mathsf{R}}_i\bar{\mathcal{V}}(\boldsymbol{x},\boldsymbol{x}_{\mathsf{R}})} = -Q^j_i(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})\underbrace{A^\alpha_j(\boldsymbol{x})\frac{\partial\bar{\mathcal{V}}}{\partial x^\alpha}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})}_{\partial_j\bar{\mathcal{V}}(\boldsymbol{x},\boldsymbol{x}_{\mathsf{R}})}, \qquad i = 1,\ldots,n. \tag{5.43}$$

From this we obtain the relation

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{V}} = \xi^i\partial_j\bar{\mathcal{V}} + \xi^i_{\mathsf{R}}\partial^{\mathsf{R}}_j\bar{\mathcal{V}} = \underbrace{(\xi^i - Q^i_j\xi^j_{\mathsf{R}})}_{\xi^i_{\mathrm{E}}}\underbrace{\partial_i\bar{\mathcal{V}}}_{\bar{f}^{\mathsf{K}}_i}. \tag{5.44}$$

Essentially the transport map gives a mean to compare the actual velocity $\boldsymbol{\xi}$ and the reference velocity $\boldsymbol{\xi}_{\mathsf{R}}$ which live on different tangent spaces, thus giving a reasonable definition for the *error velocity* $\boldsymbol{\xi}_{\mathrm{E}}$. Note that $\boldsymbol{Q}$ are the coefficients of a tensor, (see (A.87) and (A.104)), so the existence of a transport map is independent of the particular choice of coordinates.

Kinetic error energy and inertia force. Having a notion for the error velocity, we may define the kinetic error energy

$$\bar{\mathcal{T}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) = \tfrac{1}{2}\bar{M}_{ij}(\boldsymbol{x})\underbrace{(\xi^i - Q^i_k(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})\xi^k_{\mathsf{R}})}_{\xi^i_{\mathrm{E}}}\underbrace{(\xi^j - Q^j_l(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})\xi^l_{\mathsf{R}})}_{\xi^j_{\mathrm{E}}} \geq 0 \tag{5.45}$$

with a symmetric and positive definite inertia matrix $\bar{\boldsymbol{M}}(\boldsymbol{x}) \in \mathbb{SYM}^+(n)$. We associate the inertia force $\bar{\boldsymbol{f}}^{\mathsf{M}}$ with the kinetic energy through

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{T}} = \xi^i_{\mathrm{E}}\underbrace{\left(\bar{M}_{ij}\dot{\xi}^j_{\mathrm{E}} + \tfrac{1}{2}\partial_k\bar{M}_{ij}\xi^k\xi^j_{\mathrm{E}}\right)}_{\neq\bar{f}^{\mathsf{M}}_i} \overset{!}{=} \xi^i_{\mathrm{E}}\bar{f}^{\mathsf{M}}_i. \tag{5.46}$$

It is crucial to note that the marked terms are *not* covariant, so do not form a tensor. To resolve this we add the antisymmeteric coefficients $\bar{S}_{ij} = -\bar{S}_{ji}$ and define the inertia force as

$$\bar{f}^{\mathsf{M}}_i = \bar{M}_{ij}\dot{\xi}^j_{\mathrm{E}} + \tfrac{1}{2}\partial_k\bar{M}_{ij}\xi^k\xi^j_{\mathrm{E}} + \bar{S}_{ij}\xi^j_{\mathrm{E}}, \qquad i = 1,\ldots,n. \tag{5.47}$$

For the coefficients $\bar{f}_i^{\mathsf{M}}$ to form a tensor we can derive the transformation rule (under a change of coordinates $\xi^i = W_{\hat{i}}^i(\boldsymbol{x})\hat{\xi}^{\hat{i}}$ see section A.3)

$$\hat{\bar{S}}_{\hat{i}\hat{j}} = \bar{S}_{ij} W_{\hat{i}}^i W_{\hat{j}}^j + \tfrac{1}{2}\bar{M}_{ij}\big(W_{\hat{i}}^i \partial_{\hat{k}} W_{\hat{j}}^j - W_{\hat{j}}^j \partial_{\hat{k}} W_{\hat{i}}^i\big)\hat{\xi}^{\hat{k}}, \qquad \hat{i}, \hat{j} = 1, \dots, n. \tag{5.48}$$

Recall the coefficients $\bar{\Gamma}_{ijk}$ of the Levi-Civita connection for the chosen metric $\bar{\boldsymbol{M}}$:

$$\bar{\Gamma}_{ijk} = \tfrac{1}{2}\big(\partial_k \bar{M}_{ij} + \partial_j \bar{M}_{ik} - \partial_i \bar{M}_{jk} + \gamma_{ij}^s \bar{M}_{sk} + \gamma_{ik}^s \bar{M}_{sj} - \gamma_{jk}^s \bar{M}_{si}\big), \quad i, j, k = 1, \dots, n \tag{5.49}$$

and their transformation rule

$$\hat{\bar{\Gamma}}_{\hat{i}\hat{j}\hat{k}} = \bar{\Gamma}_{ijk} W_{\hat{i}}^i W_{\hat{j}}^j W_{\hat{k}}^k + \bar{M}_{ij} W_{\hat{i}}^i \partial_{\hat{k}} W_{\hat{j}}^j, \qquad \hat{i}, \hat{j}, \hat{k} = 1, \dots, n. \tag{5.50}$$

With this one definition for $\bar{S}_{ij}$ that obeys (5.48) is

$$\bar{S}_{ij} = \tfrac{1}{2}\big(\bar{\Gamma}_{ijk} - \bar{\Gamma}_{jik}\big)\xi^k, \qquad i, j = 1, \dots, n. \tag{5.51}$$

Plugging this into (5.47) we finally obtain a covariant and energy conserving inertia force as

$$f_i^{\mathrm{M}} = \bar{M}_{ij}\dot{\xi}_{\mathrm{E}}^j + \bar{\Gamma}_{ijk}\xi^k \xi_{\mathrm{E}}^j, \qquad i = 1, \dots, n. \tag{5.52}$$

Note that $\bar{\boldsymbol{f}}^{\mathsf{M}}$ from this definition can not be derived by Lagrange's formulation from the kinetic energy $\bar{\mathcal{T}}$. It is however completely determined by given metric coefficients $\bar{\boldsymbol{M}}$.

The definition of the inertia force $\bar{\boldsymbol{f}}^{\mathsf{M}}$ is not unique: For any antisymmetric *tensor* $S'_{ij} = -S'_{ji}$ the inertia force $\bar{f}_i^{\mathsf{M}} + S'_{ij}\xi_{\mathrm{E}}^j$ is covariant and energy conserving as well.

Total error energy and dissipation. Define the *total error energy* $\bar{\mathcal{W}}$ as

$$\bar{\mathcal{W}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) = \bar{\mathcal{T}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) + \bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) \geq 0. \tag{5.53}$$

The definitions of the error potential (5.42) and the kinetic error energy (5.45) imply

$$\bar{\mathcal{W}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}) = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{x}_{\mathsf{R}},\ \boldsymbol{\xi} = \boldsymbol{\xi}_{\mathsf{R}}. \tag{5.54}$$

Next we define a dissipative force $f_i^{\mathrm{D}} = \bar{D}_{ij}\xi_{\mathrm{E}}^j$, with the symmetric positive semi-definite dissipation matrix $\bar{\boldsymbol{D}}(\boldsymbol{x}) \in \mathbb{SYM}_0^+(n)$ and associate it with the change of the total energy by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{W}} = \xi_{\mathrm{E}}^i\big(\bar{f}_i^{\mathsf{M}} + \bar{f}_i^{\mathsf{K}}\big) \stackrel{!}{=} -\xi_{\mathrm{E}}^i \bar{f}_i^{\mathsf{D}}. \tag{5.55}$$

As above this implies $\bar{f}_i^{\mathsf{M}} + \bar{f}_i^{\mathsf{K}} + \bar{f}_i^{\mathsf{D}} = S''_{ij}\xi_{\mathrm{E}}^j$ where $\boldsymbol{S}''$ can be the coefficients of any skew symmetric tensor. For simplicity we chose $\boldsymbol{S}'' = \boldsymbol{0}$ and formulate the resulting controlled kinetics:

The desired error kinetics for the energy based approach have the form

$$\underbrace{\bar{M}_{ij}\dot{\xi}_{\mathrm{E}}^{j} + \bar{\Gamma}_{ijk}\xi^{k}\xi_{\mathrm{E}}^{j}}_{\bar{f}_i^{\mathsf{M}}} + \underbrace{\bar{D}_{ij}\xi_{\mathrm{E}}^{j}}_{\bar{f}_i^{\mathsf{D}}} + \underbrace{\partial_i\bar{\mathcal{V}}}_{\bar{f}_i^{\mathsf{K}}} = 0, \qquad i = 1, \ldots, n. \tag{5.56}$$

Since $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{W}}$ is only negative semidefinite, we can only conclude stability but not attractiveness. One can pursue the prove of attractiveness by adding a cross term as done in [Bullo and Murray, 1999].

5.3.2 Special cases

Euclidean space. The existing literature on control of mechanical systems uses almost exclusively minimal generalized coordinates $\boldsymbol{q} \in \mathbb{R}^n$ and the velocity coordinates $\dot{\boldsymbol{q}}$. Then the model can we written as

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} = \boldsymbol{f}^{\mathrm{A}} \tag{5.57}$$

where $C_{ij} = \Gamma_{ijk}\xi^k$ and $\boldsymbol{f}^{\mathrm{A}}$ collects the remaining forces. For a fully actuated system there exists an input transformation such that $\boldsymbol{f}^{\mathrm{A}}$ can be regarded as virtual inputs.

On the Euclidean space $\mathbb{R}^n$ it is reasonable to introduce error coordinates $\boldsymbol{q}_{\mathrm{E}} = \boldsymbol{q} - \boldsymbol{q}_{\mathsf{R}}$ and to use a quadratic error potential

$$\bar{\mathcal{V}} = \tfrac{1}{2}\boldsymbol{q}_{\mathrm{E}}^{\top}\bar{\boldsymbol{K}}\boldsymbol{q}_{\mathrm{E}}, \qquad \bar{\boldsymbol{K}} \in \mathbb{SYM}^{+}(n). \tag{5.58}$$

This error potential obviously has the transport map $\boldsymbol{Q} = \mathbf{I}_n$ and the resulting error velocity $\boldsymbol{\xi}_{\mathrm{E}} = \dot{\boldsymbol{q}}_{\mathrm{E}} = \dot{\boldsymbol{q}} - \dot{\boldsymbol{q}}_{\mathsf{R}}$. Furthermore it is reasonable to choose a constant dissipation matrix $\bar{\boldsymbol{D}} \in \mathbb{SYM}^{+}(n)$.

Joint PD-Control. Choosing the desired inertia identical to the model inertia $\bar{\boldsymbol{M}} = \boldsymbol{M}$, which also implies $\bar{\boldsymbol{C}} = \boldsymbol{C}$, yields the closed loop kinetics

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\mathrm{E}} + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_{\mathrm{E}} + \bar{\boldsymbol{D}}\dot{\boldsymbol{q}}_{\mathrm{E}} + \bar{\boldsymbol{K}}\boldsymbol{q}_{\mathrm{E}} = \mathbf{0}. \tag{5.59}$$

The resulting control law is

$$\boldsymbol{f}^{\mathrm{A}} = \boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\mathsf{R}} + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_{\mathsf{R}} - \bar{\boldsymbol{D}}\dot{\boldsymbol{q}}_{\mathrm{E}} - \bar{\boldsymbol{K}}\boldsymbol{q}_{\mathrm{E}}. \tag{5.60}$$

This approach is commonly called *joint proportional derivative controller* [Slotine and Li, 1991, sec. 9.1.1] or *augmented PD control law* [Murray et al., 1994, sec. 4.5.3], [Spong et al., 2006, sec. 8.2].

Computed torque. Choosing the desired inertia as $\bar{\boldsymbol{M}} = \mathbf{I}_n$, which implies $\bar{\boldsymbol{C}} = \mathbf{0}$ leads to the closed loop kinetics

$$\ddot{\boldsymbol{q}}_{\mathrm{E}} + \bar{\boldsymbol{D}}\dot{\boldsymbol{q}}_{\mathrm{E}} + \bar{\boldsymbol{K}}\boldsymbol{q}_{\mathrm{E}} = \mathbf{0}. \tag{5.61}$$

The resulting control law is

$$\boldsymbol{f}^{\mathrm{A}} = \boldsymbol{M}(\boldsymbol{q})\big(\ddot{\boldsymbol{q}}_{\mathsf{R}} - \bar{\boldsymbol{D}}\dot{\boldsymbol{q}}_{\mathrm{E}} - \bar{\boldsymbol{K}}\boldsymbol{q}_{\mathrm{E}}\big) + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}}. \tag{5.62}$$

This approach is commonly called *computed torque* [Murray et al., 1994, sec. 4.5.2], [Slotine and Li, 1991, sec. 9.1.2] or *inverse dynamics control* [Spong et al., 2006, sec. 8.3].

These well established approaches are contained within the derived framework (5.56). However, as argued in the introduction of this chapter, the use of the Euclidean metric (5.58) only makes sense if the configuration space is indeed an Euclidean space. Application of this approach to e.g. the rigid body orientation would lead to quite awkward motion.

5.3.3 Free rigid body

Consider a single free rigid body as extensively discussed in section 4.1. We use the position $\boldsymbol{r}$ and orientation $\boldsymbol{R}$ combined in the matrix $\mathbf{G} \in \mathbb{SE}(3)$ as configuration coordinates and the linear velocity $\boldsymbol{v} = \boldsymbol{R}^\top\dot{\boldsymbol{r}}$ and angular velocity $\boldsymbol{\omega} = \mathrm{vee}(\boldsymbol{R}^\top\dot{\boldsymbol{R}})$ combined in $[\boldsymbol{v}^\top, \boldsymbol{\omega}^\top]^\top = \boldsymbol{\xi} = \mathrm{vee}(\mathbf{G}^{-1}\dot{\mathbf{G}})$ as velocity coordinates.

Potential and transport map. As for the previous approaches, a reasonable choice (motivated from linear springs in subsection 4.1.5) for the potential energy for a rigid body is

$$\bar{\mathcal{V}} = \tfrac{1}{2}\|(\mathbf{G} - \mathbf{G}_{\mathsf{R}})^\top\|^2_{\bar{\mathbf{K}}'}, \qquad \bar{\mathbf{K}}' \in \mathbb{SYM}^+(4). \tag{5.63}$$

The time derivative of the potential is

$$\begin{aligned}\dot{\bar{\mathcal{V}}} &= \mathrm{tr}\left((\mathbf{G} - \mathbf{G}_{\mathsf{R}})\bar{\mathbf{K}}'(\mathbf{G}\,\mathrm{wed}(\boldsymbol{\xi}) - \mathbf{G}_{\mathsf{R}}\,\mathrm{wed}(\boldsymbol{\xi}_{\mathsf{R}}))^\top\right)\\ &= \mathrm{tr}\left((\mathbf{I}_4 - \mathbf{G}^{-1}\mathbf{G}_{\mathsf{R}})\bar{\mathbf{K}}'\,\mathrm{wed}(\boldsymbol{\xi} - \mathrm{Ad}_{\mathbf{G}^{-1}\mathbf{G}_{\mathsf{R}}}\boldsymbol{\xi}_{\mathsf{R}})^\top\right)\\ &= \underbrace{(\boldsymbol{\xi} - \mathrm{Ad}_{\mathbf{G}^{-1}\mathbf{G}_{\mathsf{R}}}\boldsymbol{\xi}_{\mathsf{R}})}_{\boldsymbol{\xi}_{\mathrm{E}}}{}^\top\,\underbrace{\mathrm{vee2}\left((\mathbf{I}_4 - \mathbf{G}^{-1}\mathbf{G}_{\mathsf{R}})\bar{\mathbf{K}}'\right)}_{\boldsymbol{\nabla}\bar{\mathcal{V}}}.\end{aligned} \tag{5.64}$$

Recalling the identity $\boldsymbol{\nabla}\bar{\mathcal{V}} = \partial\dot{\bar{\mathcal{V}}}/\partial\boldsymbol{\xi}$, it is evident that $\boldsymbol{Q} = \mathrm{Ad}_{\mathbf{G}^{-1}\mathbf{G}_{\mathsf{R}}}$ is a transport map for this potential. The potential $\bar{\mathcal{V}}$ and the resulting force $\bar{\boldsymbol{f}}^{\mathsf{K}} = \boldsymbol{\nabla}\bar{\mathcal{V}}$ coincides with the ones given for the previous approaches, already explicitly stated in (5.14a).

Damping and Inertia. As for the previous approaches, let the damping and inertia matrices have the structure $\bar{\boldsymbol{D}} = \mathrm{Vee}\,\bar{\mathbf{D}}'$ and $\bar{\boldsymbol{M}} = \mathrm{Vee}\,\bar{\mathbf{M}}'$ with $\bar{\mathbf{D}}', \bar{\mathbf{M}}' \in \mathbb{SYM}^+(4)$. The entries may be interpreted as controlled total mass $\bar{m}$, controlled center of mass $\bar{\boldsymbol{s}}$ etc.

as stated in (5.12). Plugging these matrices into (5.56) we find the controlled dissipation force $\bar{\boldsymbol{f}}^{\mathsf{D}}$ and controlled inertial force $\bar{\boldsymbol{f}}^{\mathsf{M}}$ as

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \underbrace{\begin{bmatrix} \bar{d}\,\mathbf{I}_3 & \bar{d}(\operatorname{wed}\bar{\boldsymbol{l}})^\top \\ \bar{d}\operatorname{wed}\bar{\boldsymbol{l}} & \bar{\boldsymbol{\Upsilon}} \end{bmatrix}}_{\bar{\boldsymbol{D}}} \underbrace{\begin{bmatrix} \boldsymbol{v}_{\mathrm{E}} \\ \boldsymbol{\omega}_{\mathrm{E}} \end{bmatrix}}_{\boldsymbol{\xi}_{\mathrm{E}}}, \tag{5.65}$$

$$\bar{\boldsymbol{f}}^{\mathsf{M}} = \underbrace{\begin{bmatrix} \bar{m}\mathbf{I}_3 & \bar{m}(\operatorname{wed}\bar{\boldsymbol{s}})^\top \\ \bar{m}\operatorname{wed}\bar{\boldsymbol{s}} & \bar{\boldsymbol{\Theta}} \end{bmatrix}}_{\bar{\boldsymbol{M}}} \underbrace{\begin{bmatrix} \dot{\boldsymbol{v}}_{\mathrm{E}} \\ \dot{\boldsymbol{\omega}}_{\mathrm{E}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}_{\mathrm{E}}} + \underbrace{\begin{bmatrix} \bar{m}\operatorname{wed}\boldsymbol{\omega} & -\bar{m}\operatorname{wed}\boldsymbol{\omega}\operatorname{wed}\bar{\boldsymbol{s}} \\ \bar{m}\operatorname{wed}\bar{\boldsymbol{s}}\operatorname{wed}\boldsymbol{\omega} & \operatorname{wed}(\operatorname{Wed}\bar{\boldsymbol{\Theta}}\,\boldsymbol{\omega}) \end{bmatrix}}_{\bar{\boldsymbol{C}}(\boldsymbol{\xi})=-\bar{\boldsymbol{C}}^\top(\boldsymbol{\xi})} \underbrace{\begin{bmatrix} \boldsymbol{v}_{\mathrm{E}} \\ \boldsymbol{\omega}_{\mathrm{E}} \end{bmatrix}}_{\boldsymbol{\xi}_{\mathrm{E}}}. \tag{5.66}$$

This result is very similar to the previous body based approach (5.22), only differing by replacing $\bar{\boldsymbol{C}}(\boldsymbol{\xi})$ with $\bar{\boldsymbol{C}}(\boldsymbol{\xi}_{\mathrm{E}})$. Consequently with this approach the closed loop dynamics are not autonomous.

An alternative transport map. It should be noted that the transport map given in (5.64) is not unique. One can check by direct calculation that the following matrix also fulfills the required relation (5.43):

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{R}^\top\boldsymbol{R}_{\mathsf{R}} & \operatorname{wed}\bar{\boldsymbol{h}}\,\boldsymbol{R}^\top\boldsymbol{R}_{\mathsf{R}} - \boldsymbol{R}^\top\boldsymbol{R}_{\mathsf{R}}\operatorname{wed}\bar{\boldsymbol{h}} \\ \mathbf{0} & \boldsymbol{R}^\top\boldsymbol{R}_{\mathsf{R}} \end{bmatrix}. \tag{5.67}$$

5.3.4 Rigid body systems

As for the previous approaches we use the rigid body structure, i.e. the configurations ${}^a_b\mathbf{G}(\boldsymbol{x})$ and velocities ${}^a_b\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{\xi}) = {}^a_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}$, as inspiration for controlled kinetics. Assigning stiffness, damping and inertia ${}^a_b\bar{\mathbf{K}}', {}^a_b\bar{\mathbf{D}}', {}^a_b\bar{\mathbf{M}}' \in \mathbb{SYM}^+_0(4)$ to each absolute and relative configuration leads to the following potential energy $\bar{\mathcal{V}}$, damping $\bar{\boldsymbol{D}}$ and inertia matrix $\bar{\boldsymbol{M}}$:

$$\bar{\mathcal{V}}(\boldsymbol{x},\boldsymbol{x}_{\mathsf{R}}) = \sum\nolimits_{a,b} \tfrac{1}{2}\|({}^a_b\mathbf{G}(\boldsymbol{x}) - {}^a_b\mathbf{G}(\boldsymbol{x}_{\mathsf{R}}))^\top\|^2_{{}^a_b\bar{\mathbf{K}}'}, \qquad {}^a_b\bar{\mathbf{K}}' \in \mathbb{SYM}^+_0(4) \tag{5.68a}$$

$$\bar{\boldsymbol{D}}(\boldsymbol{x}) = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top(\boldsymbol{x})\ \operatorname{Vee}({}^a_b\bar{\mathbf{D}}')\,{}^a_b\boldsymbol{J}(\boldsymbol{x}), \qquad {}^a_b\bar{\mathbf{D}}' \in \mathbb{SYM}^+_0(4) \tag{5.68b}$$

$$\bar{\boldsymbol{M}}(\boldsymbol{x}) = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top(\boldsymbol{x})\ \operatorname{Vee}({}^a_b\bar{\mathbf{M}}')\,{}^a_b\boldsymbol{J}(\boldsymbol{x}), \qquad {}^a_b\bar{\mathbf{M}}' \in \mathbb{SYM}^+_0(4) \tag{5.68c}$$

Note that the body matrices do not have to be positive definite, only the resulting system matrices and the potential have to be positive definite to ensure stability.

Transport map. The potential $\bar{\mathcal{V}}$ and the resulting force $\boldsymbol{\nabla}\bar{\mathcal{V}}$ are identical to the previous approaches, see e.g. (5.18b). For this energy based approach we require the existence of a transport map. The condition (5.43) for the transport map $\boldsymbol{Q}$ with the given potential (5.68a) is equivalent to

$$\sum\nolimits_{a,b} \left({}^{a}_{b}\boldsymbol{J}\boldsymbol{Q} - \mathrm{Ad}_{{}^{a}_{b}\mathsf{G}_{\mathrm{E}}^{-1}}\,{}^{a}_{b}\boldsymbol{J}_{\mathsf{R}}\right)^{\top} \mathrm{vee2}\left((\mathbf{I}_4 - {}^{a}_{a}\mathsf{G}_{\mathrm{E}}^{-1})\,{}^{a}_{b}\bar{\mathsf{K}}'\right) = \mathbf{0}. \tag{5.69}$$

with the shorthand notation ${}^{a}_{b}\mathsf{G}_{\mathrm{E}} = {}^{a}_{b}\mathsf{G}^{-1}(\boldsymbol{x}_{\mathsf{R}})\,{}^{a}_{b}\mathsf{G}_{\mathrm{E}}(\boldsymbol{x})$ and ${}^{a}_{b}\boldsymbol{J}_{\mathsf{R}} = {}^{a}_{b}\boldsymbol{J}(\boldsymbol{x}_{\mathsf{R}})$. There is no general solution for this, the transport map $\boldsymbol{Q}$ has to be computed for each example individually.

5.4 Constant reference and Linearization

Constant reference. For a constant reference configuration $\boldsymbol{x}_\mathsf{R} = const. \Rightarrow \boldsymbol{\xi}_\mathsf{R}, \dot{\boldsymbol{\xi}}_\mathsf{R} = \mathbf{0}$, the three control templates lead identical system

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}, \qquad \bar{\boldsymbol{M}}(\boldsymbol{x})\dot{\boldsymbol{\xi}} + \bar{\boldsymbol{c}}(\boldsymbol{x},\boldsymbol{\xi}) + \bar{\boldsymbol{D}}(\boldsymbol{x})\boldsymbol{\xi} + \boldsymbol{\nabla}\bar{\mathcal{V}}(\boldsymbol{x},\boldsymbol{x}_\mathsf{R}) = \mathbf{0} \tag{5.70a}$$

where

$$\bar{\boldsymbol{M}}(\boldsymbol{x}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x})\right)^\top {}^a_b\bar{\mathbf{M}}\,{}^a_b\boldsymbol{J}(\boldsymbol{x}), \tag{5.70b}$$

$$\bar{\boldsymbol{c}}(\boldsymbol{x},\boldsymbol{\xi}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x})\right)^\top \left({}^a_b\mathbf{M}\,{}^a_b\dot{\boldsymbol{J}}(\boldsymbol{x},\boldsymbol{\xi}) - \mathrm{ad}^\top_{{}^a_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}}\,{}^a_b\mathbf{M}\,{}^a_b\boldsymbol{J}(\boldsymbol{x})\right)\boldsymbol{\xi} \tag{5.70c}$$

$$\bar{\boldsymbol{D}}(\boldsymbol{x}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x})\right)^\top {}^a_b\bar{\mathbf{D}}\,{}^a_b\boldsymbol{J}(\boldsymbol{x}), \tag{5.70d}$$

$$\boldsymbol{\nabla}\bar{\mathcal{V}}(\boldsymbol{x},\boldsymbol{x}_\mathsf{R}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x})\right)^\top \mathrm{vee2}\left((\mathbf{I}_4 - {}^a_b\mathbf{G}_\mathrm{E}^{-1}(\boldsymbol{x},\boldsymbol{x}_\mathsf{R}))\,{}^a_b\bar{\mathbf{K}}'\right) \tag{5.70e}$$

The total energy $\bar{\mathcal{W}} = \frac{1}{2}\boldsymbol{\xi}^\top\bar{\boldsymbol{M}}\boldsymbol{\xi} + \bar{\mathcal{V}}$ is a Lyapunov function for this system if the system inertia matrix $\bar{\boldsymbol{M}}$ and the potential energy $\bar{\mathcal{V}}$ are positive definite, and the system dissipation matrix $\bar{\boldsymbol{D}}$ is positive semi-definite.

Linearization. Assuming that the configuration of the system is close to its reference, i.e. $\boldsymbol{x} \approx \boldsymbol{x}_\mathsf{R}$. The first order approximation (see subsection 3.3.5) of (5.70) with $\boldsymbol{\varepsilon} = \boldsymbol{A}^+(\boldsymbol{x}_\mathsf{R})(\boldsymbol{x} - \boldsymbol{x}_\mathsf{R})$ is

$$\bar{\boldsymbol{M}}_0\ddot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{D}}_0\dot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{K}}_0\boldsymbol{\varepsilon} = \mathbf{0} \tag{5.71a}$$

where

$$\bar{\boldsymbol{M}}_0 = \bar{\boldsymbol{M}}(\boldsymbol{x}_\mathsf{R}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R})\right)^\top {}^a_b\bar{\mathbf{M}}\;{}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R}), \tag{5.71b}$$

$$\bar{\boldsymbol{D}}_0 = \bar{\boldsymbol{D}}(\boldsymbol{x}_\mathsf{R}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R})\right)^\top {}^a_b\bar{\mathbf{D}}\;{}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R}), \tag{5.71c}$$

$$\bar{\boldsymbol{K}}_0 = \boldsymbol{\nabla}^2\bar{\mathcal{V}}(\boldsymbol{x}_\mathsf{R},\boldsymbol{x}_\mathsf{R}) = \sum\nolimits_{a,b} \left({}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R})\right)^\top {}^a_b\bar{\mathbf{K}}\;{}^a_b\boldsymbol{J}(\boldsymbol{x}_\mathsf{R}) \tag{5.71d}$$

5.5 Underactuated systems

The first three sections of this chapter motivated different desired closed loop dynamics which share the structure

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}, \quad \bar{\boldsymbol{M}}(\boldsymbol{x})\dot{\boldsymbol{\xi}} + \bar{\boldsymbol{b}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}}, \dot{\boldsymbol{\xi}}_{\mathsf{R}}) = \boldsymbol{0}. \tag{5.72}$$

The system model has still the form of (5.1):

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}, \quad \boldsymbol{M}(\boldsymbol{x})\dot{\boldsymbol{\xi}} + \boldsymbol{b}(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u}. \tag{5.73}$$

For a fully actuated system, i.e. $\operatorname{rank} \boldsymbol{B} = n$, the combination of (5.72) and (5.73) can be solved for the system input $\boldsymbol{u}$ yielding the actual control law. For an underactuated system, i.e. $\operatorname{rank} \boldsymbol{B} = p < n$, this is not possible.

5.5.1 Control law through static optimization

If the desired closed loop dynamics (5.72) cannot be achieved exactly, the next best thing is to get "as close as possible" while still obeying the model dynamics (5.73) This is done by computing the control input $\boldsymbol{u}$ by means of the static optimization problem

$$\begin{array}{rl} \text{minimize} & \bar{\mathcal{G}} \;=\; \frac{1}{2}\|\dot{\boldsymbol{\xi}} + \bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}}\|^2_{\bar{\boldsymbol{M}}} \\ \text{subject to} & \boldsymbol{M}\dot{\boldsymbol{\xi}} + \boldsymbol{b} = \boldsymbol{B}\boldsymbol{u}, \; \boldsymbol{u} \in \mathbb{R}^p \end{array}. \tag{5.74}$$

Mathematically we do not have to use the metric coefficients $\bar{\boldsymbol{M}}$ here for the optimization, any other symmetric positive definite matrix would serve the purpose as well. However, from the physics point of view, the terms in the norm correspond to an acceleration, so the inertia of the desired closed loop is the reasonable choice. Furthermore, from the control point of view, the additional parameters arising with a different matrix would not turn out to be really useful.

Elimination of the acceleration $\dot{\boldsymbol{\xi}} = \boldsymbol{M}^{-1}(\boldsymbol{B}\boldsymbol{u} - \boldsymbol{b})$ from (5.74) leads to

$$\begin{aligned} \bar{\mathcal{G}} &= \tfrac{1}{2}\|\boldsymbol{M}^{-1}\boldsymbol{B}\boldsymbol{u} - \underbrace{(\boldsymbol{M}^{-1}\boldsymbol{b} - \bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}})}_{\tilde{\boldsymbol{a}}}\|^2_{\bar{\boldsymbol{M}}} \\ &= \tfrac{1}{2}\boldsymbol{u}^\top \underbrace{\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}\boldsymbol{M}^{-1}\boldsymbol{B}}_{\boldsymbol{H}}\boldsymbol{u} - \boldsymbol{u}^\top\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}\tilde{\boldsymbol{a}} + \tfrac{1}{2}\tilde{\boldsymbol{a}}^\top\bar{\boldsymbol{M}}\tilde{\boldsymbol{a}} \\ &= \tfrac{1}{2}\big(\boldsymbol{u} - \underbrace{\boldsymbol{H}^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}\tilde{\boldsymbol{a}}}_{\boldsymbol{u}_0}\big)^\top\boldsymbol{H}\big(\boldsymbol{u} - \underbrace{\boldsymbol{H}^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}\tilde{\boldsymbol{a}}}_{\boldsymbol{u}_0}\big) \\ &\qquad + \tfrac{1}{2}\tilde{\boldsymbol{a}}^\top\bar{\boldsymbol{M}}\underbrace{\big(\mathbf{I}_n - \boldsymbol{M}^{-1}\boldsymbol{B}\boldsymbol{H}^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}\big)}_{\boldsymbol{P}^\perp}\tilde{\boldsymbol{a}} \\ &= \tfrac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_0\|^2_{\boldsymbol{H}} + \underbrace{\tfrac{1}{2}\|\boldsymbol{P}^\perp\tilde{\boldsymbol{a}}\|^2_{\bar{\boldsymbol{M}}}}_{\bar{\mathcal{G}}_0}. \end{aligned} \tag{5.75}$$

For the formulation of $\bar{\mathcal{G}}_0$ it is crucial to note that $\boldsymbol{P}^\perp$ is a projection matrix, which will be exploited in the next subsection.

The control law, i.e. the solution of the minimization problem, is obviously $\boldsymbol{u} = \boldsymbol{u}_0$. The resulting closed loop kinetics are

$$
\begin{aligned}
&& \boldsymbol{M}\dot{\boldsymbol{\xi}} + \boldsymbol{b} &= \boldsymbol{B}\,\overbrace{\boldsymbol{H}^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}(\boldsymbol{M}^{-1}\boldsymbol{b} - \bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}})}^{\boldsymbol{u}_0} \\
\Leftrightarrow && \dot{\boldsymbol{\xi}} &= \boldsymbol{M}^{-1}\boldsymbol{B}\boldsymbol{H}^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}(\boldsymbol{M}^{-1}\boldsymbol{b} - \bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}}) - \boldsymbol{M}^{-1}\boldsymbol{b} \\
\Leftrightarrow && \bar{\boldsymbol{M}}\dot{\boldsymbol{\xi}} + \bar{\boldsymbol{b}} &= \underbrace{\bar{\boldsymbol{M}}\boldsymbol{P}^\perp\big(\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}} - \boldsymbol{M}^{-1}\boldsymbol{b}\big)}_{\tilde{\boldsymbol{b}}}.
\end{aligned} \tag{5.76}
$$

This result means that there is an additional vector[1] $\tilde{\boldsymbol{b}}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{x}_\mathsf{R}, \boldsymbol{\xi}_\mathsf{R}, \dot{\boldsymbol{\xi}}_\mathsf{R}) \in \mathbb{R}^n$ in the closed loop which allows for the closed loop (5.76) to be realizable with the available controls.

For the special case of a fully actuated system, i.e. $\boldsymbol{B}$ is invertible, the control law simplifies to $\boldsymbol{u}_0 = \boldsymbol{B}^{-1}(\boldsymbol{b} - \boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}})$. Furthermore we have $\boldsymbol{P}^\perp = \boldsymbol{0}$ and consequently $\tilde{\boldsymbol{b}} = \boldsymbol{0}$ and $\bar{\mathcal{G}}_0 = 0$.

In general, the value $\bar{\mathcal{G}}_0$ is a measure of how much the resulting closed loop differs from the original desired system (5.72). One main goal when parameterizing the controller is to make $\bar{\mathcal{G}}_0$ as small as possible. Unfortunately the form of $\bar{\mathcal{G}}_0$ in (5.76) is not handy, mainly due to $\operatorname{rank}\boldsymbol{P}^\perp = n - p$. In the following we like to find a more handy formulation.

5.5.2 Matching condition

The matrices $\boldsymbol{P} = \boldsymbol{M}^{-1}\boldsymbol{B}\boldsymbol{H}^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}$ and its complementary $\boldsymbol{P} = \mathbf{I}_n - \boldsymbol{P}^\perp$ from (5.75) are projection matrices, i.e. $\boldsymbol{P}^2 = \boldsymbol{P}$ and $(\boldsymbol{P}^\perp)^2 = \boldsymbol{P}^\perp$. Furthermore they are orthogonal w.r.t. the inner product $\langle\cdot,\cdot\rangle_{\bar{\boldsymbol{M}}}$, i.e. $\langle \boldsymbol{P}\boldsymbol{\xi}, \boldsymbol{P}^\perp\boldsymbol{\eta}\rangle_{\bar{\boldsymbol{M}}} = 0\,\forall\,\boldsymbol{\xi},\boldsymbol{\eta} \in \mathbb{R}^n$. The image of the projector $\boldsymbol{P}$ is the subspace of $\mathbb{R}^n$ that is spanned by the columns of $\boldsymbol{M}^{-1}\boldsymbol{B}$. Consequently we have $\operatorname{rank}\boldsymbol{P} = p$ and $\operatorname{rank}\boldsymbol{P}^\perp = n-p$. So $\tilde{\boldsymbol{b}}$ lies in a $(n-p)$-dimensional subspace. The goal of this subsection is to construct a basis for this subspace.

Let $\boldsymbol{B}^\perp \in \mathbb{R}^{n-p}$ be any orthogonal complement to $\boldsymbol{B}$, i.e. $\operatorname{rank}\boldsymbol{B}^\perp = n-p$ and $\boldsymbol{B}^\top\boldsymbol{B}^\perp = \boldsymbol{0}$. The columns of $\bar{\boldsymbol{M}}^{-1}\boldsymbol{M}\boldsymbol{B}^\perp$ are orthogonal to the columns of $\boldsymbol{M}^{-1}\boldsymbol{B}$ under the metric $\langle\cdot,\cdot\rangle_{\bar{\boldsymbol{M}}}$ and span the image of $\boldsymbol{P}^\perp$. Using some basic properties of orthogonal projectors (see subsection A.1.7) we can formulate

$$
\boldsymbol{P}^\perp = \bar{\boldsymbol{M}}^{-1}\boldsymbol{M}\boldsymbol{B}^\perp\underbrace{\big((\boldsymbol{B}^\perp)^\top\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\boldsymbol{M}\boldsymbol{B}^\perp\big)^{-1}}_{=\boldsymbol{S}\,\in\,\mathbb{SYM}^+(n-p)}(\boldsymbol{B}^\perp)^\top\boldsymbol{M}, \tag{5.77}
$$

Plugging this into the expressions from the previous subsection, we find

$$
\tilde{\boldsymbol{b}} = \boldsymbol{M}\boldsymbol{B}^\perp\boldsymbol{S}\underbrace{(\boldsymbol{B}^\perp)^\top(\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}} - \boldsymbol{b})}_{=\boldsymbol{\lambda}\,\in\,\mathbb{R}^{n-p}}. \qquad \bar{\mathcal{G}}_0 = \tfrac{1}{2}\|\boldsymbol{\lambda}\|^2_{\boldsymbol{S}}. \tag{5.78}
$$

It is much simpler to analyse $\boldsymbol{\lambda}$ which has only the dimension of the underactuation $n-p$ instead of $\tilde{\boldsymbol{b}}$ which has the full dimension n of the configuration space. Though it should be stressed that the values of $\bar{\mathcal{G}}_0$ and $\tilde{\boldsymbol{b}}$ are, as derived above, independent of the choice of

[1] The coefficients $\tilde{\boldsymbol{b}}$ do indeed transform like a tensor, even though $\boldsymbol{b}$ and $\bar{\boldsymbol{b}}$ do not.

$\boldsymbol{B}^{\perp}$. The naming $\boldsymbol{\lambda}$ is because we could have derived the same expressions by applying an acceleration constraint $\boldsymbol{B}^{\perp}(\boldsymbol{M}\dot{\boldsymbol{\xi}}+\boldsymbol{b}) = \boldsymbol{0}$ to the desired closed loop and using the method of *Lagrangian multipliers*.

The best case is, of course, if we achieve

$$\boldsymbol{\lambda} = (\boldsymbol{B}^{\perp})^{\top}(\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}} - \boldsymbol{b}) = \boldsymbol{0} \qquad \Rightarrow \qquad \tilde{\boldsymbol{b}} = \boldsymbol{0},\ \bar{\mathcal{G}}_0 = 0 \tag{5.79}$$

i.e. the desired closed loop is realized exactly. An approach based on this is discussed in [Bloch et al., 2000]. A condition similar to (5.79) is therein called the *the matching condition* and is required to be fulfilled exactly. However, the examples for which this approach is demonstrated restricts to stabilization tasks $\boldsymbol{\xi}_{\mathsf{R}} = \boldsymbol{0}$ for small academic systems.

An advantage of the presented approach is that the control law $\boldsymbol{u} = \boldsymbol{u}_0$ is defined independently of whether the matching condition is fulfilled or not. Instead the quantity $\boldsymbol{\lambda}$, which we will call the *matching force* in the following, ensures that the control law is realizable.

5.5.3 Approximations

The matching force (5.79) may become very cumbersome for complex systems and might even be impossible to vanish with the given parameters. It might be instructive to analyse it for particular situations.

Zero error. Assume that the controller tracks the reference perfectly, i.e. $\boldsymbol{x} = \boldsymbol{x}_{\mathsf{R}}$ and $\boldsymbol{\xi} = \boldsymbol{\xi}_{\mathsf{R}}$. One may check that for this case the three approaches all yield $\bar{\boldsymbol{b}} = \bar{\boldsymbol{M}}\dot{\boldsymbol{\xi}}_{\mathsf{R}}$. The resulting matching force $\boldsymbol{\lambda}^{\text{ZeroError}}$ for this special case is

$$\boldsymbol{\lambda}^{\text{ZeroError}} = (\boldsymbol{B}^{\perp}(\boldsymbol{x}_{\mathsf{R}}))^{\top}(\boldsymbol{M}(\boldsymbol{x}_{\mathsf{R}})\dot{\boldsymbol{\xi}}_{\mathsf{R}} - \boldsymbol{b}(\boldsymbol{x}_{\mathsf{R}}, \boldsymbol{\xi}_{\mathsf{R}})) \tag{5.80}$$

Evidently, this is independent of the closed loop parameters, and should rather be regarded as a constraint on the *reference trajectory* $t \mapsto \boldsymbol{x}_{\mathsf{R}}(t)$. The condition $\boldsymbol{\lambda}^{\text{ZeroError}} = \boldsymbol{0}$ is essentially the model equation after elimination of the control inputs.

A very useful approach here is to formulate the reference trajectory in terms of a *flat output* [Fliess et al., 1995] of the model. The first step for a systematic construction of a flat output is commonly the elimination of the control inputs (see e.g. [Schlacher and Schöberl, 2007]) i.e. $\boldsymbol{\lambda}^{\text{ZeroError}} = \boldsymbol{0}$.

Small error. Assume that we a small error $\boldsymbol{\varepsilon} = \boldsymbol{A}^{+}(\boldsymbol{x}_{\mathsf{R}})(\boldsymbol{x} - \boldsymbol{x}_{\mathsf{R}})$ to a constant reference $\boldsymbol{\xi}_{\mathsf{R}} = \boldsymbol{0}$ as already considered in section 5.4. Then the model and the closed loop template may be approximated by

$$\boldsymbol{M}_0\ddot{\boldsymbol{\varepsilon}} + \boldsymbol{D}_0\dot{\boldsymbol{\varepsilon}} + \boldsymbol{K}_0\boldsymbol{\varepsilon} = \boldsymbol{B}(\boldsymbol{x}_{\mathsf{R}})\Delta\boldsymbol{u}, \tag{5.81a}$$

$$\bar{\boldsymbol{M}}_0\ddot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{D}}_0\dot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{K}}_0\boldsymbol{\varepsilon} = \boldsymbol{0} \tag{5.81b}$$

and the matching force $\boldsymbol{\lambda}^{\text{SmallError}}$ for this special case is

$$\begin{aligned}\boldsymbol{\lambda}^{\text{SmallError}} &= (\boldsymbol{B}^{\perp}(\boldsymbol{x}_{\mathsf{R}}))^{\top}\big(\boldsymbol{M}_0\bar{\boldsymbol{M}}_0^{-1}(\bar{\boldsymbol{D}}_0\dot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{K}}_0\boldsymbol{\varepsilon}) - (\boldsymbol{D}_0\dot{\boldsymbol{\varepsilon}} + \boldsymbol{K}_0\boldsymbol{\varepsilon})\big) \\ &= \underbrace{(\boldsymbol{B}^{\perp}(\boldsymbol{x}_{\mathsf{R}}))^{\top}\big(\boldsymbol{M}_0\bar{\boldsymbol{M}}_0^{-1}\bar{\boldsymbol{D}}_0 - \boldsymbol{D}_0\big)}_{\boldsymbol{\Lambda}_{\boldsymbol{D}}}\dot{\boldsymbol{\varepsilon}} + \underbrace{(\boldsymbol{B}^{\perp}(\boldsymbol{x}_{\mathsf{R}}))^{\top}\big(\boldsymbol{M}_0\bar{\boldsymbol{M}}_0^{-1}\bar{\boldsymbol{K}}_0 - \boldsymbol{K}_0\big)}_{\boldsymbol{\Lambda}_{\boldsymbol{K}}}\boldsymbol{\varepsilon}\end{aligned} \tag{5.82}$$

As $\boldsymbol{\varepsilon}$ and $\dot{\boldsymbol{\varepsilon}}$ can be arbitrary, the matrices $\boldsymbol{\Lambda}_{\boldsymbol{K}}$ and $\boldsymbol{\Lambda}_{\boldsymbol{D}}$ have to vanish, for $\boldsymbol{\lambda}^{\text{SmallError}}$ to vanish. For the following examples it will turn out that we can always find suitable parameters within $\bar{\boldsymbol{M}}_0$, $\bar{\boldsymbol{D}}_0$ and $\bar{\boldsymbol{K}}_0$ such that $\boldsymbol{\Lambda}_{\boldsymbol{K}} = \boldsymbol{\Lambda}_{\boldsymbol{D}} = \mathbf{0}$. Thus ensuring that at least the first order approximation of the actual matching force $\boldsymbol{\lambda}$ vanishes.

5.5.4 Systems with input constraints

In most control systems the control inputs $\boldsymbol{u}$ can not take arbitrary values, but are constrained like e.g. $-u_a^{\max} \leq u_a \leq u_a^{\max}, a = 1, \dots, p$ due to practical limitations. In general we assume that the constraints can be written as $\boldsymbol{W}\boldsymbol{u} \leq \boldsymbol{l}$ where the inequality is understood componentwise and the resulting set $\mathbb{U} = \{\boldsymbol{u} \in \mathbb{R}^p \,|\, \boldsymbol{W}\boldsymbol{u} \leq \boldsymbol{l}\}$ is assumed to be convex.

Here we can use just the same arguments as in subsection 5.5.1 to motivate a control law defined by the solution of the optimization problem

$$\begin{aligned}\text{minimize} \quad & \bar{\mathcal{G}} = \tfrac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_0\|_{\boldsymbol{H}}^2 + \bar{\mathcal{G}}_0 \\ \text{subject to} \quad & \boldsymbol{W}\boldsymbol{u} \leq \boldsymbol{l}, \ \boldsymbol{u} \in \mathbb{R}^p\end{aligned} \tag{5.83}$$

with $\boldsymbol{H}$ and $\boldsymbol{u}_0$ defined in (5.75). Given that $\boldsymbol{H} \in \mathbb{SYM}^+(p)$ is positive definite and $\mathbb{U}$ is convex, this problem has a unique solution, though it usually has to be computed numerically. For the following simulation results the MATLAB function `quadprog` was used and a C++ implementation of the Active-Set algorithm from [Nocedal and Wright, 2006, Algorithm 16.3] was used for the real-time implementation on the Multicopters.

It should be stressed that this approach does not blah stability

5.6 Summary and recipe

We have proposed three approaches for a control law for rigid body systems. Each of them formulated a sightly different template for the desired closed loop dynamics. The actual control law results from its combination with the model dynamics. For a fully actuated system the desired closed loop is achieved exactly. For an underactuated system or in the presence of input constraints one achieves closed loop dynamics that are "as close as possible" to the desired dynamics in the sense that the resulting acceleration differs the least.

The implementation of the controller is determined by the rigid body parameterization ${}^a_b\mathbf{G}(\boldsymbol{x})$, the kinematics $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}$ and the constitutive parameters ${}^a_b\mathbf{M}$, ${}^a_b\mathbf{D}$, ${}^a_b\mathbf{K}$. It is crucial to note that the resulting controlled system is invariant to the chosen coordinates $\boldsymbol{x}, \boldsymbol{\xi}$ in the same way as the system model: Though the describing equations depend explicitly on the coordinates, the resulting motion of the closed loop system is the same for any choice of coordinates. This can be validated by checking the covariance of the closed loop equations.

What does affect the motion of the controlled system are the constitutive parameters, i.e. the values within ${}^a_b\mathbf{M}$, ${}^a_b\mathbf{D}$, ${}^a_b\mathbf{K}$. These are associated with the rigid bodies and are completely independent of the system coordinates. For the energy based approach, the choice of a transport map might not be unique and consequently might also affect the motion.

THE recipe:

- Modeling
  - Choose a set of (possibly redundant) configuration coordinates $\boldsymbol{x}(t) \in \mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^\nu \,|\, \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}\}$ and minimal velocity coordinates $\boldsymbol{\xi}(t) \in \mathbb{R}^n$, $n = \dim \mathbb{X}$ that are related by the kinematics matrix $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu\times n}$ (see section 3.3):
  $$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi} \tag{5.84a}$$
  - Formulate the rigid body configurations ${}^a_b\mathbf{G}(\boldsymbol{x}) \in \mathbb{SE}(3)$, $a, b = 0, \ldots, N$ in terms of the chosen coordinates. This determines the body Jacobians (see **??**)
  $${}^a_b\boldsymbol{J} = \tfrac{\partial}{\partial \dot{\boldsymbol{x}}} \operatorname{vee}\big({}^a_b\mathbf{G}^{-1}\,{}^a_b\dot{\mathbf{G}}\big)\boldsymbol{A} \tag{5.84b}$$
  - Compute the model inertia force $\boldsymbol{f}^{\mathrm{M}} = \boldsymbol{M}\dot{\boldsymbol{\xi}} + \boldsymbol{c}$ from the body inertias ${}^0_b\mathbf{M}'$ (see subsection 4.3.3)
  $$\boldsymbol{M} = \sum\nolimits_b {}^0_b\boldsymbol{J}^\top \operatorname{Vee}({}^0_b\mathbf{M}')\,{}^0_b\boldsymbol{J}, \;\; \boldsymbol{c} = \sum\nolimits_b {}^0_b\boldsymbol{J}^\top\big(\operatorname{Vee}({}^0_b\mathbf{M}')\,{}^0_b\dot{\boldsymbol{J}} - \operatorname{ad}^\top_{{}^0_b\boldsymbol{J}\boldsymbol{\xi}} \operatorname{Vee}({}^0_b\mathbf{M}')\,{}^0_b\boldsymbol{J}\big)\boldsymbol{\xi} \tag{5.84c}$$
  - The model kinetics are the balance of the inertia force $\boldsymbol{f}^{\mathrm{M}}$, the force of control inputs $\boldsymbol{B}\boldsymbol{u}$ and whatever other forces $\boldsymbol{f}^{\mathrm{A}}$ may act on the system
  $$\boldsymbol{M}\dot{\boldsymbol{\xi}} + \underbrace{\boldsymbol{c} + \boldsymbol{f}^{\mathrm{A}}}_{\boldsymbol{b}} = \boldsymbol{B}\boldsymbol{u} \tag{5.84d}$$

- Closed loop template
  - The template is computed from the body configurations ${}^a_b\mathsf{G}$, the body Jacobians ${}^a_b\boldsymbol{J}$ and the control parameters ${}^a_b\bar{\mathsf{K}}'$, ${}^a_b\bar{\mathsf{D}}'$, ${}^a_b\bar{\mathsf{M}}'$:

$$\bar{\boldsymbol{M}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{Vee}({}^a_b\bar{\mathsf{M}}')\,{}^a_b\boldsymbol{J} \tag{5.85a}$$

$$\bar{\boldsymbol{f}}^{\mathsf{K}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\big((\mathbf{I}_4 - {}^a_b\mathsf{G}_{\mathrm{E}}^{-1})\,{}^a_b\bar{\mathsf{K}}'\big) \tag{5.85b}$$

  - particle-based approach (see subsection 5.1.3)

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\Big(\big(\operatorname{wed}({}^a_b\boldsymbol{J}(\boldsymbol{x})\boldsymbol{\xi}) - {}^a_b\mathsf{G}_{\mathrm{E}}^{-1}\operatorname{wed}({}^a_b\boldsymbol{J}(\boldsymbol{x}_{\mathsf{R}})\boldsymbol{\xi}_{\mathsf{R}})\big)\,{}^a_b\bar{\mathsf{D}}'\Big) \tag{5.85c}$$

$$\begin{aligned}\bar{\boldsymbol{c}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{vee2}\Big(\big(&\operatorname{wed}({}^a_b\dot{\boldsymbol{J}}\boldsymbol{\xi}) + \operatorname{wed}({}^a_b\boldsymbol{J}\boldsymbol{\xi})^2 \\ &- {}^a_b\mathsf{G}_{\mathrm{E}}^{-1}\big(\operatorname{wed}({}^a_b\boldsymbol{J}_{\mathsf{R}}\dot{\boldsymbol{\xi}}_{\mathsf{R}} + {}^a_b\dot{\boldsymbol{J}}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}}) + \operatorname{wed}({}^a_b\boldsymbol{J}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}})^2\big)\big)\,{}^a_b\bar{\mathsf{M}}'\Big)\end{aligned} \tag{5.85d}$$

  - body-based approach (see subsection 5.2.2)

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{Vee}({}^a_b\bar{\mathsf{D}}')\,{}^a_b\boldsymbol{\xi}_{\mathrm{E}}, \qquad {}^a_b\boldsymbol{\xi}_{\mathrm{E}} = {}^a_b\boldsymbol{J}\boldsymbol{\xi} - \operatorname{Ad}_{{}^a_b\mathsf{G}_{\mathrm{E}}^{-1}}{}^a_b\boldsymbol{J}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}}, \tag{5.85e}$$

$$\begin{aligned}\bar{\boldsymbol{c}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \big({}^a_b\bar{\mathsf{M}}({}^a_b\dot{\boldsymbol{J}}\boldsymbol{\xi} - \operatorname{Ad}_{{}^a_b\mathsf{G}_{\mathrm{E}}^{-1}}({}^a_b\boldsymbol{J}_{\mathsf{R}}\dot{\boldsymbol{\xi}}_{\mathsf{R}} + {}^a_b\dot{\boldsymbol{J}}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}}) + \operatorname{ad}_{{}^a_b\boldsymbol{\xi}_{\mathrm{E}}}\operatorname{Ad}_{{}^a_b\mathsf{G}_{\mathrm{E}}^{-1}}{}^a_b\boldsymbol{J}_{\mathsf{R}}\boldsymbol{\xi}_{\mathsf{R}}) \\ - \operatorname{ad}^\top_{{}^a_b\boldsymbol{\xi}_{\mathrm{E}}}{}^a_b\bar{\mathsf{M}}\,{}^a_b\boldsymbol{\xi}_{\mathrm{E}}\big),\end{aligned} \tag{5.85f}$$

  - energy-based approach (see subsection 5.3.4, requires the choice of a transport map $\boldsymbol{Q}$)

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \bar{\boldsymbol{D}}\boldsymbol{\xi}_{\mathrm{E}}, \quad \bar{\boldsymbol{D}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \operatorname{Vee}({}^a_b\bar{\mathsf{D}}')\,{}^a_b\boldsymbol{J}, \quad \boldsymbol{\xi}_{\mathrm{E}} = \boldsymbol{\xi} - \boldsymbol{Q}\boldsymbol{\xi}_{\mathsf{R}} \tag{5.85g}$$

$$\bar{\boldsymbol{C}} = \sum\nolimits_{a,b} {}^a_b\boldsymbol{J}^\top \big(\operatorname{Vee}({}^a_b\bar{\mathsf{M}}')\,{}^a_b\dot{\boldsymbol{J}} + {}^a_b\bar{\mathsf{C}}\,{}^a_b\boldsymbol{J}\big), \quad {}^a_b\bar{\mathsf{C}}_{pq} = \Gamma_{pqr}({}^a_b\bar{\mathsf{M}}')\,{}^a_bJ^r_k\xi^k \tag{5.85h}$$

$$\bar{\boldsymbol{c}} = \bar{\boldsymbol{C}}\boldsymbol{\xi}_{\mathrm{E}} - \bar{\boldsymbol{M}}(\boldsymbol{Q}\dot{\boldsymbol{\xi}}_{\mathsf{R}} + \dot{\boldsymbol{Q}}\boldsymbol{\xi}_{\mathsf{R}}), \tag{5.85i}$$

  - The desired closed loop kinetics are

$$\bar{\boldsymbol{M}}\dot{\boldsymbol{\xi}} + \underbrace{\bar{\boldsymbol{c}} + \bar{\boldsymbol{f}}^{\mathsf{D}} + \bar{\boldsymbol{f}}^{\mathsf{K}}}_{\bar{\boldsymbol{b}}} = \mathbf{0} \tag{5.85j}$$

- Control law:
  - For the fully actuated case, the desired closed loop is realized by

$$\boldsymbol{u} = \boldsymbol{B}^{-1}(\boldsymbol{b} - \boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}}) \tag{5.86a}$$

  - In the underactuated case, the acceleration error measured by the Gaussian constraint, is minimized by (see subsection 5.5.1)

$$\boldsymbol{u} = (\boldsymbol{B}^\top\boldsymbol{M}^{-1}\bar{\boldsymbol{M}}\boldsymbol{M}^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}^\top\boldsymbol{M}^{-1}(\bar{\boldsymbol{M}}\boldsymbol{M}^{-1}\boldsymbol{b} - \bar{\boldsymbol{b}}) \tag{5.86b}$$

    Choosing an orthogonal complement $\boldsymbol{B}^\perp$ to the input matrix $\boldsymbol{B}$, i.e. $\operatorname{rank}\boldsymbol{B}^\perp = n - p$ and $\boldsymbol{B}^\top\boldsymbol{B}^\perp = \mathbf{0}$, the residual acceleration error can be written as $\bar{\mathcal{G}}_0 = \frac{1}{2}\|\boldsymbol{\lambda}\|^2_{\boldsymbol{S}}$ where (see subsection 5.5.2)

$$\boldsymbol{\lambda} = (\boldsymbol{B}^\perp)^\top(\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{b}} - \boldsymbol{b}) = \mathbf{0}, \qquad \boldsymbol{S} = \big((\boldsymbol{B}^\perp)^\top\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\boldsymbol{M}\boldsymbol{B}^\perp\big)^{-1} \tag{5.86c}$$

    By adjusting the control parameters within $\bar{\boldsymbol{M}}$ and $\bar{\boldsymbol{b}}$ one may try to minimize $\bar{\mathcal{G}}_0$.

5.7 Examples of fully actuated systems

5.7.1 Prismatic joint

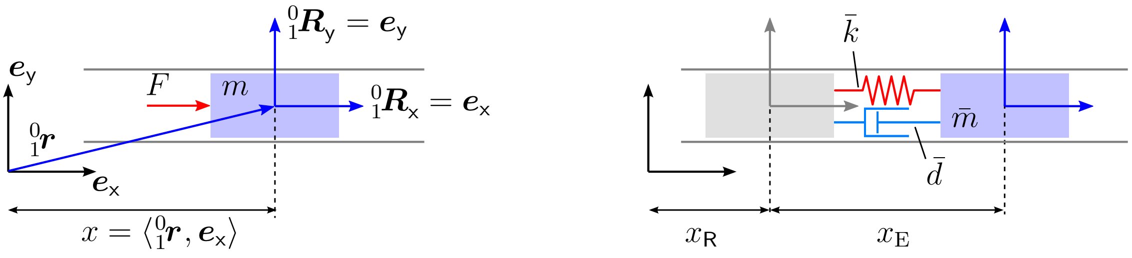


Figure 5.2: Model of a prismatic joint (left) and the closed loop (right)

Model. Probably the simplest example of a rigid body system is a single body moving in a prismatic joint, i.e. can only translate on one axis as illustrated on the left of Figure 5.2. The corresponding rigid body transformation is simply

$$^0_1\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.87}$$

With the trivial choice of the velocity coordinate $\boldsymbol{\xi} = \dot{x}$, i.e. $\boldsymbol{A} = 1$, the equation of motion is

$$m\ddot{x} = F. \tag{5.88}$$

Closed loop. Due to the geometry of the model, only the controlled total mass ^0_1m within the controlled body inertia matrix $^0_1\mathbf{M}$ contributes to the controlled kinetics and analog for the dissipation and stiffness. For the sake of readability we drop the body indices for the following examples of single bodies. So the only parameters contributing to the controlled kinetics are $\bar{m}, \bar{d}, \bar{k} \in \mathbb{R} > 0$.

For this example all three proposed control approaches are identical. With the displacement error $x_\mathrm{E} = x - x_\mathsf{R}$ the resulting energies are

$$\bar{\mathcal{V}} = \tfrac{1}{2}\bar{k}x_\mathrm{E}^2, \qquad \bar{\mathcal{R}} = \tfrac{1}{2}\bar{d}\dot{x}_\mathrm{E}^2, \qquad \bar{\mathcal{T}} = \tfrac{1}{2}\bar{m}\dot{x}_\mathrm{E}^2, \qquad \bar{\mathcal{S}} = \tfrac{1}{2}\bar{m}\ddot{x}_\mathrm{E}^2. \tag{5.89}$$

The potential has the obvious transport map $\boldsymbol{Q} = 1$ and the resulting closed loop kinetics are

$$\bar{m}\ddot{x}_\mathrm{E} + \bar{d}\dot{x}_\mathrm{E} + \bar{k}x_\mathrm{E} = 0. \tag{5.90}$$

The corresponding explicit control law is

$$F = m\ddot{x}_\mathsf{R} - \tfrac{m\bar{d}}{\bar{m}}\dot{x}_\mathrm{E} - \tfrac{m\bar{k}}{\bar{m}}x_\mathrm{E}. \tag{5.91}$$

An interpretation of the closed loop is given on the right side of Figure 5.2: The controlled body can be thought as being connected by a spring (stiffness $\bar{k}$) and a damper (viscosity $\bar{d}$) to its reference position x_{R}. The inertial force $\bar{m}\ddot{x}_{\mathrm{E}}$ reacts to the error acceleration, i.e. to the acceleration of the body relative to its reference acceleration $\ddot{z}_{\mathsf{R}}$. One could say the body has an inertia w.r.t. its reference.

5.7.2 Revolute joint

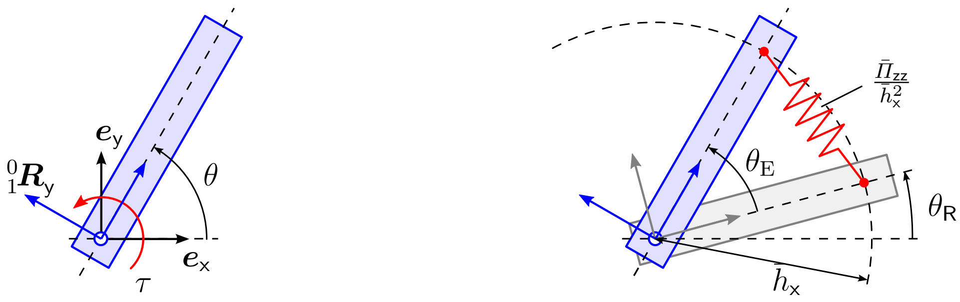


Figure 5.3: Revolute joint: rigid body constrained to rotate about an axis

Model. Another elemental case is the revolute joint, i.e. a rigid body constrained to rotate about an axis as illustrated on the left side of Figure 5.3. With the joint angle θ the rigid body configuration may be written as

$$ {}^0_1\mathbf{G} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.92} $$

With the velocity coordinate $\boldsymbol{\xi} = \dot{\theta}$ the equation of motion is

$$ \Theta_\mathsf{zz}\ddot{\theta} = \tau. \tag{5.93} $$

Potential energy. Due to the geometry of the model, only the control parameters $\bar{\Theta}_\mathsf{zz}, \bar{\Upsilon}_\mathsf{zz}, \bar{\Pi}_\mathsf{zz} \in \mathbb{R} > 0$ contribute to the closed loop kinetics. With the angle error $\theta_\mathrm{E} = \theta - \theta_\mathsf{R}$ the potential may be written as

$$ \bar{\mathcal{V}} = \bar{\Pi}_\mathsf{zz}(1 - \cos\theta_\mathrm{E}). \tag{5.94} $$

It has the obvious transport map $\boldsymbol{Q} = 1$. The potential could be realized by attaching a single linear spring with stiffness $\bar{\Pi}_\mathsf{zz}/\bar{h}_\mathsf{z}^2$ between desired configuration and actual configuration at a distance $\bar{h}_\mathsf{z}$ as illustrated on the right side of Figure 5.3. This also gives a vivid interpretation of the maximum on the potential at $\theta_\mathrm{E} = \pm\pi$.

Approach 1. The particle based approach leads to the following closed loop kinetics

$$ \bar{\Theta}_\mathsf{zz}(\ddot{\theta} - \ddot{\theta}_\mathsf{R}\cos\theta_\mathrm{E} - \dot{\theta}_\mathsf{R}^2\sin\theta_\mathrm{E}) + \bar{\Upsilon}_\mathsf{zz}(\dot{\theta} - \dot{\theta}_\mathsf{R}\cos\theta_\mathrm{E}) + \bar{\Pi}_\mathsf{zz}\sin\theta_\mathrm{E} = 0. \tag{5.95} $$

The total energy $\bar{\mathcal{W}}$ and its time derivative are

$$ \bar{\mathcal{W}} = \tfrac{1}{2}\bar{\Theta}_\mathsf{z}(\dot{\theta}^2 - 2\dot{\theta}\dot{\theta}_\mathsf{R}\cos\theta_\mathrm{E} + \dot{\theta}_\mathsf{R}^2) + \bar{\Pi}_\mathsf{z}(1 - \cos\theta_\mathrm{E}), \tag{5.96a} $$

$$ \begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{W}} = {} & -\bar{\Upsilon}_\mathsf{z}(\dot{\theta} - \dot{\theta}_\mathsf{R}\cos\theta_\mathrm{E})^2 + \bar{\Pi}_\mathsf{z}\dot{\theta}_\mathsf{R}\sin\theta_\mathrm{E}(\cos\theta_\mathrm{E} - 1) \\ & + \bar{\Theta}_\mathsf{z}\dot{\theta}_\mathsf{R}\big(\ddot{\theta}_\mathsf{R}(1 - \cos^2\theta_\mathrm{E}) + (\dot{\theta}^2 - \dot{\theta}_\mathsf{R}^2\cos\theta_\mathrm{E})\sin\theta_\mathrm{E}\big) \end{aligned} \tag{5.96b} $$

Without further assumptions on the reference trajectory $t \mapsto \theta_{\mathsf{R}}(t)$ the total energy is *not* a Lyapunov function for the closed loop. The linear approximation $\theta \approx \theta_{\mathsf{R}}$ of (5.95) has the characteristic polynomial

$$\lambda^2 + \tfrac{\bar{\Upsilon}_{\mathsf{z}}}{\Theta_{\mathsf{z}}}\lambda + \big(\tfrac{\bar{\Pi}_{\mathsf{z}}}{\Theta_{\mathsf{z}}} - \dot{\theta}_{\mathsf{R}}^2\big). \tag{5.97}$$

So even for the special case of constant reference velocity $\ddot{\theta}_{\mathsf{R}}(t) = 0$, we need $\frac{\bar{\Pi}_{\mathsf{z}}}{\Theta_{\mathsf{z}}} > \dot{\theta}_{\mathsf{R}}^2$ to ensure local stability.

Approach 2 & 3. For this example the body-based and energy-based approaches lead to identical energies and closed loop kinetics:

$$\bar{\mathcal{R}} = \tfrac{1}{2}\bar{\Upsilon}_{\mathsf{zz}}\dot{\theta}_{\mathrm{E}}^2, \qquad \bar{\mathcal{T}} = \tfrac{1}{2}\bar{\Theta}_{\mathsf{zz}}\dot{\theta}_{\mathrm{E}}^2, \qquad \bar{\Theta}_{\mathsf{zz}}\ddot{\theta}_{\mathrm{E}} + \bar{\Upsilon}_{\mathsf{zz}}\dot{\theta}_{\mathrm{E}} + \bar{\Pi}_{\mathsf{zz}}\sin\theta_{\mathrm{E}} = 0. \tag{5.98}$$

The total energy $\bar{\mathcal{W}} = \bar{\mathcal{T}} + \bar{\mathcal{V}}$, $\dot{\bar{\mathcal{W}}} = -2\bar{\mathcal{R}}$ con be used to conclude that the system converges for *almost* all initial conditions $(\theta_{\mathrm{E}}(0), \dot{\theta}_{\mathrm{E}}(0))$. The remaining initial condition $\theta_{\mathrm{E}}(0) = \pm\pi$ and $\dot{\theta}_{\mathrm{E}}(0) = 0$ is instable, see Figure 5.4. As a physical interpretation: the controlled dynamics coincide with the dynamics of a damped physical pendulum.

Linear control. Since the model (5.93) is a linear differential equation, the following linear closed loop equation might also be reasonable

$$\bar{\Theta}_{\mathsf{zz}}\ddot{\theta}_{\mathrm{E}} + \bar{\Upsilon}_{\mathsf{zz}}\dot{\theta}_{\mathrm{E}} + \bar{\Pi}_{\mathsf{zz}}\theta_{\mathrm{E}} = 0. \tag{5.99}$$

The difference between the closed loop (5.98) and (5.99) may be visualized by the corresponding phase plots, see Figure 5.4: The linear control law leads to non-smooth phase curves at $\theta = \pm\pi$, which is the consequence of the linear design for a system whose configuration space is actually $\mathbb{S}^1 \ncong \mathbb{R}$. See [Konz and Rudolph, 2016, sec. 1.2] for a deeper discussion.

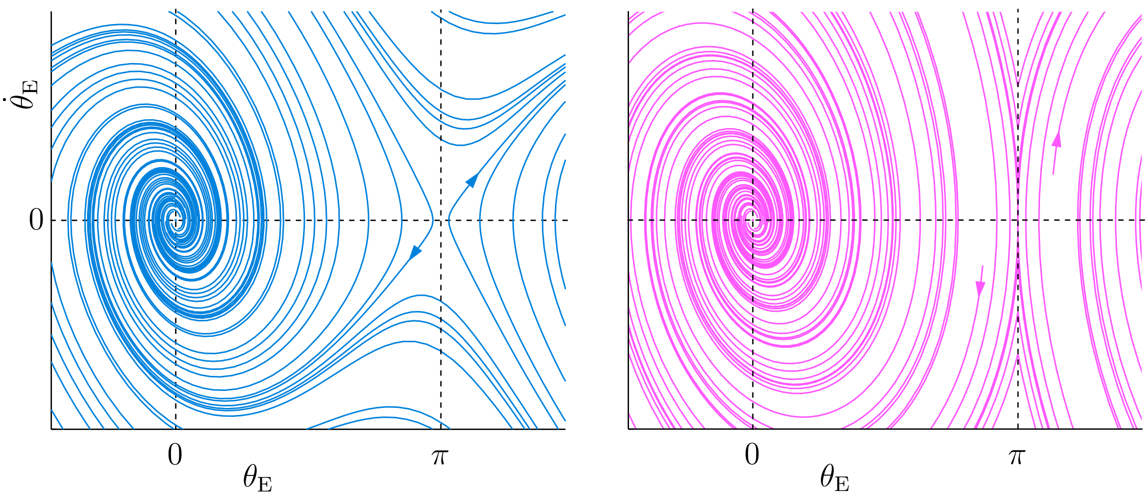


Figure 5.4: Phase plot for (5.98), left, and for (5.99), right

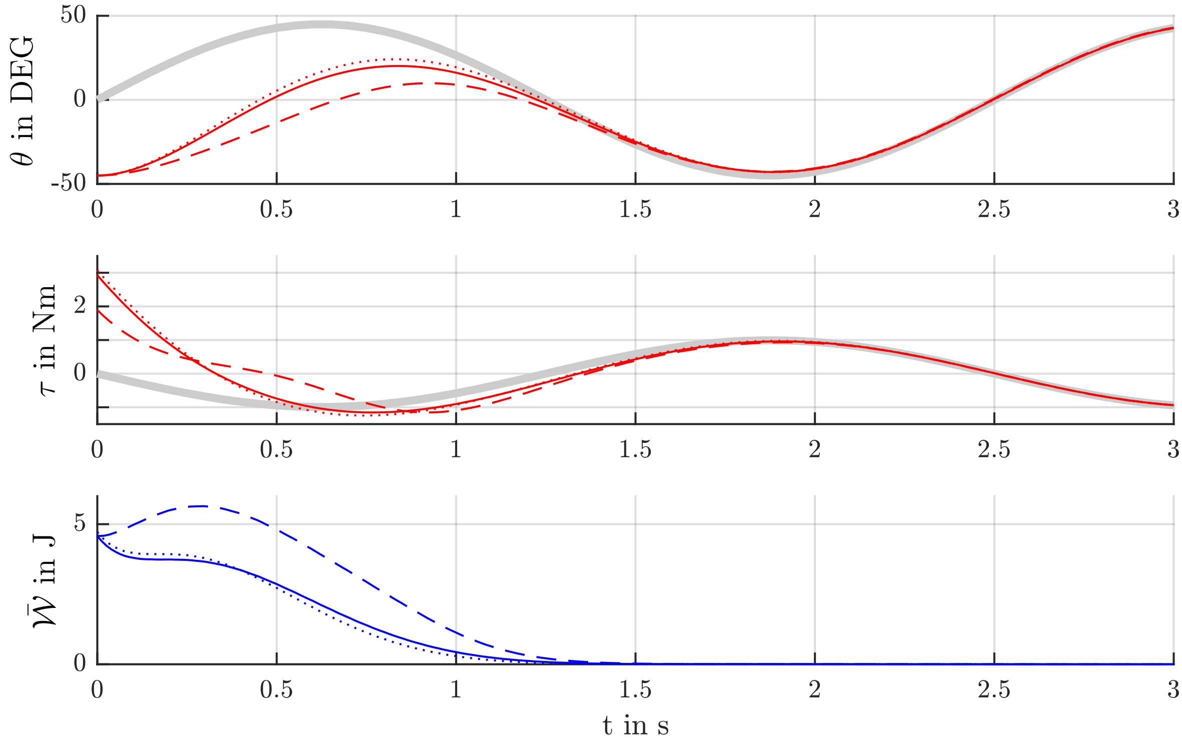


Figure 5.5: Simulation result for the revolute joint: gray line: reference, dashed line: particle-based approach, solid line: energy-based approach, dotted line: linear approach

Simulation results. Figure 5.5 shows simulation results comparing the three different approaches (5.95), (5.98) and (5.99) tracking a reference $\theta_{\mathsf{R}}(t) = \frac{\pi}{2}\sin(\frac{2\pi}{2.5}t)$. Evidently, all approaches fulfill the control objective, i.e. the joint angle θ converges to its reference θ_{R}.

5.7.3 Rigid body orientation

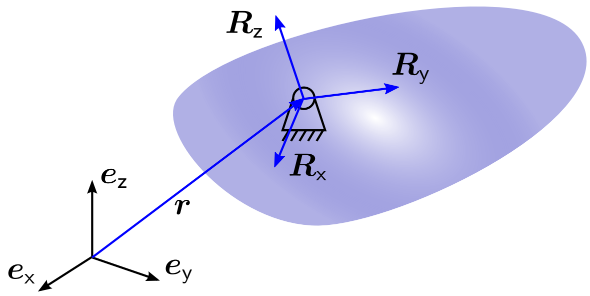


Figure 5.6: rigid body fixed at one point

Model. Consider a rigid body fixed at one point $\boldsymbol{r} = const.$ as illustrated in Figure 5.6. Its orientation may be parameterized by the coefficients of the rotation matrix $\boldsymbol{R} = [\boldsymbol{R}_{\mathsf{x}}, \boldsymbol{R}_{\mathsf{y}}, \boldsymbol{R}_{\mathsf{z}}] \in \mathbb{SO}(3)$. With the angular velocity $\boldsymbol{\omega} = \mathrm{Vee}(\boldsymbol{R}^\top \dot{\boldsymbol{R}})$ as velocity coordinates, the inertia matrix $\boldsymbol{\Theta}$ and the control torques $\boldsymbol{\tau}$ about the body fixed axes, the equations of motion may be written as

$$\dot{\boldsymbol{R}} = \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega}) \qquad \boldsymbol{\Theta}\dot{\boldsymbol{\omega}} + \mathrm{wed}(\boldsymbol{\omega})\boldsymbol{\Theta}\boldsymbol{\omega} = \boldsymbol{\tau}. \tag{5.100}$$

Potential energy. Only the parameters $\bar{\boldsymbol{\Theta}}$, $\bar{\boldsymbol{\Upsilon}}$ and $\bar{\boldsymbol{\Pi}}$ contribute to the closed loop kinetics. Using the attitude error $\boldsymbol{R}_{\mathrm{E}} = \boldsymbol{R}_{\mathsf{R}}^\top \boldsymbol{R}$ the error potential, its differential and Hessian are

$$\bar{\mathcal{V}} = \mathrm{tr}\left(\mathrm{Wed}(\bar{\boldsymbol{\Pi}})(\mathbf{I}_3 - \boldsymbol{R}_{\mathrm{E}})\right), \tag{5.101a}$$
$$\bar{\boldsymbol{f}}^{\mathsf{K}} = \boldsymbol{\nabla}\bar{\mathcal{V}} = \mathrm{vee2}(\mathrm{Wed}(\bar{\boldsymbol{\Pi}})\boldsymbol{R}_{\mathrm{E}}), \tag{5.101b}$$
$$(\boldsymbol{\nabla}^2\bar{\mathcal{V}})|_{\boldsymbol{R}=\boldsymbol{R}_{\mathsf{R}}} = \bar{\boldsymbol{\Pi}}. \tag{5.101c}$$

The potential has the transport map $\boldsymbol{Q} = \boldsymbol{R}_{\mathrm{E}}^\top$, so the velocity error for the energy based approach is $\boldsymbol{\omega}_{\mathrm{E}} = \boldsymbol{\omega} - \boldsymbol{R}_{\mathrm{E}}^\top \boldsymbol{\omega}_{\mathsf{R}}$ which coincides with the body velocity error.

Particle-based approach. The particle-based approach (5.18) leads to

$$\begin{aligned}\bar{\boldsymbol{f}}^{\mathsf{M}} &= \bar{\boldsymbol{\Theta}}\dot{\boldsymbol{\omega}} - \mathrm{Vee}(\boldsymbol{R}_{\mathrm{E}}^\top \mathrm{Wed}(\bar{\boldsymbol{\Theta}}))\boldsymbol{R}_{\mathrm{E}}^\top \dot{\boldsymbol{\omega}}_{\mathsf{R}} \\ &\qquad + \mathrm{wed}(\boldsymbol{\omega})\bar{\boldsymbol{\Theta}}\boldsymbol{\omega} + \mathrm{vee2}(\mathrm{Wed}(\bar{\boldsymbol{\Theta}})\,\mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}})^2 \boldsymbol{R}_{\mathrm{E}})\end{aligned} \tag{5.102a}$$
$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \bar{\boldsymbol{\Upsilon}}\boldsymbol{\omega} - \mathrm{Vee}(\boldsymbol{R}_{\mathrm{E}}^\top \mathrm{Wed}(\bar{\boldsymbol{\Upsilon}}))\boldsymbol{R}_{\mathrm{E}}^\top \boldsymbol{\omega}_{\mathsf{R}} \tag{5.102b}$$

Body-based approach. The body-based approach (5.26) leads to

$$\bar{\boldsymbol{f}}^{\mathsf{M}} = \bar{\boldsymbol{\Theta}}\dot{\boldsymbol{\omega}}_{\mathrm{E}} + \mathrm{wed}(\boldsymbol{\omega}_{\mathrm{E}})\bar{\boldsymbol{\Theta}}\boldsymbol{\omega}_{\mathrm{E}} \tag{5.103a}$$
$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \bar{\boldsymbol{\Upsilon}}\boldsymbol{\omega}_{\mathrm{E}}. \tag{5.103b}$$

The corresponding control law coincides with the one proposed in [Koditschek, 1989]. The total energy $\bar{\mathcal{W}} = \frac{1}{2}\boldsymbol{\omega}_{\mathrm{E}}^\top \bar{\boldsymbol{\Theta}}\boldsymbol{\omega}_{\mathrm{E}} + \bar{\mathcal{V}}$ with $\dot{\bar{\mathcal{W}}} = -\boldsymbol{\omega}_{\mathrm{E}}^\top \bar{\boldsymbol{\Upsilon}}\boldsymbol{\omega}_{\mathrm{E}}$ serves as a Lyapunov function for the closed loop.

Energy-based approach. For the energy-based approach (5.56) leads to

$$\bar{\boldsymbol{f}}^{\mathsf{M}} = \bar{\boldsymbol{\Theta}}\dot{\boldsymbol{\omega}}_{\mathrm{E}} + \operatorname{wed}(\operatorname{Wed}(\bar{\boldsymbol{\Theta}})\boldsymbol{\omega})\boldsymbol{\omega}_{\mathrm{E}} \tag{5.104a}$$

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \bar{\boldsymbol{\Upsilon}}\boldsymbol{\omega}_{\mathrm{E}}. \tag{5.104b}$$

The corresponding control law coincides with one proposed in [Bullo and Murray, 1999]. The total energy is the same as the one for the body based approach. The two approaches only differ in the gyroscopic terms.

Linearization. For a small error to the reference we have

$$\boldsymbol{R} = \boldsymbol{R}_{\mathsf{R}} + \boldsymbol{R}_{\mathsf{R}} \operatorname{wed}(\boldsymbol{\varepsilon}), \qquad \boldsymbol{\omega} = \boldsymbol{\omega}_{\mathsf{R}} + \dot{\boldsymbol{\varepsilon}}, \qquad \bar{\boldsymbol{\Theta}}\ddot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{\Upsilon}}\dot{\boldsymbol{\varepsilon}} + \bar{\boldsymbol{\Pi}}\boldsymbol{\varepsilon} = 0. \tag{5.105}$$

5.7.4 Planar rigid body

A planar rigid body is a free rigid body in two dimensional space, i.e. it can translate in two dimensions and rotate about an perpendicular axis as illustrated in Figure 5.7. The model equations as well as the closed loop equations could be directly derived from the three dimensional rigid body by setting e.g. $v_{\mathsf{z}} = 0$, $\omega_{\mathsf{x}} = \omega_{\mathsf{y}} = 0$ and removing the trivial equations. However it might be still instructive to display the resulting equations.

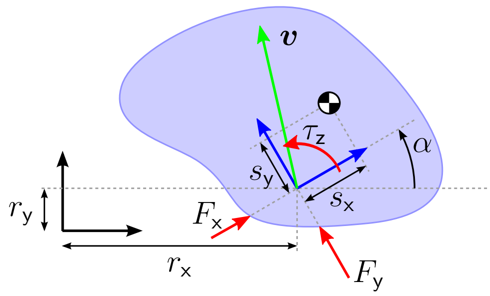


Figure 5.7: model of the planar rigid body

Coordinates and kinematics. As configuration coordinates $\boldsymbol{x}$ we use the position $r_{\mathsf{x}}, r_{\mathsf{y}}$ and the sine s_α and cosine c_α of the angle α. Consequently we have to impose the constraint $\mathrm{c}_\alpha^2 + \mathrm{s}_\alpha^2 - 1 = 0$ on the configuration coordinates. As velocity coordinates $\boldsymbol{\xi}$ we use the components $v_{\mathsf{x}}, v_{\mathsf{y}}$ of the translational velocity w.r.t. the body fixed frame as illustrated in Figure 5.7 and the angular velocity $\omega_{\mathsf{z}} = \dot{\alpha}$. This kinematic relation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\begin{bmatrix} r_{\mathsf{x}} \\ r_{\mathsf{y}} \\ \mathrm{s}_\alpha \\ \mathrm{c}_\alpha \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} \mathrm{c}_\alpha & -\mathrm{s}_\alpha & 0 \\ \mathrm{s}_\alpha & \mathrm{c}_\alpha & 0 \\ 0 & 0 & \mathrm{c}_\alpha \\ 0 & 0 & -\mathrm{s}_\alpha \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} v_{\mathsf{x}} \\ v_{\mathsf{y}} \\ \omega_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{\xi}} \tag{5.106}$$

The rigid body configuration ${}^0_1\mathbf{G}$ and the resulting body Jacobian ${}^0_1\boldsymbol{J}$ w.r.t. the chosen velocity coordinates are

$${}^0_1\mathbf{G} = \begin{bmatrix} \mathrm{c}_\alpha & -\mathrm{s}_\alpha & 0 & r_{\mathsf{x}} \\ \mathrm{s}_\alpha & \mathrm{c}_\alpha & 0 & r_{\mathsf{y}} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^0_1\boldsymbol{J} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.107}$$

Kinetic equation. Let the rigid body have the total mass m, the moment of inertia Θ_{z} and the coordinates $s_{\mathsf{x}}, s_{\mathsf{y}}$ of the center of mass w.r.t. the body fixed frame. As control input consider the forces $F_{\mathsf{x}}, F_{\mathsf{y}}$ and the torque τ_{z} as displayed in Figure 5.7. The resulting kinetic equation is

$$\underbrace{\begin{bmatrix} m & 0 & -ms_\mathsf{y} \\ 0 & m & ms_\mathsf{x} \\ -ms_\mathsf{y} & ms_\mathsf{x} & \Theta_\mathsf{z} \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} \dot{v}_\mathsf{x} \\ \dot{v}_\mathsf{y} \\ \dot{\omega}_\mathsf{z} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}} + \begin{bmatrix} -m(v_\mathsf{y} + s_\mathsf{x}\omega_\mathsf{z})\omega_\mathsf{z} \\ m(v_\mathsf{x} - s_\mathsf{y}\omega_\mathsf{z})\omega_\mathsf{z} \\ m(s_\mathsf{x}v_\mathsf{x} + s_\mathsf{y}v_\mathsf{y})\omega_\mathsf{z} \end{bmatrix} = \underbrace{\begin{bmatrix} F_\mathsf{x} \\ F_\mathsf{y} \\ \tau_\mathsf{z} \end{bmatrix}}_{\boldsymbol{u}}. \tag{5.108}$$

Control parameters. For the controlled kinetics we chose the following non-zero parameters

$$ {}^0_1\bar{m}, {}^0_1\bar{d}, {}^0_1\bar{k} \in \mathbb{R}^+, \quad {}^0_1\bar{s}_\mathsf{x}, {}^0_1\bar{s}_\mathsf{y}, {}^0_1\bar{l}_\mathsf{x}, {}^0_1\bar{l}_\mathsf{y}, {}^0_1\bar{h}_\mathsf{x}, {}^0_1\bar{h}_\mathsf{y} \in \mathbb{R}, \quad {}^0_1\bar{\Theta}_\mathsf{z}, {}^0_1\bar{\Upsilon}_\mathsf{z}, {}^0_1\bar{\Pi}_\mathsf{z} \in \mathbb{R}^+. \tag{5.109}$$

Since all parameters are associated with the configuration ${}^0_1\mathbf{G}$, we drop the indices in the following, i.e. $\bar{m} = {}^0_1\bar{m}$.

Potential. The potential, resulting from the chosen parameters (5.109), and its derivatives are

$$\begin{aligned} \bar{\mathcal{V}} &= \tfrac{1}{2}\bar{k}(r_\mathsf{x} - r_\mathsf{xR})^2 + \tfrac{1}{2}\bar{k}(r_\mathsf{y} - r_\mathsf{yR})^2 + \bar{\Pi}_\mathsf{z}(1 - \mathrm{c}_{\alpha_\mathrm{E}}) \\ &\quad + \bar{k}\bar{h}_\mathsf{x}(\mathrm{c}_\alpha - \mathrm{c}_{\alpha_\mathsf{R}})(r_\mathsf{x} - r_\mathsf{xR}) - \bar{k}\bar{h}_\mathsf{y}(\mathrm{s}_\alpha - \mathrm{s}_{\alpha_\mathsf{R}})(r_\mathsf{x} - r_\mathsf{xR}) \\ &\quad + \bar{k}\bar{h}_\mathsf{y}(\mathrm{c}_\alpha - \mathrm{c}_{\alpha_\mathsf{R}})(r_\mathsf{y} - r_\mathsf{yR}) - \bar{k}\bar{h}_\mathsf{x}(\mathrm{s}_\alpha - \mathrm{s}_{\alpha_\mathsf{R}})(r_\mathsf{y} - r_\mathsf{yR}) \end{aligned} \tag{5.110a}$$

$$\boldsymbol{\nabla}\bar{\mathcal{V}} = \begin{bmatrix} \bar{k}(\mathrm{c}_\alpha(r_\mathsf{x} - r_\mathsf{xR}) + \mathrm{s}_\alpha(r_\mathsf{y} - r_\mathsf{yR}) + \bar{h}_\mathsf{x}(1 - \mathrm{c}_{\alpha_\mathrm{E}}) - \bar{h}_\mathsf{y}\mathrm{s}_{\alpha_\mathrm{E}}) \\ \bar{k}(-\mathrm{s}_\alpha(r_\mathsf{x} - r_\mathsf{xR}) + \mathrm{c}_\alpha(r_\mathsf{y} - r_\mathsf{yR}) + \bar{h}_\mathsf{x}\mathrm{s}_{\alpha_\mathrm{E}} + \bar{h}_\mathsf{y}(1 - \mathrm{c}_{\alpha_\mathrm{E}})) \\ \bar{k}((\bar{h}_\mathsf{x}\mathrm{c}_\alpha + \bar{h}_\mathsf{y}\mathrm{s}_\alpha)(r_\mathsf{y} - r_\mathsf{yR}) - (\bar{h}_\mathsf{y}\mathrm{c}_\alpha + \bar{h}_\mathsf{x}\mathrm{s}_\alpha)(r_\mathsf{x} - r_\mathsf{xR})) + \bar{\Pi}_\mathsf{z}\mathrm{s}_{\alpha_\mathrm{E}} \end{bmatrix} \tag{5.110b}$$

$$\boldsymbol{\nabla}^2\bar{\mathcal{V}}|_\mathsf{R} = \begin{bmatrix} \bar{k} & 0 & -\bar{k}\bar{h}_\mathsf{y} \\ 0 & \bar{k} & \bar{k}\bar{h}_\mathsf{x} \\ -\bar{k}\bar{h}_\mathsf{y} & \bar{k}\bar{h}_\mathsf{x} & \bar{\Pi}_\mathsf{z} \end{bmatrix} \tag{5.110c}$$

The sine and cosine of the angle error $\alpha - \alpha_\mathsf{R}$ are introduced just for readability

$$\mathrm{c}_{\alpha_\mathrm{E}} = \mathrm{c}_\alpha\mathrm{c}_{\alpha_\mathsf{R}} + \mathrm{s}_\alpha\mathrm{s}_{\alpha_\mathsf{R}} = \cos(\alpha - \alpha_\mathsf{R}), \qquad \mathrm{s}_{\alpha_\mathrm{E}} = \mathrm{s}_\alpha\mathrm{c}_{\alpha_\mathsf{R}} - \mathrm{c}_\alpha\mathrm{s}_{\alpha_\mathsf{R}} = \sin(\alpha - \alpha_\mathsf{R}). \tag{5.111}$$

From the Hessian $\boldsymbol{\nabla}^2\bar{\mathcal{V}}|_\mathsf{R}$ at the critical point $\boldsymbol{x} = \boldsymbol{x}_\mathsf{R}$ one can see that (local) positive definiteness requires $\bar{\Pi}_\mathsf{z} > \bar{k}(\bar{h}_\mathsf{x}^2 + \bar{h}_\mathsf{y}^2)$. We will encounter the analog requirement for the controlled moment of inertia $\bar{\Theta}_\mathsf{z}$ and damping $\bar{\Upsilon}_\mathsf{z}$.

A transport map for (5.110a) is given by[2]

$$\underbrace{\begin{bmatrix} v_\mathrm{xE} \\ v_\mathrm{yE} \\ \omega_\mathrm{zE} \end{bmatrix}}_{\boldsymbol{\xi}_\mathrm{E}} = \underbrace{\begin{bmatrix} v_\mathsf{x} \\ v_\mathsf{y} \\ \omega_\mathsf{z} \end{bmatrix}}_{\boldsymbol{\xi}} - \underbrace{\begin{bmatrix} \mathrm{c}_{\alpha_\mathrm{E}} & \mathrm{s}_{\alpha_\mathrm{E}} & \mathrm{s}_\alpha(r_\mathsf{x} - r_\mathsf{xR}) - \mathrm{c}_\alpha(r_\mathsf{y} - r_\mathsf{yR}) \\ -\mathrm{s}_{\alpha_\mathrm{E}} & \mathrm{c}_{\alpha_\mathrm{E}} & \mathrm{c}_\alpha(r_\mathsf{x} - r_\mathsf{xR}) + \mathrm{s}_\alpha(r_\mathsf{y} - r_\mathsf{yR}) \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{Q}} \underbrace{\begin{bmatrix} v_\mathsf{xR} \\ v_\mathsf{yR} \\ \omega_\mathsf{zR} \end{bmatrix}}_{\boldsymbol{\xi}_\mathsf{R}}. \tag{5.112}$$

[2] An alternative transport map corresponding to (5.67) is

$$\boldsymbol{Q} = \begin{bmatrix} \mathrm{c}_{\alpha_\mathrm{E}} & \mathrm{s}_{\alpha_\mathrm{E}} & \bar{h}_\mathsf{x}\mathrm{s}_{\alpha_\mathrm{E}} - \bar{h}_\mathsf{y}(\mathrm{c}_{\alpha_\mathrm{E}} - 1) \\ -\mathrm{s}_{\alpha_\mathrm{E}} & \mathrm{c}_{\alpha_\mathrm{E}} & \bar{h}_\mathsf{x}(\mathrm{c}_{\alpha_\mathrm{E}} - 1) + \bar{h}_\mathsf{y}\mathrm{s}_{\alpha_\mathrm{E}} \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.113}$$

Particle-based approach. The damping and inertia force using the particle based approach are:

$$
\begin{aligned}
\bar{\boldsymbol{f}}^{\mathsf{D}} = {} & \underbrace{\begin{bmatrix} \bar{d} & 0 & -\bar{d}\,\bar{l}_{\mathsf{y}} \\ 0 & \bar{d} & \bar{d}\,\bar{l}_{\mathsf{x}} \\ -\bar{d}\,\bar{l}_{\mathsf{y}} & \bar{d}\,\bar{l}_{\mathsf{x}} & \bar{\Upsilon}_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\begin{bmatrix} v_{\mathsf{x}} \\ v_{\mathsf{y}} \\ \omega_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{\xi}} \\
& - \begin{bmatrix} \bar{d}\mathrm{c}_{\alpha_{\mathrm{E}}} & \bar{d}\mathrm{s}_{\alpha_{\mathrm{E}}} & \bar{d}(\bar{l}_{\mathsf{x}}\mathrm{s}_{\alpha_{\mathrm{E}}} - \bar{l}_{\mathsf{y}}\mathrm{c}_{\alpha_{\mathrm{E}}}) \\ -\bar{d}\mathrm{s}_{\alpha_{\mathrm{E}}} & \bar{d}\mathrm{c}_{\alpha_{\mathrm{E}}} & \bar{d}(\bar{l}_{\mathsf{x}}\mathrm{c}_{\alpha_{\mathrm{E}}} + \bar{l}_{\mathsf{y}}\mathrm{s}_{\alpha_{\mathrm{E}}}) \\ -\bar{d}(\bar{l}_{\mathsf{x}}\mathrm{s}_{\alpha_{\mathrm{E}}} + \bar{l}_{\mathsf{y}}\mathrm{c}_{\alpha_{\mathrm{E}}}) & \bar{d}(\bar{l}_{\mathsf{x}}\mathrm{c}_{\alpha_{\mathrm{E}}} - \bar{l}_{\mathsf{y}}\mathrm{s}_{\alpha_{\mathrm{E}}}) & \bar{\Upsilon}_{\mathsf{z}}\mathrm{c}_{\alpha_{\mathrm{E}}} \end{bmatrix} \underbrace{\begin{bmatrix} v_{\mathsf{xR}} \\ v_{\mathsf{yR}} \\ \omega_{\mathsf{zR}} \end{bmatrix}}_{\boldsymbol{\xi}_{\mathsf{R}}},
\end{aligned} \tag{5.114a}
$$

$$
\begin{aligned}
\bar{\boldsymbol{f}}^{\mathsf{M}} = {} & \underbrace{\begin{bmatrix} \bar{m} & 0 & -\bar{m}\,\bar{s}_{\mathsf{y}} \\ 0 & \bar{m} & \bar{m}\,\bar{s}_{\mathsf{x}} \\ -\bar{m}\,\bar{s}_{\mathsf{y}} & \bar{m}\,\bar{s}_{\mathsf{x}} & \bar{\Upsilon}_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\begin{bmatrix} \dot{v}_{\mathsf{x}} \\ \dot{v}_{\mathsf{y}} \\ \dot{\omega}_{\mathsf{z}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}} + \begin{bmatrix} -\bar{m}(v_{\mathsf{y}} + \bar{s}_{\mathsf{x}}\omega_{\mathsf{z}})\omega_{\mathsf{z}} \\ \bar{m}(v_{\mathsf{x}} - \bar{s}_{\mathsf{y}}\omega_{\mathsf{z}})\omega_{\mathsf{z}} \\ \bar{m}(\bar{s}_{\mathsf{x}}v_{\mathsf{x}} + \bar{s}_{\mathsf{y}}v_{\mathsf{y}})\omega_{\mathsf{z}} \end{bmatrix} \\
& - \begin{bmatrix} \bar{m}\mathrm{c}_{\alpha_{\mathrm{E}}} & \bar{m}\mathrm{s}_{\alpha_{\mathrm{E}}} & \bar{m}(\bar{s}_{\mathsf{x}}\mathrm{s}_{\alpha_{\mathrm{E}}} - \bar{s}_{\mathsf{y}}\mathrm{c}_{\alpha_{\mathrm{E}}}) \\ -\bar{m}\mathrm{s}_{\alpha_{\mathrm{E}}} & \bar{m}\mathrm{c}_{\alpha_{\mathrm{E}}} & \bar{m}(\bar{s}_{\mathsf{x}}\mathrm{c}_{\alpha_{\mathrm{E}}} + \bar{s}_{\mathsf{y}}\mathrm{s}_{\alpha_{\mathrm{E}}}) \\ -\bar{m}(\bar{s}_{\mathsf{x}}\mathrm{s}_{\alpha_{\mathrm{E}}} + \bar{s}_{\mathsf{y}}\mathrm{c}_{\alpha_{\mathrm{E}}}) & \bar{m}(\bar{s}_{\mathsf{x}}\mathrm{c}_{\alpha_{\mathrm{E}}} - \bar{s}_{\mathsf{y}}\mathrm{s}_{\alpha_{\mathrm{E}}}) & \bar{\Upsilon}_{\mathsf{z}}\mathrm{c}_{\alpha_{\mathrm{E}}} \end{bmatrix} \underbrace{\begin{bmatrix} \dot{v}_{\mathsf{xR}} \\ \dot{v}_{\mathsf{yR}} \\ \dot{\omega}_{\mathsf{zR}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}_{\mathsf{R}}} \\
& - \begin{bmatrix} -\bar{m}((v_{\mathsf{yR}} + \bar{s}_{\mathsf{x}}\omega_{\mathsf{zR}})\mathrm{c}_{\alpha_{\mathrm{E}}} - (v_{\mathsf{xR}} - \bar{s}_{\mathsf{y}}\omega_{\mathsf{zR}})\mathrm{s}_{\alpha_{\mathrm{E}}})\omega_{\mathsf{zR}} \\ \bar{m}((v_{\mathsf{yR}} + \bar{s}_{\mathsf{x}}\omega_{\mathsf{zR}})\mathrm{s}_{\alpha_{\mathrm{E}}} + (v_{\mathsf{xR}} - \bar{s}_{\mathsf{y}}\omega_{\mathsf{zR}})\mathrm{c}_{\alpha_{\mathrm{E}}})\omega_{\mathsf{zR}} \\ \bar{m}((\bar{s}_{\mathsf{x}}v_{\mathsf{xR}} + \bar{s}_{\mathsf{y}}v_{\mathsf{yR}})\mathrm{c}_{\alpha_{\mathrm{E}}} - (\bar{s}_{\mathsf{y}}v_{\mathsf{xR}} - \bar{s}_{\mathsf{x}}v_{\mathsf{yR}})\mathrm{s}_{\alpha_{\mathrm{E}}})\omega_{\mathsf{zR}} + \bar{\Theta}_{\mathsf{z}}\mathrm{s}_{\alpha_{\mathrm{E}}}\omega_{\mathsf{zR}}^2 \end{bmatrix}.
\end{aligned} \tag{5.114b}
$$

The corresponding total energy as defined in (5.5), is not a Lyapunov function for the closed loop.

Body-based approach. The damping and inertia force using the body-based approach are:

$$
\bar{\boldsymbol{f}}^{\mathsf{D}} = \underbrace{\begin{bmatrix} \bar{d} & 0 & -\bar{d}\,\bar{l}_{\mathsf{y}} \\ 0 & \bar{d} & \bar{d}\,\bar{l}_{\mathsf{x}} \\ -\bar{d}\,\bar{l}_{\mathsf{y}} & \bar{d}\,\bar{l}_{\mathsf{x}} & \bar{\Upsilon}_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\begin{bmatrix} v_{\mathsf{x}\mathrm{E}} \\ v_{\mathsf{y}\mathrm{E}} \\ \omega_{\mathsf{z}\mathrm{E}} \end{bmatrix}}_{\boldsymbol{\xi}_{\mathrm{E}}} \tag{5.115a}
$$

$$
\bar{\boldsymbol{f}}^{\mathsf{M}} = \underbrace{\begin{bmatrix} \bar{m} & 0 & -\bar{m}\bar{s}_{\mathsf{y}} \\ 0 & \bar{m} & \bar{m}\bar{s}_{\mathsf{x}} \\ -\bar{m}\bar{s}_{\mathsf{y}} & \bar{m}\bar{s}_{\mathsf{x}} & \bar{\Theta}_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} \dot{v}_{\mathsf{x}\mathrm{E}} \\ \dot{v}_{\mathsf{y}\mathrm{E}} \\ \dot{\omega}_{\mathsf{z}\mathrm{E}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}_{\mathrm{E}}} + \begin{bmatrix} 0 & -\bar{m}\omega_{\mathsf{z}\mathrm{E}} & -\bar{m}\bar{s}_{\mathsf{x}}\omega_{\mathsf{z}\mathrm{E}} \\ \bar{m}\omega_{\mathsf{z}\mathrm{E}} & 0 & -\bar{m}\bar{s}_{\mathsf{y}}\omega_{\mathsf{z}\mathrm{E}} \\ \bar{m}\bar{s}_{\mathsf{x}}\omega_{\mathsf{z}\mathrm{E}} & \bar{m}\bar{s}_{\mathsf{y}}\omega_{\mathsf{z}\mathrm{E}} & 0 \end{bmatrix} \begin{bmatrix} v_{\mathsf{x}\mathrm{E}} \\ v_{\mathsf{y}\mathrm{E}} \\ \omega_{\mathsf{z}\mathrm{E}} \end{bmatrix} \tag{5.115b}
$$

where the velocity error $\boldsymbol{\xi}_{\mathrm{E}}$ was defined in (5.112). The total energy $\bar{\mathcal{W}} = \frac{1}{2}\boldsymbol{\xi}_{\mathrm{E}}^{\top}\boldsymbol{M}\boldsymbol{\xi}_{\mathrm{E}} + \bar{\mathcal{V}}$ is a Lyapunov function for the closed loop.

Energy-based approach. The damping and inertia forces using the energy-based approach are:

$$\bar{\boldsymbol{f}}^{\mathsf{D}} = \underbrace{\begin{bmatrix} \bar{d} & 0 & -\bar{d}\,\bar{l}_{\mathsf{y}} \\ 0 & \bar{d} & \bar{d}\,\bar{l}_{\mathsf{x}} \\ -\bar{d}\,\bar{l}_{\mathsf{y}} & \bar{d}\,\bar{l}_{\mathsf{x}} & \bar{\Upsilon}_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\begin{bmatrix} v_{\mathrm{xE}} \\ v_{\mathrm{yE}} \\ \omega_{\mathrm{zE}} \end{bmatrix}}_{\boldsymbol{\xi}_{\mathrm{E}}}, \tag{5.116a}$$

$$\bar{\boldsymbol{f}}^{\mathsf{M}} = \underbrace{\begin{bmatrix} \bar{m} & 0 & -\bar{m}\bar{s}_{\mathsf{y}} \\ 0 & \bar{m} & \bar{m}\bar{s}_{\mathsf{x}} \\ -\bar{m}\bar{s}_{\mathsf{y}} & \bar{m}\bar{s}_{\mathsf{x}} & \bar{\Theta}_{\mathsf{z}} \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} \dot{v}_{\mathrm{xE}} \\ \dot{v}_{\mathrm{yE}} \\ \dot{\omega}_{\mathrm{zE}} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}_{\mathrm{E}}} + \begin{bmatrix} 0 & -\bar{m}\omega_{\mathsf{z}} & -\bar{m}\bar{s}_{\mathsf{x}}\omega_{\mathsf{z}} \\ \bar{m}\omega_{\mathsf{z}} & 0 & -\bar{m}\bar{s}_{\mathsf{y}}\omega_{\mathsf{z}} \\ \bar{m}\bar{s}_{\mathsf{x}}\omega_{\mathsf{z}} & \bar{m}\bar{s}_{\mathsf{y}}\omega_{\mathsf{z}} & 0 \end{bmatrix} \begin{bmatrix} v_{\mathrm{xE}} \\ v_{\mathrm{yE}} \\ \omega_{\mathrm{zE}} \end{bmatrix}. \tag{5.116b}$$

The total energy $\bar{\mathcal{W}} = \frac{1}{2}\boldsymbol{\xi}_{\mathrm{E}}^{\top}\boldsymbol{M}\boldsymbol{\xi}_{\mathrm{E}} + \bar{\mathcal{V}}$ coincides with the total energy for the body-based approach and is a Lyapunov function for this closed loop as well. Note that the two approaches do differ in the gyroscopic terms, so do lead to different solutions of the closed loop dynamics.

Simulation result. In subsection 5.7.5 we will give and discuss a simulation result for the special parameter choice $\bar{s}_{\mathsf{x}} = \bar{l}_{\mathsf{x}} = \bar{h}_{\mathsf{x}} = 0$ and $\bar{s}_{\mathsf{y}} = \bar{l}_{\mathsf{y}} = \bar{h}_{\mathsf{y}} = 0$.

5.7.5 Decoupling of translational and rotational motion

The closed loop equations for a (free) three dimensional rigid body were given in subsection 5.1.2, subsection 5.2.1 and subsection 5.3.3. The reduced equations for the planar case were given in subsection 5.7.4. For most applications we like to *decouple* the translational and the rotational motion of the body.

Observing the closed loop equations one can see immediately that the coupling terms vanish if $\bar{\boldsymbol{s}} = \bar{\boldsymbol{l}} = \bar{\boldsymbol{h}} = \boldsymbol{0}$, i.e. the the chosen body fixed point $\boldsymbol{r}$ coincides with the center of mass, damping and stiffness. Then the rotational dynamics are identical to the closed loop given in subsection 5.7.3 (subsection 5.7.2 for the planar case), so are indeed decoupled/independent from the translational motion.

Translational dynamics. For the translational dynamics the situation is more difficult: Introduce $\boldsymbol{e} = \boldsymbol{r} - \boldsymbol{r}_{\mathsf{R}}$ as the components of the position error w.r.t. the inertial frame and $\boldsymbol{r}_{\mathrm{E}} = \boldsymbol{R}_{\mathsf{R}}^{\top}(\boldsymbol{r} - \boldsymbol{r}_{\mathsf{R}})$ as the components w.r.t. the reference frame. The translational dynamics for the different approaches and given transport map are equivalent to

$$\begin{aligned} &\text{particle-based:} && \bar{m}\ddot{\boldsymbol{e}} + \bar{d}\dot{\boldsymbol{e}} + \bar{k}\boldsymbol{e} = \boldsymbol{0} && (5.117a)\\ &\text{body-based:} && \bar{m}\ddot{\boldsymbol{r}}_{\mathrm{E}} + \bar{d}\dot{\boldsymbol{r}}_{\mathrm{E}} + \bar{k}\boldsymbol{r}_{\mathrm{E}} = \boldsymbol{0} && (5.117b)\\ &\text{energy-based:} && \bar{m}\big(\ddot{\boldsymbol{r}}_{\mathrm{E}} + \mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}})\dot{\boldsymbol{r}}_{\mathrm{E}}\big) + \bar{d}\dot{\boldsymbol{r}}_{\mathrm{E}} + \bar{k}\boldsymbol{r}_{\mathrm{E}} = \boldsymbol{0} && (5.117c) \end{aligned}$$

Translational energy. For the rigid body we can split the total energy $\bar{\mathcal{W}} = \bar{\mathcal{W}}_{\boldsymbol{r}} + \bar{\mathcal{W}}_{\boldsymbol{R}}$ into a part associated with the position $\bar{\mathcal{W}}_{\boldsymbol{r}}$ and one associated with the orientation $\bar{\mathcal{W}}_{\boldsymbol{R}}$. The rotational energies for the corresponding approaches coincide with the ones given in subsection 5.7.3 (subsection 5.7.2 for the planar case). The translational energies and their derivatives are

$$\begin{aligned} &\text{particle-based:} && \bar{\mathcal{W}}_{\boldsymbol{r}} = \tfrac{1}{2}\bar{k}\|\boldsymbol{e}\|^2 + \tfrac{1}{2}\bar{m}\|\dot{\boldsymbol{e}}\|^2, && \dot{\bar{\mathcal{W}}}_{\boldsymbol{r}} = -\bar{d}\|\dot{\boldsymbol{e}}\|^2 && (5.118a)\\ &\text{body-based:} && \bar{\mathcal{W}}_{\boldsymbol{r}} = \tfrac{1}{2}\bar{k}\|\boldsymbol{r}_{\mathrm{E}}\|^2 + \tfrac{1}{2}\bar{m}\|\dot{\boldsymbol{r}}_{\mathrm{E}}\|^2, && \dot{\bar{\mathcal{W}}}_{\boldsymbol{r}} = -\bar{d}\|\dot{\boldsymbol{r}}_{\mathrm{E}}\|^2 && (5.118b)\\ &\text{energy-based:} && \bar{\mathcal{W}}_{\boldsymbol{r}} = \tfrac{1}{2}\bar{k}\|\boldsymbol{r}_{\mathrm{E}}\|^2 + \tfrac{1}{2}\bar{m}\|\dot{\boldsymbol{r}}_{\mathrm{E}}\|^2, && \dot{\bar{\mathcal{W}}}_{\boldsymbol{r}} = -\bar{d}\|\dot{\boldsymbol{r}}_{\mathrm{E}}\|^2 && (5.118c) \end{aligned}$$

For their comparison note that

$$\|\boldsymbol{e}\| = \|\boldsymbol{r}_{\mathrm{E}}\|, \qquad \|\dot{\boldsymbol{e}}\| = \|\dot{\boldsymbol{r}}_{\mathrm{E}} + \mathrm{wed}(\boldsymbol{\omega}_{\mathsf{R}})\boldsymbol{r}_{\mathrm{E}}\|. \tag{5.119}$$

The crucial observation is that for all approaches the translational dynamics and energy are indeed independent of the actual orientation $\boldsymbol{R}$ and its velocity $\boldsymbol{\omega}$, but for some approaches they depend on their *reference* $\boldsymbol{R}_{\mathsf{R}}$ and $\boldsymbol{\omega}_{\mathsf{R}}$. For a constant reference orientation $\boldsymbol{R}_{\mathsf{R}} = const.$ and consequently $\boldsymbol{\omega}_{\mathsf{R}} = \boldsymbol{0}$ all four approaches are equivalent. Furthermore it is worth noting that the error dynamics as well as the energies are invariant to the reference trajectory $t \mapsto \boldsymbol{r}_{\mathsf{R}}(t)$ for the position.

Simulation. The difference between these cases will be discussed on simulation results for the simpler, yet as illustrative, example of a planar rigid body: The reference configuration is $r_{\mathsf{xR}}(t) = r_{\mathsf{yR}}(t) = 0$ and $\alpha_{\mathsf{R}}(t) = \pi t$ which yields the constant reference velocity $\boldsymbol{\xi}_{\mathsf{R}}(t) = [0, 0, \pi]$. The control parameters are set to $\bar{m} = 1$, $\bar{d} = 4$, $\bar{k} = 4$ (neglecting the units). The roots of the characteristic polynomial of (5.117c) are $\lambda \approx \{-0.5 \pm 0.6\mathrm{i}, -3.5 \pm 3.7\mathrm{i}\}$, resulting from the control parameters as well as the constant angular velocity $\omega_{\mathsf{zR}} = \pi$. The characteristic polynomial for the other approaches is independent of the reference trajectory and has a quadruple root at $\lambda = -2$.

Figure 5.8 shows the simulation result for the initial conditions $r_{\mathsf{x}}(0) = 0, r_{\mathsf{y}}(0) = 1$, $\alpha(0) = 0$ and $\boldsymbol{\xi}(0) = \mathbf{0}$. Observing from the inertial frame, top left of Figure 5.8, for approach 2 the body follows a straight line to its reference position, whereas for the other approaches spiral around it. Observing from the reference frame, top right of Figure 5.8, for approach 3 the body follows a direct path, given the initial velocity. The middle graph in Figure 5.8 shows the evolution of the translational energy $\bar{\mathcal{W}}_{\boldsymbol{r}}$. The difference in the initial values results from $\dot{\boldsymbol{e}}(0) = \mathbf{0}$, but $\dot{\boldsymbol{r}}_{\mathrm{E}}(0) \neq \mathbf{0}$. The bottom graph in Figure 5.8 shows the evolution of the euclidean distance $\|\boldsymbol{e}\| = \|\boldsymbol{r}_{\mathrm{E}}\|$. The rate of convergence for approach 2 and 3 are the same as could be expected from having the same characteristic polynomial.

Even though the energy based approach with the transport map from (5.64) might be mathematically the most elegant solution, its simulation result is not intuitive. Which approach is most desirable, depends given application. For indoor robots (like the multicopters discussed in the next chapter) it is probably most desirable if it corrects its position error following a straight line in the inertial frame.

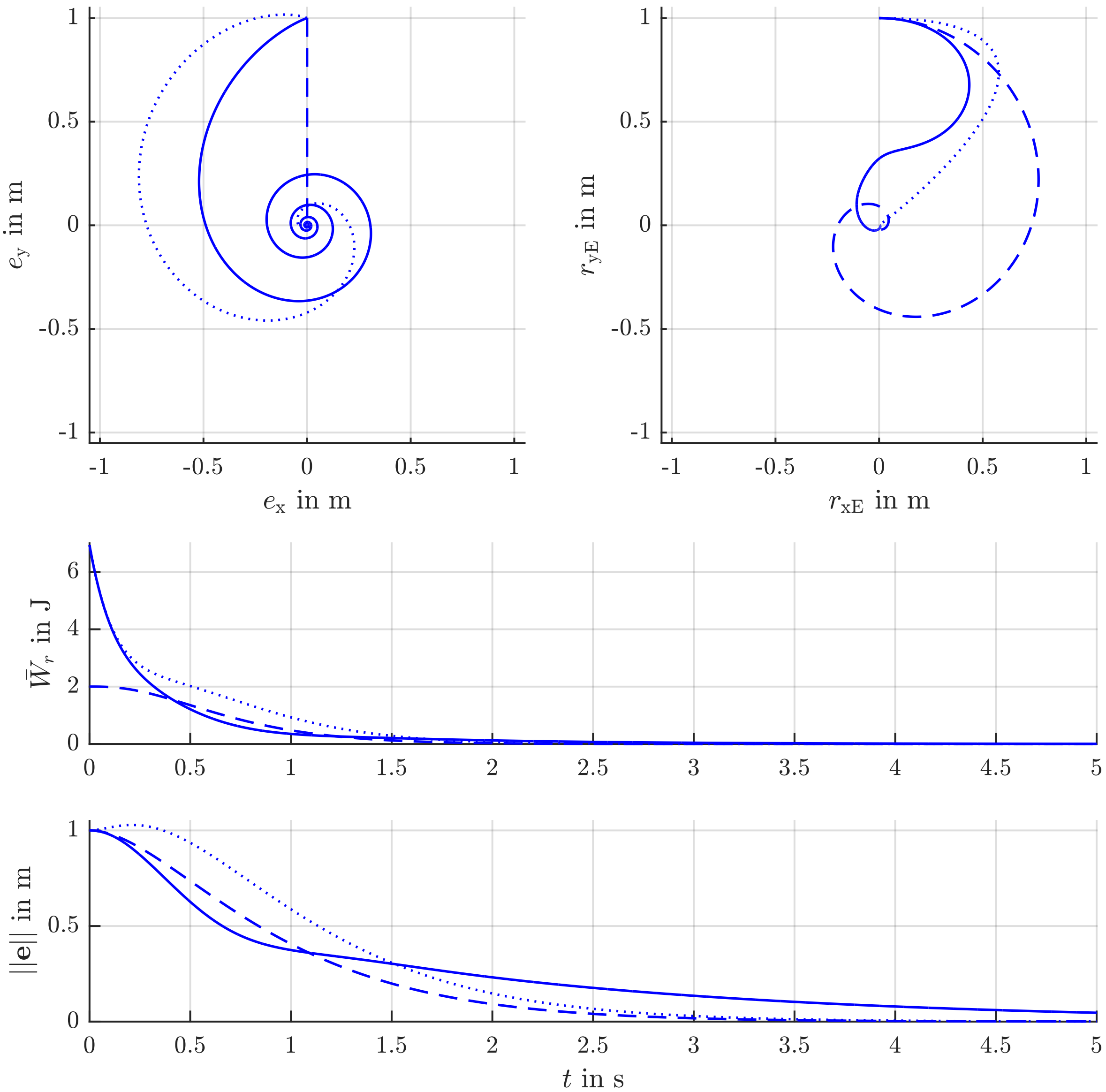


Figure 5.8: Simulation result for the planar rigid body. The solid line: energy-based approach with $\boldsymbol{Q} = \mathrm{Ad}_{\mathbf{G}^{-1}\mathbf{G}_{\mathrm{R}}}$, dashed line: particle-based approach, dotted line: body-based approach.

5.7.6 SCARA robot

As a simple example of a multi-body system we consider a SCARA robot as displayed in Figure 5.9. For the sake of demonstration we neglect the vertical axis and the tool orientation. The remaining two axis are sufficient to position a tool (red point in Figure 5.9) in the workspace (green shaded area).

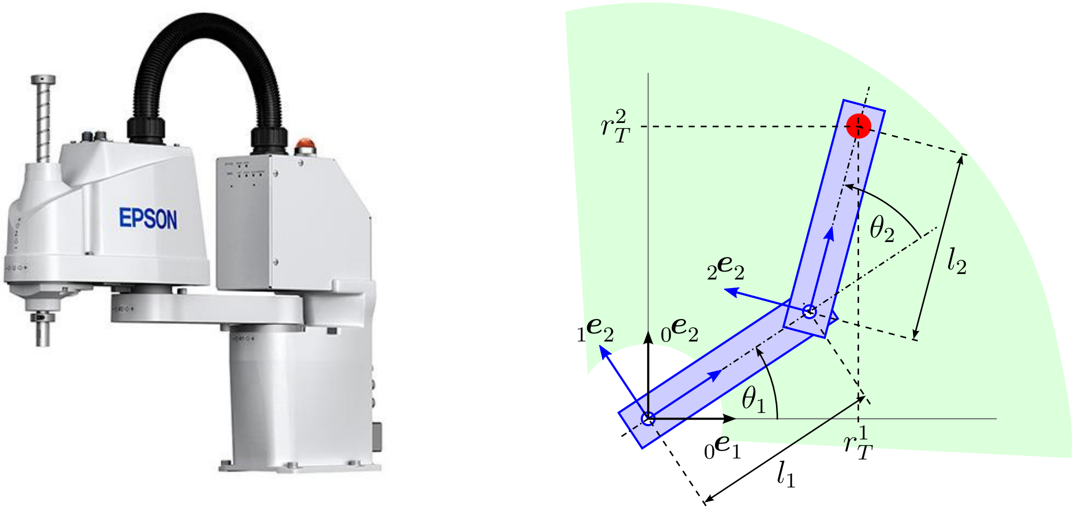


Figure 5.9: A Scara robot and its mechanical model (image www.epson.com)

Model. The model consists of two rigid bodies constraint by two revolute joints. A reasonable choice of coordinates are the relative joint angles $\boldsymbol{x} = [\theta_1, \theta_2]^\top$ and their derivatives $\boldsymbol{\xi} = [\dot\theta_1, \dot\theta_2]^\top$. The rigid body configurations are

$$
{}^0_1\mathbf{G} = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^0_1\boldsymbol{J} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{5.120a}
$$

$$
{}^1_2\mathbf{G} = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 & 0 & l_1 \\ \sin\theta_2 & \cos\theta_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^1_2\boldsymbol{J} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{5.120b}
$$

Let ${}_1\Theta_\mathrm{z}$ be the moment of inertia of the first body about the first joint and l_1 be the distance between the two joints. The second body has the mass ${}_2m$, the center of mass $({}_2s_\mathrm{x}, {}_2s_\mathrm{y})$ and the moment of inertia ${}_2\Theta_\mathrm{z}$ about the second joint. The control forces are the joint torques $\boldsymbol{u} = [\tau_1, \tau_2]^\top$. Overall, the equations of motion for the SCARA robot are

$$
\begin{bmatrix} {}_1\Theta_\mathrm{z} + {}_2\Theta_\mathrm{z} + {}_2m l_1^2 + 2a(\theta_2) & {}_2\Theta_\mathrm{z} + a(\theta_2) \\ {}_2\Theta_\mathrm{z} + a(\theta_2) & {}_2\Theta_\mathrm{z} \end{bmatrix} \begin{bmatrix} \ddot\theta_1 \\ \ddot\theta_2 \end{bmatrix} + \begin{bmatrix} a'(\theta_2)(2\dot\theta_1\dot\theta_2 + \dot\theta_2^2) \\ -a'(\theta_2)\dot\theta_1^2 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} \tag{5.121}
$$

where

$$a(\theta_2) = {}_2ml_1({}_2s_\mathsf{x}\cos\theta_2 - {}_2s_\mathsf{y}\sin\theta_2), \qquad a'(\theta_2) = -{}_2ml_1({}_2s_\mathsf{x}\sin\theta_2 + {}_2s_\mathsf{y}\cos\theta_2). \qquad (5.122)$$

Controller parameterization 1. In the following we will discuss two different controller parameterizations for the SCARA. For the first parameterization the non-zero parameters are

$${}^0_1\bar{\mathit{\Pi}}_\mathsf{z}, {}^0_1\bar{\mathit{\Upsilon}}_\mathsf{z}, {}^0_1\bar{\mathit{\Theta}}_\mathsf{z},\ {}^1_2\bar{\mathit{\Pi}}_\mathsf{z}, {}^1_2\bar{\mathit{\Upsilon}}_\mathsf{z}, {}^1_2\bar{\mathit{\Theta}}_\mathsf{z} \in \mathbb{R} > 0. \qquad (5.123)$$

These parameters are directly associated with the errors $\theta_{i\mathrm{E}} = \theta_i - \theta_{i\mathsf{R}}, i = 1, 2$ of the joint angles. The resulting potential is

$$\bar{\mathcal{V}} = {}^0_1\bar{\mathit{\Pi}}_\mathsf{z}\big(1 - \cos\theta_{1\mathrm{E}}\big) + {}^1_2\bar{\mathit{\Pi}}_\mathsf{z}\big(1 - \cos\theta_{2\mathrm{E}}\big). \qquad (5.124)$$

and obeys the transport map $\boldsymbol{Q} = \mathbf{I}_2$.

The resulting controlled kinetics for the body and energy-based approach are

$${}^{i-1}_{\ \ i}\bar{\mathit{\Theta}}_\mathsf{z}\ddot{\theta}_{i\mathrm{E}} + {}^{i-1}_{\ \ i}\bar{\mathit{\Upsilon}}_\mathsf{z}\dot{\theta}_{i\mathrm{E}} + {}^{i-1}_{\ \ i}\bar{\mathit{\Pi}}_\mathsf{z}\sin\theta_{i\mathrm{E}} = 0, \qquad i = 1, 2. \qquad (5.125)$$

The controlled kinetics for the particle based approach yield

$${}^{i-1}_{\ \ i}\bar{\mathit{\Theta}}_\mathsf{z}\big(\ddot{\theta}_i - \ddot{\theta}_{i\mathsf{R}}\cos\theta_{i\mathrm{E}} - \dot{\theta}^2_{i\mathsf{R}}\sin\theta_{i\mathrm{E}}\big) + {}^{i-1}_{\ \ i}\bar{\mathit{\Upsilon}}_\mathsf{z}\big(\dot{\theta}_i - \dot{\theta}_{i\mathsf{R}}\cos\theta_{i\mathrm{E}}\big) + {}^{i-1}_{\ \ i}\bar{\mathit{\Pi}}_\mathsf{z}\sin\theta_{i\mathrm{E}} = 0, \quad i = 1, 2. \qquad (5.126)$$

With this parameterization the controlled kinetics coincide with two copies of the kinetics of the revolute joint discussed in subsection 5.7.2.

Controller parameterization 2. The non-zero parameters for another interesting parameterization of the controller are

$${}^0_2\bar{k}, {}^0_2\bar{d}, {}^0_2\bar{m} \in \mathbb{R} > 0, \quad {}^0_2\bar{h}_\mathsf{x} = {}^0_2\bar{l}_\mathsf{x} = {}^0_2\bar{s}_\mathsf{x} = l_2. \qquad (5.127)$$

The resulting potential can be written as

$$\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = \tfrac{1}{2}\,{}^0_2\bar{k}\,\|\boldsymbol{r}_\mathsf{T}(\boldsymbol{x}) - \boldsymbol{r}_\mathsf{T}(\boldsymbol{x}_\mathsf{R})\|^2 \qquad \boldsymbol{r}_\mathsf{T}(\boldsymbol{x}) = \begin{bmatrix} l_1\cos\theta_1 + l_2\cos(\theta_1+\theta_2) \\ l_1\sin\theta_1 + l_2\sin(\theta_1+\theta_2) \end{bmatrix}, \qquad (5.128)$$

where $\boldsymbol{r}_\mathsf{T}$ is the position of the tool as illustrated in Figure 5.9. Using the *tool position error* $\boldsymbol{e}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = \boldsymbol{r}_\mathsf{T}(\boldsymbol{x}) - \boldsymbol{r}_\mathsf{T}(\boldsymbol{x}_\mathsf{R})$ as error coordinates, we can apply the rule from (B.46) to compute the transport map as

$$\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = \big(\boldsymbol{\nabla}\boldsymbol{r}_\mathsf{T}(\boldsymbol{x})\big)^{-1}\,\boldsymbol{\nabla}\boldsymbol{r}_\mathsf{T}(\boldsymbol{x}_\mathsf{R}). \qquad (5.129)$$

The determinant of the differential $\det \boldsymbol{\nabla}\boldsymbol{r}_\mathsf{T}(\boldsymbol{x}) = \sin\theta_2$ reflects the well known singularity of the SCARA inverse kinematics, see e.g. [Murray et al., 1994, example 3.6].

The closed loop kinetics for the particle and energy-based approach in terms of the model coordinates $\boldsymbol{x}$ and the error velocity $\boldsymbol{\xi}_\mathrm{E} = \boldsymbol{\xi} - \boldsymbol{Q}\boldsymbol{\xi}_\mathrm{R}$ are

$$\begin{aligned}
&\underbrace{{}^0_2\bar{m}\begin{bmatrix} l_1^2 + 2l_1l_2\cos\theta_2 + l_2^2 & l_1l_2\cos\theta_2 + l_2^2 \\ l_1l_2\cos\theta_2 + l_2^2 & l_2^2 \end{bmatrix}}_{\bar{\boldsymbol{M}}}\dot{\boldsymbol{\xi}}_\mathrm{E} + {}^0_2\bar{m}l_1l_2\sin\theta_2\begin{bmatrix} -\dot\theta_2 & -\dot\theta_1 - \dot\theta_2 \\ \dot\theta_1 & 0 \end{bmatrix}\boldsymbol{\xi}_\mathrm{E} \\
&\qquad + \underbrace{{}^0_2\bar{d}\begin{bmatrix} l_1^2 + 2l_1l_2\cos\theta_2 + l_2^2 & l_1l_2\cos\theta_2 + l_2^2 \\ l_1l_2\cos\theta_2 + l_2^2 & l_2^2 \end{bmatrix}}_{\bar{\boldsymbol{D}}}\boldsymbol{\xi}_\mathrm{E} \\
&+ \underbrace{{}^0_2\bar{k}\begin{bmatrix} l_1^2\sin\theta_{1\mathrm{E}} + l_1l_2(\sin(\theta_{1\mathrm{E}} - \theta_{2\mathrm{R}}) + \sin(\theta_{1\mathrm{E}} + \theta_2)) + l_2^2\sin(\theta_{1\mathrm{E}} + \theta_{2\mathrm{E}}) \\ l_1l_2(\sin(\theta_{1\mathrm{E}} + \theta_2) - \sin(\theta_2)) + l_2^2\sin(\theta_{1\mathrm{E}} + \theta_{2\mathrm{E}}) \end{bmatrix}}_{\boldsymbol{\nabla}\bar{\mathcal{V}}} = \mathbf{0}.
\end{aligned} \tag{5.130a}$$

In terms of the tool position error $\boldsymbol{e}$ this is equivalent to the much simpler equation

$${}^0_2\bar{m}\ddot{\boldsymbol{e}} + {}^0_2\bar{d}\dot{\boldsymbol{e}} + {}^0_2\bar{k}\boldsymbol{e} = \mathbf{0}. \tag{5.130b}$$

With the body-based approach we get a similar closed loop that is not displayed or discussed here.

As mentioned above, the transport map $\boldsymbol{Q}$ contains terms with $1/\sin\theta_2$. Fortunately these terms cancel out in $\bar{\boldsymbol{M}}\boldsymbol{Q}$ and $\bar{\boldsymbol{D}}\boldsymbol{Q}$ in the closed loop equation (5.130a), so this singularity actually does not hurt in practice. This could also be expected since the particle based approach, which does not rely on the transport map, leads to the same closed loop.

A singularity that does hurt, is the inertia matrix with $\det\bar{\boldsymbol{M}} = ({}^0_2\bar{m}l_1l_2\sin\theta_2)^2$. This means that one can not compute the control law if $\sin\theta_2 = 0$. Recalling the mechanical model of the SCARA Figure 5.9, this singularity is evident from a geometric point of view: If $\sin\theta_2 = 0$ the tool can only move in a tangential direction to the boundary of the workspace but not radial. However, it should be stressed that this singularity is not a consequence of unsuitable configuration coordinates $\boldsymbol{x} = [\theta_1, \theta_2]^\top$. It is rather an intrinsic one resulting from forcing dynamics suitable for $\mathbb{R}^2$ on a system that has the configuration space $\mathbb{S}^2$.

Simulation result. Figure 5.10 and Figure 5.11 show a simulation results for the SCARA robot with the two proposed parameterizations. The robot starts in a rather random initial configuration. The reference configuration is constant till $t = 1\,\mathrm{s}$, then follows a straight line for the tool position till $t = 4\,\mathrm{s}$ and remains constant thereafter.

For both parameterizations the controlled total energy $\bar{\mathcal{W}}$ converges. The crucial difference between the two parameterizations is, though the tool position $\boldsymbol{r}_\mathrm{T}$ tracks its reference in both cases, the joint angles θ_1, θ_2 do not for the second parameterization. The reason for this is best understood when looking at the controlled potential energy $\bar{\mathcal{V}}$ illustrated in Figure 5.12: For parameterization 2 the potential has two minima $\bar{\mathcal{V}} = 0$ for which the tool is at its reference position, but with different joint angles. This holds for any tool position except the ones on the boundary of the workspace where $\theta_2 = 0$ or $\theta_2 = \pi$.

Which of the two parameterizations is "better" probably depends on the practical control task: If the actual joint configuration (θ_1, θ_2) matters then the control parameters associated with them, i.e. ${}^0_1\bar{\Pi}_\mathsf{z}, {}^0_1\bar{\Upsilon}_\mathsf{z}, \ldots$, are more suited for the control design. If one is only interested in the tool position $\boldsymbol{r}_\mathsf{T}$, then the parameters of the parameterization 2 are useful.

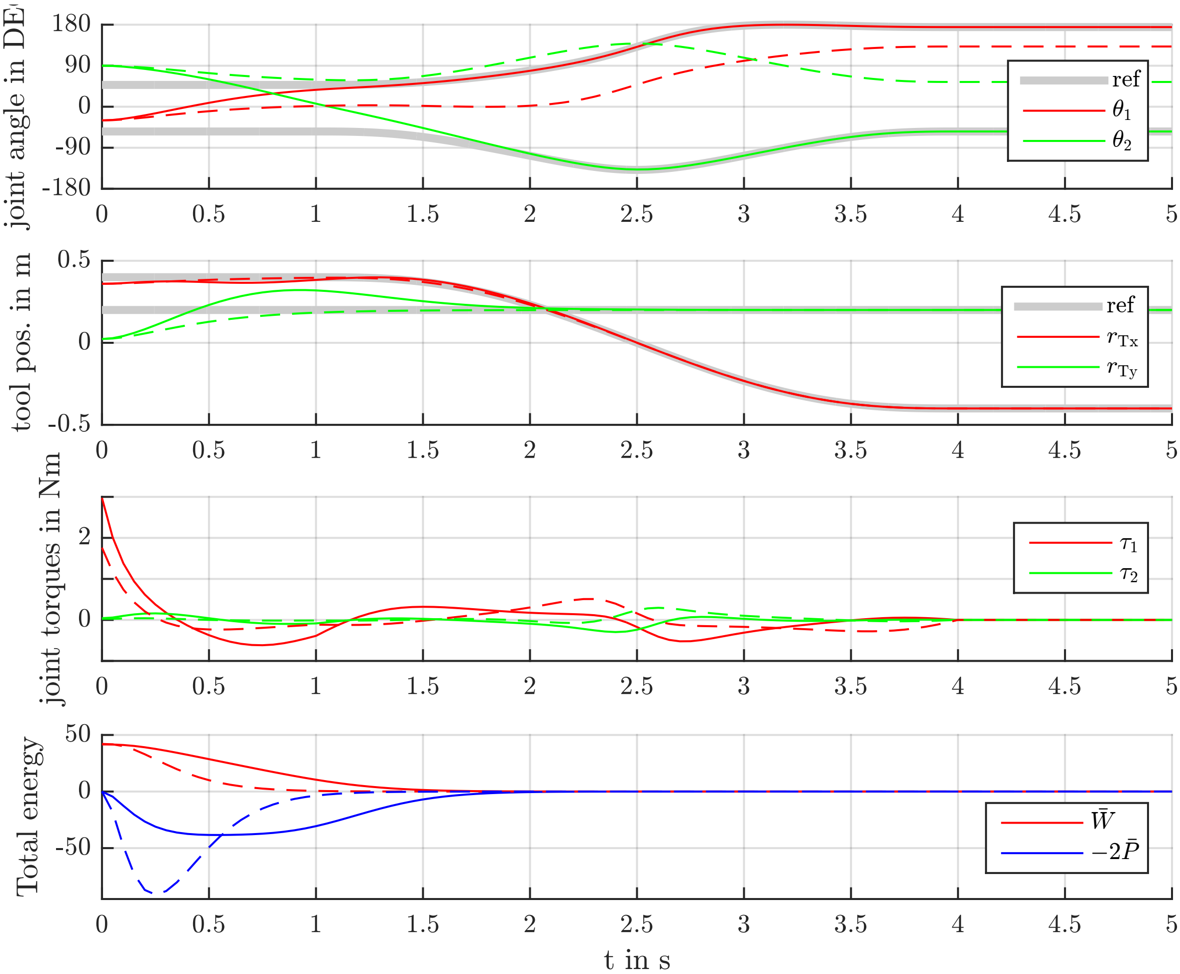


Figure 5.10: Simulation result for the SCARA with parameterization 1 (solid lines) and 2 (dashed lines)

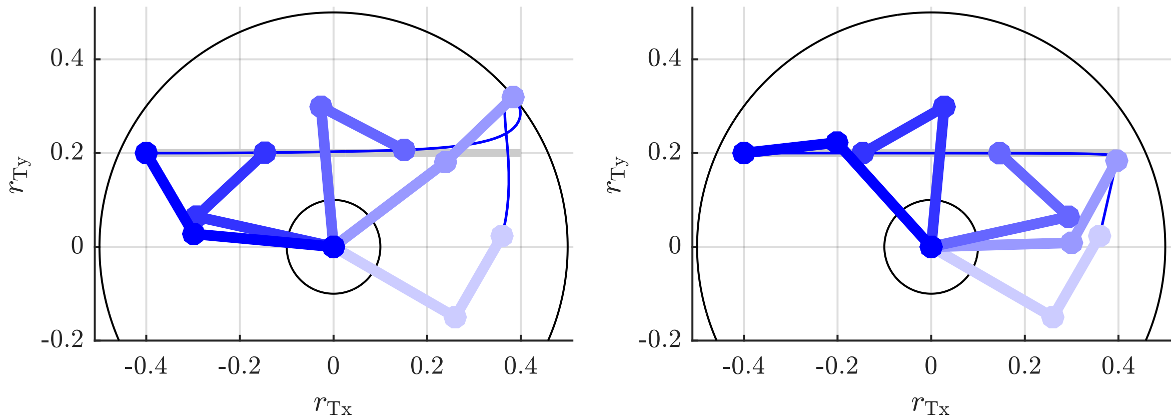


Figure 5.11: Snapshots for the simulation result for the SCARA with parameterization 1 (left) and 2 (right)

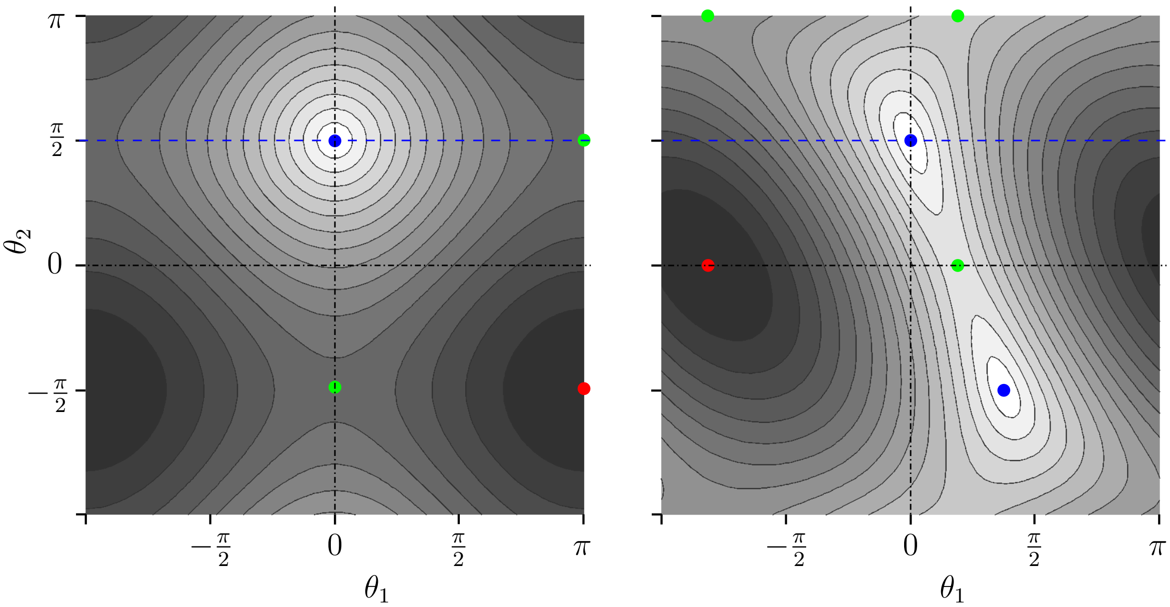


Figure 5.12: The controlled potential energy $\mathcal{V}$ for parameterization 1 (left) and 2 (right) for $\theta_{1\mathsf{R}} = 0$, $\theta_{2\mathsf{R}} = \frac{\pi}{2}$. Blue dots are minima, red are maxima and green are saddle points.

5.7.7 Robot arm

As a more complex multibody system we consider a robot arm as illustrated in Figure 5.13. For this example the model equations and the resulting closed loop equations become quite cumbersome and are not displayed explicitly. However this displays some benefits of the proposed control approach: One does not have to look at e.g. the actual system inertia matrix but only at the much less cumbersome body inertia matrices to conclude e.g. stability of the closed loop.

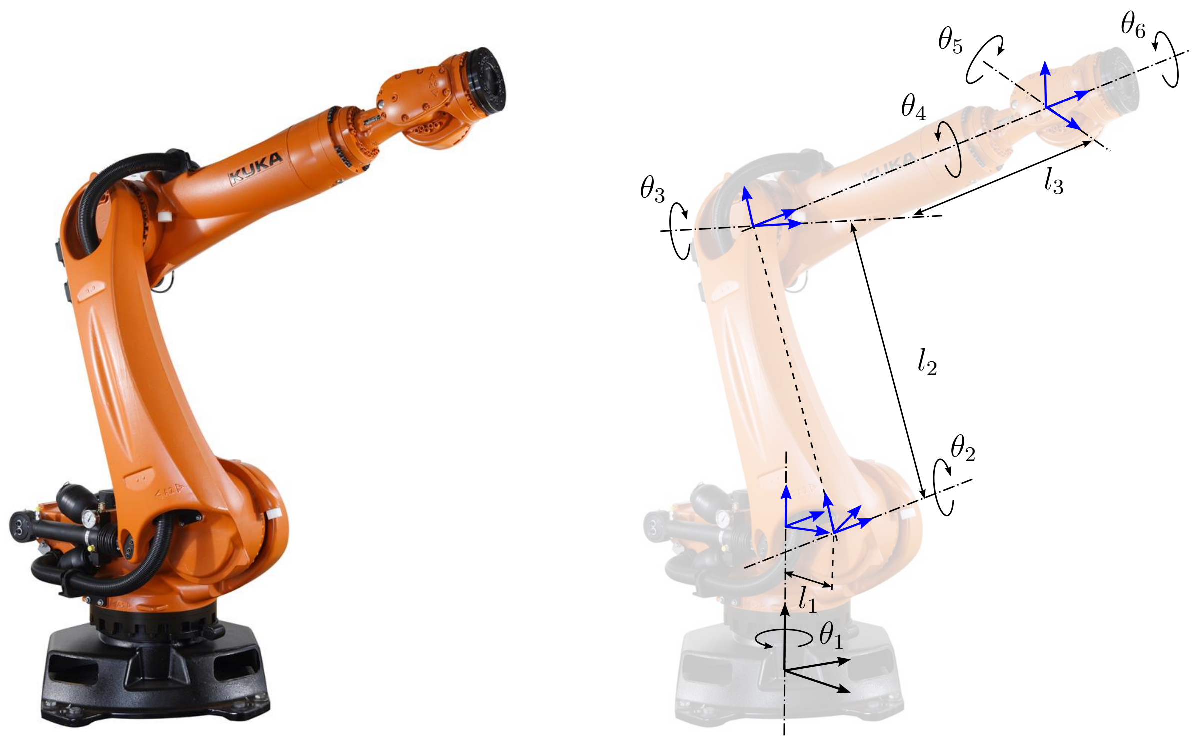


Figure 5.13: A model of a robot arm (background image from www.kuka.de)

Model. A reasonable parameterization of the system are the joint angles $\boldsymbol{x} = [\theta_1, \ldots, \theta_6]^\top$ as minimal configuration coordinates and trivial kinematics, i.e. $\boldsymbol{\xi} = \dot{\boldsymbol{x}}$. The body configurations can be computed from the following relative transformations

$$
{}^0_1\mathbf{G} = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^1_2\mathbf{G} = \begin{bmatrix} \cos\theta_2 & 0 & \sin\theta_2 & l_1 \\ 0 & 1 & 0 & 0 \\ -\sin\theta_2 & 0 & \cos\theta_2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},
$$
$$
{}^2_3\mathbf{G} = \begin{bmatrix} \cos\theta_3 & 0 & \sin\theta_3 & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta_3 & 0 & \cos\theta_3 & l_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^3_4\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & l_3 \\ 0 & \cos\theta_4 & -\sin\theta_4 & 0 \\ 0 & \sin\theta_4 & \cos\theta_4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},
$$
$$
{}^4_5\mathbf{G} = \begin{bmatrix} \cos\theta_5 & 0 & \sin\theta_5 & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta_5 & 0 & \cos\theta_5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^5_6\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_6 & -\sin\theta_6 & 0 \\ 0 & \sin\theta_6 & \cos\theta_6 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.131}
$$

This together with the body inertia matrices and the gravity coefficients $\boldsymbol{a}_{\mathsf{G}}$ and the control forces, $\boldsymbol{u} = [\tau_1, \ldots, \tau_6]^\top$ determines the equations of motion.

Controller parameterization 1: Joint space control. Like above we consider two different sets of controller parameterizations: For the first case, the nonzero control parameters are

$$\begin{aligned} {}^0_1\bar{\mathit{\Pi}}_{\mathsf{zz}}, {}^1_2\bar{\mathit{\Pi}}_{\mathsf{yy}}, {}^2_3\bar{\mathit{\Pi}}_{\mathsf{yy}}, {}^3_4\bar{\mathit{\Pi}}_{\mathsf{xx}}, {}^4_5\bar{\mathit{\Pi}}_{\mathsf{yy}}, {}^5_6\bar{\mathit{\Pi}}_{\mathsf{xx}} &\in \mathbb{R} > 0 \\ {}^0_1\bar{\mathit{\Upsilon}}_{\mathsf{zz}}, {}^1_2\bar{\mathit{\Upsilon}}_{\mathsf{yy}}, {}^2_3\bar{\mathit{\Upsilon}}_{\mathsf{yy}}, {}^3_4\bar{\mathit{\Upsilon}}_{\mathsf{xx}}, {}^4_5\bar{\mathit{\Upsilon}}_{\mathsf{yy}}, {}^5_6\bar{\mathit{\Upsilon}}_{\mathsf{xx}} &\in \mathbb{R} > 0 \\ {}^0_1\bar{\mathit{\Theta}}_{\mathsf{zz}}, {}^1_2\bar{\mathit{\Theta}}_{\mathsf{yy}}, {}^2_3\bar{\mathit{\Theta}}_{\mathsf{yy}}, {}^3_4\bar{\mathit{\Theta}}_{\mathsf{xx}}, {}^4_5\bar{\mathit{\Theta}}_{\mathsf{yy}}, {}^5_6\bar{\mathit{\Theta}}_{\mathsf{xx}} &\in \mathbb{R} > 0 \end{aligned} \tag{5.132}$$

A transport map for the resulting potential energy is $\boldsymbol{Q} = \mathbf{I}_6$. The resulting closed loop kinetics are 6 decoupled equations identical to the ones for the SCARA (5.126) resp. (5.126).

Controller parameterization 2: Work space control. As a second case consider: For many applications the task of the robot arm is to control the position and orientation of a tool mounted at the end of its kinematic chain. This tool might have a particularly meaningful center point (TCP) and and principle axis. Let the configuration ${}^6_7\mathsf{G} = \mathit{const}.$ capture these tool specific parameters, for example the tip position and direction of a welding electrode as shown in Figure 5.14.

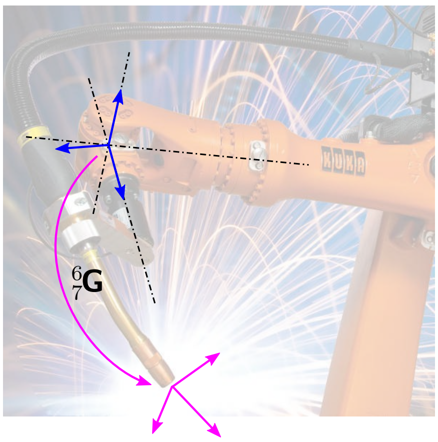


Figure 5.14: Welding tool attached to the robot arm (background image from www.kuka.de)

For this example it could be useful to control the tool as if it is a free rigid body (with its center of mass, damping and stiffness at the TCP) and not care about the particular mechanism that is used to give it this degree of freedom. This is achieved by the following nonzero control parameters

$${}^0_7\bar{k}, {}^0_7\bar{d}, {}^0_7\bar{m} \in \mathbb{R}^+, \qquad {}^0_7\bar{\boldsymbol{\Pi}}, {}^0_7\bar{\boldsymbol{\Upsilon}}, {}^0_7\bar{\boldsymbol{\Theta}} \in \mathbb{SYM}^+(3). \tag{5.133}$$

The resulting potential and corresponding transport map are

$$\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) = \tfrac{1}{2} \| (({}^0_7\mathsf{G}(\boldsymbol{x}_{\mathsf{R}}))^{-1}\, {}^0_7\mathsf{G}(\boldsymbol{x}) - \mathbf{I}_4)^\top \|^2_{{}^0_7\bar{\mathsf{K}}'}, \quad \boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) = \left({}^7_0\boldsymbol{J}(\boldsymbol{x})\right)^{-1} {}^7_0\boldsymbol{J}(\boldsymbol{x}_{\mathsf{R}}). \tag{5.134}$$

The resulting closed loop dynamics of the robot arm may be written by plugging the absolute tool configuration ${}^{0}_{7}\mathbf{G}(\boldsymbol{x})$ and its reference ${}^{0}_{7}\mathbf{G}(\boldsymbol{x}_{\mathsf{R}})$ into the dynamics of a single rigid body for either of the three proposed approaches (5.14), (5.21) or (5.65).

The determinant of the transport map is

$$\det \boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) = \frac{\det {}^{0}_{7}\boldsymbol{J}(\boldsymbol{x}_{\mathsf{R}})}{\det {}^{0}_{7}\boldsymbol{J}(\boldsymbol{x})}, \quad \det {}^{0}_{7}\boldsymbol{J}(\boldsymbol{x}) = -l_2 l_3 (l_1 + l_2 \sin\theta_2 + l_3 \cos(\theta_2 + \theta_3)) \cos\theta_3 \sin\theta_5 \tag{5.135}$$

If the term in the brackets vanishes means that the wrist lies on the axis of θ_1 and $\cos\theta_3 = 0$ is the case if the arm is completely straight which is the singularity we already encountered with the SCARA robot. The last three axis with angles θ_4, θ_5, θ_6 can be regarded as Euler angles in the sequence XYX and $\sin\theta_5 = 0$ is their singularity. Comparing this to the motivation example in section 3.2 we have the same problem but the other way around: The Euler angles are an absolutely appropriate choice of coordinates since the mechanism is realized like this. Consequently the configuration manifold of this part is $\mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{S}^1$ and we are assigning a control that was designed for $\mathbb{SO}(3)$.

Conclusion. The behavior of the two different parameterizations are quite analog to the two parameterizations of the SCARA robot. Which one is more suitable depends on the actual control task. Furthermore, the two presented parameterizations are just two special cases of which dynamics can be achieved with the more general approach of control of this work.

5.8 Examples of underactuated systems

5.8.1 Two masses connected by a spring

In order to illustrate the control approach for underactuated systems we consider the minimal example: Two bodies in prismatic joints connected by a linear spring but where only one is directly actuated by the force F as illustrated in Figure 5.15.

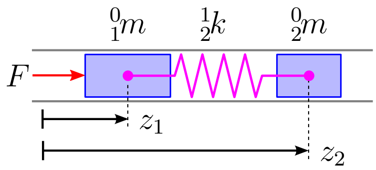


Figure 5.15: Model of two bodies connected by a spring

Model. We choose the absolute positions of the bodies as configuration coordinates $\boldsymbol{x} = [z_1, z_2]^\top$ and their derivative as velocity coordinates $\boldsymbol{\xi} = \dot{\boldsymbol{x}}$. With this the body configurations are

$$ {}^0_1\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & z_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad {}^0_2\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & z_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.136} $$

With the total mass ${}^0_1m$, ${}^0_2m$ of the individual bodies and the spring stiffness ${}^1_2k$ the resulting equations of motion may be written as

$$ \underbrace{\begin{bmatrix} {}^0_1m & 0 \\ 0 & {}^0_2m \end{bmatrix}}_{\boldsymbol{M}} \begin{bmatrix} \ddot{z}_1 \\ \ddot{z}_2 \end{bmatrix} + \underbrace{\begin{bmatrix} {}^1_2k & -{}^1_2k \\ -{}^1_2k & {}^1_2k \end{bmatrix}}_{\boldsymbol{K}} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\boldsymbol{B}} F. \tag{5.137} $$

Desired closed loop. We assume general body inertia ${}^0_1\mathbf{M}, \ldots$, damping and stiffness. The three proposed control approaches lead to identical desired closed loop dynamics:

$$ \bar{\boldsymbol{M}}\ddot{\boldsymbol{e}} + \bar{\boldsymbol{D}}\dot{\boldsymbol{e}} + \bar{\boldsymbol{K}}\boldsymbol{e} = \boldsymbol{0}, \qquad \boldsymbol{e} = \boldsymbol{x} - \boldsymbol{x}_{\mathsf{R}} \tag{5.138} $$

where

$$ \bar{\boldsymbol{M}} = \begin{bmatrix} {}^0_1\bar{m} + {}^1_2\bar{m} & -{}^1_2\bar{m} \\ -{}^1_2\bar{m} & {}^0_2\bar{m} + {}^1_2\bar{m} \end{bmatrix}, \quad \bar{\boldsymbol{D}} = \begin{bmatrix} {}^0_1\bar{d} + {}^1_2\bar{d} & -{}^1_2\bar{d} \\ -{}^1_2\bar{d} & {}^0_2\bar{d} + {}^1_2\bar{d} \end{bmatrix}, \quad \bar{\boldsymbol{K}} = \begin{bmatrix} {}^0_1\bar{k} + {}^1_2\bar{k} & -{}^1_2\bar{k} \\ -{}^1_2\bar{k} & {}^0_2\bar{k} + {}^1_2\bar{k} \end{bmatrix}. \tag{5.139} $$

The corresponding potential $\bar{\mathcal{V}} = \frac{1}{2}\boldsymbol{e}^\top\bar{\boldsymbol{K}}\boldsymbol{e}$ has the obvious transport map $\boldsymbol{Q} = \mathbf{I}_2$.

Matching. The matching condition (5.79) for this example may be written as

$$\boldsymbol{\lambda} = (\boldsymbol{B}^{\perp})^{\top}(\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}(-\bar{\boldsymbol{M}}\ddot{\boldsymbol{x}}_{\mathsf{R}} + \bar{\boldsymbol{D}}(\dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_{\mathsf{R}}) + \bar{\boldsymbol{K}}(\boldsymbol{x} - \boldsymbol{x}_{\mathsf{R}})) - \boldsymbol{K}\boldsymbol{x}) = \boldsymbol{0}. \tag{5.140}$$

Since this equation is linear in the system coordinates we can separate the into

$$(\boldsymbol{B}^{\perp})^{\top}(\boldsymbol{M}\ddot{\boldsymbol{x}}_{\mathsf{R}} + \boldsymbol{K}\boldsymbol{x}_{\mathsf{R}}) = \boldsymbol{0}. \tag{5.141a}$$

$$(\boldsymbol{B}^{\perp})^{\top}\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}\bar{\boldsymbol{D}} = \boldsymbol{0}, \tag{5.141b}$$

$$(\boldsymbol{B}^{\perp})^{\top}\boldsymbol{M}\bar{\boldsymbol{M}}^{-1}(\bar{\boldsymbol{K}} - \boldsymbol{K}) = \boldsymbol{0}, \tag{5.141c}$$

Choosing $\boldsymbol{B}^{\perp} = [0,1]^{\top}$ this is explicitly

$$\ddot{z}_{2\mathsf{R}} + \varpi(z_{2\mathsf{R}} - z_{1\mathsf{R}}) = 0 \tag{5.142a}$$

$$\begin{cases} {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{d} - {}^{1}_{2}\bar{m}\,{}^{0}_{1}\bar{d} = 0 \\ {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{d} + ({}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m})\,{}^{0}_{2}\bar{d} = 0 \end{cases} \tag{5.142b}$$

$$\begin{cases} {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{k} - {}^{1}_{2}\bar{m}\,{}^{0}_{1}\bar{k} = ({}^{0}_{1}\bar{m}\,{}^{0}_{2}\bar{m} + {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{m} + {}^{0}_{2}\bar{m}\,{}^{1}_{2}\bar{m})\varpi \\ {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{k} + ({}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m})\,{}^{0}_{2}\bar{k} = ({}^{0}_{1}\bar{m}\,{}^{0}_{2}\bar{m} + {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{m} + {}^{0}_{2}\bar{m}\,{}^{1}_{2}\bar{m})\varpi \end{cases} \tag{5.142c}$$

where $\varpi = {}^{1}_{2}k/{}^{0}_{2}m$ is the sole model parameter relevant for the matching condition. The first part (5.142a) is a constraint on the reference trajectory as it is independent of tunable parameters. It can be resolved by acknowledging that z_2 is a *flat output* of the system and planing the reference trajectory accordingly, i.e.

$$z_{1\mathsf{R}} = z_{2\mathsf{R}} - \ddot{z}_{2\mathsf{R}}/\varpi. \tag{5.143}$$

The other conditions can be resolved by setting

$${}^{1}_{2}\bar{k} = \tfrac{{}^{1}_{2}\bar{m}}{{}^{0}_{1}\bar{m}}\,{}^{0}_{1}\bar{k} + \tfrac{{}^{0}_{1}\bar{m}\,{}^{0}_{2}\bar{m} + {}^{0}_{1}\bar{m}\,{}^{1}_{2}\bar{m} + {}^{0}_{2}\bar{m}\,{}^{1}_{2}\bar{m}}{{}^{0}_{1}\bar{m}}\varpi, \quad {}^{0}_{2}\bar{k} = -\tfrac{{}^{1}_{2}\bar{m}}{{}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m}}\,{}^{0}_{1}\bar{k}, \quad {}^{1}_{2}\bar{d} = \tfrac{{}^{1}_{2}\bar{m}}{{}^{0}_{1}\bar{m}}\,{}^{0}_{1}\bar{d}, \quad {}^{0}_{2}\bar{d} = -\tfrac{{}^{1}_{2}\bar{m}}{{}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m}}\,{}^{0}_{1}\bar{d}, \tag{5.144}$$

which leaves the 5 tuning parameters ${}^{0}_{1}\bar{k}$, ${}^{0}_{1}\bar{d}$, ${}^{0}_{1}\bar{m}$, ${}^{0}_{2}\bar{m}$ and ${}^{1}_{2}\bar{m}$.

The resulting control law is

$$\begin{aligned} F = {}^{0}_{1}m\,\ddot{z}_{1\mathsf{R}} + {}^{1}_{2}k(z_{1\mathsf{R}} - z_{2\mathsf{R}}) + \big({}^{1}_{2}k + \tfrac{{}^{0}_{1}m({}^{1}_{2}\bar{m}\,{}^{1}_{2}k + {}^{0}_{2}m({}^{0}_{1}\bar{k} - {}^{1}_{2}\bar{k}))}{{}^{0}_{2}m({}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m})}\big)e_1 \\ - \big({}^{1}_{2}k - \tfrac{{}^{0}_{1}m({}^{1}_{2}\bar{m}\,{}^{1}_{2}k - {}^{0}_{2}m\,{}^{1}_{2}\bar{k})}{{}^{0}_{2}m({}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m})}\big)e_2 - \tfrac{{}^{0}_{1}m({}^{0}_{1}\bar{d} + {}^{1}_{2}\bar{d})}{{}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m}}\dot{e}_1 + \tfrac{{}^{0}_{1}m\,{}^{1}_{2}\bar{d}}{{}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m}}\dot{e}_2. \end{aligned} \tag{5.145}$$

Pole placement. Tuning the design parameters under the given matching conditions might not be intuitive for this example. To resolve this we can fall back to the classical approach of placing the eigenvalues of the closed loop system (5.138). Taking into account the matching condition (5.144), the characteristic polynomial of (5.138) is

$$\tfrac{\det(\bar{\boldsymbol{M}}\lambda^2 + \bar{\boldsymbol{D}}\lambda + \bar{\boldsymbol{K}})}{\det \bar{\boldsymbol{M}}} = \lambda^4 + \underbrace{\tfrac{{}^{0}_{1}\bar{d}}{{}^{0}_{1}\bar{m}}}_{p_3}\lambda^3 + \underbrace{\tfrac{{}^{0}_{1}\bar{k} + ({}^{0}_{1}\bar{m} + {}^{0}_{2}\bar{m})\varpi}{{}^{0}_{1}\bar{m}}}_{p_2}\lambda^2 + \underbrace{\tfrac{\varpi\,{}^{0}_{1}\bar{d}}{{}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m}}}_{p_1}\lambda + \underbrace{\tfrac{\varpi\,{}^{0}_{1}\bar{k}}{{}^{0}_{1}\bar{m} + {}^{1}_{2}\bar{m}}}_{p_0}. \tag{5.146}$$

This can be solved for

$${}^{0}_{1}\bar{k} = \tfrac{{}^{1}_{2}\bar{m}\,p_0 p_3}{\varpi p_3 - p_1}, \qquad {}^{0}_{1}\bar{d} = \tfrac{{}^{1}_{2}\bar{m}\,p_1 p_3}{\varpi p_3 - p_1}, \qquad {}^{0}_{1}\bar{m} = \tfrac{{}^{1}_{2}\bar{m}\,p_1}{\varpi p_3 - p_1}, \qquad {}^{0}_{2}\bar{m} = \tfrac{{}^{1}_{2}\bar{m}(p_1 p_2 - p_0 p_3 - \varpi p_1)}{\varpi(\varpi p_3 - p_1)}. \tag{5.147a}$$

and ${}^{1}_{2}\bar{m} \in \mathbb{R} \neq 0$. Choosing any Hurwitz polynomial for the coefficients p_i guarantees the asymptotic stability of the closed loop. In order to conclude $\bar{\boldsymbol{M}} > 0$, $\bar{\boldsymbol{D}} \geq 0$ and $\bar{\boldsymbol{K}} > 0$ from the Hurwitz criterion $(p_0, p_1, p_2, p_3, p_1p_2 - p_0p_3, p_1p_2p_3 - p_1^2 - p_0p_3^2 > 0)$ we need $\operatorname{sign} {}^{1}_{2}\bar{m} = \operatorname{sign}(\varpi p_3 - p_1)$.

Conclusions. The resulting controller is equivalent to one that could be designed by standard linear state-feedback methods. However, this approach here might give some *physical* insight to the resulting closed loop system. For example that the closed loop system must have an inertial coupling ($\frac{1}{2}\bar{m} \neq 0$) of the two bodies, if one wants to tune all 4 poles.

5.8.2 PVTOL

The planar vertical take off landing aircraft (PVTOL), Figure 5.16, is a common benchmark problem discussed in e.g. [Hauser et al., 1992] or [Fliess et al., 1999].

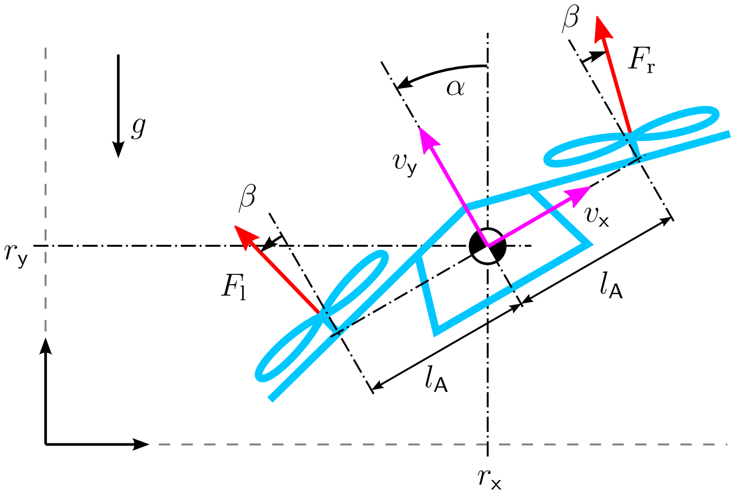


Figure 5.16: Model of the PVTOL

Model. We choose the position (r_x, r_y) of the center of mass and the tilt angle α as configuration coordinates. So the body configuration is

$$
{}^0_1\mathbf{G} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 & r_x \\ \sin\alpha & \cos\alpha & 0 & r_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.148}
$$

Note that in these coordinates the gravity acceleration has the components $\boldsymbol{a}_G = [0, -g, 0]^\top$, so the potential energy due to gravity is $\mathcal{V} = mgr_y$. In contrast to the sources mentioned above, we use the coefficients (v_x, v_y) of the absolute velocity w.r.t. the body fixed frame and the angular velocity ω_z of the body tilt as velocity coordinates. The kinematic relation is

$$
\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\begin{bmatrix} r_x \\ r_y \\ \alpha \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} v_x \\ v_y \\ \omega_z \end{bmatrix}}_{\boldsymbol{\xi}}. \tag{5.149}
$$

Let m be the total mass of the PVTOL and Θ_z the moment of inertia around the center of mass. The propeller thrusts are F_r, F_l according to Figure 5.16. Overall, we have the kinetic equation

$$
\underbrace{\begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & \Theta_z \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} \dot{v}_x \\ \dot{v}_y \\ \dot{\omega}_z \end{bmatrix}}_{\dot{\boldsymbol{\xi}}} + \underbrace{\begin{bmatrix} m(g\sin\alpha - v_y\omega_z) \\ m(g\cos\alpha + v_x\omega_z) \\ 0 \end{bmatrix}}_{\boldsymbol{b}} = \underbrace{\begin{bmatrix} \sin\beta & -\sin\beta \\ \cos\beta & \cos\beta \\ l_A\cos\beta & -l_A\cos\beta \end{bmatrix}}_{\boldsymbol{B}} \underbrace{\begin{bmatrix} F_r \\ F_l \end{bmatrix}}_{\boldsymbol{u}}. \tag{5.150}
$$

Reference trajectory. For the following we chose the left complement $\boldsymbol{B}^{\perp} = [1, 0, -\frac{\sin\beta}{l\cos\beta}]$. The condition for the reference from (5.80) for this example is

$$\lambda^{\text{ZeroError}} = m\big(\dot{v}_{\mathsf{xR}} - \underbrace{\tfrac{\Theta_{\mathsf{z}}\sin\beta}{ml_{\mathsf{A}}\cos\beta}}_{\varepsilon}\dot{\omega}_{\mathsf{zR}} - \omega_{\mathsf{zR}}v_{\mathsf{yR}} + g\sin\alpha_{\mathsf{R}}\big) = 0 \tag{5.151}$$

This can be fulfilled by parameterizing the configuration through the flat output $y_{1\mathsf{R}} = r_{\mathsf{xR}} - \varepsilon\sin\alpha_{\mathsf{R}}$, $y_{2\mathsf{R}} = r_{\mathsf{yR}} + \varepsilon\cos\alpha_{\mathsf{R}}$ (see e.g. [Fliess et al., 1999]), i.e.

$$r_{\mathsf{xR}} = y_{1\mathsf{R}} - \varepsilon\frac{\ddot{y}_{1\mathsf{R}}}{\sqrt{\ddot{y}_{1\mathsf{R}}^2 + (\ddot{y}_{1\mathsf{R}} + g)^2}}, \tag{5.152}$$

$$r_{\mathsf{yR}} = y_{2\mathsf{R}} - \varepsilon\frac{\ddot{y}_{2\mathsf{R}} - g}{\sqrt{\ddot{y}_{1\mathsf{R}}^2 + (\ddot{y}_{1\mathsf{R}} + g)^2}}, \tag{5.153}$$

$$\alpha_{\mathsf{R}} = \operatorname{atan2}(\ddot{y}_{2\mathsf{R}}, \ddot{y}_{1\mathsf{R}} + g). \tag{5.154}$$

Note that this parameterization fails if $\ddot{y}_{1\mathsf{R}} = \ddot{y}_{1\mathsf{R}} + g = 0$, i.e. the body is in free fall.

Closed loop. The PVTOL model is just a planar rigid body with particular actuation. So its closed loop template coincides with the one for the planar rigid body from subsection 5.7.4, but for symmetry reasons we choose the parameters $\bar{h}_{\mathsf{x}} = \bar{l}_{\mathsf{x}} = \bar{s}_{\mathsf{x}} = 0$.

Matching. The matching force $\boldsymbol{\lambda}$ from (5.78) with the orthogonal complement from above takes a rather cumbersome form and is not given explicitly here. Instead we will investigate its linear approximation about any reference trajectory with $\alpha_{\mathsf{R}} = 0$: The matrices of the linearized model and desired closed loop are

$$\boldsymbol{M}_0 = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & J \end{bmatrix}, \qquad \bar{\boldsymbol{M}}_0 = \begin{bmatrix} \bar{m} & 0 & -\bar{m}\bar{s}_{\mathsf{y}} \\ 0 & \bar{m} & 0 \\ -\bar{m}\bar{s}_{\mathsf{y}} & 0 & \bar{\Theta}_{\mathsf{z}} + \bar{m}\bar{s}_{\mathsf{y}}^2 \end{bmatrix}, \tag{5.155a}$$

$$\boldsymbol{D}_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \bar{\boldsymbol{D}}_0 = \begin{bmatrix} \bar{d} & 0 & -\bar{d}\bar{l}_{\mathsf{y}} \\ 0 & \bar{d} & 0 \\ -\bar{d}\bar{l}_{\mathsf{y}} & 0 & \bar{\Upsilon}_{\mathsf{z}} + \bar{d}\bar{l}_{\mathsf{y}}^2 \end{bmatrix}, \tag{5.155b}$$

$$\boldsymbol{K}_0 = \begin{bmatrix} 0 & 0 & mg \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \bar{\boldsymbol{K}}_0 = \begin{bmatrix} \bar{k} & 0 & -\bar{k}\bar{h}_{\mathsf{y}} \\ 0 & \bar{k} & 0 \\ -\bar{k}\bar{h}_{\mathsf{y}} & 0 & \bar{\Pi}_{\mathsf{z}} + \bar{k}\bar{h}_{\mathsf{y}}^2 \end{bmatrix}. \tag{5.155c}$$

The conditions for $(\boldsymbol{B}^{\perp})^{\top}\big(\boldsymbol{M}_0\bar{\boldsymbol{M}}_0^{-1}\bar{\boldsymbol{D}}_0 - \boldsymbol{D}_0\big) = \boldsymbol{0}$ and $(\boldsymbol{B}^{\perp})^{\top}\big(\boldsymbol{M}_0\bar{\boldsymbol{M}}_0^{-1}\bar{\boldsymbol{K}}_0 - \boldsymbol{K}_0\big) = \boldsymbol{0}$ for the linearized matching force from (5.82) to vanish are equivalent to

$$\bar{k}\big(\bar{\Theta}_{\mathsf{z}} - \bar{m}(\bar{h}_{\mathsf{y}} - \bar{s}_{\mathsf{y}})(\bar{s}_{\mathsf{y}} - \varepsilon)\big) = 0, \tag{5.156a}$$

$$\bar{m}\bar{\Pi}_{\mathsf{z}}(\bar{s}_{\mathsf{y}} - \varepsilon) - \bar{k}\bar{h}_{\mathsf{y}}\big(\bar{\Theta}_{\mathsf{z}} - \bar{m}(\bar{h}_{\mathsf{y}} - \bar{s}_{\mathsf{y}})(\bar{s}_{\mathsf{y}} - \varepsilon)\big) = \bar{\Theta}_{\mathsf{z}}\bar{m}g, \tag{5.156b}$$

$$\bar{d}\big(\bar{\Theta}_{\mathsf{z}} - \bar{m}(\bar{l}_{\mathsf{y}} - \bar{s}_{\mathsf{y}})(\bar{s}_{\mathsf{y}} - \varepsilon)\big) = 0, \tag{5.156c}$$

$$\bar{m}\bar{\Upsilon}_{\mathsf{z}}(\bar{s}_{\mathsf{y}} - \varepsilon) - \bar{d}\bar{l}_{\mathsf{y}}\big(\bar{\Theta}_{\mathsf{z}} - \bar{m}(\bar{l}_{\mathsf{y}} - \bar{s}_{\mathsf{y}})(\bar{s}_{\mathsf{y}} - \varepsilon)\big) = 0. \tag{5.156d}$$

One solution for this is

$$\bar{\Theta}_{\mathsf{z}} = \bar{m}(\bar{h}_{\mathsf{y}} - \bar{s}_{\mathsf{y}})(\bar{s}_{\mathsf{y}} - \varepsilon), \qquad \bar{l}_{\mathsf{y}} = \bar{h}_{\mathsf{y}}, \qquad \bar{\Upsilon}_{\mathsf{z}} = 0, \qquad \bar{\Pi}_{\mathsf{z}} = \bar{m}g(\bar{h}_{\mathsf{y}} - \bar{s}_{\mathsf{y}}) \tag{5.157}$$

which leaves the parameters $\bar{k}, \bar{d}, \bar{m}, \bar{h}_{\mathsf{y}}, \bar{s}_{\mathsf{y}}$ for tuning. The resulting matching force is

$$\tilde{\boldsymbol{b}} = \frac{\bar{m}(\bar{h}_{\mathsf{y}} - \bar{s}_{\mathsf{y}})}{m(\bar{h}_{\mathsf{y}} - \varepsilon)} \begin{bmatrix} 1 \\ 0 \\ -\varepsilon \end{bmatrix} \lambda \tag{5.158}$$

where λ for the corresponding approach is

$$\lambda^{\text{ParticleBased}} = m(a_{\mathsf{yR}} - \varepsilon\omega_{\mathsf{zR}}^2)\sin\alpha_{\mathrm{E}} \tag{5.159a}$$
$$\lambda^{\text{BodyBased}} = m(a_{\mathsf{yR}}\sin\alpha_{\mathrm{E}} - r_{\mathrm{xE}}\omega_{\mathsf{zR}}^2 + 2v_{\mathrm{yE}}\omega_{\mathsf{zR}} + (\varepsilon(1-\cos\alpha_{\mathrm{E}}) - r_{\mathrm{yE}})\dot{\omega}_{\mathsf{zR}}) \tag{5.159b}$$
$$\lambda^{\text{EnergyBased}} = m(a_{\mathsf{yR}}\sin\alpha_{\mathrm{E}} - r_{\mathrm{xE}}\omega_{\mathsf{zR}}^2 + \ v_{\mathrm{yE}}\omega_{\mathsf{zR}} + (\varepsilon(1-\cos\alpha_{\mathrm{E}}) - r_{\mathrm{yE}})\dot{\omega}_{\mathsf{zR}}) \tag{5.159c}$$

where $a_{\mathsf{yR}} = \dot{v}_{\mathsf{yR}} + v_{\mathsf{xR}}\omega_{\mathsf{zR}} + g(\cos\alpha_{\mathsf{R}} - 1)$.

Tuning. For the energies to be positive (semi) definite, we need $\bar{\boldsymbol{M}} > 0$, $\bar{\boldsymbol{D}} \geq 0$ and $\bar{\mathcal{V}} > 0$ which means for the remaining tuning parameters

$$\bar{k}, \bar{d}, \bar{m} > 0, \qquad \bar{h}_{\mathsf{y}} > \bar{s}_{\mathsf{y}} > \varepsilon. \tag{5.160}$$

The characteristic polynomial of the linearized system is

$$\begin{aligned} &\det(\bar{\boldsymbol{M}}\lambda^2 + \bar{\boldsymbol{D}}\lambda + \bar{\boldsymbol{K}})/\det\bar{\boldsymbol{M}} \\ &\quad = (\lambda^2 + \tfrac{\bar{d}}{\bar{m}}\lambda + \tfrac{\bar{k}}{\bar{m}})\big(\lambda^4 + \tfrac{\bar{d}(\bar{h}_{\mathsf{y}}-\varepsilon)}{\bar{m}(\bar{s}_{\mathsf{y}}-\varepsilon)}\lambda^3 + \tfrac{\bar{k}(\bar{h}_{\mathsf{y}}-\varepsilon)+\bar{m}g}{\bar{m}(\bar{s}_{\mathsf{y}}-\varepsilon)}\lambda^2 + \tfrac{\bar{d}g}{\bar{m}(\bar{s}_{\mathsf{y}}-\varepsilon)}\lambda + \tfrac{\bar{k}g}{\bar{m}(\bar{s}_{\mathsf{y}}-\varepsilon)}\big) \end{aligned} \tag{5.161}$$

Set a desired polynomial of forth degree $\lambda^4 + p_3\lambda^3 + p_2\lambda^2 + p_1\lambda + p_0$ we get the parameters

$$\bar{k} = \frac{\bar{m}\,p_0p_1}{p_1p_2 - p_0p_3}, \quad \bar{h}_{\mathsf{y}} = \varepsilon + \frac{g\,p_3}{p_1}, \quad \bar{\Theta}_{\mathsf{z}} = \frac{\bar{m}g^2(p_1p_2p_3 - p_1^2 - p_0p_3^2)}{(p_1p_2 - p_0p_3)^2}, \tag{5.162}$$
$$\bar{d} = \frac{\bar{m}\,p_1^2}{p_1p_2 - p_0p_3}, \quad \bar{s}_{\mathsf{y}} = \varepsilon + \frac{g\,p_1}{p_1p_2 - p_0p_3}, \quad \bar{\Pi}_{\mathsf{z}} = \frac{\bar{m}g^2(p_1p_2p_3 - p_1^2 - p_0p_3)}{p_1(p_1p_2 - p_0p_3)}, \tag{5.163}$$

Note that the Hurwitz criterion ($p_0, p_1, p_2, p_3, p_1p_2 - p_0p_3, p_1p_2p_3 - p_1^2 - p_0p_3^2 > 0$) implies the positive definiteness of the inertia $\bar{\boldsymbol{M}}$ and stiffness matrix $\bar{\boldsymbol{K}}$. Even though the damping matrix $\bar{\boldsymbol{D}}$ is only positive semidefinite the *local* attractiveness of the nonlinear system can be concluded by Lyapunov's indirect method.

Mechanical interpretation.

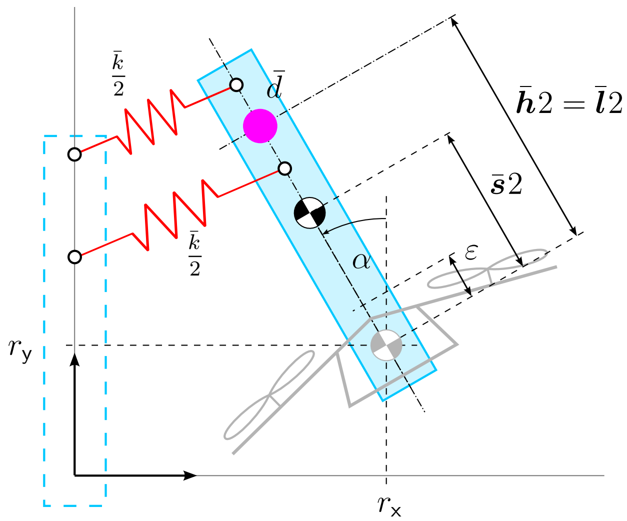


Figure 5.17: Interpretation of the controlled PVTOL as a mechanical system

5.8.3 Quadcopter

Consider a Quadcopter Figure 5.18

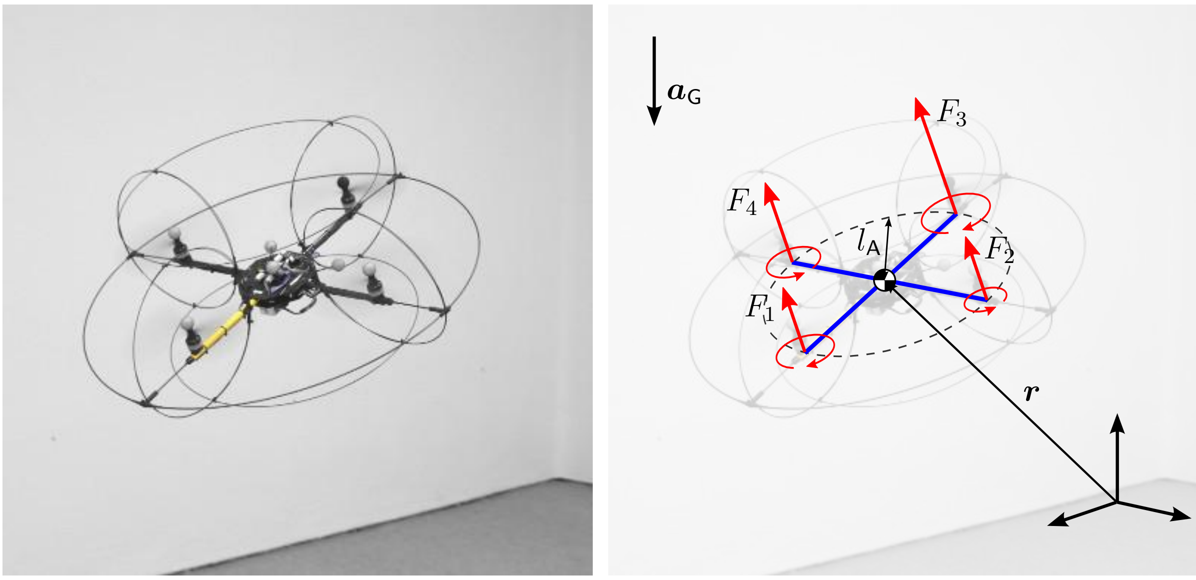


Figure 5.18: Model of the Quadcopter

Model. Kinematics

$$\dot{\boldsymbol{r}} = \boldsymbol{R}\boldsymbol{v}, \quad \dot{\boldsymbol{R}} = \boldsymbol{R}\,\mathrm{wed}(\boldsymbol{\omega}) \tag{5.164}$$

Kinetics

$$\underbrace{\begin{bmatrix} m & 0 & 0 & 0 & 0 & 0 \\ 0 & m & 0 & 0 & 0 & 0 \\ 0 & 0 & m & 0 & 0 & 0 \\ 0 & 0 & 0 & \Theta_\mathsf{y} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Theta_\mathsf{x} & 0 \\ 0 & 0 & 0 & 0 & 0 & \Theta_\mathsf{z} \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} \dot{v}_\mathsf{x} \\ \dot{v}_\mathsf{y} \\ \dot{v}_\mathsf{z} \\ \dot{\omega}_\mathsf{x} \\ \dot{\omega}_\mathsf{y} \\ \dot{\omega}_\mathsf{z} \end{bmatrix}}_{\dot{\boldsymbol{\xi}}} + \underbrace{\begin{bmatrix} m(v_\mathsf{y}\omega_\mathsf{z} - v_\mathsf{z}\omega_\mathsf{y} + R^\mathsf{z}_\mathsf{x} g) \\ m(v_\mathsf{x}\omega_\mathsf{z} - v_\mathsf{z}\omega_\mathsf{x} + R^\mathsf{z}_\mathsf{y} g) \\ m(v_\mathsf{y}\omega_\mathsf{x} - v_\mathsf{x}\omega_\mathsf{y} + R^\mathsf{z}_\mathsf{z} g) \\ (\Theta_\mathsf{z} - \Theta_\mathsf{y})\omega_\mathsf{y}\omega_\mathsf{z} \\ (\Theta_\mathsf{x} - \Theta_\mathsf{z})\omega_\mathsf{x}\omega_\mathsf{z} \\ (\Theta_\mathsf{y} - \Theta_\mathsf{x})\omega_\mathsf{x}\omega_\mathsf{y} \end{bmatrix}}_{\boldsymbol{b}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & l_\mathsf{A} & 0 & -l_\mathsf{A} \\ -l_\mathsf{A} & 0 & l_\mathsf{A} & 0 \\ -b_\mathsf{F} & b_\mathsf{F} & -b_\mathsf{F} & b_\mathsf{F} \end{bmatrix}}_{\boldsymbol{B}} \underbrace{\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix}}_{\boldsymbol{u}} \tag{5.165}$$

Closed loop template. The quadcopter is just a free rigid body with a particular actuation, so the closed loop templates coincide with the ones given in subsection 5.1.2, subsection 5.2.1 and subsection 5.3.3. Due to symmetry considerations we set $\bar{s}_\mathsf{x} = \bar{s}_\mathsf{y} = 0$, $\bar{\Theta}_\mathsf{xx} = \bar{\Theta}_\mathsf{yy}$, $\bar{\Theta}_\mathsf{xy} = \bar{\Theta}_\mathsf{xz} = \bar{\Theta}_\mathsf{yz} = 0$ and analog for the stiffness and damping parameters.

Matching An obvious left complement for $\boldsymbol{B}$ is

$$\boldsymbol{B}^{\perp} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{5.166}$$

As the actual matching force $\boldsymbol{\lambda}$ is extremely cumbersome, we will consider its first order approximation first: The reference part is

$$\boldsymbol{\lambda}^{\text{ZeroError}} = m \begin{bmatrix} \dot{v}_{\mathsf{xR}} + v_{\mathsf{zR}}\omega_{\mathsf{yR}} - v_{\mathsf{yR}}\omega_{\mathsf{zR}} + R^{\mathsf{z}}_{\mathsf{xR}} g \\ \dot{v}_{\mathsf{yR}} + v_{\mathsf{xR}}\omega_{\mathsf{zR}} - v_{\mathsf{zR}}\omega_{\mathsf{xR}} + R^{\mathsf{z}}_{\mathsf{yR}} g \end{bmatrix} \tag{5.167}$$

which is just the first two coefficients of (5.165). Combining this with the kinematic relation $\dot{\boldsymbol{x}}_{\mathsf{R}} = \boldsymbol{A}(\boldsymbol{x}_{\mathsf{R}})\boldsymbol{\xi}_{\mathsf{R}}$ .. flat output.. see [Konz and Rudolph, 2013]

With the linearized system matrices

$$\boldsymbol{M}_0 = \begin{bmatrix} m & 0 & 0 & 0 & 0 & 0 \\ 0 & m & 0 & 0 & 0 & 0 \\ 0 & 0 & m & 0 & 0 & 0 \\ 0 & 0 & 0 & \Theta_{\mathsf{y}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Theta_{\mathsf{x}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \Theta_{\mathsf{z}} \end{bmatrix}, \quad \bar{\boldsymbol{M}}_0 = \begin{bmatrix} \bar{m} & 0 & 0 & 0 & \bar{m}\bar{s}_{\mathsf{z}} & 0 \\ 0 & \bar{m} & 0 & -\bar{m}\bar{s}_{\mathsf{z}} & 0 & 0 \\ 0 & 0 & \bar{m} & 0 & 0 & 0 \\ 0 & -\bar{m}\bar{s}_{\mathsf{z}} & 0 & \bar{\Theta}_{\mathsf{y}} + \bar{m}\bar{s}_{\mathsf{z}}^2 & 0 & 0 \\ \bar{m}\bar{s}_{\mathsf{z}} & 0 & 0 & 0 & \bar{\Theta}_{\mathsf{x}} + \bar{m}\bar{s}_{\mathsf{z}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bar{\Theta}_{\mathsf{z}} \end{bmatrix},$$

$$\boldsymbol{D}_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \bar{\boldsymbol{D}}_0 = \begin{bmatrix} \bar{d} & 0 & 0 & 0 & \bar{d}\bar{l}_{\mathsf{z}} & 0 \\ 0 & \bar{d} & 0 & -\bar{d}\bar{l}_{\mathsf{z}} & 0 & 0 \\ 0 & 0 & \bar{d} & 0 & 0 & 0 \\ 0 & -\bar{d}\bar{l}_{\mathsf{z}} & 0 & \bar{\Upsilon}_{\mathsf{y}} + \bar{d}\bar{l}_{\mathsf{z}}^2 & 0 & 0 \\ \bar{d}\bar{l}_{\mathsf{z}} & 0 & 0 & 0 & \bar{\Upsilon}_{\mathsf{x}} + \bar{d}\bar{l}_{\mathsf{z}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bar{\Upsilon}_{\mathsf{z}} \end{bmatrix},$$

$$\boldsymbol{K}_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & -mg & 0 \\ 0 & 0 & 0 & mg & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \bar{\boldsymbol{K}}_0 = \begin{bmatrix} \bar{k} & 0 & 0 & 0 & \bar{k}\bar{h}_{\mathsf{z}} & 0 \\ 0 & \bar{k} & 0 & -\bar{k}\bar{h}_{\mathsf{z}} & 0 & 0 \\ 0 & 0 & \bar{k} & 0 & 0 & 0 \\ 0 & -\bar{k}\bar{h}_{\mathsf{z}} & 0 & \bar{\Pi}_{\mathsf{y}} + \bar{k}\bar{h}_{\mathsf{z}}^2 & 0 & 0 \\ \bar{k}\bar{h}_{\mathsf{z}} & 0 & 0 & 0 & \bar{\Pi}_{\mathsf{x}} + \bar{k}\bar{h}_{\mathsf{z}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bar{\Pi}_{\mathsf{z}} \end{bmatrix}.$$

the linearized matching condition is

$$\bar{k}(\bar{\Theta}_{\mathsf{x}} - \bar{m}\bar{s}_{\mathsf{z}}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})) = 0 \tag{5.168a}$$
$$\bar{m}(\bar{\Theta}_{\mathsf{x}} g - \bar{\Pi}_{\mathsf{x}}\bar{s}_{\mathsf{z}}) + \bar{k}\bar{h}_{\mathsf{z}}(\bar{\Theta}_{\mathsf{x}} - \bar{m}\bar{s}_{\mathsf{z}}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})) = 0 \tag{5.168b}$$
$$\bar{d}(\bar{\Theta}_{\mathsf{x}} - \bar{m}\bar{s}_{\mathsf{z}}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})) = 0 \tag{5.168c}$$
$$-\bar{m}\bar{s}_{\mathsf{z}}\bar{\Upsilon}_{\mathsf{x}} + \bar{d}\bar{l}_{\mathsf{z}}(\bar{\Theta}_{\mathsf{x}} - \bar{m}\bar{s}_{\mathsf{z}}(\bar{l}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})) = 0 \tag{5.168d}$$

These terms vanish for

$$\bar{l}_{\mathsf{z}} = \bar{h}_{\mathsf{z}}, \qquad \bar{\Theta}_{\mathsf{x}} = \bar{m}\bar{s}_{\mathsf{z}}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}}), \qquad \bar{\Upsilon}_{\mathsf{x}} = 0, \qquad \bar{\Pi}_{\mathsf{x}} = \bar{m}g(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}}). \tag{5.169}$$

Even with these the ...

$$\tilde{\boldsymbol{b}} = \frac{1}{m\bar{h}_\mathsf{z}} \begin{bmatrix} \bar{\Theta}_\mathsf{z}\omega_\mathsf{x}\omega_\mathsf{z} - \frac{1}{2}\bar{\Pi}_\mathsf{z}(R^\mathsf{x}_\mathsf{z} + R^\mathsf{z}_\mathsf{x}) \\ \bar{\Theta}_\mathsf{z}\omega_\mathsf{y}\omega_\mathsf{z} - \frac{1}{2}\bar{\Pi}_\mathsf{z}(R^\mathsf{y}_{\mathsf{z}\mathrm{E}} + R^\mathsf{z}_{\mathsf{y}\mathrm{E}}) \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.170}$$

5.8.4 Bicopter

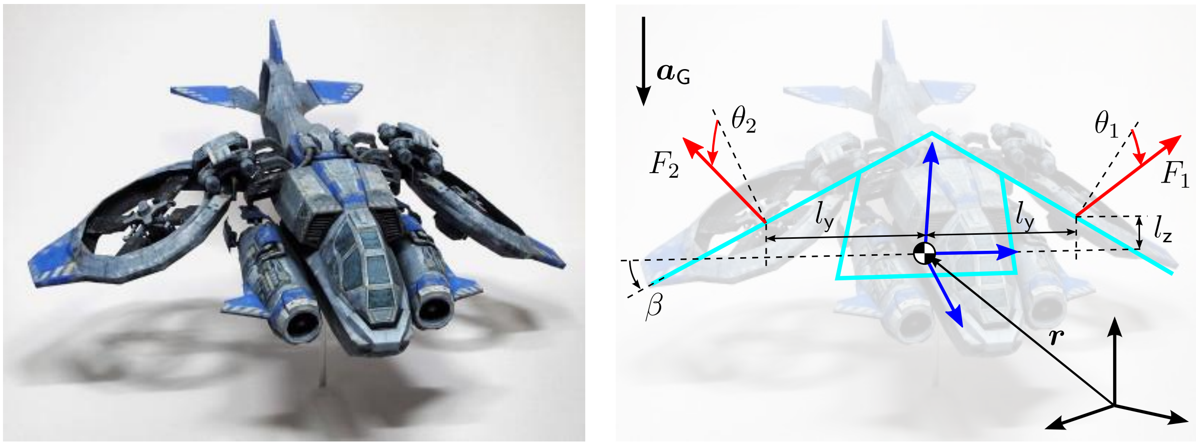


Figure 5.19: Model of a bicopter (background image: www.poppaper.net/80155164809)

Equations of motion. The bicopter considered here is a single rigid body with two tiltable propellers as illustrated in Figure 5.18. With the same coordinates as for the previous examples, the equations of motion are identical as well up to the generalized force from the propellers

$$\boldsymbol{f}^{\mathrm{U}} = \underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & \sin\beta_{\mathsf{F}} & 0 & -\sin\beta_{\mathsf{F}} \\ 0 & \cos\beta_{\mathsf{F}} & 0 & \cos\beta_{\mathsf{F}} \\ 0 & l_{\mathsf{y}}\cos\beta_{\mathsf{F}} - l_{\mathsf{z}}\sin\beta_{\mathsf{F}} & 0 & -l_{\mathsf{y}}\cos\beta_{\mathsf{F}} + l_{\mathsf{z}}\sin\beta_{\mathsf{F}} \\ l_{\mathsf{z}} & 0 & l_{\mathsf{z}} & 0 \\ -l_{\mathsf{y}} & 0 & l_{\mathsf{y}} & 0 \end{bmatrix}}_{\boldsymbol{B}} \underbrace{\begin{bmatrix} F_1\sin\theta_1 \\ F_1\cos\theta_1 \\ F_2\sin\theta_2 \\ F_2\cos\theta_2 \end{bmatrix}}_{\boldsymbol{u}}. \tag{5.171}$$

The transformation of the actual control inputs F_1, F_2 and θ_1, θ_2 to a auxiliary input $\boldsymbol{u}$ is used to achieve the linear form $\boldsymbol{f}^{\mathrm{U}} = \boldsymbol{B}\boldsymbol{u}$. Within the input constraints $2\,\mathrm{N} \leq F_i \leq 14\,\mathrm{N}$, $-30^\circ \leq \theta_i \leq 30^\circ$, $i = 1, 2$ this transformation is bijective. To account for the original constraints in transformed input $\boldsymbol{u} \in \mathbb{R}^4$ a convex approximation illustrated in Figure 5.20 is used. These constraints can be written in the required form $\boldsymbol{W}\boldsymbol{u} \leq \boldsymbol{l}$.

Closed loop template. Mechanically the bicopter is just a free rigid body, so the control template from (??) is reasonable. Due to symmetry of the mechanical model we set the center of mass on the body fixed vertical axis and set the principle axis of inertia to be parallel to the body fixed axis, i.e. $\bar{s}_{\mathsf{x}} = \bar{s}_{\mathsf{y}} = 0$ and $\bar{\boldsymbol{\Theta}} = \mathrm{diag}(\bar{\Theta}_{\mathsf{x}}, \bar{\Theta}_{\mathsf{y}}, \bar{\Theta}_{\mathsf{z}})$ and the same for damping and stiffness.

Matching

$$\boldsymbol{B}^{\perp} = \begin{bmatrix} l_{\mathsf{z}} & 0 & 0 & 0 & -1 & 0 \\ 0 & l_{\mathsf{y}}\cos\beta_{\mathsf{F}} - l_{\mathsf{z}}\sin\beta_{\mathsf{F}} & 0 & \sin\beta_{\mathsf{F}} & 0 & 0 \end{bmatrix} \tag{5.172}$$

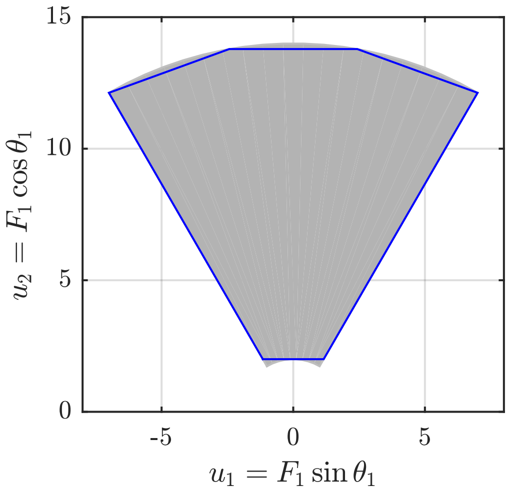


Figure 5.20: Approximation of the Bicopter input constraints

constraints on parameters

$$\bar{l}_{\mathsf{z}} = \bar{h}_{\mathsf{z}}, \tag{5.173}$$

$$\bar{\Theta}_{\mathsf{x}} = \bar{m}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})(\bar{s}_{\mathsf{z}} - \varepsilon_{\mathsf{y}}), \qquad \bar{\Upsilon}_{\mathsf{x}} = 0, \qquad \bar{\Pi}_{\mathsf{x}} = \bar{m}g(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}}), \tag{5.174}$$

$$\bar{\Theta}_{\mathsf{y}} = \bar{m}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})(\bar{s}_{\mathsf{z}} - \varepsilon_{\mathsf{x}}), \qquad \bar{\Upsilon}_{\mathsf{y}} = 0, \qquad \bar{\Pi}_{\mathsf{y}} = \bar{m}g(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}}). \tag{5.175}$$

where

$$\varepsilon_{\mathsf{x}} = -\frac{\Theta_{\mathsf{y}}}{m l_{\mathsf{z}}}, \qquad \varepsilon_{\mathsf{y}} = \frac{\Theta_{\mathsf{x}} \sin\beta_{\mathsf{F}}}{m(l_{\mathsf{y}}\cos\beta_{\mathsf{F}} - l_{\mathsf{z}}\sin\beta_{\mathsf{F}})}. \tag{5.176}$$

This leaves the tuning parameters $\bar{m}, \bar{d}, \bar{k}, \bar{s}_{\mathsf{z}}, \bar{h}_{\mathsf{z}}$ and $\bar{\Theta}_{\mathsf{z}}, \bar{\Upsilon}_{\mathsf{z}}, \bar{\Pi}_{\mathsf{z}}$.

The remaining matching force is

$$\tilde{\boldsymbol{f}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & -\varepsilon_{\mathsf{y}} \\ \varepsilon_{\mathsf{x}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\bar{h}_{\mathsf{z}} - \varepsilon_{\mathsf{x}}}\left(\bar{\Theta}_{\mathsf{z}} - \bar{m}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})(\frac{\Theta_{\mathsf{x}} - \Theta_{\mathsf{z}}}{\Theta_{\mathsf{y}}}\varepsilon_{\mathsf{x}} - \varepsilon_{\mathsf{y}})\omega_{\mathsf{x}}\omega_{\mathsf{z}} - \frac{\bar{\Pi}_{\mathsf{z}}}{2}(R^{\mathsf{x}}_{\mathsf{z}} + R^{\mathsf{z}}_{\mathsf{x}})\right) \\ \frac{1}{\bar{h}_{\mathsf{z}} - \varepsilon_{\mathsf{y}}}\left(\bar{\Theta}_{\mathsf{z}} - \bar{m}(\bar{h}_{\mathsf{z}} - \bar{s}_{\mathsf{z}})(\frac{\Theta_{\mathsf{y}} - \Theta_{\mathsf{z}}}{\Theta_{\mathsf{x}}}\varepsilon_{\mathsf{y}} - \varepsilon_{\mathsf{x}})\omega_{\mathsf{x}}\omega_{\mathsf{z}} - \frac{\bar{\Pi}_{\mathsf{z}}}{2}(R^{\mathsf{y}}_{\mathsf{z}} + R^{\mathsf{z}}_{\mathsf{y}})\right) \end{bmatrix} \tag{5.177}$$

Simulation result.

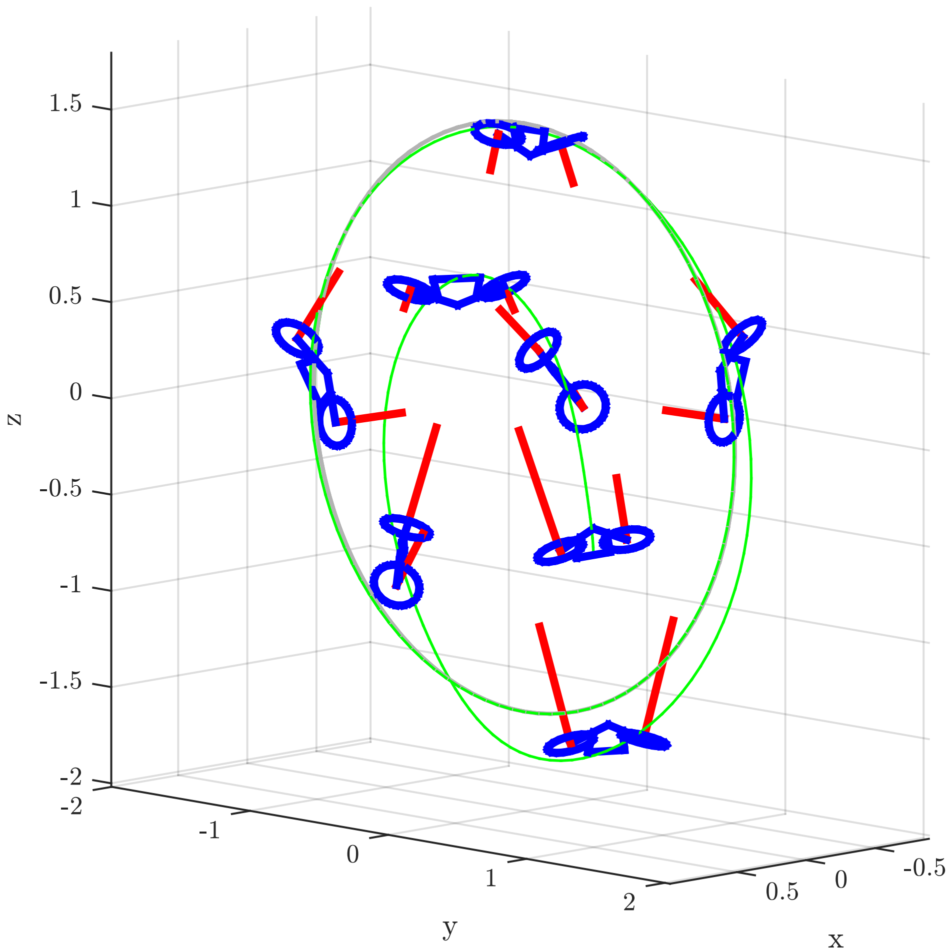


Figure 5.21: Simulation result for the Bicopter

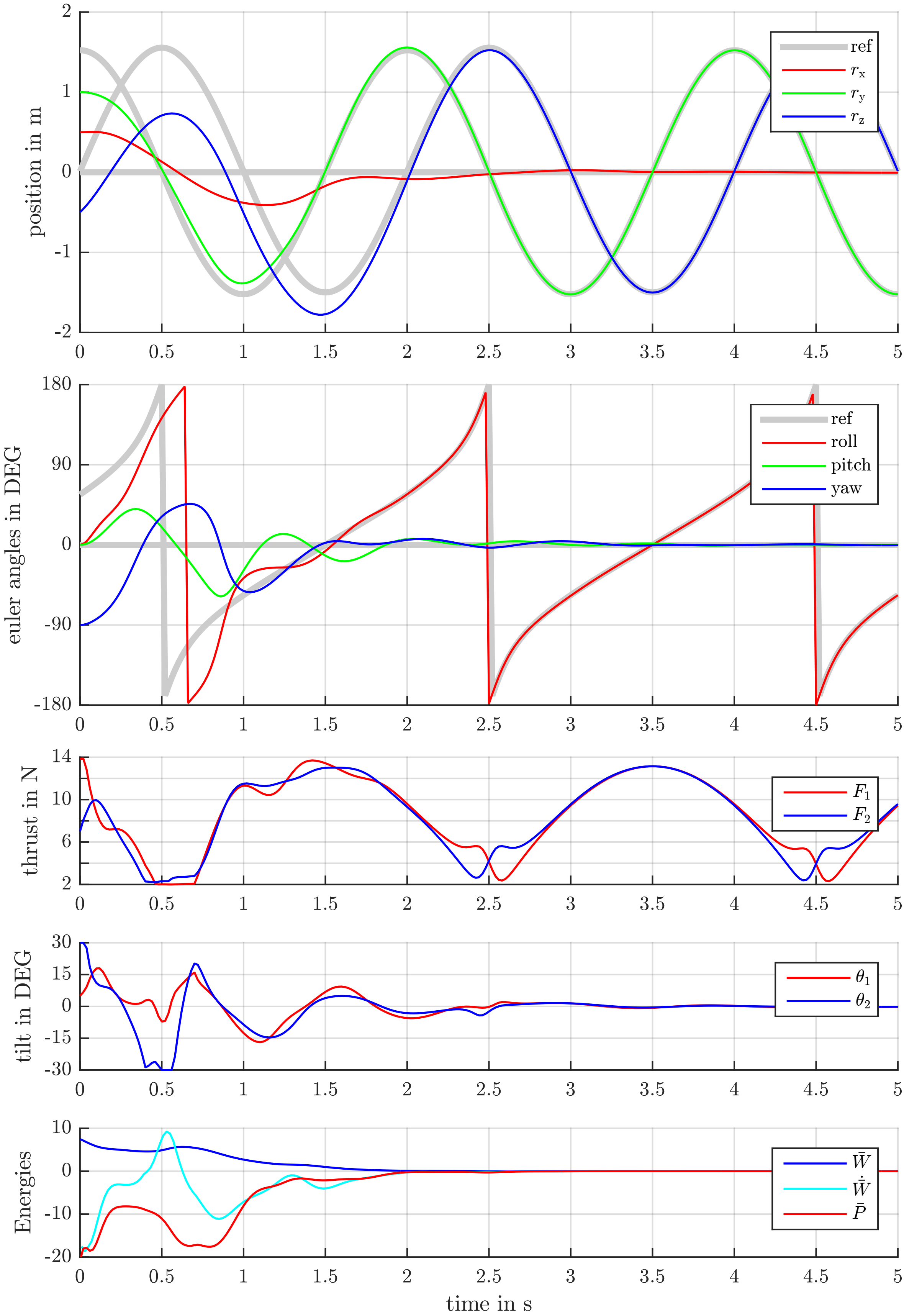


Figure 5.22: Simulation result for the Bicopter

Appendix A

Some math

A.1 Linear Algebra

A.1.1 Matrix sets

Define the following sets of real matrices that are frequently used in the work:

$$\begin{aligned}
&\text{(symmetric)} & \mathbb{SYM}(n) &= \{\boldsymbol{A} \in \mathbb{R}^{n\times n} \,|\, \boldsymbol{A} = \boldsymbol{A}^\top\}, && \text{(A.1a)}\\
&\text{(symmetric, pos. def.)} & \mathbb{SYM}^+(n) &= \{\boldsymbol{A} \in \mathbb{SYM}(n) \,|\, \boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} > 0 \,\forall\, \boldsymbol{x} \in \mathbb{R}^n\backslash\{\boldsymbol{0}\}\}, && \text{(A.1b)}\\
&\text{(sym., pos. semi-def.)} & \mathbb{SYM}_0^+(n) &= \{\boldsymbol{A} \in \mathbb{SYM}(n) \,|\, \boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} \geq 0 \,\forall\, \boldsymbol{x} \in \mathbb{R}^n\backslash\{\boldsymbol{0}\}\}, && \text{(A.1c)}
\end{aligned}$$

$$\begin{aligned}
&\text{(unit sphere)} & \mathbb{S}^n &= \{\boldsymbol{a} \in \mathbb{R}^n \,|\, \boldsymbol{a}^\top\boldsymbol{a} = 1\}, && \text{(A.1d)}\\
&\text{(orthogonal)} & \mathbb{O}(n) &= \{\boldsymbol{A} \in \mathbb{R}^{n\times n} \,|\, \boldsymbol{A}^{-1} = \boldsymbol{A}^\top\}, && \text{(A.1e)}\\
&\text{(special orthogonal)} & \mathbb{SO}(n) &= \{\boldsymbol{R} \in \mathbb{O}(n) \,|\, \det \boldsymbol{R} = +1\}, && \text{(A.1f)}\\
&\text{(special Euclidean)} & \mathbb{SE}(n) &= \left\{ \begin{bmatrix} \boldsymbol{R} & \boldsymbol{r} \\ \boldsymbol{0} & 1 \end{bmatrix} \,\middle|\, \boldsymbol{r} \in \mathbb{R}^n, \boldsymbol{R} \in \mathbb{SO}(n) \right\}, && \text{(A.1g)}
\end{aligned}$$

$$\begin{aligned}
&\text{(skew symmetric)} & \mathfrak{so}(n) &= \{\boldsymbol{\Omega} \in \mathbb{R}^{n\times n} \,|\, \boldsymbol{\Omega}^\top = -\boldsymbol{\Omega}\}, && \text{(A.1h)}\\
& & \mathfrak{se}(n) &= \left\{ \begin{bmatrix} \boldsymbol{\Omega} & \boldsymbol{v} \\ \boldsymbol{0} & 0 \end{bmatrix} \,|\, \boldsymbol{\Omega} \in \mathfrak{so}(n), \boldsymbol{v} \in \mathbb{R}^n \right\}. && \text{(A.1i)}
\end{aligned}$$

A.1.2 Trace

The *trace* of a quadratic matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is the sum of its diagonal entries

$$\operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} A_{ii} \tag{A.2}$$

Some properties of the trace are

$$\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n},\lambda\in\mathbb{R}\ :\qquad \operatorname{tr}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{tr}\boldsymbol{A}+\operatorname{tr}\boldsymbol{B}, \tag{A.3a}$$

$$\operatorname{tr}(\lambda\boldsymbol{A})=\lambda\operatorname{tr}\boldsymbol{A}, \tag{A.3b}$$

$$\operatorname{tr}\boldsymbol{A}^\top=\operatorname{tr}\boldsymbol{A}, \tag{A.3c}$$

$$\boldsymbol{A}\in\mathbb{R}^{n\times m},\boldsymbol{B}\in\mathbb{R}^{m\times n}\ :\qquad \operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\operatorname{tr}(\boldsymbol{B}\boldsymbol{A}), \tag{A.3d}$$

$$\boldsymbol{A},\boldsymbol{P}\in\mathbb{R}^{n\times n},\det\boldsymbol{P}\neq 0\ :\qquad \operatorname{tr}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=\operatorname{tr}\boldsymbol{A}. \tag{A.3e}$$

$$\boldsymbol{K}\in\mathbb{SYM}(n),\boldsymbol{\Omega}\in\mathfrak{so}(n)\ :\qquad \operatorname{tr}(\boldsymbol{K}\boldsymbol{\Omega})=0. \tag{A.3f}$$

For matrices $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}, n\geq 2$ we may define the bijective mapping

$$\boldsymbol{B}=\operatorname{tr}(\boldsymbol{A})\mathbf{I}_n-\boldsymbol{A}\qquad\Leftrightarrow\qquad \boldsymbol{A}=\tfrac{1}{n-1}\operatorname{tr}(\boldsymbol{B})\mathbf{I}_n-\boldsymbol{B}. \tag{A.4}$$

which is due to $\operatorname{tr}\boldsymbol{B}=(n-1)\operatorname{tr}\boldsymbol{A}$.

A.1.3 Vee & wedge operatos

Define

$$\operatorname{wed}:\ \mathbb{R}^3\to\mathfrak{so}(3)\ :\ \begin{bmatrix}\omega_1\\\omega_2\\\omega_3\end{bmatrix}\mapsto\begin{bmatrix}0&-\omega_3&\omega_2\\\omega_3&0&-\omega_1\\-\omega_2&\omega_1&0\end{bmatrix}, \tag{A.5a}$$

$$\operatorname{wed}:\ \mathbb{R}^6\to\mathfrak{se}(3)\ :\ \begin{bmatrix}\boldsymbol{v}\\\boldsymbol{\omega}\end{bmatrix}\mapsto\begin{bmatrix}\operatorname{wed}\boldsymbol{\omega}&\boldsymbol{v}\\\mathbf{0}&0\end{bmatrix}, \tag{A.5b}$$

$$\operatorname{Vee}:\mathbb{R}^{3\times 3}\to\mathbb{R}^{3\times 3}\ :\ \boldsymbol{A}\mapsto\operatorname{tr}(\boldsymbol{A})\mathbf{I}_3-\boldsymbol{A}, \tag{A.6a}$$

$$\operatorname{Vee}:\mathbb{R}^{4\times 4}\to\mathbb{R}^{6\times 6}\ :\ \begin{bmatrix}\boldsymbol{A}&\boldsymbol{b}\\\boldsymbol{c}^\top&d\end{bmatrix}\mapsto\begin{bmatrix}d\mathbf{I}_3&(\operatorname{wed}\boldsymbol{b})^\top\\\operatorname{wed}\boldsymbol{c}&\operatorname{Vee}\boldsymbol{A}\end{bmatrix}, \tag{A.6b}$$

$$\boldsymbol{A},\boldsymbol{P}\in\mathbb{R}^{n\times n},\det\boldsymbol{P}\neq 0\ :\qquad \operatorname{Vee}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=\boldsymbol{P}^{-1}\operatorname{Vee}(\boldsymbol{A})\boldsymbol{P} \tag{A.7}$$

A.1.4 Inner product

Inner product. For matrices $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times m}$ and a symmetric, positive definite matrix $\boldsymbol{K}\in\mathbb{SYM}^+(n)$, define an *inner product* as

$$\langle\boldsymbol{A},\boldsymbol{B}\rangle_{\boldsymbol{K}}=\operatorname{tr}(\boldsymbol{A}^\top\boldsymbol{K}\boldsymbol{B}). \tag{A.8}$$

For $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\in\mathbb{R}^{n\times m}$ and $\lambda\in\mathbb{R}$ we have the basic properties

(linearity) $$\langle\lambda\boldsymbol{A},\boldsymbol{B}\rangle_{\boldsymbol{K}}=\lambda\langle\boldsymbol{A},\boldsymbol{B}\rangle_{\boldsymbol{K}}=\langle\boldsymbol{A},\lambda\boldsymbol{B}\rangle_{\boldsymbol{K}}, \tag{A.9a}$$

$$\langle\boldsymbol{A}+\boldsymbol{C},\boldsymbol{B}\rangle_{\boldsymbol{K}}=\langle\boldsymbol{A},\boldsymbol{B}\rangle_{\boldsymbol{K}}+\langle\boldsymbol{C},\boldsymbol{B}\rangle_{\boldsymbol{K}} \tag{A.9b}$$

(symmetry) $$\langle\boldsymbol{B},\boldsymbol{A}\rangle_{\boldsymbol{K}}=\langle\boldsymbol{A},\boldsymbol{B}\rangle_{\boldsymbol{K}} \tag{A.9c}$$

(positive definiteness) $$\langle\boldsymbol{A},\boldsymbol{A}\rangle_{\boldsymbol{K}}\geq 0,\quad\langle\boldsymbol{A},\boldsymbol{A}\rangle_{\boldsymbol{K}}=0\ \Leftrightarrow\ \boldsymbol{A}=\mathbf{0}. \tag{A.9d}$$

Setting $\boldsymbol{K}=\mathbf{I}_n$ in the definition (A.8) is called the *Frobenius inner product* in [Horn and Johnson, 1985, sec. 5.2] or *Hilbert-Schmidt inner product* in [Hall, 2015, sec. A.6]. Furthermore, for $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times 1}$ it coincides with the common *dot product*.

Norm. The induced norm is

$$\|\boldsymbol{A}\|_{\boldsymbol{K}} = \sqrt{\langle \boldsymbol{A}, \boldsymbol{A}\rangle_{\boldsymbol{K}}}. \tag{A.10}$$

For $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times m}$ and $\lambda \in \mathbb{R}$ we have the basic properties

$$\begin{aligned}
&\text{(triangle inequality)} && \|\boldsymbol{A}+\boldsymbol{B}\|_{\boldsymbol{K}} \leq \|\boldsymbol{A}\|_{\boldsymbol{K}} + \|\boldsymbol{B}\|_{\boldsymbol{K}} && \text{(A.11a)}\\
&\text{(absolute homogeneity)} && \|\lambda\boldsymbol{A}\|_{\boldsymbol{K}} = |\lambda|\,\|\boldsymbol{A}\|_{\boldsymbol{K}}, && \text{(A.11b)}\\
&\text{(positive definiteness)} && \|\boldsymbol{A}\|_{\boldsymbol{K}} \geq 0, \quad \|\boldsymbol{A}\|_{\boldsymbol{K}} = 0 \ \Leftrightarrow\ \boldsymbol{A} = \boldsymbol{0}. && \text{(A.11c)}
\end{aligned}$$

Metric. The induced metric is

$$d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{B}) = \|\boldsymbol{A}-\boldsymbol{B}\|_{\boldsymbol{K}}. \tag{A.12}$$

For $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{n\times m}$ and $\lambda \in \mathbb{R}$ we have the basic properties

$$\begin{aligned}
&\text{(triangle inequality)} && d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{B}) \leq d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{C}) + d_{\boldsymbol{K}}(\boldsymbol{B}, \boldsymbol{C}) && \text{(A.13a)}\\
&\text{(symmetry)} && d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{B}) = d_{\boldsymbol{K}}(\boldsymbol{B}, \boldsymbol{A}), && \text{(A.13b)}\\
&\text{(positive definiteness)} && d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{B}) \geq 0, \quad d_{\boldsymbol{K}}(\boldsymbol{A}, \boldsymbol{B}) = 0 \ \Leftrightarrow\ \boldsymbol{A} = \boldsymbol{B}. && \text{(A.13c)}
\end{aligned}$$

Translation of particular arguments. For $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times n}$ and $\boldsymbol{R} \in \mathbb{SO}(n)$ we have the properties

$$\begin{aligned}
\langle(\boldsymbol{R}\mathbf{X}_1)^\top, (\boldsymbol{R}\mathbf{X}_2)^\top\rangle_{\boldsymbol{K}} &= \langle\mathbf{X}_1^\top, \mathbf{X}_2^\top\rangle_{\boldsymbol{K}}, && \text{(A.14a)}\\
\langle(\mathbf{X}_1\boldsymbol{R})^\top, (\mathbf{X}_2\boldsymbol{R})^\top\rangle_{\boldsymbol{K}} &= \langle\mathbf{X}_1^\top, \mathbf{X}_2^\top\rangle_{\boldsymbol{R}\boldsymbol{K}\boldsymbol{R}^\top} && \text{(A.14b)}
\end{aligned}$$

For $\boldsymbol{\Xi}_1 = \left[\begin{smallmatrix}\mathbf{X}_1 & \mathbf{x}_1\\ \mathbf{0} & 0\end{smallmatrix}\right]$, $\boldsymbol{\Xi}_2 = \left[\begin{smallmatrix}\mathbf{X}_2 & \mathbf{x}_2\\ \mathbf{0} & 0\end{smallmatrix}\right]$ with $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times n}$ and $\mathbf{G} \in \mathbb{SE}(n)$ we have

$$\begin{aligned}
\langle(\mathbf{G}\boldsymbol{\Xi}_1)^\top, (\mathbf{G}\boldsymbol{\Xi}_2)^\top\rangle_{\boldsymbol{K}} &= \langle\boldsymbol{\Xi}_1^\top, \boldsymbol{\Xi}_2^\top\rangle_{\boldsymbol{K}}, && \text{(A.15a)}\\
\langle(\boldsymbol{\Xi}_1\mathbf{G})^\top, (\boldsymbol{\Xi}_2\mathbf{G})^\top\rangle_{\boldsymbol{K}} &= \langle\boldsymbol{\Xi}_1^\top, \boldsymbol{\Xi}_2^\top\rangle_{\mathbf{G}\boldsymbol{K}\mathbf{G}^\top} && \text{(A.15b)}
\end{aligned}$$

Note that this case includes in particular $\boldsymbol{\Xi}_1, \boldsymbol{\Xi}_2 \in \mathfrak{se}(n)$.

Derivative. This might be useful for the following

$$\mathcal{W} = \tfrac{1}{2}\|(\operatorname{wed}(\boldsymbol{\xi}) + \mathbf{X})^\top\|_{\boldsymbol{K}}^2 = \tfrac{1}{2}\boldsymbol{\xi}^\top\boldsymbol{K}\boldsymbol{\xi} + \boldsymbol{\xi}^\top \operatorname{vee2}(\mathbf{X}\boldsymbol{K}) + \tfrac{1}{2}\operatorname{tr}(\mathbf{X}\boldsymbol{K}\mathbf{X}^\top) \tag{A.16}$$

$$\frac{\partial\mathcal{W}}{\partial\boldsymbol{\xi}} = \operatorname{Vee}(\boldsymbol{K})\boldsymbol{\xi} + \operatorname{vee2}(\mathbf{X}\boldsymbol{K}) = \operatorname{vee2}((\operatorname{wed}(\boldsymbol{\xi}) + \mathbf{X})\boldsymbol{K}) \tag{A.17}$$

$$\frac{\partial^2\mathcal{W}}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}} = \operatorname{Vee}(\boldsymbol{K}) \tag{A.18}$$

A.1.5 Matrix decompositions

Singular value decomposition (SVD). Any matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$ can be decomposed into $\boldsymbol{A} = \boldsymbol{X\Sigma Y}^\top$, where $\boldsymbol{X} \in \mathbb{O}(n)$, $\boldsymbol{Y} \in \mathbb{O}(m)$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times m}$ with $\Sigma_{ii} \geq 0, i = 1,\dots,\min(n,m), \Sigma_{ij} = 0, i \neq j$.

- The columns of $\boldsymbol{X}$ are eigenvectors of $\boldsymbol{AA}^\top = \boldsymbol{X}(\boldsymbol{\Sigma\Sigma}^\top)\boldsymbol{X}^\top$.
- The columns of $\boldsymbol{Y}$ are eigenvectors of $\boldsymbol{A}^\top\boldsymbol{A} = \boldsymbol{Y}(\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma})\boldsymbol{Y}^\top$.

It is common practice to place the singular values in descending order, i.e. $\Sigma_{ii} \geq \Sigma_{jj}, i > j$, thus making $\boldsymbol{\Sigma}$ unique. The matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are unique up to orthogonal transformations of the subspaces of each singular value and the kernel and co-kernel of $\boldsymbol{A}$.

Polar decomposition. Any square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ can be decomposed into $\boldsymbol{A} = \boldsymbol{UK}$, where $\boldsymbol{U} \in \mathbb{O}(n)$, $\boldsymbol{K} \in \mathbb{SYM}_0^+(n)$. The matrix $\boldsymbol{K}$ is uniquely defined while $\boldsymbol{U}$ is only unique if $\boldsymbol{A}$ is invertible. The same holds for the decomposition $\boldsymbol{A} = \boldsymbol{LV}$, where $\boldsymbol{V} \in \mathbb{O}(n)$, $\boldsymbol{L} \in \mathbb{SYM}_0^+(n)$. The relation to the singular value decomposition $\boldsymbol{A} = \boldsymbol{X\Sigma Y}^\top$ is

$$\boldsymbol{U} = \boldsymbol{XY}^\top, \quad \boldsymbol{K} = \boldsymbol{Y\Sigma Y}^\top \qquad \text{and} \qquad \boldsymbol{L} = \boldsymbol{X\Sigma X}^\top, \quad \boldsymbol{V} = \boldsymbol{XY}^\top. \tag{A.19}$$

Special polar decomposition. Consider the SVD $\boldsymbol{A} = \boldsymbol{X\Sigma Y}^\top$ and define

$$\bar{\boldsymbol{X}} = \boldsymbol{X}\,\mathrm{diag}(1,\dots,1,\det\boldsymbol{X}) \in \mathbb{SO}(n), \tag{A.20a}$$
$$\bar{\boldsymbol{Y}} = \boldsymbol{Y}\,\mathrm{diag}(1,\dots,1,\det\boldsymbol{Y}) \in \mathbb{SO}(n), \tag{A.20b}$$
$$\bar{\boldsymbol{\Sigma}} = \mathrm{diag}(\sigma_1,\dots,\sigma_{n-1},\det\boldsymbol{X}\det\boldsymbol{Y}\sigma_n). \tag{A.20c}$$

where σ_n is the smallest singular value. The matrix $\bar{\boldsymbol{\Sigma}}$ is indifferent in general, but since we may have only flipped the smallest singular value, the matrix $\hat{\boldsymbol{\Sigma}} = \mathrm{Vee}(\bar{\boldsymbol{\Sigma}})$ is positive semidefinite. With this we may write

$$\boldsymbol{A} = \bar{\boldsymbol{X}}\underbrace{\bar{\boldsymbol{Y}}^\top\bar{\boldsymbol{Y}}}_{\mathbf{I}_n}\mathrm{Wed}(\hat{\boldsymbol{\Sigma}})\bar{\boldsymbol{Y}}^\top = \underbrace{\bar{\boldsymbol{X}}\bar{\boldsymbol{Y}}^\top}_{\boldsymbol{U}\in\mathbb{SO}(n)}\ \mathrm{Wed}\Big(\underbrace{\bar{\boldsymbol{X}}\hat{\boldsymbol{\Sigma}}\bar{\boldsymbol{X}}^\top}_{\boldsymbol{K}\in\mathbb{SYM}_0^+(n)}\Big) \tag{A.21a}$$
$$= \bar{\boldsymbol{X}}\,\mathrm{Wed}(\hat{\boldsymbol{\Sigma}})\underbrace{\bar{\boldsymbol{X}}^\top\bar{\boldsymbol{X}}}_{\mathbf{I}_n}\bar{\boldsymbol{Y}}^\top = \mathrm{Wed}\Big(\underbrace{\bar{\boldsymbol{X}}\hat{\boldsymbol{\Sigma}}\bar{\boldsymbol{X}}^\top}_{\boldsymbol{L}\in\mathbb{SYM}_0^+(n)}\Big)\underbrace{\bar{\boldsymbol{X}}\bar{\boldsymbol{Y}}^\top}_{\boldsymbol{V}\in\mathbb{SO}(n)} \tag{A.21b}$$

Since this only involves proper rotations this is potentially useful in physics where we cant have reflections. This decomposition is not established in the literature.

A.1.6 Moore-Penrose pseudoinverse

For any matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ there exists a unique *Moore-Penrose inverse*, or *pseudoinverse*, $\boldsymbol{A}^+ \in \mathbb{R}^{n\times m}$ determined by the following conditions [Penrose, 1955, Theo. 1]:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} &= \boldsymbol{A}, && \text{(A.22a)}\\ \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ &= \boldsymbol{A}^+, && \text{(A.22b)}\\ (\boldsymbol{A}\boldsymbol{A}^+)^\top &= \boldsymbol{A}\boldsymbol{A}^+, && \text{(A.22c)}\\ (\boldsymbol{A}^+\boldsymbol{A})^\top &= \boldsymbol{A}^+\boldsymbol{A}. && \text{(A.22d)} \end{aligned}$$

Some useful identities with the pseudoinverse are

$$\begin{aligned} & (\boldsymbol{A}^\top)^+ = (\boldsymbol{A}^+)^\top, && \text{(A.23a)}\\ & \operatorname{rank}\boldsymbol{A} = \operatorname{rank}\boldsymbol{A}^+, && \text{(A.23b)}\\ & \boldsymbol{A}^+ = \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{A}^\top)^+ = (\boldsymbol{A}^\top\boldsymbol{A})^+\boldsymbol{A}^\top, && \text{(A.23c)}\\ \operatorname{rank}\boldsymbol{A} = m \quad \Rightarrow \quad & \boldsymbol{A}^+ = \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}, && \text{(A.23d)}\\ \operatorname{rank}\boldsymbol{A} = n \quad \Rightarrow \quad & \boldsymbol{A}^+ = (\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top, && \text{(A.23e)}\\ \operatorname{rank}\boldsymbol{A} = m = n \quad \Rightarrow \quad & \boldsymbol{A}^+ = \boldsymbol{A}^{-1}, && \text{(A.23f)}\\ & (\boldsymbol{A}\boldsymbol{B})^+ = (\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{B})^+(\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^+)^+. && \text{(A.23g)} \end{aligned}$$

With a singular value decomposition $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{Y}^\top$, the pseudoinverse can be stated as $\boldsymbol{A}^+ = \boldsymbol{Y}\boldsymbol{\Sigma}^+\boldsymbol{X}^\top$ where $\boldsymbol{\Sigma}^+$ is the transpose of $\boldsymbol{\Sigma}$ with its nonzero diagonal entries inverted.

Linear equations. For the linear system

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{A.24}$$

with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$ to have any solutions for $\boldsymbol{x} \in \mathbb{R}^n$ we need $\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{b} = \boldsymbol{b}$. Then the solutions have the form

$$\boldsymbol{x} = \boldsymbol{A}^+\boldsymbol{b} + (\mathbf{I}_n - \boldsymbol{A}^+\boldsymbol{A})\boldsymbol{\xi},\ \ \boldsymbol{\xi} \in \mathbb{R}^n. \tag{A.25}$$

The solution is unique if $\operatorname{rank}\boldsymbol{A} = n$ which implies $\boldsymbol{A}^+\boldsymbol{A} = \mathbf{I}_n$. The solution with minimal Euclidean norm is $\boldsymbol{x} = \boldsymbol{A}^+\boldsymbol{b}$. Furthermore, $\boldsymbol{z} = \boldsymbol{A}^+\boldsymbol{b}$ minimizes the Euclidean norm $\|\boldsymbol{A}\boldsymbol{z} - \boldsymbol{b}\|$, even if no solutions exist.

A.1.7 Orthogonal projectors

A *projector* is a square matrix $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ that is idempotent $\boldsymbol{P}^2 = \boldsymbol{P}$. For each projector there is a *complementary projector* $\boldsymbol{P}^\perp = \mathbf{I}_n - \boldsymbol{P}$ and we have $\operatorname{rank}\boldsymbol{P} + \operatorname{rank}\boldsymbol{P}^\perp = n$. The image of $\boldsymbol{P}$ is the kernel of $\boldsymbol{P}^\perp$ and vice versa.

For a given inner product $\langle\cdot,\cdot\rangle_{\boldsymbol{M}}$, a projector is called *orthogonal* iff $\langle\boldsymbol{P}\boldsymbol{\xi}, \boldsymbol{P}^\perp\boldsymbol{\eta}\rangle_{\boldsymbol{M}} = 0\,\forall\,\boldsymbol{\xi},\boldsymbol{\eta} \in \mathbb{R}^n$ or equivalently $\boldsymbol{M}\boldsymbol{P} = \boldsymbol{P}^\top\boldsymbol{M}$. If $\boldsymbol{P}$ is orthogonal so is $\boldsymbol{P}^\perp$.

For a given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, the unique orthogonal projector $\boldsymbol{P}_{\boldsymbol{A}}$ into the image of $\boldsymbol{A}$ can be constructed as follows: Let $\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{A}\boldsymbol{\xi}$, $\boldsymbol{\xi} \in \mathbb{R}^m$:

$$\min_{\boldsymbol{\xi} \in \mathbb{R}^m} J = \|\boldsymbol{A}\boldsymbol{\xi} - \boldsymbol{x}\|_{\boldsymbol{M}}^2 \qquad \Rightarrow \qquad \frac{\partial J}{\partial \boldsymbol{\xi}} = \boldsymbol{0} \qquad \Leftrightarrow \qquad \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}\, \boldsymbol{\xi} = \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{x} \tag{A.26}$$

Though there might be several solutions for $\boldsymbol{\xi}$ if $\operatorname{rank} \boldsymbol{A} < n$, the solution for $\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^+ \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{x}$ is unique. Consequently we have

$$\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^+ \boldsymbol{A}^\top \boldsymbol{M}. \tag{A.27}$$

Consider the matrices $\boldsymbol{A} \in \mathbb{R}^{n \times a}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times b}$ with $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{B} = \boldsymbol{0}$ and $\operatorname{rank} \boldsymbol{A} + \operatorname{rank} \boldsymbol{B} = n$. This implies that the direct sum of their images is $\mathbb{R}^n$. Furthermore, their associated orthogonal projectors $\boldsymbol{P}_{\boldsymbol{A}}$ and $\boldsymbol{P}_{\boldsymbol{B}}$ are complementary, i.e. $\boldsymbol{P}_{\boldsymbol{A}}^\perp = \boldsymbol{P}_{\boldsymbol{B}}$. This implies the identity

$$\underbrace{\boldsymbol{A}(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^+ \boldsymbol{A}^\top \boldsymbol{M}}_{\boldsymbol{P}_{\boldsymbol{A}}} + \underbrace{\boldsymbol{B}(\boldsymbol{B}^\top \boldsymbol{M} \boldsymbol{B})^+ \boldsymbol{B}^\top \boldsymbol{M}}_{\boldsymbol{P}_{\boldsymbol{B}}} = \mathbf{I}_n. \tag{A.28}$$

For the special case $\boldsymbol{M} = \mathbf{I}_n$ and using (A.23c), this is

$$\boldsymbol{A}\boldsymbol{A}^+ + \boldsymbol{B}\boldsymbol{B}^+ = \mathbf{I}_n. \tag{A.29}$$

A.2 Differential geometry

A.2.1 Embedded manifold

Inverse function theorem. [Dontchev and Rockafellar, 2010, Theo. 1A.1]: Let $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in a neighborhood of $\bar{\boldsymbol{x}}$ and let $\bar{\boldsymbol{y}} = \boldsymbol{f}(\bar{\boldsymbol{x}})$. If $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\bar{\boldsymbol{x}})$ is nonsingular, then there exists an continuous differentiable inverse $\boldsymbol{g}$ for some neighborhood $\mathbb{Y} \subset \mathbb{R}^n$ of $\bar{\boldsymbol{y}}$. Its Jacobian satisfies

$$\tfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}(\boldsymbol{y}) = \big(\tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{g}(\boldsymbol{y}))\big)^{-1}, \quad \boldsymbol{y} \in \mathbb{Y}. \tag{A.30}$$

Implicit function theorem. [Dontchev and Rockafellar, 2010, Theo. 1B.1]: Let $\boldsymbol{f} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in a neighborhood of $(\bar{\boldsymbol{p}}, \bar{\boldsymbol{x}})$ and let $\boldsymbol{f}(\bar{\boldsymbol{p}}, \bar{\boldsymbol{x}}) = \boldsymbol{0}$. If $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\bar{\boldsymbol{p}}, \bar{\boldsymbol{x}})$ is nonsingular, then there exists an continuous differentiable function $\boldsymbol{g} : \mathbb{R}^p \to \mathbb{R}^n$ for some neighborhood $\mathbb{P} \subset \mathbb{R}^d$ of $\bar{\boldsymbol{p}}$ such that

$$\boldsymbol{f}(\boldsymbol{p}, \boldsymbol{g}(\boldsymbol{p})) = \boldsymbol{0}, \quad \boldsymbol{p} \in \mathbb{P}. \tag{A.31}$$

Its Jacobian satisfies

$$\tfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{p}}(\boldsymbol{p}) = -\big(\tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{p}, \boldsymbol{g}(\boldsymbol{p}))\big)^{-1} \tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{p}}(\boldsymbol{p}, \boldsymbol{g}(\boldsymbol{p})), \quad \boldsymbol{p} \in \mathbb{P}. \tag{A.32}$$

Embedded manifold. Consider the set $\mathbb{X}$ defined by

$$\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^\nu \,|\, \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}\} \tag{A.33}$$

with smooth functions $\phi^\kappa(\boldsymbol{x}) = 0, \kappa = 1, \dots c$. Let their Jacobian have constant rank

$$\Phi^\kappa_\alpha(\boldsymbol{x}) = \frac{\partial \phi^\kappa}{\partial x^\alpha}(\boldsymbol{x}), \qquad \operatorname{rank} \boldsymbol{\Phi}(\boldsymbol{x}) = \nu - n = \mathit{const.}\ \forall\ \boldsymbol{x} \in \mathbb{X}. \tag{A.34}$$

Due to the rank condition, the implicit function theorem guarantees the existence of a local chart around each point $\boldsymbol{x} \in \mathbb{X}$. Consequently $\mathbb{X}$ is a n dimensional *embedded submanifold* of $\mathbb{R}^\nu$. If there exits a global chart, then $\mathbb{X}$ is homeomorphic to $\mathbb{R}^n$ and the manifold is *linear*. If this is not the case, the manifold is *nonlinear*.

Tangent space. Consider a parametrized curve $\boldsymbol{x} : \mathbb{R} \to \mathbb{X} : t \mapsto \boldsymbol{x}(t)$. Since we have $\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}$, we also have

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}(\boldsymbol{x}(t)) = \boldsymbol{\Phi}(\boldsymbol{x}(t))\tfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) = \boldsymbol{0}, \tag{A.35}$$

which has to hold for any curve through $\boldsymbol{x}$. The set of tangent vectors $\boldsymbol{v} = \dot{\boldsymbol{x}}$ of arbitrary curves through a particular point $\boldsymbol{x} \in \mathbb{X}$ is the *tangent space*

$$\mathsf{T}_{\boldsymbol{x}}\mathbb{X} = \{\boldsymbol{v} \in \mathbb{R}^\nu \,|\, \boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{v} = \boldsymbol{0}\}. \tag{A.36}$$

Since $\operatorname{rank} \boldsymbol{\Phi} = \nu - n$ there is a matrix $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu \times n}$ with $\boldsymbol{\Phi}\boldsymbol{A} \equiv 0$, $\operatorname{rank} \boldsymbol{A} = n$ with which we can give an explicit statement of the tangent space

$$\mathsf{T}_{\boldsymbol{x}}\mathbb{X} = \{\boldsymbol{v} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi} \,|\, \boldsymbol{\xi} \in \mathbb{R}^n\} \tag{A.37}$$

i.e. the column space of the matrix $\boldsymbol{A}$.

Linear algebra guarantees the existence of a matrix $\boldsymbol{A}(\boldsymbol{x})$ at each point $\boldsymbol{x} \in \mathbb{X}$ and the implicit function theorem guarantees that there is a neighborhood in which $\boldsymbol{A}$ is smooth. However, there is no guarantee for $\boldsymbol{A}$ to be smooth over all $\mathbb{X}$, i.e. to be global. In the most important case for this work $\mathbb{X} = \mathbb{SO}(3)$, thought the manifold is nonlinear, there is a global matrix $\boldsymbol{A}$, see Example 8. The conjecture is that this is true for any *Lie group*. In other cases, e.g. the 2-sphere $\mathbb{X} = \mathbb{S}^2$, the *hairy-ball theorem*, e.g. [Poincaré, 1885], states that no global matrix $\boldsymbol{A}$ can exist. In these cases one may resort to having overlapping, local patches for $\boldsymbol{A}$.

Some identities. Consider the matrices $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu\times n}, \nu \geq n$ and $\boldsymbol{\Phi}(\boldsymbol{x}) \in \mathbb{R}^{c\times\nu}, c \geq \nu - n$ with

$$\operatorname{rank} \boldsymbol{A} = n, \qquad \operatorname{rank} \boldsymbol{\Phi} = \nu - n, \qquad \boldsymbol{\Phi}\boldsymbol{A} = \mathbf{0}. \tag{A.38}$$

Let $\boldsymbol{Y}(\boldsymbol{x}) \in \mathbb{R}^{n\times\nu}$ and $\boldsymbol{\Psi}(\boldsymbol{x}) \in \mathbb{R}^{\nu\times c}$ be matrices that fulfill the identities

$$\boldsymbol{Y}\boldsymbol{A} = \mathbf{I}_n, \qquad \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{\Psi}\boldsymbol{\Phi} = \mathbf{I}_\nu. \tag{A.39}$$

Recalling the properties of the Moore-Penrose-pseudoinverse (A.22) and the identity (A.28), one possible solution to this is

$$\boldsymbol{Y} = \boldsymbol{A}^+, \qquad \boldsymbol{\Psi} = \boldsymbol{\Phi}^+. \tag{A.40}$$

A.2.2 Vectors

Basis vectors. In differential geometry it is common to identify vectors and differential operators. The standard basis for $\mathsf{T}\mathbb{R}^\nu$ may be written as $\boldsymbol{e}_\alpha = \frac{\partial}{\partial x^\alpha}, \alpha = 1, \ldots, \nu$. For the sake of readability we adopt the notation of [Frankel, 1997] and denote them as $\frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^\alpha}$ when they should be regarded as basis vectors. Using the relations from the previous paragraph, we define the basis vectors $\boldsymbol{\partial}_i$ for the tangent space $\mathsf{T}_{\boldsymbol{x}}\mathbb{X}$ and the complementary vectors $\boldsymbol{\partial}_\kappa^\perp$ that span $(\mathsf{T}_{\boldsymbol{x}}\mathbb{X})^\perp$ as

$$\boldsymbol{\partial}_i = A_i^\alpha(\boldsymbol{x}) \frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^\alpha}, \qquad i = 1, \ldots, n, \tag{A.41a}$$

$$\boldsymbol{\partial}_\kappa^\perp = \Psi_\kappa^\alpha(\boldsymbol{x}) \frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^\alpha}, \qquad \kappa = 1, \ldots, c, \tag{A.41b}$$

$$\frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^\alpha} = Y_\alpha^i(\boldsymbol{x})\boldsymbol{\partial}_i + \Phi_\alpha^\kappa(\boldsymbol{x})\boldsymbol{\partial}_\kappa^\perp, \qquad \alpha = 1, \ldots, \nu. \tag{A.41c}$$

Directional derivative. Using the basis vectors we may write any tangent vector $\boldsymbol{v} \in \mathsf{T}_{\boldsymbol{x}}\mathbb{X}$ as $\boldsymbol{v} = v^i\boldsymbol{\partial}_i$. The derivative of a function $f : \mathbb{X} \to \mathbb{R}$ in the direction of a tangent vector $\boldsymbol{v} \in \mathsf{T}_{\boldsymbol{x}}\mathbb{X}$ is

$$\boldsymbol{v}(f) = v^i\partial_i f = v^i A_i^\alpha \frac{\partial f}{\partial x^\alpha}. \tag{A.42}$$

Note that for computing $\partial f/\partial x^\alpha$, f needs to be defined not only on $\mathbb{X}$, but also on an (infinitesimal) neighborhood in the normal directions. This is usually the case if the function is expressed in terms of the redundant coordinates $\boldsymbol{x}$. However, we could avoid this by introducing minimal, local coordinates $\boldsymbol{x} = \boldsymbol{\chi}(\boldsymbol{q}), \boldsymbol{q} \in \mathbb{R}^n$ and taking the derivatives in their directions. The crucial point is that the result, i.e. $\boldsymbol{v}(f)$ is *invariant* to the used coordinates. See the corresponding paragraphs in subsection A.3.2 and subsection A.3.3 for the explicit transformation rules.

Vector field. For the given manifold $\mathbb{X}$, a vector field is the assignment $\boldsymbol{v} \in \mathsf{T}_{\boldsymbol{x}}\mathbb{X}$ to each point $\boldsymbol{x} \in \mathbb{X}$. The collection of vector fields on $\mathbb{X}$ is commonly denoted by $\mathfrak{X}(\mathbb{X})$.

Lie bracket. The *Lie bracket* is a map $[\cdot,\cdot] : \mathfrak{X}(\mathbb{X}) \times \mathfrak{X}(\mathbb{X}) \to \mathfrak{X}(\mathbb{X})$. For $\boldsymbol{u}, \boldsymbol{v} \in \mathfrak{X}(\mathbb{X})$ and a function $f : \mathbb{X} \to \mathbb{R}$ it yields

$$\begin{aligned}
[\boldsymbol{u}, \boldsymbol{v}]\,(f) &= \boldsymbol{u}(\boldsymbol{v}(f)) - \boldsymbol{v}(\boldsymbol{u}(f)) \\
&= u^i \partial_i (v^j \partial_j f) - v^j \partial_j (u^i \partial_i f) \\
&= (u^i \partial_i v^j - v^i \partial_i u^j)\partial_j f + u^i v^j (\partial_i \partial_j f - \partial_j \partial_i f)
\end{aligned} \tag{A.43}$$

Whereas derivatives like $\partial/\partial x^\alpha$ do commute, this is not true for ∂_i. Instead we have

$$\begin{aligned}
\partial_i \partial_j f - \partial_j \partial_i f &= A_i^\alpha \frac{\partial}{\partial x^\alpha}\left(A_j^\beta \frac{\partial f}{\partial x^\beta}\right) - A_j^\beta \frac{\partial}{\partial x^\beta}\left(A_i^\alpha \frac{\partial f}{\partial x^\alpha}\right) \\
&= A_i^\alpha \frac{\partial A_j^\beta}{\partial x^\alpha}\frac{\partial f}{\partial x^\beta} - A_j^\beta \frac{\partial A_i^\alpha}{\partial x^\beta}\frac{\partial f}{\partial x^\alpha} + A_i^\alpha A_j^\beta \underbrace{\left(\frac{\partial^2 f}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2 f}{\partial x^\beta \partial x^\alpha}\right)}_{0} \\
&= \left(A_i^\alpha \frac{\partial A_j^\beta}{\partial x^\alpha} - A_j^\alpha \frac{\partial A_i^\beta}{\partial x^\alpha}\right) \overbrace{\left(Y_\beta^k \partial_k f + \Phi_\beta^\kappa \partial_\kappa^\perp f\right)}^{\frac{\partial f}{\partial x^\beta}} \\
&= \underbrace{\left(\partial_i A_j^\beta - \partial_j A_i^\beta\right) Y_\beta^k}_{\gamma_{ij}^k} \partial_k f + \underbrace{\left(\frac{\partial^2 \phi^\kappa}{\partial x^\beta \partial x^\alpha} - \frac{\partial^2 \phi^\kappa}{\partial x^\alpha \partial x^\beta}\right)}_{0} A_i^\alpha A_j^\beta \partial_\kappa^\perp f
\end{aligned} \tag{A.44}$$

Since this holds for any function f we can give the coordinate expression for the *Lie-bracket* as

$$\left[u^i \boldsymbol{\partial}_i, v^j \boldsymbol{\partial}_j\right] = \left(u^i \partial_i v^k - v^i \partial_i u^k + \gamma_{ij}^k u^i v^j\right)\boldsymbol{\partial}_k \tag{A.45}$$

We call its coefficients γ_{ij}^k, defined by $[\boldsymbol{\partial}_i, \boldsymbol{\partial}_j] = \gamma_{ij}^k \boldsymbol{\partial}_k$, the *commutation coefficients*. Using the identities (A.39) we may also give alternative formulations

$$\gamma_{ij}^k = Y_\alpha^k \left(\partial_i A_j^\alpha - \partial_j A_i^\alpha\right) = A_i^\alpha \partial_j Y_\alpha^k - A_j^\alpha \partial_i Y_\alpha^k = \left(\frac{\partial Y_\alpha^k}{\partial x^\beta} - \frac{\partial Y_\beta^k}{\partial x^\alpha}\right) A_j^\beta A_i^\alpha \tag{A.46}$$

For the Lie bracket and consequently for its coefficients, we have the antisymmetry and Jacobi identity:

$$[\boldsymbol{u}, \boldsymbol{v}] = -[\boldsymbol{v}, \boldsymbol{u}] \quad \Leftrightarrow \quad \gamma_{ij}^k = -\gamma_{ji}^k \tag{A.47}$$

$$[\boldsymbol{u}, [\boldsymbol{v}, \boldsymbol{w}]] + [\boldsymbol{w}, [\boldsymbol{u}, \boldsymbol{v}]] + [\boldsymbol{v}, [\boldsymbol{w}, \boldsymbol{u}]] = \boldsymbol{0} \quad \Leftrightarrow \quad \gamma_{il}^s \gamma_{jk}^l + \gamma_{kl}^s \gamma_{ij}^l + \gamma_{jl}^s \gamma_{ki}^l = 0. \tag{A.48}$$

The commutation coefficients are invariant to a change in the coordinates, see subsection A.3.2. They do, however, depend on the basis $\boldsymbol{\partial}_i$ and do consequently *not* form a tensor themselves. See subsection A.3.3 for the transformation rule. Nevertheless, the Lie bracket itself is, of course, coordinate independent.

Connection and covariant derivative. The partial derivatives $\partial_j v^i$ of the coefficients of a vector do *not* yield a tensor. A more sophisticated concept is required to establish a notion for the change of a vector.

A *connection* is a map $\boldsymbol{\nabla} : \mathfrak{X}(\mathbb{X}) \times \mathfrak{X}(\mathbb{X}) \to \mathfrak{X}(\mathbb{X})$ defined through the following properties [Boothby, 1986, Def. VII.3.1]:

$$\boldsymbol{\nabla}_{f\boldsymbol{v}+g\boldsymbol{w}}\boldsymbol{u} = f\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} + g\boldsymbol{\nabla}_{\boldsymbol{w}}\boldsymbol{u}, \tag{A.49a}$$

$$\boldsymbol{\nabla}_{\boldsymbol{v}}(f\,\boldsymbol{u}) = f\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} + \boldsymbol{v}(f)\boldsymbol{u}, \tag{A.49b}$$

$$\boldsymbol{\nabla}_{\boldsymbol{v}}(\boldsymbol{u}+\boldsymbol{w}) = \boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} + \boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{w} \tag{A.49c}$$

for all $f, g : \mathbb{X} \to \mathbb{R}$ and $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathfrak{X}(\mathbb{X})$. Using basis vectors we have

$$\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} = \boldsymbol{\nabla}_{v^j\boldsymbol{\partial}_j}(u^i\boldsymbol{\partial}_i) = v^j\big((\partial_j u^i)\boldsymbol{\partial}_i + u^i\,\underbrace{\boldsymbol{\nabla}_{\boldsymbol{\partial}_j}\boldsymbol{\partial}_i}_{\Gamma^k_{ij}\boldsymbol{\partial}_k}\big) = \big(\underbrace{\partial_j u^i + \Gamma^i_{kj}u^k}_{\nabla_j u^i}\big)v^j\boldsymbol{\partial}_i. \tag{A.50}$$

Here we have introduced the *connection coefficients* Γ^i_{kj}, defined by $\boldsymbol{\nabla}_{\boldsymbol{\partial}_j}\boldsymbol{\partial}_i = \Gamma^k_{ij}\boldsymbol{\partial}_k$. For the coordinate expression (A.50) to make sense, it has to be the same in any basis: Consider another set of basis vectors $\hat{\boldsymbol{\partial}}_{\hat{i}}, \hat{i} = 1, \dots, n$ and the relations

$$\boldsymbol{u} = u^i\boldsymbol{\partial}_i = \hat{u}^{\hat{i}}\hat{\boldsymbol{\partial}}_{\hat{i}}, \qquad u^i = W^i_{\hat{i}}\hat{u}^{\hat{i}}, \quad \boldsymbol{\partial}_i = Z^{\hat{i}}_i\hat{\boldsymbol{\partial}}_{\hat{i}}, \quad \boldsymbol{Z} = \boldsymbol{W}^{-1}. \tag{A.51}$$

Then the following has to hold

$$\begin{aligned}\boldsymbol{\nabla}_{v^j\boldsymbol{\partial}_j}(u^i\boldsymbol{\partial}_i) &= \big(\partial_j u^i + \Gamma^i_{kj}u^k\big)v^j\boldsymbol{\partial}_i = \big(Z^{\hat{l}}_j\hat{\partial}_{\hat{l}}(W^i_{\hat{k}}\hat{u}^{\hat{k}}) + \Gamma^i_{kj}W^k_{\hat{k}}\hat{u}^{\hat{k}}\big)W^j_{\hat{j}}\hat{v}^{\hat{j}}Z^{\hat{i}}_i\hat{\boldsymbol{\partial}}_{\hat{i}}\\ &= \big(\hat{\partial}_{\hat{j}}\hat{u}^{\hat{i}} + \underbrace{Z^{\hat{i}}_i\big(\hat{\partial}_{\hat{j}}W^i_{\hat{k}} + \Gamma^i_{kj}W^k_{\hat{k}}W^j_{\hat{j}}\big)}_{\hat{\Gamma}^{\hat{i}}_{\hat{k}\hat{j}}}\hat{u}^{\hat{k}}\big)\hat{v}^{\hat{j}}\hat{\boldsymbol{\partial}}_{\hat{i}} = \boldsymbol{\nabla}_{\hat{v}^{\hat{j}}\hat{\boldsymbol{\partial}}_{\hat{j}}}(\hat{u}^{\hat{i}}\hat{\boldsymbol{\partial}}_{\hat{i}}).\end{aligned} \tag{A.52}$$

This implies the transformation law for the connection coefficients. From this it should be clear that the connection coefficients Γ^k_{ij} do *not* form a tensor despite being indexed in the same way. The connection does transform covariantly, thus $\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u}$ is also called the *covariant derivative* of $\boldsymbol{u}$ along $\boldsymbol{v}$.

In the special case of coordinate basis vectors the transformation law (A.52) simplifies to

$$\partial_i = \tfrac{\partial}{\partial q^i},\ \hat{\partial}_{\hat{i}} = \tfrac{\partial}{\partial \hat{q}^{\hat{i}}} \quad\Rightarrow\quad W^i_{\hat{i}} = \tfrac{\partial q^i}{\partial \hat{q}^{\hat{i}}},\ Z^{\hat{i}}_i = \tfrac{\partial \hat{q}^{\hat{i}}}{\partial q^i},\ \hat{\Gamma}^{\hat{i}}_{\hat{k}\hat{j}} = \tfrac{\partial \hat{q}^{\hat{i}}}{\partial q^i}\Big(\tfrac{\partial^2 q^i}{\partial \hat{q}^{\hat{j}}\partial \hat{q}^{\hat{k}}} + \Gamma^i_{kj}\tfrac{\partial q^k}{\partial \hat{q}^{\hat{k}}}\tfrac{\partial q^j}{\partial \hat{q}^{\hat{j}}}\Big) \tag{A.53}$$

which can be found in e.g. [Spivak, 1999, Vol. 2, p. 221] or [Abraham and Marsden, 1978, p. 145].

A.2.3 Riemannian geometry

Riemannian metric. A *Riemannian metric* is a (bilinear, symmetric and positive definite) inner product $\langle\cdot,\cdot\rangle : \mathsf{T}_{\boldsymbol{x}}\mathbb{X} \times \mathsf{T}_{\boldsymbol{x}}\mathbb{X} \to \mathbb{R}$ that may vary smoothly over the manifold $\mathbb{X}$. In terms of basis vectors we have

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle u^i \boldsymbol{\partial}_i, v^j \boldsymbol{\partial}_j\rangle = u^i v^j \underbrace{\langle \boldsymbol{\partial}_i, \boldsymbol{\partial}_j\rangle}_{M_{ij}} \tag{A.54}$$

The symmetry and positive definiteness of the inner product imply symmetry and positive definiteness of the *metric coefficients* $\boldsymbol{M} \in \mathbb{SYM}^+(n)$. A differentiable manifold equipped with a Riemannian metric is called a *Riemannian manifold*.

Levi-Civita connection. For a Riemannian manifold there exists a unique connection $\boldsymbol{\nabla}$, commonly called the *Levi-Civita connection*, determined by the additional properties

$$\boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{v}\rangle \tag{A.55a}$$
$$[\boldsymbol{u}, \boldsymbol{v}] = \boldsymbol{\nabla}_{\boldsymbol{u}} \boldsymbol{v} - \boldsymbol{\nabla}_{\boldsymbol{v}} \boldsymbol{u}. \tag{A.55b}$$

In [Boothby, 1986, Theo. VII.3.3] this is called the *fundamental theorem of Riemannian geometry*.

Combination of permutations of (A.55a) and (A.55b) yield

$$\begin{aligned}
\langle \boldsymbol{u}, \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{v}\rangle &= \boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{u}, \boldsymbol{v}\rangle \\
&= \boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle [\boldsymbol{w}, \boldsymbol{u}], \boldsymbol{v}\rangle - \langle \boldsymbol{\nabla}_{\boldsymbol{u}} \boldsymbol{w}, \boldsymbol{v}\rangle \\
&= \boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle [\boldsymbol{w}, \boldsymbol{u}], \boldsymbol{v}\rangle - \boldsymbol{u}\langle \boldsymbol{w}, \boldsymbol{v}\rangle - \langle \boldsymbol{w}, \boldsymbol{\nabla}_{\boldsymbol{u}} \boldsymbol{v}\rangle \\
&= \boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle [\boldsymbol{w}, \boldsymbol{u}], \boldsymbol{v}\rangle - \boldsymbol{u}\langle \boldsymbol{w}, \boldsymbol{v}\rangle + \langle \boldsymbol{w}, [\boldsymbol{u}, \boldsymbol{v}]\rangle - \langle \boldsymbol{w}, \boldsymbol{\nabla}_{\boldsymbol{v}} \boldsymbol{u}\rangle \\
&= \boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle [\boldsymbol{w}, \boldsymbol{u}], \boldsymbol{v}\rangle - \boldsymbol{u}\langle \boldsymbol{w}, \boldsymbol{v}\rangle + \langle \boldsymbol{w}, [\boldsymbol{u}, \boldsymbol{v}]\rangle + \boldsymbol{v}\langle \boldsymbol{w}, \boldsymbol{u}\rangle - \langle \boldsymbol{\nabla}_{\boldsymbol{v}} \boldsymbol{w}, \boldsymbol{u}\rangle \\
&= \boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle [\boldsymbol{w}, \boldsymbol{u}], \boldsymbol{v}\rangle - \boldsymbol{u}\langle \boldsymbol{w}, \boldsymbol{v}\rangle + \langle \boldsymbol{w}, [\boldsymbol{u}, \boldsymbol{v}]\rangle + \boldsymbol{v}\langle \boldsymbol{w}, \boldsymbol{u}\rangle + \langle [\boldsymbol{v}, \boldsymbol{w}], \boldsymbol{u}\rangle \\
&\quad - \langle \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{v}, \boldsymbol{u}\rangle \\
\Leftrightarrow \quad \langle \boldsymbol{u}, \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{v}\rangle &= \tfrac{1}{2}\big(\boldsymbol{w}\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \boldsymbol{v}\langle \boldsymbol{u}, \boldsymbol{w}\rangle - \boldsymbol{u}\langle \boldsymbol{v}, \boldsymbol{w}\rangle \\
&\quad + \langle \boldsymbol{w}, [\boldsymbol{u}, \boldsymbol{v}]\rangle + \langle \boldsymbol{v}, [\boldsymbol{u}, \boldsymbol{w}]\rangle - \langle \boldsymbol{u}, [\boldsymbol{v}, \boldsymbol{w}]\rangle\big).
\end{aligned} \tag{A.56}$$

This result can also be found in [Abraham and Marsden, 1978, proof of Theorem 2.7.6], but is evaluated in terms of coordinate basis vectors so that the Lie brackets cancel. Recall the previous definitions for the commutation coefficients $[\boldsymbol{\partial}_i, \boldsymbol{\partial}_j] = \gamma^k_{ij}\boldsymbol{\partial}_k$, the connection coefficients $\boldsymbol{\nabla}_{\boldsymbol{\partial}_j}\boldsymbol{\partial}_i = \mathit{\Gamma}^k_{ij}\boldsymbol{\partial}_k$ and the metric coefficients $\langle \boldsymbol{\partial}_i, \boldsymbol{\partial}_j\rangle = M_{ij}$. Evaluation of (A.56) with the basis vectors yields

$$\langle \boldsymbol{\partial}_i, \boldsymbol{\nabla}_{\boldsymbol{\partial}_k}\boldsymbol{\partial}_j\rangle = M_{is}\mathit{\Gamma}^s_{jk} = \mathit{\Gamma}_{ijk} = \tfrac{1}{2}\big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \gamma^s_{ij}M_{sk} + \gamma^s_{ik}M_{sj} - \gamma^s_{jk}M_{si}\big). \tag{A.57}$$

where we also defined the completely covariant[1] connection coefficients $\mathit{\Gamma}_{ijk}$. Note that the coordinate expressions for (A.55a) and (A.55b) are

$$\partial_k M_{ij} = \mathit{\Gamma}_{ijk} + \mathit{\Gamma}_{jik} \tag{A.58a}$$
$$\gamma^k_{ij} = \mathit{\Gamma}^k_{ji} - \mathit{\Gamma}^k_{ij} \tag{A.58b}$$

[1]Covariant refers to the placement of the indices, $\mathit{\Gamma}_{ijk}$ is not a tensor just as $\mathit{\Gamma}^i_{jk}$ is not.

The only source known to the author that states the result in (A.57) is [Misner et al., 1973, eq. 8.24], though still restricted to minimal configuration coordinates. Most books state the coefficients of the Levi-Civita connection restricted to the case of coordinate basis vector in which case they are called the *Christoffel symbols* (e.g. [Frankel, 1997, sec. 9.2]):

$$\partial_i = \tfrac{\partial}{\partial q^i} \quad \Rightarrow \quad \gamma^k_{ij} = 0, \quad \mathit{\Gamma}_{ijk} = \tfrac{1}{2}\Big(\tfrac{\partial M_{ij}}{\partial q^k} + \tfrac{\partial M_{ik}}{\partial q^j} - \tfrac{\partial M_{jk}}{\partial q^i}\Big) = \mathit{\Gamma}_{ikj} \tag{A.59}$$

In [Lurie, 2002, eq. 4.10.9] only the first three terms of (A.57) are introduced as the "generalized Christoffel symbols" in the context of minimal coordinates and non-coordinate basis vectors. This might be misleading since these quantities do not obey the transformation rule (A.52), so do not define a connection.

Gradient. The gradient vector of a scalar function f on a Riemannian manifold is defined as [Frankel, 1997, sec. 2.1d]

$$\boldsymbol{v}(f) = \langle \boldsymbol{v}, \mathrm{grad}(f) \rangle \tag{A.60}$$

with an arbitrary vector $\boldsymbol{v} \in \mathsf{T}_{\boldsymbol{x}}\mathbb{X}$. With M^{ij} being the coefficients of the inverse $\boldsymbol{M}^{-1}$ of the metric coefficients we may write it as

$$\mathbf{grad} = M^{ij}\, \boldsymbol{\partial}_j. \tag{A.61}$$

Hessian. The Hessian of a scalar function f on a Riemannian manifold may be defined as (e.g. [Bullo and Lewis, 2004, sec. 6.1.4])

$$(\mathrm{Hess}\, f)(\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{u}, \boldsymbol{\nabla}_{\boldsymbol{v}} \mathbf{grad} f \rangle. \tag{A.62}$$

With the property (A.55a) of the Levi-Civita connection $\boldsymbol{\nabla}$ we may formulate its coefficients H_{ij} as

$$H_{ij} = \langle \boldsymbol{\partial}_i, \boldsymbol{\nabla}_{\boldsymbol{\partial}_j} \mathbf{grad} f \rangle = \partial_j \langle \boldsymbol{\partial}_i, \mathbf{grad} f \rangle - \langle \boldsymbol{\nabla}_{\boldsymbol{\partial}_j} \boldsymbol{\partial}_i, \mathbf{grad} f \rangle = \partial_j(\partial_i f) - \mathit{\Gamma}^k_{ij} \partial_k f \tag{A.63}$$

Note that the second property (A.55b) of the Levi-Civita connection implies the symmetry $H_{ij} = H_{ji}$ of the Hessian.

A.2.4 The induced metric

For the Euclidean space $\mathbb{R}^\nu$ we have the standard metric

$$\langle \frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^\alpha}, \frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^\beta} \rangle = \delta_{\alpha\beta}, \qquad \delta_{\alpha\beta} = \begin{cases} 1 \ : & \alpha = \beta \\ 0 \ : & \alpha \neq \beta \end{cases} \tag{A.64}$$

The corresponding *induced metric* for the tangent space is

$$\langle \boldsymbol{\partial}_i, \boldsymbol{\partial}_j \rangle = A^\alpha_i \delta_{\alpha\beta} A^\beta_j = M_{ij} \tag{A.65}$$

As rank $\boldsymbol{A} = n$ we have $\boldsymbol{Y} = \boldsymbol{A}^{+} = (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}$ and $\boldsymbol{M}^{-1} = \boldsymbol{Y}\boldsymbol{Y}^{\top}$. The coefficients of the Levi-Civita connection with the induced metric may be written as

$$
\begin{aligned}
\Gamma_{ijk} &= \tfrac{1}{2}\big(\partial_k(A^\alpha_i\delta_{\alpha\beta}A^\beta_j) + \partial_j(A^\alpha_i\delta_{\alpha\beta}A^\beta_k) - \partial_i(A^\alpha_j\delta_{\alpha\beta}A^\beta_k) \\
&\quad + Y^s_\upsilon\big(\partial_i A^\upsilon_j - \partial_j A^\upsilon_i\big)(A^\alpha_s\delta_{\alpha\beta}A^\beta_k) + Y^s_\upsilon\big(\partial_i A^\upsilon_k - \partial_k A^\upsilon_i\big)(A^\alpha_s\delta_{\alpha\beta}A^\beta_j) \\
&\quad - Y^s_\upsilon\big(\partial_j A^\upsilon_k - \partial_k A^\upsilon_j\big)(A^\alpha_s\delta_{\alpha\beta}A^\beta_i)\big). \\
&= \tfrac{1}{2}\big(\partial_k A^\upsilon_i\,\underbrace{(\delta^\alpha_\upsilon - Y^s_\upsilon A^\alpha_s)\delta_{\alpha\beta}A^\beta_j}_{0} + \partial_k A^\upsilon_j\,\underbrace{(\delta^\alpha_\upsilon + Y^s_\upsilon A^\alpha_s)\delta_{\alpha\beta}A^\beta_i}_{2\delta_{\alpha\beta}A^\beta_i} + \partial_j A^\upsilon_i\,\underbrace{(\delta^\alpha_\upsilon - Y^s_\upsilon A^\alpha_s)\delta_{\alpha\beta}A^\beta_k}_{0} \\
&\quad + \partial_j A^\upsilon_k\,\underbrace{(\delta^\alpha_\upsilon - Y^s_\upsilon A^\alpha_s)\delta_{\alpha\beta}A^\beta_i}_{0} - \partial_i A^\upsilon_j\,\underbrace{(\delta^\alpha_\upsilon - Y^s_\upsilon A^\alpha_s)\delta_{\alpha\beta}A^\beta_k}_{0} - \partial_i A^\upsilon_k\,\underbrace{(\delta^\alpha_\upsilon - Y^s_\upsilon A^\alpha_s)\delta_{\alpha\beta}A^\beta_j}_{0}\big) \\
&= A^\alpha_i\delta_{\alpha\beta}\partial_k A^\beta_j && \text{(A.66a)} \\
\Gamma^i_{jk} &= M^{is}\Gamma_{sjk} = Y^i_\alpha\partial_k A^\alpha_j && \text{(A.66b)}
\end{aligned}
$$

Consider a parameterized curve $\boldsymbol{x} : \mathbb{R} \mapsto \mathbb{X}$ with $\dot{x}^\alpha = A^\alpha_i\xi^i$. The velocity vector is

$$
\dot{\boldsymbol{x}} = \dot{x}^\alpha\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}x^\alpha} = A^\alpha_i\xi^i\big(Y^j_\alpha\boldsymbol{\partial}_j + \Phi^\kappa_\alpha\boldsymbol{\partial}^\perp_\kappa\big) = \xi^i\boldsymbol{\partial}_i. \tag{A.67}
$$

The acceleration vector is

$$
\begin{aligned}
\ddot{\boldsymbol{x}} = \ddot{x}^\alpha\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}x^\alpha} &= \big(A^\alpha_i\dot{\xi}^i + \partial_k A^\alpha_i\xi^k\xi^i\big)\big(Y^j_\alpha\boldsymbol{\partial}_j + \Phi^\kappa_\alpha\boldsymbol{\partial}^\perp_\kappa\big) \\
&= \big(\dot{\xi}^i + Y^i_\alpha\partial_k A^\alpha_j\xi^j\xi^k\big)\boldsymbol{\partial}_j + \partial_k A^\alpha_i\xi^k\xi^i\Phi^\kappa_\alpha\boldsymbol{\partial}^\perp_\kappa
\end{aligned} \tag{A.68}
$$

A.2.5 Riemannian geometry and Lagrangian mechanics.

Consider a mechanical system parameterized by the config. coordinates $(x^1, \dots, x^\nu)(t) \in \mathbb{X}$ and the velocity coordinates $(\xi^1, \dots, \xi^n)(t) \in \mathbb{R}^n$ related by $\dot{x}^\alpha = A^\alpha_i\xi^i$. Define the velocity *vector* as $\boldsymbol{v} = \xi^i\boldsymbol{\partial}_i$. The correspondence between Riemannian geometry and Lagrangian mechanics is established by relating the metric coefficients to the kinetic energy:

$$
\mathcal{T} = \tfrac{1}{2}M_{ij}\xi^i\xi^j = \tfrac{1}{2}\langle\boldsymbol{v}, \boldsymbol{v}\rangle. \tag{A.69}
$$

This determines the (Levi-Civita) connection coefficients Γ_{ijk} and we have

$$
\langle\boldsymbol{\partial}_i, \boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{v}\rangle = M_{ij}\big(\partial_k\xi^j + \Gamma^j_{lk}\xi^l\big)\xi^k = M_{ij}\dot{\xi}^j + \Gamma_{ijk}\xi^j\xi^k = f^{\mathrm{M}}_i, \qquad i = 1\dots n \tag{A.70}
$$

i.e. the coefficients f^{M}_i of the generalized inertia force are the projections of the covariant derivative $\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{v}$ of the velocity vector along itself to the tangent vectors $\boldsymbol{\partial}_i$. Equivalently we could say f^{M}_i are the coefficients of the co-vector corresponding to $\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{v}$.

The conservative force of a potential $\mathcal{V}(\boldsymbol{x})$ are $f^{\mathrm{V}}_i = \langle\boldsymbol{\partial}_i, \mathbf{grad}\mathcal{V}\rangle$.

Overall, we may give the equation of motion of a conservative, autonomous mechanical system in the vector form

$$
\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{v} + \mathbf{grad}\mathcal{V} = \mathbf{0}. \tag{A.71}
$$

A.2.6 Relation to Euclidean geometry.

In the previous subsection on Euclidean geometry we considered a special choice of coordinates, such that the kinetic energy $\mathcal{T}$ resembles the Euclidean metric and the inertia matrix $M_{ij} = A_i^\alpha \delta_{\alpha\beta} A_j^\beta$ is the *induced metric* for the tangent space. In total we have

$$M_{ij} = A_i^\alpha \delta_{\alpha\beta} A_j^\beta, \qquad \varGamma_{ijk} = A_i^\alpha \delta_{\alpha\beta} \partial_k A_j^\beta \tag{A.72a}$$

$$M^{ij} = (\boldsymbol{A}^+)^i_\alpha \delta^{\alpha\beta} (\boldsymbol{A}^+)^j_\beta, \qquad \varGamma^i_{jk} = M^{is} \varGamma_{sjk} = (\boldsymbol{A}^+)^i_\alpha \partial_k A_j^\alpha \tag{A.72b}$$

The computations here might take some extra lines but it is all pretty straight forward when expressing the commutation symbols (??) with the pseudo inverse and using the defining criteria of the pseudo inverse.

Probably remove section on Euclidean geometry and put the conclusions here.

The Levi-Civita connection resembles the result we would get from Euclidean geometry but *without requiring an isometric embedding* Due to the arbitrary choice of configuration coordinates the standard metric for the (embedded) configuration space is meaningless for the mechanics.

A.3 Transformation rules

We use a parameterization with (possibly redundant) configuration coordinates $\boldsymbol{x} \in \mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^\nu \,|\, \boldsymbol{\phi}(\boldsymbol{x})\}$ and velocity coordinates $\boldsymbol{\xi} \in \mathbb{R}^n$ that are related by $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi}$. The configuration coordinates $\boldsymbol{x}$ and the velocity coordinates $\boldsymbol{\xi}$ can be transformed independently from each other. The following subsections summarize the resulting coordinate expressions and state the transformation rules.

A.3.1 Recap coordinate expressions

Kinematics

$$\xi^i = Y^i_\alpha \dot{x}^\alpha, \qquad \dot{x}^\alpha = A_i^\alpha \xi^i \tag{A.73}$$

Directional derivative

$$\partial_i = A_i^\alpha \frac{\partial}{\partial x^\alpha}, \qquad \partial_i \partial_j - \partial_j \partial_i = \gamma^k_{ij} \partial_k \tag{A.74}$$

commutation coefficients

$$\gamma^k_{ij} = \left(\frac{\partial Y^k_\alpha}{\partial x^\beta} - \frac{\partial Y^k_\beta}{\partial x^\alpha} \right) A_j^\beta A_i^\alpha = A_i^\alpha \partial_j Y^k_\alpha - A_j^\alpha \partial_i Y^k_\alpha = Y^k_\alpha \big(\partial_i A_j^\alpha - \partial_j A_i^\alpha \big) \tag{A.75}$$

With the Riemannian metric $\boldsymbol{M}$ we have the coefficients $\varGamma$ of the Levi-Civita connection

$$M_{is} \varGamma^s_{jk} = \varGamma_{ijk} = \tfrac{1}{2} \big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \gamma^s_{ij} M_{sk} + \gamma^s_{ik} M_{sj} - \gamma^s_{jk} M_{si} \big) \tag{A.76}$$

Riemann curvature tensor

$$R^i_{jkl} = \partial_k \varGamma^i_{jl} - \partial_l \varGamma^i_{jk} + \varGamma^i_{sk} \varGamma^s_{jl} - \varGamma^i_{sl} \varGamma^s_{jk} - \varGamma^i_{js} \gamma^s_{kl} \tag{A.77a}$$

$$M_{is} R^s_{jkl} = R_{ijkl} = \partial_k \varGamma_{ijl} - \partial_l \varGamma_{ijk} - \varGamma_{sik} \varGamma^s_{jl} + \varGamma_{sil} \varGamma^s_{jk} - \varGamma_{ijs} \gamma^s_{kl} \tag{A.77b}$$

A.3.2 Change of configuration coordinates

Consider a different parameterization of the configuration space

$$\mathbb{X} \cong \{\hat{\boldsymbol{x}} \in \mathbb{R}^{\hat{\nu}} \,|\, \hat{\boldsymbol{\phi}}(\hat{\boldsymbol{x}}) = 0\}. \tag{A.78}$$

Let the relation to the original coordinates be

$$x^\alpha = \chi^\alpha(\hat{\boldsymbol{x}}), \qquad \alpha = 1, \ldots, \nu, \tag{A.79a}$$

$$\hat{x}^{\hat{\alpha}} = \hat{\chi}^{\hat{\alpha}}(\boldsymbol{x}), \qquad \hat{\alpha} = 1, \ldots, \hat{\nu}. \tag{A.79b}$$

Note that the dimension of the embedding space $\hat{\nu}$ can be different to the original dimension ν, only $\hat{\nu} - \operatorname{rank} \frac{\partial \hat{\boldsymbol{\phi}}}{\partial \hat{\boldsymbol{x}}} = \dim \mathbb{X} = n$ has to hold. Consequently the Jacobi matrices $\frac{\partial \boldsymbol{\chi}}{\partial \hat{\boldsymbol{x}}}$ and $\frac{\partial \hat{\boldsymbol{\chi}}}{\partial \boldsymbol{x}}$ are rectangular. Nevertheless we have the identities

$$\frac{\partial \chi^\alpha}{\partial \hat{x}^{\hat{\alpha}}} \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\beta} = \delta^\alpha_\beta, \qquad \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha} \frac{\partial \chi^\alpha}{\partial \hat{x}^{\hat{\beta}}} = \delta^{\hat{\alpha}}_{\hat{\beta}} \tag{A.80}$$

Kinematic matrices. We choose the *same basis* $\boldsymbol{\xi}$ as with the original coordinates, i.e.

$$\xi^i = Y^i_\alpha(\boldsymbol{x})\dot{x}^\alpha = Y^i_\alpha(\boldsymbol{\chi}(\hat{\boldsymbol{x}}))\tfrac{\partial \chi^\alpha}{\partial \hat{x}^{\hat{\alpha}}}(\hat{\boldsymbol{x}})\,\dot{\hat{x}}^{\hat{\alpha}} = \hat{Y}^i_{\hat{\alpha}}(\hat{\boldsymbol{x}})\dot{\hat{x}}^{\hat{\alpha}}, \qquad i = 1, \ldots, n. \tag{A.81}$$

The explicit dependencies should be clear from the context, so they will be dropped for the following.

$$\hat{Y}^i_{\hat{\alpha}} = Y^i_\alpha \tfrac{\partial \chi^\alpha}{\partial \hat{x}^{\hat{\alpha}}}, \qquad \hat{A}^{\hat{\alpha}}_i = \tfrac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha} A^\alpha_i, \qquad Y^i_\alpha = Y^i_{\hat{\alpha}} \tfrac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha}, \qquad A^\alpha_i = \tfrac{\partial \chi^\alpha}{\partial \hat{x}^{\hat{\alpha}}} \hat{A}^{\hat{\alpha}}_i. \tag{A.82}$$

Directional derivative. Consider the function $\mathcal{V}(\boldsymbol{x}) = \mathcal{V}(\boldsymbol{\chi}(\hat{\boldsymbol{x}})) =: \hat{\mathcal{V}}(\hat{\boldsymbol{x}})$. Its differential is invariant to the change of coordinates, i.e. $\partial_i \hat{\mathcal{V}}(\hat{\boldsymbol{x}}) = \big(\partial_i \mathcal{V}\big)(\boldsymbol{\chi}(\hat{\boldsymbol{x}}))$, since

$$\partial_i \hat{\mathcal{V}} = \frac{\partial \hat{\mathcal{V}}}{\partial \hat{x}^{\hat{\alpha}}} \hat{A}^{\hat{\alpha}}_i = \frac{\partial \mathcal{V}}{\partial x^\alpha} \frac{\partial \chi^\alpha}{\partial \hat{x}^{\hat{\alpha}}} \hat{A}^{\hat{\alpha}}_i = \frac{\partial \mathcal{V}}{\partial x^\alpha} A^\alpha_i = \partial_i \mathcal{V}. \tag{A.83}$$

Commutation coefficients. The Commutation coefficients are invariant, i.e. $\hat{\gamma}^k_{ij}(\hat{\boldsymbol{x}}) = \gamma^k_{ij}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}))$, to the change of coordinates:

$$\begin{aligned}
\hat{\gamma}^k_{ij} &= \hat{Y}^k_{\hat{\alpha}} \big(\partial_i \hat{A}^{\hat{\alpha}}_j - \partial_j \hat{A}^{\hat{\alpha}}_i\big) \\
&= Y^k_\beta \frac{\partial \chi^\beta}{\partial \hat{x}^{\hat{\alpha}}} \Big(\partial_i \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha} A^\alpha_j + \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha} \partial_i A^\alpha_j - \partial_j \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha} A^\alpha_i - \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha} \partial_j A^\alpha_i\Big) \\
&= Y^k_\beta \underbrace{\frac{\partial \chi^\beta}{\partial \hat{x}^{\hat{\alpha}}} \frac{\partial \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha}}_{\delta^\beta_\alpha} \big(\partial_i A^\alpha_j - \partial_j A^\alpha_i\big) + Y^k_\beta \frac{\partial \chi^\beta}{\partial \hat{x}^{\hat{\alpha}}} \underbrace{\Big(\frac{\partial^2 \hat{\chi}^{\hat{\alpha}}}{\partial x^\beta \partial x^\alpha} - \frac{\partial^2 \hat{\chi}^{\hat{\alpha}}}{\partial x^\alpha \partial x^\beta}\Big)}_{0} A^\beta_j A^\alpha_i = \gamma^k_{ij}.
\end{aligned} \tag{A.84}$$

Connection coefficients. Since the differential $\partial_k M_{ij}$ of the inertia matrix and the commutation coefficients γ^k_{ij} are invariant to the change of configuration coordinates, so are the connection coefficients, i.e. $\hat{\Gamma}_{ijk}(\hat{\boldsymbol{x}}) = \Gamma_{ijk}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}))$.

Lagrange's equation. Consider the Lagrangian $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\xi}, t) = \mathcal{L}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}), \boldsymbol{\xi}, t) =: \hat{\mathcal{L}}(\hat{\boldsymbol{x}}, \boldsymbol{\xi}, t)$ in terms of the transformed configuration coordinates. We have $\partial_i \hat{\mathcal{L}}(\hat{\boldsymbol{x}}) = \partial_i \mathcal{L}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}))$ and $\hat{\gamma}^k_{ij}(\hat{\boldsymbol{x}}) = \gamma^k_{ij}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}))$ and $\partial\hat{\mathcal{L}}/\partial\xi^i = \partial\mathcal{L}/\partial\xi^i$. Consequently Lagrange's equation is invariant to the change in configuration coordinates, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\hat{\mathcal{L}}}{\partial\xi^i} + \hat{\gamma}^k_{ij}\xi^j\frac{\partial\hat{\mathcal{L}}}{\partial\xi^k} - \hat{A}^{\hat{\alpha}}_i\frac{\partial\hat{\mathcal{L}}}{\partial\hat{x}^{\hat{\alpha}}} = \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\xi^i} + \gamma^k_{ij}\xi^j\frac{\partial\mathcal{L}}{\partial\xi^k} - A^\alpha_i\frac{\partial\mathcal{L}}{\partial x^\alpha}\right)\Big|_{\boldsymbol{x}=\boldsymbol{\chi}(\hat{\boldsymbol{x}})}. \tag{A.85}$$

Transport map. Consider the function $\mathcal{V}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) = \mathcal{V}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}), \boldsymbol{\chi}(\hat{\boldsymbol{x}}_\mathsf{R})) =: \hat{\mathcal{V}}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}_\mathsf{R})$ with the transport map

$$\tfrac{\partial\mathcal{V}}{\partial x^\alpha_\mathsf{R}} A^\alpha_i(\boldsymbol{x}_\mathsf{R}) = Q^j_i(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\tfrac{\partial\mathcal{V}}{\partial x^\alpha} A^\alpha_j(\boldsymbol{x}) \tag{A.86}$$

The invariance, i.e. $\hat{\boldsymbol{Q}}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}_\mathsf{R}) = \boldsymbol{Q}(\boldsymbol{\chi}(\hat{\boldsymbol{x}}), \boldsymbol{\chi}(\hat{\boldsymbol{x}}_\mathsf{R}))$, is a direct consequence of the invariance of the directional derivative:

$$\begin{aligned}\tfrac{\partial\hat{\mathcal{V}}}{\partial\hat{x}^{\hat{\alpha}}_\mathsf{R}}\hat{A}^{\hat{\alpha}}_i(\hat{\boldsymbol{x}}_\mathsf{R}) = \tfrac{\partial\mathcal{V}}{\partial x^\alpha_\mathsf{R}} A^\alpha_i(\boldsymbol{\chi}(\hat{\boldsymbol{x}}_\mathsf{R})) &= Q^j_i(\boldsymbol{\chi}(\hat{\boldsymbol{x}}), \boldsymbol{\chi}(\hat{\boldsymbol{x}}_\mathsf{R}))\tfrac{\partial\mathcal{V}}{\partial x^\alpha} A^\alpha_j(\boldsymbol{\chi}(\boldsymbol{x}))\\ &= \underbrace{Q^j_i(\boldsymbol{\chi}(\hat{\boldsymbol{x}}), \boldsymbol{\chi}(\hat{\boldsymbol{x}}_\mathsf{R}))}_{\hat{Q}^j_i(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}_\mathsf{R})}\tfrac{\partial\hat{\mathcal{V}}}{\partial\hat{x}^{\hat{\alpha}}}\hat{A}^{\hat{\alpha}}_j(\hat{\boldsymbol{x}}).\end{aligned} \tag{A.87}$$

A.3.3 Change of velocity coordinates / basis

Consider we change the velocity coordinates, but leave the configuration coordinates $\boldsymbol{x}$ unchanged:

$$\xi^i = W^i_{\hat{i}}(\boldsymbol{x})\hat{\xi}^{\hat{i}}, \qquad \hat{\xi}^{\hat{i}} = Z^{\hat{i}}_i(\boldsymbol{x})\xi^i, \qquad i, \hat{i} = 1, \ldots, n, \tag{A.88}$$

with the transformation matrix $\boldsymbol{W}(\boldsymbol{x}) = \boldsymbol{Z}^{-1}(\boldsymbol{x}) \in \mathbb{R}^{n\times n}$. We have the kinematic relation

$$\dot{x}^\alpha = A^\alpha_i(\boldsymbol{x})\xi^i = \underbrace{A^\alpha_i(\boldsymbol{x})W^i_{\hat{i}}(\boldsymbol{x})}_{\hat{A}^\alpha_{\hat{i}}(\boldsymbol{x})}\hat{\xi}^{\hat{i}} \tag{A.89}$$

Kinematic matrices.

$$\hat{Y}^{\hat{i}}_\alpha = Z^{\hat{i}}_i Y^i_\alpha, \qquad \hat{A}^\alpha_{\hat{i}} = A^\alpha_i W^i_{\hat{i}}, \qquad Y^i_\alpha = W^i_{\hat{i}}\hat{Y}^{\hat{i}}_\alpha, \qquad A^\alpha_i = \hat{A}^\alpha_{\hat{i}} Z^{\hat{i}}_i \tag{A.90}$$

Directional derivative.

$$\partial_{\hat{i}} V = \tfrac{\partial V}{\partial x^\alpha}\hat{A}^\alpha_{\hat{i}} = \tfrac{\partial V}{\partial x^\alpha} A^\alpha_i W^i_{\hat{i}} = W^i_{\hat{i}}\partial_i V \tag{A.91}$$

Commutation coefficients.

$$\begin{aligned}
\hat{\gamma}^{\hat{k}}_{\hat{i}\hat{j}} &= \hat{Y}^{\hat{k}}_{\alpha}\big(\partial_{\hat{i}}\hat{A}^{\alpha}_{\hat{j}} - \partial_{\hat{j}}\hat{A}^{\alpha}_{\hat{i}}\big) \\
&= Z^{\hat{k}}_{k}Y^{k}_{\alpha}\big(\partial_{\hat{i}}A^{\alpha}_{j}W^{j}_{\hat{j}} + A^{\alpha}_{j}\partial_{\hat{i}}W^{j}_{\hat{j}} - \partial_{\hat{j}}A^{\alpha}_{i}W^{i}_{\hat{i}} - A^{\alpha}_{i}\partial_{\hat{j}}W^{i}_{\hat{i}}\big) \\
&= Z^{\hat{k}}_{k}\big(W^{i}_{\hat{i}}W^{j}_{\hat{j}}\underbrace{Y^{k}_{\alpha}(\partial_{i}A^{\alpha}_{j} - \partial_{j}A^{\alpha}_{i})}_{\gamma^{k}_{ij}} + \underbrace{Y^{k}_{\alpha}A^{\alpha}_{j}}_{\delta^{k}_{j}}\partial_{\hat{i}}W^{j}_{\hat{j}} - \underbrace{Y^{k}_{\alpha}A^{\alpha}_{i}}_{\delta^{k}_{i}}\partial_{\hat{j}}W^{i}_{\hat{i}}\big) \\
&= Z^{\hat{k}}_{k}\big(W^{i}_{\hat{i}}W^{j}_{\hat{j}}\gamma^{k}_{ij} + \partial_{\hat{i}}W^{k}_{\hat{j}} - \partial_{\hat{j}}W^{k}_{\hat{i}}\big)
\end{aligned} \tag{A.92}$$

Connection coefficients.

$$\hat{\Gamma}^{\hat{i}}_{\hat{k}\hat{j}} = Z^{\hat{i}}_{i}\big(\Gamma^{i}_{kj}W^{k}_{\hat{k}}W^{j}_{\hat{j}} + \hat{\partial}_{\hat{j}}W^{i}_{\hat{k}}\big) \tag{A.93}$$

Lagrange's equation. Consider the Lagrangian $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\xi}, t) = \mathcal{L}(\boldsymbol{x}, \boldsymbol{W}(\boldsymbol{x})\hat{\boldsymbol{\xi}}, t) =: \hat{\mathcal{L}}(\boldsymbol{x}, \hat{\boldsymbol{\xi}}, t)$ in terms of the transformed velocity coordinates. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\hat{\mathcal{L}}}{\partial\hat{\xi}^{\hat{i}}} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\xi^{i}}W^{i}_{\hat{i}} + \frac{\partial\mathcal{L}}{\partial\xi^{i}}\partial_{j}W^{i}_{\hat{i}}\xi^{j}, \tag{A.94a}$$

$$\hat{\gamma}^{\hat{k}}_{\hat{i}\hat{j}}\hat{\xi}^{\hat{j}}\frac{\partial\hat{\mathcal{L}}}{\partial\hat{\xi}^{\hat{k}}} = \big(W^{i}_{\hat{i}}W^{j}_{\hat{j}}\gamma^{k}_{ij} + W^{s}_{\hat{i}}\partial_{s}W^{k}_{\hat{j}} - W^{s}_{\hat{j}}\partial_{s}W^{k}_{\hat{i}}\big)Z^{\hat{j}}_{l}\xi^{l}\frac{\partial\mathcal{L}}{\partial\xi^{k}} \tag{A.94b}$$

$$\partial_{\hat{i}}\hat{\mathcal{L}} = W^{i}_{\hat{i}}\left(\partial_{i}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial\xi^{k}}\partial_{i}W^{k}_{\hat{j}}Z^{\hat{j}}_{j}\xi^{j}\right) \tag{A.94c}$$

Adding these three terms together we find that Lagrange's equations are covariant

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\hat{\mathcal{L}}}{\partial\hat{\xi}^{\hat{i}}} + \hat{\gamma}^{\hat{k}}_{\hat{i}\hat{j}}\hat{\xi}^{\hat{j}}\frac{\partial\hat{\mathcal{L}}}{\partial\hat{\xi}^{\hat{k}}} - \hat{A}^{\alpha}_{\hat{i}}\frac{\partial\hat{\mathcal{L}}}{\partial x^{\alpha}} = W^{i}_{\hat{i}}\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\xi^{i}} + \gamma^{k}_{ij}\xi^{j}\frac{\partial\mathcal{L}}{\partial\xi^{k}} - A^{\alpha}_{i}\frac{\partial\mathcal{L}}{\partial x^{\alpha}}\right). \tag{A.95}$$

System inertia matrix / metric The kinetic energy $\mathcal{T}$ must be the same in terms of both velocity coordinates, so

$$\mathcal{T} = \tfrac{1}{2}\hat{M}_{\hat{i}\hat{j}}\hat{\xi}^{\hat{i}}\hat{\xi}^{\hat{j}} = \tfrac{1}{2}M_{ij}\xi^{i}\xi^{j} \tag{A.96}$$

we get the relation

$$\hat{M}_{\hat{i}\hat{j}} = M_{ij}W^{i}_{\hat{i}}W^{j}_{\hat{j}}, \qquad \hat{M}^{\hat{i}\hat{j}} = M^{ij}Z^{\hat{i}}_{i}Z^{\hat{j}}_{j}. \tag{A.97}$$

The differential of the inertia matrix transforms as

$$\partial_{\hat{k}}\hat{M}_{\hat{i}\hat{j}} = W^{i}_{\hat{i}}W^{j}_{\hat{j}}W^{k}_{\hat{k}}\partial_{k}M_{ij} + (W^{j}_{\hat{j}}\partial_{\hat{k}}W^{i}_{\hat{i}} + W^{i}_{\hat{i}}\partial_{\hat{k}}W^{j}_{\hat{j}})M_{ij}. \tag{A.98}$$

Levi-Civita connection coefficients It is useful to first consider the transformation rule for the commutation coefficients with lowered indices

$$\gamma_{\hat{i}\hat{j}\hat{k}} = \gamma^{\hat{l}}_{\hat{i}\hat{j}}\hat{M}_{\hat{l}\hat{k}} = W^i_{\hat{i}}W^j_{\hat{j}}W^k_{\hat{k}}\gamma_{ijk} + \big(W^i_{\hat{k}}\partial_{\hat{i}}W^j_{\hat{j}} - W^j_{\hat{k}}\partial_{\hat{j}}W^i_{\hat{i}}\big)M_{ij} \tag{A.99}$$

Utilize (A.98) and (A.99) to compute the transformation rule for the connection coefficients

$$\begin{aligned}
\hat{\Gamma}_{\hat{i}\hat{j}\hat{k}} &= \tfrac{1}{2}\Big(\partial_{\hat{k}}\hat{M}_{\hat{i}\hat{j}} + \partial_{\hat{j}}\hat{M}_{\hat{i}\hat{k}} - \partial_{\hat{k}}\hat{M}_{\hat{j}\hat{k}} + \hat{\gamma}_{\hat{i}\hat{j}\hat{k}} + \hat{\gamma}_{\hat{i}\hat{k}\hat{j}} - \hat{\gamma}_{\hat{j}\hat{k}\hat{i}}\Big)\\
&= W^i_{\hat{i}}W^j_{\hat{j}}W^k_{\hat{k}}\underbrace{\tfrac{1}{2}\big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \hat{\gamma}_{ijk} + \hat{\gamma}_{ikj} - \hat{\gamma}_{jki}\big)}_{\Gamma_{ijk}}\\
&\qquad + \tfrac{1}{2}\big(W^j_{\hat{j}}\partial_{\hat{k}}W^i_{\hat{i}} + W^i_{\hat{i}}\partial_{\hat{k}}W^j_{\hat{j}} + W^i_{\hat{k}}\partial_{\hat{i}}W^j_{\hat{j}} - W^j_{\hat{k}}\partial_{\hat{j}}W^i_{\hat{i}}\\
&\qquad\quad + W^j_{\hat{k}}\partial_{\hat{j}}W^i_{\hat{i}} + W^i_{\hat{i}}\partial_{\hat{j}}W^j_{\hat{k}} + W^i_{\hat{j}}\partial_{\hat{i}}W^j_{\hat{k}} - W^j_{\hat{j}}\partial_{\hat{k}}W^i_{\hat{i}}\\
&\qquad\quad - W^j_{\hat{k}}\partial_{\hat{i}}W^i_{\hat{j}} - W^i_{\hat{j}}\partial_{\hat{i}}W^j_{\hat{k}} - W^i_{\hat{i}}\partial_{\hat{j}}W^j_{\hat{k}} + W^j_{\hat{i}}\partial_{\hat{k}}W^i_{\hat{j}}\big)M_{ij}\\
&= W^i_{\hat{i}}W^j_{\hat{j}}W^k_{\hat{k}}\tfrac{1}{2}\big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \hat{\gamma}_{ijk} + \hat{\gamma}_{ikj} - \hat{\gamma}_{jki}\big)\\
&\qquad + \tfrac{1}{2}\big(W^i_{\hat{i}}(\partial_{\hat{k}}W^j_{\hat{j}} + \partial_{\hat{j}}W^j_{\hat{k}}) + W^i_{\hat{j}}(\partial_{\hat{k}}W^j_{\hat{i}} - \partial_{\hat{i}}W^j_{\hat{k}}) + W^i_{\hat{k}}(\partial_{\hat{j}}W^j_{\hat{i}} - \partial_{\hat{i}}W^j_{\hat{j}})\\
&\qquad + W^j_{\hat{k}}\big(\partial_{\hat{i}}W^i_{\hat{j}} - \partial_{\hat{j}}W^i_{\hat{i}}\big) + W^j_{\hat{j}}\big(\partial_{\hat{i}}W^i_{\hat{k}} - \partial_{\hat{k}}W^i_{\hat{i}}\big) - W^j_{\hat{i}}\big(\partial_{\hat{j}}W^i_{\hat{k}} - \partial_{\hat{k}}W^i_{\hat{j}}\big)\big)M_{ij}\\
&= W^i_{\hat{i}}W^j_{\hat{j}}W^k_{\hat{k}}\Gamma_{ijk} + W^i_{\hat{i}}\partial_{\hat{k}}W^j_{\hat{j}}M_{ij}
\end{aligned} \tag{A.100}$$

and

$$\hat{\Gamma}^{\hat{i}}_{\hat{j}\hat{k}} = \hat{M}^{\hat{i}\hat{s}}\hat{\Gamma}_{\hat{s}\hat{j}\hat{k}} = M^{il}Z^{\hat{i}}_iZ^{\hat{s}}_l\big(W^s_{\hat{s}}W^j_{\hat{j}}W^k_{\hat{k}}\Gamma_{sjk} + W^s_{\hat{s}}\partial_{\hat{k}}W^j_{\hat{j}}M_{sj}\big) = Z^{\hat{i}}_i\big(W^j_{\hat{j}}W^k_{\hat{k}}\Gamma^i_{jk} + \partial_{\hat{k}}W^i_{\hat{j}}\big) \tag{A.101}$$

Covariant derivative. The partial derivatives of the coefficients of a vector field v do *not* transform like a tensor

$$\partial_{\hat{j}}\hat{v}^{\hat{i}} = W^j_{\hat{j}}(Z^{\hat{i}}_i\partial_j v^i + \partial_j Z^{\hat{i}}_i v^i) \tag{A.102}$$

but the coefficients of the covariant derivative do

$$\begin{aligned}
\nabla_{\hat{j}}\hat{v}^{\hat{i}} = \partial_{\hat{j}}\hat{v}^{\hat{i}} + \hat{\Gamma}^{\hat{i}}_{\hat{k}\hat{j}}\hat{v}^{\hat{k}} &= W^j_{\hat{j}}(Z^{\hat{i}}_i\partial_j v^i + \partial_j Z^{\hat{i}}_i v^i) + Z^{\hat{i}}_i(W^k_{\hat{k}}W^j_{\hat{j}}\Gamma^i_{kj} + W^j_{\hat{j}}\partial_j W^i_{\hat{k}})Z^{\hat{k}}_l v^l\\
&= W^j_{\hat{j}}Z^{\hat{i}}_i\underbrace{\big(\partial_j v^i + \Gamma^i_{kj}v^k\big)}_{\nabla_j v^i} + W^j_{\hat{j}}\underbrace{\big(\partial_j Z^{\hat{i}}_i - \partial_j Z^{\hat{i}}_i\big)}_{0}v^i
\end{aligned} \tag{A.103}$$

Transport map. The covariance of the transport map $\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})$ follows directly from the covariance of the directional derivatives:

$$\begin{aligned}
&& \partial^\mathsf{R}_{\hat{i}}\mathcal{V}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) &= -Q^{\hat{j}}_{\hat{i}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\partial_{\hat{j}}\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\\
\Leftrightarrow && W^i_{\hat{i}}(\boldsymbol{x}_\mathsf{R})\partial^\mathsf{R}_i\mathcal{V}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) &= -\hat{Q}^{\hat{j}}_{\hat{i}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})W^j_{\hat{j}}(\boldsymbol{x})\partial_j\mathcal{V}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})\\
\Leftrightarrow && \partial^\mathsf{R}_i\mathcal{V}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R}) &= -\underbrace{Z^{\hat{i}}_i(\boldsymbol{x}_\mathsf{R})\hat{Q}^{\hat{j}}_{\hat{i}}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})W^j_{\hat{j}}(\boldsymbol{x})}_{Q^j_i(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})}\partial_j\mathcal{V}(\boldsymbol{x}, \boldsymbol{x}_\mathsf{R})
\end{aligned} \tag{A.104}$$

TBD. For the scalar quantity $\mathcal{T} = \frac{1}{2} M_{ij} v^i v^j$ we are looking for a *(co-)vector* p with which $\dot{\mathcal{T}} = v^i p_i$ along a curve $t \mapsto \boldsymbol{x}(t)$ with $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi}$. Taking the time derivative yields

$$\dot{\mathcal{T}} = v^i \underbrace{\left(M_{ij}\dot{v}^j + \tfrac{1}{2}\partial_k M_{ij}\xi^k v^j\right)}_{\text{not a tensor}} = v^i \underbrace{\left(M_{ij}\dot{v}^j + (\tfrac{1}{2}\partial_k M_{ij}\xi^k + S_{ij})v^j\right)}_{=p_i}. \tag{A.105}$$

Since the first term is not a tensor, we add some skew symmetric coefficients $S_{ij} = -S_{ji}$ that cancel out in the multiplication but may be used to make p a tensor. From the transformation rule $\hat{p}_{\hat{i}} = W^i_{\hat{i}} p_i$ we can derive a transformation rule for S:

$$\begin{aligned}
\hat{p}_{\hat{i}} &= \hat{M}_{\hat{i}\hat{j}}\dot{\hat{v}}^{\hat{j}} + \left(\tfrac{1}{2}\hat{\partial}_{\hat{k}}\hat{M}_{\hat{i}\hat{j}}\hat{\xi}^{\hat{k}} + \hat{S}_{\hat{i}\hat{j}}\right)\hat{v}^{\hat{j}} \\
&= W^i_{\hat{i}} M_{il} W^l_{\hat{j}} \tfrac{\mathrm{d}}{\mathrm{d}t}\left(Z^{\hat{j}}_j v^j\right) + \left(\tfrac{1}{2}\partial_k\left(W^i_{\hat{i}} M_{il} W^l_{\hat{j}}\right)\xi^k + \hat{S}_{\hat{i}\hat{j}}\right) Z^{\hat{j}}_j v^j \\
&= W^i_{\hat{i}} M_{il} W^l_{\hat{j}}\left(Z^{\hat{j}}_j \dot{v}^j + \partial_k Z^{\hat{j}}_j \xi^k v^j\right) \\
&\quad + \left(\tfrac{1}{2}\left(\partial_k W^i_{\hat{i}} M_{il} W^l_{\hat{j}} + W^i_{\hat{i}}\partial_k M_{il} W^l_{\hat{j}} + W^i_{\hat{i}} M_{il}\partial_k W^l_{\hat{j}}\right)\xi^k + \hat{S}_{\hat{i}\hat{j}}\right) Z^{\hat{j}}_j v^j \\
&= W^i_{\hat{i}} M_{ij}\dot{v}^j - W^i_{\hat{i}} M_{il}\partial_k W^l_{\hat{j}} Z^{\hat{j}}_j \xi^k v^j \\
&\quad + \left(\tfrac{1}{2}\left(\partial_k W^i_{\hat{i}} M_{ij} + W^i_{\hat{i}}\partial_k M_{ij} + W^i_{\hat{i}} M_{il}\partial_k W^l_{\hat{j}} Z^{\hat{j}}_j\right)\xi^k + \hat{S}_{\hat{i}\hat{j}} Z^{\hat{j}}_j\right) v^j \\
&= W^i_{\hat{i}}\Big(M_{ij}\dot{v}^j + \big(\tfrac{1}{2}\partial_k M_{ij}\xi^k + \underbrace{Z^{\hat{l}}_i \hat{S}_{\hat{l}\hat{j}} Z^{\hat{j}}_j + \tfrac{1}{2}\left(M_{lj} Z^{\hat{l}}_i \partial_k W^l_{\hat{l}} - M_{il} Z^{\hat{l}}_j \partial_k W^l_{\hat{l}}\right)\xi^k}_{S_{ij}}\big) v^j\Big)
\end{aligned} \tag{A.106}$$

or equivalently

$$\hat{S}_{\hat{i}\hat{j}} = W^i_{\hat{i}} W^j_{\hat{j}} S_{ij} + \tfrac{1}{2} M_{ij}\left(W^i_{\hat{i}}\partial_k W^j_{\hat{j}} - W^j_{\hat{j}}\partial_k W^i_{\hat{i}}\right)\xi^k. \tag{A.107}$$

Recalling the transformation rule (A.93) of the connection coefficients Γ^i_{jk} (not necessarily the Levi-Civita connection), this may be fulfilled by setting

$$S_{ij} = \left(M_{is}\Gamma^s_{jk} - M_{js}\Gamma^s_{ik}\right)\xi^k + C_{ij} \tag{A.108}$$

where $C_{ij} = -C_{ji}$ may be any skew symmetric *tensor*.

Identification of the kinetic energy $\mathcal{T} = \frac{1}{2}\langle \boldsymbol{v}, \boldsymbol{v}\rangle$ with the Riemannian metric and using the defining property of the Levi-Civita connection (A.55a), yields

$$\dot{\mathcal{T}} = \tfrac{1}{2}\boldsymbol{\xi}\langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{\nabla}_{\boldsymbol{\xi}}\boldsymbol{v}\rangle = v^i M_{ij}\left(\partial_k v^j + \Gamma^j_{lk} v^l\right)\xi^k = v^i \underbrace{\left(M_{ij}\dot{v}^j + \Gamma_{ijk}\xi^k v^j\right)}_{p_i} \tag{A.109}$$

This is a special case of what was derived above where Γ^s_{jk} are the coefficients of the Levi-Civita connection and $C_{ij} = 0$.

A.4 Static optimization

A.4.1 Nonlinear programming

We are interested in the extrema of a function $\mathcal{V} : \mathbb{X} \to \mathbb{R}$ with $\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^\nu \mid \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0}\}$ and $\boldsymbol{\phi} : \mathbb{R}^\nu \to \mathbb{R}^c$ is sufficiently smooth. This can be regarded as a nonlinear programming problem with equality constraints

$$\min_{\boldsymbol{x}\in\mathbb{X}} \mathcal{V}(\boldsymbol{x}) \qquad \Leftrightarrow \qquad \begin{array}{ll} \min\limits_{\boldsymbol{x}\in\mathbb{R}^\nu} & \mathcal{V}(\boldsymbol{x}) \\ \text{s. t.} & \boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{0} \end{array} \tag{A.110}$$

Solution using multipliers. A solution can be found in any textbook on nonlinear optimization e.g. [Luenberger and Ye, 2015, chap. 11] or [Güler, 2010, Theorem 9.4 & 9.21]: Using *Lagrange multipliers* $\boldsymbol{\lambda} \in \mathbb{R}^c$, define the axillary function

$$\tilde{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{\lambda}) = \mathcal{V}(\boldsymbol{x}) + \lambda_\kappa \phi^\kappa(\boldsymbol{x}). \tag{A.111}$$

A stationary point $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}})$ of $\tilde{\mathcal{V}}$ is defined by (necessary condition)

$$0 = \frac{\partial \tilde{\mathcal{V}}}{\partial x^\alpha}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}}) = \frac{\partial \mathcal{V}}{\partial x^\alpha}(\bar{\boldsymbol{x}}) + \bar{\lambda}_\kappa \frac{\partial \phi^\kappa}{\partial x^\alpha}(\bar{\boldsymbol{x}}), \qquad \alpha = 1, \ldots, \nu, \tag{A.112a}$$

$$0 = \frac{\partial \tilde{\mathcal{V}}}{\partial \bar{\lambda}_\kappa}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}}) = \phi^\kappa(\bar{\boldsymbol{x}}), \qquad \kappa = 1, \ldots, p. \tag{A.112b}$$

It is a local minimum if for all $\boldsymbol{z} \in \mathbb{R}^\nu$ that satisfy the orthogonality condition

$$\frac{\partial \phi^\kappa}{\partial x^\alpha}(\bar{\boldsymbol{x}}) z^\alpha = 0, \qquad \kappa = 1, \ldots, p \tag{A.113}$$

the following condition holds (sufficient condition)

$$0 < z^\alpha \frac{\partial^2 \tilde{\mathcal{V}}}{\partial x^\alpha \partial x^\beta}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\lambda}}) z^\beta = z^\alpha \left(\frac{\partial^2 \mathcal{V}}{\partial x^\alpha \partial x^\beta}(\bar{\boldsymbol{x}}) + \bar{\lambda}_\kappa \frac{\partial^2 \phi^\kappa}{\partial x^\alpha \partial x^\beta}(\bar{\boldsymbol{x}}) \right) z^\beta. \tag{A.114}$$

If the sign is reversed the stationary point is a local maximum.

Basis for the tangent space. Assume that $n = \dim \mathbb{X} = \nu - \operatorname{rank} \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}(\boldsymbol{x})$ is constant for all $\boldsymbol{x} \in \mathbb{X}$ and we can choose a basis $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{\nu\times n}$, $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}} \boldsymbol{A} \equiv \boldsymbol{0}$, $\operatorname{rank} \boldsymbol{A} = n$. With this we can eliminate the multipliers $\bar{\boldsymbol{\lambda}}$ from the necessary condition (A.112) and restate it as

$$0 = A_i^\alpha(\bar{\boldsymbol{x}}) \frac{\partial \mathcal{V}}{\partial x^\alpha}(\bar{\boldsymbol{x}}) = (\partial_i \mathcal{V})(\bar{\boldsymbol{x}}), \qquad i = 1, \ldots, n \tag{A.115a}$$

$$0 = \phi^\kappa(\bar{\boldsymbol{x}}), \qquad \kappa = 1, \ldots, p. \tag{A.115b}$$

Furthermore, all solutions to the orthogonality condition (A.113) can be stated as $\boldsymbol{z} = \boldsymbol{A}(\bar{\boldsymbol{x}})\boldsymbol{\zeta}, \boldsymbol{\zeta} \in \mathbb{R}^n$. Plugging this into (A.114) the sufficient condition is $\boldsymbol{\zeta}^\top \boldsymbol{K} \boldsymbol{\zeta} > 0 \, \forall \boldsymbol{\zeta} \in \mathbb{R}^n$ where

$$\begin{aligned} K_{ij}(\bar{\boldsymbol{x}}) &= A_i^\alpha(\bar{\boldsymbol{x}}) \left(\frac{\partial^2 \mathcal{V}}{\partial x^\alpha \partial x^\beta}(\bar{\boldsymbol{x}}) + \bar{\lambda}_\kappa \frac{\partial^2 \phi^\kappa}{\partial x^\alpha \partial x^\beta}(\bar{\boldsymbol{x}}) \right) A_j^\beta(\bar{\boldsymbol{x}}) \\ &= A_i^\alpha(\bar{\boldsymbol{x}}) \frac{\partial}{\partial x^\alpha} \left(A_j^\beta \frac{\partial \mathcal{V}}{\partial x^\beta} \right)(\bar{\boldsymbol{x}}) + \bar{\lambda}_\kappa A_i^\alpha(\bar{\boldsymbol{x}}) \frac{\partial}{\partial x^\alpha} \Bigg(\underbrace{A_j^\beta \frac{\partial \phi^\kappa}{\partial x^\beta}}_{0} \Bigg)(\bar{\boldsymbol{x}}) \\ &\qquad\qquad\qquad - A_i^\alpha(\bar{\boldsymbol{x}}) \frac{\partial A_j^\beta}{\partial x^\alpha}(\bar{\boldsymbol{x}}) \Bigg(\underbrace{\frac{\partial \mathcal{V}}{\partial x^\beta}(\bar{\boldsymbol{x}}) + \bar{\lambda}_\kappa \frac{\partial \phi^\kappa}{\partial x^\beta}(\bar{\boldsymbol{x}})}_{0} \Bigg) \\ &= \big(\partial_i \partial_j \mathcal{V}\big)(\bar{\boldsymbol{x}}). \end{aligned} \tag{A.116}$$

So the sufficient condition for the critical point $\bar{\boldsymbol{x}}$ to be a minimum (maximum) is equivalent to the matrix $\boldsymbol{K} = \boldsymbol{K}^\top$ to be positive (negative) definite.

A.4.2 Quadratic programming

A special case of (A.110) is the quadratic programming problem with linear equality constraints:

$$\begin{aligned} \min_{\boldsymbol{v}\in\mathbb{R}^\nu} \quad & \tfrac{1}{2}\boldsymbol{v}^\top \boldsymbol{M}\boldsymbol{v} + \boldsymbol{c}^\top \boldsymbol{v} \\ \text{s. t.} \quad & \boldsymbol{Z}\boldsymbol{v} = \boldsymbol{b} \end{aligned} \tag{A.117}$$

where $\boldsymbol{M} \in \mathbb{SYM}^+(\nu)$, $\boldsymbol{Z} \in \mathbb{R}^{c\times\nu}$ and $\boldsymbol{v}, \boldsymbol{c} \in \mathbb{R}^\nu$, $\boldsymbol{b} \in \mathbb{R}^c$. The necessary condition from (A.112) for $\bar{\boldsymbol{v}}$ to be a critical point is

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{Z}^\top \\ \boldsymbol{Z} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{v}} \\ \bar{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{c} \\ \boldsymbol{b} \end{bmatrix}. \tag{A.118}$$

The sufficient condition for $\bar{\boldsymbol{v}}$ to be a minimum is

$$\forall\, \boldsymbol{z} \in \mathbb{R}^\nu,\ \boldsymbol{Z}\boldsymbol{z} = \boldsymbol{0} \ :\ \boldsymbol{z}^\top \boldsymbol{M}\boldsymbol{z} > 0. \tag{A.119}$$

Since $\boldsymbol{M}$ is positive definite this is always fulfilled.

Explicit solution. Taking a weighted sum of the two rows yields

$$\underbrace{\boldsymbol{Z}\boldsymbol{M}^{-1}\boldsymbol{Z}^\top}_{\boldsymbol{S}} \bar{\boldsymbol{\lambda}} = -\underbrace{\boldsymbol{Z}\boldsymbol{M}^{-1}\boldsymbol{c} + \boldsymbol{b}}_{\boldsymbol{d}} \tag{A.120}$$

with the Schur complement $\boldsymbol{S}$ and $\operatorname{rank} \boldsymbol{S} = \operatorname{rank} \boldsymbol{Z}$. For the following we use the (Moore-Penrose) pseudoinverse [Penrose, 1955]. The identity of the projectors $\boldsymbol{S}^+\boldsymbol{S} = \boldsymbol{S}\boldsymbol{S}^+ = \boldsymbol{Z}\boldsymbol{Z}^+$ can be derived using the basic properties of the pseudoinverse and projection matrices. A solution of (A.120) exists if and only if $\boldsymbol{S}\boldsymbol{S}^+\boldsymbol{d} = \boldsymbol{d}$. This is equivalent to

$$\boldsymbol{0} = \boldsymbol{S}\boldsymbol{S}^+\boldsymbol{d} - \boldsymbol{d} = \boldsymbol{Z}\boldsymbol{Z}^+(\boldsymbol{Z}\boldsymbol{M}^{-1}\boldsymbol{c} + \boldsymbol{b}) - (\boldsymbol{Z}\boldsymbol{M}^{-1}\boldsymbol{c} + \boldsymbol{b}) = \boldsymbol{Z}\boldsymbol{Z}^+\boldsymbol{b} - \boldsymbol{b} \tag{A.121}$$

which is the condition for the constraint equation $\boldsymbol{Z}\boldsymbol{v} = \boldsymbol{b}$ to have a solution. The explicit solution(s) for the multipliers is

$$\bar{\boldsymbol{\lambda}} = -\boldsymbol{S}^+(\boldsymbol{Z}\boldsymbol{M}^{-1}\boldsymbol{c} + \boldsymbol{b}) + (\mathbf{I} - \boldsymbol{S}^+\boldsymbol{S})\boldsymbol{\mu}, \quad \boldsymbol{\mu} \in \mathbb{R}^c. \tag{A.122}$$

Plugging this into the original equation we get

$$\bar{\boldsymbol{v}} = \boldsymbol{M}^{-1}\boldsymbol{Z}^\top\boldsymbol{S}^+\boldsymbol{b} - (\boldsymbol{M}^{-1} - \boldsymbol{M}^{-1}\boldsymbol{Z}^\top\boldsymbol{S}^+\boldsymbol{Z}\boldsymbol{M}^{-1})\boldsymbol{c}. \tag{A.123}$$

It is crucial to note that the solution $\bar{\boldsymbol{v}}$ is unique even though there might be several solutions for the multipliers $\bar{\boldsymbol{\lambda}}$. This is because the terms with $\boldsymbol{\mu}$ cancel out, since

$$\boldsymbol{Z}^\top(\mathbf{I} - \boldsymbol{S}^+\boldsymbol{S}) = \boldsymbol{Z}^\top(\mathbf{I} - \boldsymbol{Z}\boldsymbol{Z}^+) = (\boldsymbol{Z} - \boldsymbol{Z}\boldsymbol{Z}^+\boldsymbol{Z})^\top = \boldsymbol{0}. \tag{A.124}$$

Udwadia-Kalaba equation. Since $\boldsymbol{M}$ is symmetric and positive definite, its eigendecomposition can be written as $\boldsymbol{M} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^\top$ with $\boldsymbol{W} \in \mathbb{SO}(\nu)$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_\nu)$. Define $\boldsymbol{M}^{1/2} = \boldsymbol{W}\boldsymbol{\Lambda}^{1/2}\boldsymbol{W}^\top$ with $\boldsymbol{\Lambda}^{1/2} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_\nu})$ and analogous for $\boldsymbol{M}^{-1/2}$. Let $\bar{\boldsymbol{Z}} = \boldsymbol{Z}\boldsymbol{M}^{-1/2}$ to find $\boldsymbol{S} = \bar{\boldsymbol{Z}}\bar{\boldsymbol{Z}}^\top$ and $\boldsymbol{S}^+ = (\bar{\boldsymbol{Z}}^+)^\top\bar{\boldsymbol{Z}}^+$. Substituting this into (A.123) yields

$$\bar{\boldsymbol{v}} = \boldsymbol{M}^{-1/2}\bar{\boldsymbol{Z}}^+\boldsymbol{b} - (\boldsymbol{M}^{-1} - \boldsymbol{M}^{-1/2}\bar{\boldsymbol{Z}}^+\bar{\boldsymbol{Z}}\boldsymbol{M}^{-1/2})\boldsymbol{c}. \tag{A.125}$$

This form was proposed in [Udwadia and Kalaba, 2002]

Choosing a basis. By choosing a matrix $\boldsymbol{A} \in \mathbb{R}^{\nu\times n}$ with $\boldsymbol{Z}\boldsymbol{A} \equiv \boldsymbol{0}$, rank $\boldsymbol{A} = n$ we can state all solutions to the constraint as

$$\boldsymbol{v} = \boldsymbol{Z}^+\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\xi},\ \ \boldsymbol{\xi} \in \mathbb{R}^n. \tag{A.126}$$

The remaining *unconstrained* problem is

$$\min_{\boldsymbol{\xi}\in\mathbb{R}^n} \ \tfrac{1}{2}\boldsymbol{\xi}^\top\boldsymbol{A}^\top\boldsymbol{M}\boldsymbol{A}\boldsymbol{\xi} + (\boldsymbol{c} + \boldsymbol{M}\boldsymbol{Z}^+\boldsymbol{b})^\top\boldsymbol{A}\boldsymbol{\xi} \tag{A.127}$$

with the solution

$$\bar{\boldsymbol{\xi}} = -(\boldsymbol{A}^\top\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^\top(\boldsymbol{c} + \boldsymbol{M}\boldsymbol{Z}^+\boldsymbol{b}) \tag{A.128}$$

So the minimum of the original problem is

$$\bar{\boldsymbol{v}} = \big(\mathbf{I}_\nu - \boldsymbol{A}(\boldsymbol{A}^\top\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^\top\boldsymbol{M}\big)\boldsymbol{Z}^+\boldsymbol{b} - \boldsymbol{A}(\boldsymbol{A}^\top\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^\top\boldsymbol{c} \tag{A.129}$$

A.4.3 Blah

Consider the quadratic optimization problem with linear inequality constraints

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & \tfrac{1}{2}x^\top G x + x^\top g \\ \text{subject to} \quad & A\,x \leq b \end{aligned} \tag{A.130}$$

Problem normalization. Since the matrix G is symmetric positive definite, it can be decomposed into $G = L^\top L$. We transform the coordinates and the matrices by

$$\bar{x} = L\,x, \qquad \bar{g} = L^{-\top}g, \qquad \bar{A} = A\,L^{-1} \tag{A.131}$$

The optimization problem in the transformed coordinates is

$$\begin{aligned} \underset{\bar{x}\in\mathbb{R}^n}{\text{minimize}} \quad & \tfrac{1}{2}\bar{x}^\top\bar{x} + \bar{x}^\top\bar{g} \\ \text{subject to} \quad & \bar{A}\,\bar{x} \leq b \end{aligned} \tag{A.132}$$

The decomposition is done off line. This normalization also leads to better numerical properties in the online implementation as analyzed in [Gould et al., 2001].

Subproblem solution. The Active-Set method requires the solution of equality constrained subproblems, where $\bar{x} = \bar{x}_k + s$ and $\bar{A}_k$ is a subset of the rows of $\bar{A}$:

$$\begin{aligned} \underset{s\in\mathbb{R}^n}{\text{minimize}} \quad & \tfrac{1}{2}s^\top s + s^\top(\bar{x}_k + \bar{g}) \\ \text{subject to} \quad & \bar{A}_k s = 0 \end{aligned} \tag{A.133}$$

The solution using the concept of the Lagrange-multipliers is

$$\begin{bmatrix} \mathbf{I}_n & \bar{A}_k^\top \\ \bar{A}_k & 0 \end{bmatrix} \begin{bmatrix} s \\ \mu \end{bmatrix} + \begin{bmatrix} \bar{x}_k + \bar{g} \\ 0 \end{bmatrix} = 0 \tag{A.134}$$

For a matrix of this form we have

$$\begin{bmatrix} \mathbf{I} & Z^\top \\ Z & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I} - Z^\top SZ & Z^\top S \\ SZ & -S \end{bmatrix}, \qquad S = (ZZ^\top)^{-1} = S^\top \geq 0 \tag{A.135}$$

So the subproblem solution (s, μ) can be computed economically from

$$h = -\bar{x}_k - \bar{g}, \qquad H = \bar{A}_k \bar{A}_k^\top, \qquad H\mu = \bar{A}_k h, \qquad s = h - \bar{A}_k^\top \mu. \tag{A.136}$$

A.4.4 Attitude potential

We are interested in the extrema of the function

$$\mathcal{V}(\boldsymbol{R}) = -\operatorname{tr}(\boldsymbol{P}\boldsymbol{R}), \quad \boldsymbol{R} \in \mathbb{SO}(3) \tag{A.137}$$

with the constant parameter $\boldsymbol{P} \in \mathbb{R}^{3\times 3}$. Similar functions appear in the context of attitude control [Koditschek, 1989] or in the so-called Wahba's problem [Wahba, 1965].

Coordinate transformation. Consider a singular value decomposition $\boldsymbol{P} = \boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{Y}^\top$ with $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{O}(3)$ and $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \sigma_2, \sigma_3)$, $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$. Define

$$\bar{\boldsymbol{X}} = \boldsymbol{X} \operatorname{diag}(1, 1, \det \boldsymbol{X}) \in \mathbb{SO}(3), \tag{A.138a}$$
$$\bar{\boldsymbol{Y}} = \boldsymbol{Y} \operatorname{diag}(1, 1, \det \boldsymbol{Y}) \in \mathbb{SO}(3), \tag{A.138b}$$
$$\bar{\boldsymbol{\Sigma}} = \operatorname{diag}(\sigma_1, \sigma_2, \bar{\sigma}_3), \ \ \bar{\sigma}_3 = \det \boldsymbol{X} \det \boldsymbol{Y} \sigma_3 \tag{A.138c}$$

which yields a decomposition $\boldsymbol{P} = \bar{\boldsymbol{X}}\bar{\boldsymbol{\Sigma}}\bar{\boldsymbol{Y}}^\top$ with proper rotations. Using the cyclic permutation property of the trace we get

$$\mathcal{V}(\boldsymbol{R}) = -\operatorname{tr}(\bar{\boldsymbol{X}}\bar{\boldsymbol{\Sigma}}\bar{\boldsymbol{Y}}^\top\boldsymbol{R}) = -\operatorname{tr}(\bar{\boldsymbol{\Sigma}}\underbrace{\bar{\boldsymbol{Y}}^\top\boldsymbol{R}\bar{\boldsymbol{X}}}_{\bar{\boldsymbol{R}}}) =: \bar{\mathcal{V}}(\bar{\boldsymbol{R}}) \tag{A.139}$$

Since the SVD is not unique in general, the transformed function $\bar{\mathcal{V}}$ is neither. However, since the coordinate transformation $\boldsymbol{R} = \bar{\boldsymbol{Y}}\bar{\boldsymbol{R}}\bar{\boldsymbol{X}}^\top$ is bijective, no information is lost here. The SVD is also the most robust approach for the numerical solution of the problem [Markley, 1988].

Critical points. The first and second differential of the transformed function are

$$\nabla\bar{\mathcal{V}}(\bar{\boldsymbol{R}}) = \mathrm{vee2}(\bar{\boldsymbol{\Sigma}}\bar{\boldsymbol{R}}), \tag{A.140}$$

$$\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}) = \mathrm{Vee}(\bar{\boldsymbol{\Sigma}}\bar{\boldsymbol{R}})^\top. \tag{A.141}$$

So, for a critical point $\bar{\boldsymbol{R}}_0 : \nabla\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_0) = \mathbf{0}$ we need the matrix $\bar{\boldsymbol{\Sigma}}\bar{\boldsymbol{R}}_0$ to be symmetric. For the following it will be useful to substitute the entries/eigenvalues of $\mathrm{Vee}(\bar{\boldsymbol{\Sigma}}) = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) = \boldsymbol{\Lambda}$ as

$$\left.\begin{aligned} \lambda_1 &= \sigma_2 + \bar{\sigma}_3, \\ \lambda_2 &= \bar{\sigma}_3 + \sigma_1, \\ \lambda_3 &= \sigma_1 + \sigma_2 \end{aligned}\right\} \quad \Leftrightarrow \quad \left\{\begin{aligned} \sigma_1 &= \tfrac{1}{2}(\lambda_2 + \lambda_3 - \lambda_1), \\ \sigma_2 &= \tfrac{1}{2}(\lambda_3 + \lambda_1 - \lambda_2), \\ \bar{\sigma}_3 &= \tfrac{1}{2}(\lambda_1 + \lambda_2 - \lambda_3) \end{aligned}\right. \tag{A.142}$$

Note that $\sigma_1 \geq \sigma_2 \geq |\bar{\sigma}_3| \geq 0$ implies $\lambda_3 \geq \lambda_2 \geq \lambda_1 \geq 0$. Depending on the constellation of the eigenvalues we have the following critical points:

- Distinct eigenvalues: $\lambda_3 > \lambda_2 > \lambda_1 > 0$: We have the critical points

$$\begin{aligned}
&\bar{\boldsymbol{R}}_0 = \mathbf{I}_3 \ : \\
&\quad \mathcal{V}(\bar{\boldsymbol{R}}_0) = -\tfrac{\lambda_1}{2} - \tfrac{\lambda_2}{2} - \tfrac{\lambda_3}{2}, \quad \mathrm{eig}(\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_0)) = \{\lambda_3, \lambda_2, \lambda_1\} && \text{(A.143a)} \\
&\bar{\boldsymbol{R}}_1 = \mathrm{diag}(1, -1, -1) \ : \\
&\quad \mathcal{V}(\bar{\boldsymbol{R}}_1) = \tfrac{3\lambda_1 - \lambda_2 - \lambda_3}{2}, \quad \mathrm{eig}(\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_1)) = \{\lambda_3 - \lambda_1, \lambda_2 - \lambda_1, -\lambda_1\} && \text{(A.143b)} \\
&\bar{\boldsymbol{R}}_2 = \mathrm{diag}(-1, 1, -1) \ : \\
&\quad \mathcal{V}(\bar{\boldsymbol{R}}_2) = \tfrac{3\lambda_2 - \lambda_1 - \lambda_3}{2}, \quad \mathrm{eig}(\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_2)) = \{\lambda_3 - \lambda_2, \lambda_1 - \lambda_2, -\lambda_2\} && \text{(A.143c)} \\
&\bar{\boldsymbol{R}}_3 = \mathrm{diag}(-1, -1, 1) \ : \\
&\quad \mathcal{V}(\bar{\boldsymbol{R}}_3) = \tfrac{3\lambda_3 - \lambda_1 - \lambda_2}{2}, \quad \mathrm{eig}(\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_3)) = \{\lambda_2 - \lambda_3, \lambda_1 - \lambda_3, -\lambda_3\} && \text{(A.143d)}
\end{aligned}$$

  so $\bar{\boldsymbol{R}}_0$ is a minimum, $\bar{\boldsymbol{R}}_1$ and $\bar{\boldsymbol{R}}_2$ are saddle points, and $\bar{\boldsymbol{R}}_3$ is a maximum.

- Double eigenvalue: $\lambda_3 > \lambda_2 = \lambda_1 > 0$: We have a minimum at $\bar{\boldsymbol{R}}_0$, a maximum at $\bar{\boldsymbol{R}}_3$ and a saddle on the circular manifold

$$\begin{aligned}
&\bar{\boldsymbol{R}}_4 = \begin{bmatrix} -c & s & 0 \\ s & c & 0 \\ 0 & 0 & -1 \end{bmatrix}, \ c^2 + s^2 = 1 \ : \\
&\quad \mathcal{V}(\bar{\boldsymbol{R}}_4) = \lambda_1 - \tfrac{\lambda_3}{2}, \quad \mathrm{eig}(\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_4)) = \{\lambda_3 - \lambda_1, 0, -\lambda_1\} && \text{(A.143e)}
\end{aligned}$$

  which includes the points $\bar{\boldsymbol{R}}_1$ and $\bar{\boldsymbol{R}}_2$.

- Double eigenvalue: $\lambda_3 = \lambda_2 > \lambda_1 > 0$: Analog to above we have a minimum at $\bar{\boldsymbol{R}}_0$, a saddle at $\bar{\boldsymbol{R}}_4$ and a maximum on the circular manifold

$$\begin{aligned}
&\bar{\boldsymbol{R}}_5 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & c & s \\ 0 & s & -c \end{bmatrix}, \ c^2 + s^2 = 1 \ : \\
&\quad \mathcal{V}(\bar{\boldsymbol{R}}_5) = \lambda_2 - \tfrac{\lambda_1}{2}, \quad \mathrm{eig}(\nabla^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_5)) = \{0, \lambda_1 - \lambda_2, -\lambda_2\} && \text{(A.143f)}
\end{aligned}$$

  which includes the points $\bar{\boldsymbol{R}}_2$ and $\bar{\boldsymbol{R}}_3$.

- Triple eigenvalue: $\lambda_3 = \lambda_2 = \lambda_1 > 0$: Minimum at $\bar{\boldsymbol{R}}_0$ and a maximum on the spherical manifold

$$\bar{\boldsymbol{R}}_6 = \begin{bmatrix} q_\mathsf{x}^2 - q_\mathsf{y}^2 - q_\mathsf{z}^2 & 2q_\mathsf{x}q_\mathsf{y} & 2q_\mathsf{x}q_\mathsf{z} \\ 2q_\mathsf{x}q_\mathsf{y} & q_\mathsf{y}^2 - q_\mathsf{x}^2 + q_\mathsf{z}^2 & 2q_\mathsf{y}q_\mathsf{z} \\ 2q_\mathsf{x}q_\mathsf{z} & 2q_\mathsf{y}q_\mathsf{z} & q_\mathsf{z}^2 - q_\mathsf{x}^2 - q_\mathsf{y}^2 \end{bmatrix}, \ q_\mathsf{x}^2 + q_\mathsf{y}^2 + q_\mathsf{z}^2 = 1 :$$
$$\mathcal{V}(\bar{\boldsymbol{R}}_6) = \tfrac{\lambda_1}{2}, \quad \mathrm{eig}(\boldsymbol{\nabla}^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_6)) = \{0, 0, -\lambda_1\} \tag{A.143g}$$

which includes the points $\bar{\boldsymbol{R}}_1$, $\bar{\boldsymbol{R}}_2$ and $\bar{\boldsymbol{R}}_3$ and the circles $\bar{\boldsymbol{R}}_4$ and $\bar{\boldsymbol{R}}_5$. It corresponds to a 180° rotation about an arbitrary axis $[q_\mathsf{x}, q_\mathsf{y}, q_\mathsf{z}]^\top \in \mathbb{S}^2$.

- One zero eigenvalue: $\lambda_3 > \lambda_2 > \lambda_1 = 0$: We have a minimum on the circular manifold

$$\bar{\boldsymbol{R}}_7 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}, \ c^2 + s^2 = 1 \ :$$
$$\mathcal{V}(\bar{\boldsymbol{R}}_7) = -\tfrac{\lambda_2+\lambda_3}{2}, \qquad \mathrm{eig}(\boldsymbol{\nabla}^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_7)) = \{\lambda_3, \lambda_2, 0\} \tag{A.143h}$$

which includes $\bar{\boldsymbol{R}}_0$ and $\bar{\boldsymbol{R}}_1$. Furthermore we have a saddle point at $\bar{\boldsymbol{R}}_2$ and a maximum at $\bar{\boldsymbol{R}}_3$.

- Double eigenvalue and zero eigenvalue: $\lambda_3 = \lambda_2 > \lambda_1 = 0$: We have a minimum on $\bar{\boldsymbol{R}}_7$ and a maximum on $\bar{\boldsymbol{R}}_5$.

- Two zero eigenvalues: $\lambda_3 > \lambda_2 = \lambda_1 = 0$: We have a minimum on the spherical manifold

$$\bar{\boldsymbol{R}}_8 = \begin{bmatrix} q_\mathsf{w}^2 + q_\mathsf{x}^2 - q_\mathsf{y}^2 & 2q_\mathsf{x}q_\mathsf{y} & 2q_\mathsf{w}q_\mathsf{y} \\ 2q_\mathsf{x}q_\mathsf{y} & q_\mathsf{w}^2 - q_\mathsf{x}^2 + q_\mathsf{y}^2 & -2q_\mathsf{w}q_\mathsf{x} \\ -2q_\mathsf{w}q_\mathsf{x} & 2q_\mathsf{w}q_\mathsf{x} & q_\mathsf{w}^2 - q_\mathsf{x}^2 - q_\mathsf{y}^2 \end{bmatrix}, \ q_\mathsf{w}^2 + q_\mathsf{x}^2 + q_\mathsf{y}^2 = 1 :$$
$$\mathcal{V}(\bar{\boldsymbol{R}}_8) = -\tfrac{\lambda_3}{2}, \qquad \mathrm{eig}(\boldsymbol{\nabla}^2\bar{\mathcal{V}}(\bar{\boldsymbol{R}}_8)) = \{\lambda_3, 0, 0\} \tag{A.143i}$$

which includes $\bar{\boldsymbol{R}}_0$, $\bar{\boldsymbol{R}}_1$ and $\bar{\boldsymbol{R}}_2$ and corresponds to an arbitrary rotation about an axis $[q_\mathsf{x}, q_\mathsf{y}, 0]^\top$. Furthermore we have a maximum at $\bar{\boldsymbol{R}}_3$.

- All zero Eigenvalues: $\lambda_3 = \lambda_2 = \lambda_1 = 0$: for this we have $\bar{\boldsymbol{\Sigma}} = \boldsymbol{P} = \boldsymbol{0}$ and the function is $\mathcal{V} = 0$.

We may conclude that the function $\bar{\mathcal{V}}$ has a minimum at $\bar{\boldsymbol{R}}_0 = \mathbf{I}_3$ and a maximum at $\bar{\boldsymbol{R}}_3 = \mathrm{diag}(-1, -1, 1)$, though they may not be strict. The minimum is strict if and only if $\lambda_1 > 0$. The maximum is strict if and only if $\lambda_3 > \lambda_2$.

It should also be noted that the results of this paragraph would be much more "symmetric" if we would not have required the descending order of the singular values σ_i. This did however reduce the number of cases to distinguish.

Original coordinates. The original function $\mathcal{V}$ has a minimum at $\boldsymbol{R}_0 = \bar{\boldsymbol{Y}}\bar{\boldsymbol{X}}^\top$. The minimum $\boldsymbol{R}_0$ is strict, if, and only if, $\lambda_i > 0, i = 1,2,3$ or equivalently if $\boldsymbol{K}$ is positive definite:

$$\boldsymbol{K} = \boldsymbol{\nabla}^2\mathcal{V}(\boldsymbol{R}_0) = \mathrm{Vee}(\boldsymbol{P}\boldsymbol{R}_0) = \bar{\boldsymbol{X}}\boldsymbol{\Lambda}\bar{\boldsymbol{X}}^\top. \tag{A.144}$$

Substracting the minimal value $\mathcal{V}(\boldsymbol{R}_0)$ from the function we obtain

$$\begin{aligned}
\hat{\mathcal{V}}(\boldsymbol{R}) &= \mathcal{V}(\boldsymbol{R}) - \mathcal{V}(\boldsymbol{R}_0) \\
&= -\operatorname{tr}(\bar{\boldsymbol{X}}\,\mathrm{Wed}(\boldsymbol{\Lambda})\bar{\boldsymbol{Y}}^\top\boldsymbol{R}) + \tfrac{1}{2}\operatorname{tr}(\boldsymbol{\Lambda}) \\
&= -\operatorname{tr}(\bar{\boldsymbol{X}}\,\mathrm{Wed}(\boldsymbol{\Lambda})\bar{\boldsymbol{X}}^\top\boldsymbol{R}_0^\top\boldsymbol{R}) + \operatorname{tr}(\mathrm{Wed}(\boldsymbol{K})) \\
&= \operatorname{tr}\big(\mathrm{Wed}(\boldsymbol{K})(\mathbf{I}_3 - \boldsymbol{R}_0^\top\boldsymbol{R})\big) && \text{(A.145a)} \\
&= \tfrac{1}{2}\operatorname{tr}\big(\mathrm{Wed}(\boldsymbol{K})(\boldsymbol{R}-\boldsymbol{R}_0)^\top(\boldsymbol{R}-\boldsymbol{R}_0)\big) \\
&= \tfrac{1}{2}\|\boldsymbol{R}-\boldsymbol{R}_0\|^2_{\mathrm{Wed}(\boldsymbol{K})}. && \text{(A.145b)}
\end{aligned}$$

We have the properties

$$\begin{aligned}
&\boldsymbol{K} \geq 0 && \Leftrightarrow && \hat{\mathcal{V}}(\boldsymbol{R}) \geq 0 && \text{(A.146a)} \\
&\boldsymbol{K} > 0 && \Leftrightarrow && \hat{\mathcal{V}}(\boldsymbol{R}) \geq 0 \ \wedge\ \hat{\mathcal{V}}(\boldsymbol{R}) = 0 \Leftrightarrow \boldsymbol{R} = \boldsymbol{R}_0. && \text{(A.146b)}
\end{aligned}$$

The form (A.145a) is called the *navigation function* for $\mathbb{SO}(3)$ in [Koditschek, 1989]. From its properties $\hat{\mathcal{V}}$ is an $\mathbb{SO}(3)$ analogon to $\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)^\top\boldsymbol{K}(\boldsymbol{x}-\boldsymbol{x}_0), \boldsymbol{x}, \boldsymbol{x}_0 \in \mathbb{R}^3$.

A.5 Attitude parameterizations

A rotation matrix $\boldsymbol{R} \in \mathbb{SO}(3)$ and its angular velocity $\boldsymbol{\omega} = \mathrm{vee}(\boldsymbol{R}^\top\dot{\boldsymbol{R}})$ is one possible representation of an attitude. The following presents some other popular representations.

A.5.1 Axis-angle representation

The (unit) axis $\boldsymbol{a} = [a_\mathsf{x}, a_\mathsf{y}, a_\mathsf{z}]^\top \in \mathbb{S}^2$, angle $\theta \in [0,\pi]$ representation are related to a rotation matrix by Rodrigues' rotation formula (e.g. [Murray et al., 1994, sec. 2.2])

$$\begin{aligned}
\boldsymbol{R} &= \exp(\theta\,\mathrm{wed}(\boldsymbol{a})) = \mathbf{I}_3 + \mathrm{s}_\theta\,\mathrm{wed}(\boldsymbol{a}) + (1-\mathrm{c}_\theta)\ \mathrm{wed}(\boldsymbol{a})^2 \\
&= \begin{bmatrix} (1-\mathrm{c}_\theta)a_\mathsf{x}^2 + \mathrm{c}_\theta & -\mathrm{s}_\theta a_\mathsf{z} + (1-\mathrm{c}_\theta)a_\mathsf{y}a_\mathsf{x} & \mathrm{s}_\theta a_\mathsf{y} + (1-\mathrm{c}_\theta)a_\mathsf{z}a_\mathsf{x} \\ \mathrm{s}_\theta a_\mathsf{z} + (1-\mathrm{c}_\theta)a_\mathsf{y}a_\mathsf{x} & (1-\mathrm{c}_\theta)a_\mathsf{y}^2 + \mathrm{c}_\theta & -\mathrm{s}_\theta a_\mathsf{x} + (1-\mathrm{c}_\theta)a_\mathsf{z}a_\mathsf{y} \\ -\mathrm{s}_\theta a_\mathsf{y} + (1-\mathrm{c}_\theta)a_\mathsf{z}a_\mathsf{x} & \mathrm{s}_\theta a_\mathsf{x} + (1-\mathrm{c}_\theta)a_\mathsf{z}a_\mathsf{y} & (1-\mathrm{c}_\theta)a_\mathsf{z}^2 + \mathrm{c}_\theta \end{bmatrix}
\end{aligned} \tag{A.147}$$

$$\begin{gathered}
\omega = \boldsymbol{a}\dot{\theta} + (\mathrm{s}_\theta\mathbf{I}_3 - (1-\mathrm{c}_\theta)\,\mathrm{wed}(\boldsymbol{a}))\dot{\boldsymbol{a}} \\
\begin{bmatrix}\omega_\mathsf{x}\\ \omega_\mathsf{y}\\ \omega_\mathsf{z}\\ 0\end{bmatrix} = \underbrace{\begin{bmatrix} \mathrm{s}_\theta & (1-\mathrm{c}_\theta)a_\mathsf{z} & -(1-\mathrm{c}_\theta)a_\mathsf{y} & a_\mathsf{x} \\ -(1-\mathrm{c}_\theta)a_\mathsf{z} & \mathrm{s}_\theta & (1-\mathrm{c}_\theta)a_\mathsf{x} & a_\mathsf{y} \\ (1-\mathrm{c}_\theta)a_\mathsf{y} & -(1-\mathrm{c}_\theta)a_\mathsf{x} & \mathrm{s}_\theta & a_\mathsf{z} \\ a_\mathsf{x} & a_\mathsf{y} & a_\mathsf{z} & 0 \end{bmatrix}}_{\boldsymbol{Y}_\square} \begin{bmatrix}\dot{a}_\mathsf{x}\\ \dot{a}_\mathsf{y}\\ \dot{a}_\mathsf{z}\\ \dot{\theta}\end{bmatrix}
\end{gathered} \tag{A.148}$$

Note the relations

$$\mathrm{tr}(\boldsymbol{R}) = 2\mathrm{c}_\theta + 1, \quad \mathrm{vee2}(\boldsymbol{R}) = 2\mathrm{s}_\theta \boldsymbol{a}. \tag{A.149}$$

With this we may formulate the inverse relation as

$$\theta = \mathrm{atan2}(\|\mathrm{vee2}(\boldsymbol{R})\|, \mathrm{tr}(\boldsymbol{R}) - 1), \quad \boldsymbol{a} = \frac{\mathrm{vee2}(\boldsymbol{R})}{\|\mathrm{vee2}(\boldsymbol{R})\|}, \tag{A.150}$$

$$\begin{bmatrix} \dot{a}_\mathsf{x} \\ \dot{a}_\mathsf{y} \\ \dot{a}_\mathsf{z} \\ \dot{\theta} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\mathrm{s}_\theta(1-a_\mathsf{x}^2)}{2(1-\mathrm{c}_\theta)} & -\frac{(\mathrm{c}_\theta+1)a_\mathsf{y}a_\mathsf{x}+\mathrm{s}_\theta a_\mathsf{z}}{2\mathrm{s}_\theta} & -\frac{(\mathrm{c}_\theta+1)a_\mathsf{z}a_\mathsf{x}-\mathrm{s}_\theta a_\mathsf{y}}{2\mathrm{s}_\theta} & a_\mathsf{x} \\ -\frac{(\mathrm{c}_\theta+1)a_\mathsf{y}a_\mathsf{x}-\mathrm{s}_\theta a_\mathsf{z}}{2\mathrm{s}_\theta} & \frac{\mathrm{s}_\theta(1-a_\mathsf{y}^2)}{2(1-\mathrm{c}_\theta)} & -\frac{\mathrm{s}_\theta a_\mathsf{x}+(\mathrm{c}_\theta+1)a_\mathsf{z}a_\mathsf{y}}{2\mathrm{s}_\theta} & a_\mathsf{y} \\ -\frac{(\mathrm{c}_\theta+1)a_\mathsf{z}a_\mathsf{x}+\mathrm{s}_\theta a_\mathsf{y}}{2\mathrm{s}_\theta} & -\frac{-\mathrm{s}_\theta a_\mathsf{x}+(\mathrm{c}_\theta+1)a_\mathsf{z}a_\mathsf{y}}{2\mathrm{s}_\theta} & \frac{\mathrm{s}_\theta(1-a_\mathsf{z}^2)}{2(1-\mathrm{c}_\theta)} & a_\mathsf{z} \\ a_\mathsf{x} & a_\mathsf{y} & a_\mathsf{z} & -\mathrm{s}_\theta \end{bmatrix}}_{\boldsymbol{Y}_\square^{-1}} \begin{bmatrix} \omega_\mathsf{x} \\ \omega_\mathsf{y} \\ \omega_\mathsf{z} \\ 0 \end{bmatrix} \tag{A.151}$$

Note that the axis a is undefined for $\mathrm{s}_\theta = 0$, i.e. $\theta = k\pi$, $k \in \mathbb{Z}$.

Potential:

$$\mathcal{V} = \mathrm{tr}(\boldsymbol{K}'(\mathbf{I}_3 - \boldsymbol{R})) = (1 - \mathrm{c}_\theta)\boldsymbol{a}^\top \boldsymbol{K}\boldsymbol{a}. \tag{A.152}$$

So for the case $\boldsymbol{K} = \mathbf{I}_3$, the metric is related to the angle by $d^2(\boldsymbol{R}, \mathbf{I}_3) = 1 - \cos(\theta)$.

A.5.2 Unit quaternion

Instead of working with the imaginary units i, j, k for the quaternions we can consider them as a quadruple $\boldsymbol{q} = [q_\mathsf{w}, q_\mathsf{x}, q_\mathsf{y}, q_\mathsf{z}]^\top \in \mathbb{S}^3 = \{\boldsymbol{q} \in \mathbb{R}^4 \mid \|\boldsymbol{q}\|^2 = 1\}$. It is common to call q_w the scalar and $\boldsymbol{q}_\mathsf{xyz} = [q_\mathsf{x}, q_\mathsf{y}, q_\mathsf{z}]^\top$ the vector component of the quaternion. The relation to axis-angle representation is (e.g. [Murray et al., 1994, sec. 2.3])

$$q_\mathsf{w} = \cos\tfrac{\theta}{2}, \quad \boldsymbol{q}_\mathsf{xyz} = \boldsymbol{a}\sin\tfrac{\theta}{2} \tag{A.153}$$

Combining this with Rodrigues' rotation formula (A.147) and the half-angle formulas, we find

$$\begin{aligned} \boldsymbol{R} &= ((q_\mathsf{w})^2 - \|\boldsymbol{q}_\mathsf{xyz}\|^2)\mathbf{I}_3 + 2\boldsymbol{q}_\mathsf{xyz}(\boldsymbol{q}_\mathsf{xyz})^\top + 2q_\mathsf{w}\,\mathrm{wed}(\boldsymbol{q}_\mathsf{xyz}) \\ &= \begin{bmatrix} 1 - 2(q_\mathsf{y}^2 + q_\mathsf{z}^2) & 2(q_\mathsf{y}q_\mathsf{x} - q_\mathsf{w}q_\mathsf{z}) & 2(q_\mathsf{z}q_\mathsf{x} + q_\mathsf{w}q_\mathsf{y}) \\ 2(q_\mathsf{y}q_\mathsf{x} + q_\mathsf{w}q_\mathsf{z}) & 1 - 2(q_\mathsf{x}^2 + q_\mathsf{z}^2) & 2(q_\mathsf{y}q_\mathsf{z} - q_\mathsf{w}q_\mathsf{x}) \\ 2(q_\mathsf{z}q_\mathsf{x} - q_\mathsf{w}q_\mathsf{y}) & 2(q_\mathsf{y}q_\mathsf{z} + q_\mathsf{w}q_\mathsf{x}) & 1 - 2(q_\mathsf{x}^2 + q_\mathsf{y}^2) \end{bmatrix} \end{aligned} \tag{A.154}$$

$$\begin{bmatrix} 0 \\ \omega_\mathsf{x} \\ \omega_\mathsf{y} \\ \omega_\mathsf{z} \end{bmatrix} = 2 \underbrace{\begin{bmatrix} q_\mathsf{w} & q_\mathsf{x} & q_\mathsf{y} & q_\mathsf{z} \\ -q_\mathsf{x} & q_\mathsf{w} & q_\mathsf{z} & -q_\mathsf{y} \\ -q_\mathsf{y} & -q_\mathsf{z} & q_\mathsf{w} & q_\mathsf{x} \\ -q_\mathsf{z} & q_\mathsf{y} & -q_\mathsf{x} & q_\mathsf{w} \end{bmatrix}}_{\widehat{q}^\top} \begin{bmatrix} \dot{q}_\mathsf{w} \\ \dot{q}_\mathsf{x} \\ \dot{q}_\mathsf{y} \\ \dot{q}_\mathsf{z} \end{bmatrix} \tag{A.155}$$

Note that $\boldsymbol{q}$ and $-\boldsymbol{q}$ represent the same rotation matrix. For the inversion we restrict to the case $q_{\mathsf{w}} \geq 0$ and get

$$q_{\mathsf{w}} = \tfrac{1}{2}\sqrt{1+\operatorname{tr}\boldsymbol{R}}\,, \quad \boldsymbol{q}_{\mathsf{xyz}} = \tfrac{1}{2q_{\mathsf{w}}}\boldsymbol{R}^{\vee} \tag{A.156}$$

$$\begin{bmatrix}\dot{q}_{\mathsf{w}}\\ \dot{q}_{\mathsf{x}}\\ \dot{q}_{\mathsf{y}}\\ \dot{q}_{\mathsf{z}}\end{bmatrix} = \tfrac{1}{2}\underbrace{\begin{bmatrix}q_{\mathsf{w}} & -q_{\mathsf{x}} & -q_{\mathsf{y}} & -q_{\mathsf{z}}\\ q_{\mathsf{x}} & q_{\mathsf{w}} & -q_{\mathsf{z}} & q_{\mathsf{y}}\\ q_{\mathsf{y}} & q_{\mathsf{z}} & q_{\mathsf{w}} & -q_{\mathsf{x}}\\ q_{\mathsf{z}} & -q_{\mathsf{y}} & q_{\mathsf{x}} & q_{\mathsf{w}}\end{bmatrix}}_{\widehat{q}}\begin{bmatrix}0\\ \omega_{\mathsf{x}}\\ \omega_{\mathsf{y}}\\ \omega_{\mathsf{z}}\end{bmatrix} \tag{A.157}$$

For the kinematic relation we have $\det\widehat{q} = \|q\|^4 = 1$. The given geometric relation (A.156) is singular for $q_{\mathsf{w}} = 0$. This could be resolved by constructing a relation with a different combination of the diagonal elements, resulting in a different singularity, see [Horn, 1987]. A more numerically stable relation is

$$\begin{bmatrix}q_{\mathsf{w}}\\ q_{\mathsf{x}}\\ q_{\mathsf{y}}\\ q_{\mathsf{z}}\end{bmatrix} = \tfrac{1}{2}\begin{bmatrix}\sqrt{1+R^{\mathsf{x}}_{\mathsf{x}}+R^{\mathsf{y}}_{\mathsf{y}}+R^{\mathsf{z}}_{\mathsf{z}}}\\ \sqrt{1+R^{\mathsf{x}}_{\mathsf{x}}-R^{\mathsf{y}}_{\mathsf{y}}-R^{\mathsf{z}}_{\mathsf{z}}}\,\operatorname{sign}(R^{\mathsf{z}}_{\mathsf{y}}-R^{\mathsf{y}}_{\mathsf{z}})\\ \sqrt{1-R^{\mathsf{x}}_{\mathsf{x}}+R^{\mathsf{y}}_{\mathsf{y}}-R^{\mathsf{z}}_{\mathsf{z}}}\,\operatorname{sign}(R^{\mathsf{x}}_{\mathsf{z}}-R^{\mathsf{z}}_{\mathsf{x}})\\ \sqrt{1-R^{\mathsf{x}}_{\mathsf{x}}-R^{\mathsf{y}}_{\mathsf{y}}+R^{\mathsf{z}}_{\mathsf{z}}}\,\operatorname{sign}(R^{\mathsf{y}}_{\mathsf{x}}-R^{\mathsf{x}}_{\mathsf{y}})\end{bmatrix}. \tag{A.158}$$

Quaternion algebra. The quaternion multiplication (sometimes called the Hamilton product) can be written as a matrix multiplication (the wedge operator was introduced in (A.157))

$$\boldsymbol{q}_1 \cdot \boldsymbol{q}_2 = \widehat{\boldsymbol{q}}_1\boldsymbol{q}_2 \tag{A.159}$$

The conjugate of a quaternion is

$$\boldsymbol{q}^* = [q_{\mathsf{w}}, (-\boldsymbol{q}_{\mathsf{xyz}})^\top]^\top, \qquad \boldsymbol{q}^*\cdot\boldsymbol{q} = \boldsymbol{q}\cdot\boldsymbol{q}^* = [1,0,0,0]^\top. \tag{A.160}$$

We have the following relations for rotation matrices

$$\boldsymbol{R}(\boldsymbol{q}_1)\boldsymbol{R}(\boldsymbol{q}_2) = \boldsymbol{R}(\boldsymbol{q}_1\cdot\boldsymbol{q}_2), \qquad \big(\boldsymbol{R}(\boldsymbol{q})\big)^\top = \boldsymbol{R}(\boldsymbol{q}^*). \tag{A.161}$$

A.5.3 Euler angles

The rotation matrix corresponding to the Euler angles in the *roll* $\alpha \in (-\pi,\pi]$, *pitch* $\beta \in [-\pi/2,\pi/2]$, *yaw* $\varphi \in (-\pi,\pi]$ convention is

$$\boldsymbol{R} = \boldsymbol{R}_{\mathsf{Z}}(\varphi)\boldsymbol{R}_{\mathsf{Y}}(\beta)\boldsymbol{R}_{\mathsf{X}}(\alpha) = \begin{bmatrix}\mathrm{c}_\varphi\mathrm{c}_\beta & \mathrm{c}_\varphi\mathrm{s}_\beta\mathrm{s}_\alpha-\mathrm{s}_\varphi\mathrm{c}_\alpha & \mathrm{c}_\varphi\mathrm{s}_\beta\mathrm{c}_\alpha+\mathrm{s}_\varphi\mathrm{s}_\alpha\\ \mathrm{s}_\varphi\mathrm{c}_\beta & \mathrm{s}_\varphi\mathrm{s}_\beta\mathrm{s}_\alpha+\mathrm{c}_\varphi\mathrm{c}_\alpha & \mathrm{s}_\varphi\mathrm{s}_\beta\mathrm{c}_\alpha-\mathrm{c}_\varphi\mathrm{s}_\alpha\\ -\mathrm{s}_\beta & \mathrm{c}_\beta\mathrm{s}_\alpha & \mathrm{c}_\beta\mathrm{c}_\alpha\end{bmatrix} \tag{A.162}$$

$$\begin{bmatrix}\omega_{\mathsf{x}}\\ \omega_{\mathsf{y}}\\ \omega_{\mathsf{z}}\end{bmatrix} = \begin{bmatrix}1 & 0 & -\mathrm{s}_\beta\\ 0 & \mathrm{c}_\alpha & \mathrm{c}_\beta\mathrm{s}_\alpha\\ 0 & -\mathrm{s}_\alpha & \mathrm{c}_\beta\mathrm{c}_\alpha\end{bmatrix}\begin{bmatrix}\dot{\alpha}\\ \dot{\beta}\\ \dot{\varphi}\end{bmatrix}. \tag{A.163}$$

The inverse relation is

$$\varphi = \operatorname{atan2}(R^{\mathsf{y}}_{\mathsf{x}}, R^{\mathsf{x}}_{\mathsf{x}}), \quad \beta = -\arcsin R^{\mathsf{z}}_{\mathsf{x}}, \quad \alpha = \operatorname{atan2}(R^{\mathsf{z}}_{\mathsf{y}}, R^{\mathsf{z}}_{\mathsf{z}}) \tag{A.164}$$

$$\begin{bmatrix}\dot\alpha\\ \dot\beta\\ \dot\varphi\end{bmatrix} = \begin{bmatrix}1 & \frac{\mathrm{s}_\beta \mathrm{s}_\alpha}{\mathrm{c}_\beta} & \frac{\mathrm{c}_\alpha \mathrm{s}_\beta}{\mathrm{c}_\beta}\\ 0 & \mathrm{c}_\alpha & -\mathrm{s}_\alpha\\ 0 & \frac{\mathrm{s}_\alpha}{\mathrm{c}_\beta} & \frac{\mathrm{c}_\alpha}{\mathrm{c}_\beta}\end{bmatrix}\begin{bmatrix}\omega_{\mathsf{x}}\\ \omega_{\mathsf{y}}\\ \omega_{\mathsf{z}}\end{bmatrix}. \tag{A.165}$$

Note that the kinematic relation is singular for $\mathrm{c}_\beta = 0$, i.e. $\beta = \pm\frac{\pi}{2}$.

Quadcopter decomposition. First define

$$\boldsymbol{R}_{\mathsf{PZ}} \;:\; \mathbb{S}^2 \to \mathbb{SO}(3), \;\begin{bmatrix}p_{\mathsf{x}}\\ p_{\mathsf{y}}\\ p_{\mathsf{z}}\end{bmatrix} \mapsto \begin{bmatrix}1-\frac{p_{\mathsf{x}}^2}{1+p_{\mathsf{z}}} & -\frac{p_{\mathsf{x}}p_{\mathsf{y}}}{1+p_{\mathsf{z}}} & p_{\mathsf{x}}\\ -\frac{p_{\mathsf{x}}p_{\mathsf{y}}}{1+p_{\mathsf{z}}} & 1-\frac{p_{\mathsf{y}}^2}{1+p_{\mathsf{z}}} & p_{\mathsf{y}}\\ -p_{\mathsf{x}} & -p_{\mathsf{y}} & p_{\mathsf{z}}\end{bmatrix}, \tag{A.166}$$

Consider the decomposition by

$$\boldsymbol{R} = \boldsymbol{R}_{\mathsf{Z}}(\varphi)\boldsymbol{R}_{\mathsf{PZ}}(p) = \begin{bmatrix}\mathrm{c}_\varphi - \frac{p_{\mathsf{x}}(\mathrm{c}_\varphi p_{\mathsf{x}} - \mathrm{s}_\varphi p_{\mathsf{y}})}{1+p_{\mathsf{z}}} & -\mathrm{s}_\varphi - \frac{p_{\mathsf{y}}(\mathrm{c}_\varphi p_{\mathsf{x}} - \mathrm{s}_\varphi p_{\mathsf{y}})}{1+p_{\mathsf{z}}} & \mathrm{c}_\varphi p_{\mathsf{x}} - \mathrm{s}_\varphi p_{\mathsf{y}}\\ \mathrm{s}_\varphi - \frac{p_{\mathsf{x}}(\mathrm{s}_\varphi p_{\mathsf{x}} + \mathrm{c}_\varphi p_{\mathsf{y}})}{1+p_{\mathsf{z}}} & \mathrm{c}_\varphi - \frac{p_{\mathsf{y}}(\mathrm{s}_\varphi p_{\mathsf{x}} + \mathrm{c}_\varphi p_{\mathsf{y}})}{1+p_{\mathsf{z}}} & \mathrm{s}_\varphi p_{\mathsf{x}} + \mathrm{c}_\varphi p_{\mathsf{y}}\\ -p_{\mathsf{x}} & -p_{\mathsf{y}} & p_{\mathsf{z}}\end{bmatrix} \tag{A.167}$$

$$\begin{bmatrix}\omega_{\mathsf{x}}\\ \omega_{\mathsf{y}}\\ \omega_{\mathsf{z}}\\ 0\end{bmatrix} = \begin{bmatrix}0 & -1 & \frac{p_{\mathsf{y}}}{1+p_{\mathsf{z}}} & -p_{\mathsf{x}}\\ 1 & 0 & -\frac{p_{\mathsf{x}}}{1+p_{\mathsf{z}}} & -p_{\mathsf{y}}\\ \frac{p_{\mathsf{y}}}{1+p_{\mathsf{z}}} & -\frac{p_{\mathsf{x}}}{1+p_{\mathsf{z}}} & 0 & p_{\mathsf{z}}\\ p_{\mathsf{x}} & p_{\mathsf{y}} & p_{\mathsf{z}} & 0\end{bmatrix}\begin{bmatrix}\dot p_{\mathsf{x}}\\ \dot p_{\mathsf{y}}\\ \dot p_{\mathsf{z}}\\ \dot\varphi\end{bmatrix} \tag{A.168}$$

inverse relation

$$\varphi = \operatorname{atan2}(R^{\mathsf{y}}_{\mathsf{x}} - R^{\mathsf{x}}_{\mathsf{y}}, R^{\mathsf{x}}_{\mathsf{x}} + R^{\mathsf{y}}_{\mathsf{y}}), \quad p = [-R^{\mathsf{z}}_{\mathsf{x}}, -R^{\mathsf{z}}_{\mathsf{y}}, R^{\mathsf{z}}_{\mathsf{z}}]^\top \tag{A.169}$$

$$\begin{bmatrix}\dot p_{\mathsf{x}}\\ \dot p_{\mathsf{y}}\\ \dot p_{\mathsf{z}}\\ \dot\varphi\end{bmatrix} = \begin{bmatrix}0 & p_{\mathsf{z}} & p_{\mathsf{y}} & \frac{p_{\mathsf{x}}}{1+p_{\mathsf{z}}}\\ -p_{\mathsf{z}} & 0 & -p_{\mathsf{x}} & \frac{p_{\mathsf{y}}}{1+p_{\mathsf{z}}}\\ p_{\mathsf{y}} & -p_{\mathsf{x}} & 0 & 1\\ -\frac{p_{\mathsf{x}}}{1+p_{\mathsf{z}}} & -\frac{p_{\mathsf{y}}}{1+p_{\mathsf{z}}} & 1 & 0\end{bmatrix}\begin{bmatrix}\omega_{\mathsf{x}}\\ \omega_{\mathsf{y}}\\ \omega_{\mathsf{z}}\\ 0\end{bmatrix} \tag{A.170}$$

Note that

$$\boldsymbol{R}_{\mathsf{Z}}(\varphi)\boldsymbol{R}_{\mathsf{PZ}}(p) = \boldsymbol{R}_{\mathsf{PZ}}(p')\boldsymbol{R}_{\mathsf{Z}}(\varphi) \quad \text{where} \quad p' = \boldsymbol{R}_{\mathsf{Z}}(\varphi)p \tag{A.171}$$

A.5.4 Minimal rotation matrix

For two given unit vectors $a, b \in \mathbb{S}^2$ for a rotation matrix $\boldsymbol{R}_{\mathsf{AB}} \in \mathbb{SO}(3)$ such that $b = \boldsymbol{R}_{\mathsf{AB}}a$. Since this only fixes 2 of the 3 degrees of freedom of the rotation matrix we impose the additional requirement that we want the solution which has the minimal angle. This unique solution is

$$\boldsymbol{R}_{\mathsf{AB}} = \mathbf{I}_3 + \mathrm{s}_\alpha \widehat{z} + (1-\mathrm{c}_\alpha)\widehat{z}^2, \qquad z = \tfrac{a\times b}{\|a\times b\|}, \quad \mathrm{c}_\alpha = \langle a, b\rangle, \quad \mathrm{s}_\alpha = \|a\times b\| \tag{A.172}$$

$$= \mathbf{I}_3 + \widehat{c} + s\widehat{c}^2, \qquad c = a\times b, \quad s = (1+\langle a, b\rangle)^{-1} \tag{A.173}$$

The angular velocity $\boldsymbol{\omega} = (\boldsymbol{R}^\top \dot{\boldsymbol{R}})^\vee$ can be expressed by

$$\boldsymbol{\omega} = s\big((a+b) \times (\dot{b} - \dot{a}) - 2\langle c, \dot{b}\rangle a\big), \tag{A.174}$$

$$\begin{aligned}\dot{\boldsymbol{\omega}} = s\big((a+b) \times (\ddot{b} - \ddot{a}) + 2\big(\dot{a} \times \dot{b} + \langle \dot{a} \times \dot{b}, b\rangle a - \langle c, \ddot{b}\rangle a - \langle c, \dot{b}\rangle \dot{a}\big) \\ - (\langle \dot{a}, b\rangle + \langle a, \dot{b}\rangle)\boldsymbol{\omega}\big),\end{aligned} \tag{A.175}$$

A.6 Constraint stabilization

In practice we might run into trouble with flawed initial conditions $\phi(x_0) \neq 0$ or with the flaws of numerical integration. To counter this we can add a *stabilization term* $-\Lambda\boldsymbol{\phi}$ to the kinematic equation:

$$\dot{\boldsymbol{x}} = \underbrace{[\boldsymbol{A}\ \boldsymbol{\Psi}]}_{\boldsymbol{A}_\square} \begin{bmatrix} \boldsymbol{\xi} \\ -\Lambda\boldsymbol{\phi} \end{bmatrix}. \tag{A.176}$$

This results in

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi} = \tfrac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\dot{\boldsymbol{x}} = \underbrace{\tfrac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\boldsymbol{A}}_{0}\boldsymbol{\xi} - \underbrace{\tfrac{\partial \phi}{\partial x}\boldsymbol{\Psi}}_{\mathbf{I}_{n-\nu}}\Lambda\phi = -\Lambda\boldsymbol{\phi} \tag{A.177}$$

i.e. for an appropriate $\Lambda \in \mathbb{R}^{(\nu-n)\times(\nu-n)}$ the error to the geometric constraint converges exponentially. For the example of the rigid body orientation (A.176) can also be written in the probably more intuitive matrix form

$$\dot{\boldsymbol{R}} = \boldsymbol{R}(\widehat{\boldsymbol{\omega}} - \Lambda(\boldsymbol{R}^\top\boldsymbol{R} - \mathbf{I}_3)) \tag{A.178}$$

with a symmetric positive definite matrix $\Lambda = \Lambda^\top \in \mathbb{R}^{3\times3} > 0$.

The resulting EoM with Lagrangian multipliers (3.83) only consider the derivatives of the constraints, e.g. $\ddot{\psi} = 0$ or $\dot{\eta} = 0$. As before, for the case of bad initial conditions or the errors of numerical integration, we should add a stabilization term: In the case of kinematic constraints this is

$$\underbrace{\tfrac{\partial \eta}{\partial \boldsymbol{\xi}}}_{\boldsymbol{Z}}\dot{\boldsymbol{\xi}} = \underbrace{-\boldsymbol{\nabla}\eta\,\boldsymbol{\xi} - \tfrac{\partial \eta}{\partial t}}_{\boldsymbol{b}}\ \underbrace{-\Lambda\eta(\boldsymbol{x},\boldsymbol{\xi},t)}_{\text{stabilization}} \qquad \Leftrightarrow \qquad \dot{\eta} + \Lambda\eta = 0 \tag{A.179a}$$

and for a geometric constraints

$$\underbrace{\boldsymbol{\nabla}\psi}_{\boldsymbol{Z}}\dot{\boldsymbol{\xi}} = \underbrace{-\tfrac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\nabla}\psi)\boldsymbol{\xi}}_{\boldsymbol{b}}\ \underbrace{-\Lambda_1\boldsymbol{\nabla}\psi(\boldsymbol{x})\,\boldsymbol{\xi} - \Lambda_0\psi(\boldsymbol{x})}_{\text{stabilization}} \qquad \Leftrightarrow \qquad \ddot{\psi} + \Lambda_1\dot{\psi} + \Lambda_0\psi = 0. \tag{A.179b}$$

consult [Bremer, 2008, sec. 5.2.4]

A.7 Alternative from for the body Jacobian

Rigid body coordinates. For the configuration of the frame with index b w.r.t. the frame with index a with we use the configuration coordinates

$$ {}^{a}_{b}\mathbf{x} = \left[{}^{a}_{b}r^{\mathsf{x}}, {}^{a}_{b}r^{\mathsf{y}}, {}^{a}_{b}r^{\mathsf{z}}, {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}}, {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}}, {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}}, {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}}, {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}}, {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}}, {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}}, {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}}, {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}}, \right]. \tag{A.180} $$

constrained by

$$ \underbrace{\begin{bmatrix} ({}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}})^2 + ({}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}})^2 + ({}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}})^2 - 1 \\ ({}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}})^2 + ({}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}})^2 + ({}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}})^2 - 1 \\ ({}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}})^2 + ({}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}})^2 + ({}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}})^2 - 1 \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} + {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} + {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} + {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} + {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} + {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} + {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} + {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} + {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} - {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} - {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} - {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}}\,{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}}\,{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} - 1 \end{bmatrix}}_{\boldsymbol{\phi}({}^{a}_{b}\mathbf{x})} = \mathbf{0}. \tag{A.181} $$

We use the velocity coordinates ${}^{a}_{b}\boldsymbol{\xi}$ with the kinematic relation

$$ \frac{\mathrm{d}}{\mathrm{d}t} \underbrace{\begin{bmatrix} {}^{a}_{b}r^{\mathsf{x}} \\ {}^{a}_{b}r^{\mathsf{y}} \\ {}^{a}_{b}r^{\mathsf{z}} \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} \\ {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} \\ {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} \\ {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} \\ {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} \\ {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} \\ {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} \end{bmatrix}}_{{}^{a}_{b}\mathbf{x}} = \underbrace{\begin{bmatrix} {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} & {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} & 0 & 0 & 0 \\ {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} & 0 & 0 & 0 \\ {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} & {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} \\ 0 & 0 & 0 & 0 & -{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} \\ 0 & 0 & 0 & 0 & -{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} \\ 0 & 0 & 0 & {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} & 0 & -{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} \\ 0 & 0 & 0 & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} & 0 & -{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} \\ 0 & 0 & 0 & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} & 0 & -{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} \\ 0 & 0 & 0 & -{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} & 0 \\ 0 & 0 & 0 & -{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} & 0 \\ 0 & 0 & 0 & -{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} & 0 \end{bmatrix}}_{\mathbf{A}({}^{a}_{b}\mathbf{x})} \underbrace{\begin{bmatrix} {}^{a}_{b}v^{\mathsf{x}} \\ {}^{a}_{b}v^{\mathsf{y}} \\ {}^{a}_{b}v^{\mathsf{z}} \\ {}^{a}_{b}\omega^{\mathsf{x}} \\ {}^{a}_{b}\omega^{\mathsf{y}} \\ {}^{a}_{b}\omega^{\mathsf{z}} \end{bmatrix}}_{{}^{a}_{b}\boldsymbol{\xi}}. \tag{A.182} $$

The inverse relation is ${}^{a}_{b}\boldsymbol{\xi} = \mathbf{Y}({}^{a}_{b}\mathbf{x})\,{}^{a}_{b}\dot{\mathbf{x}}$ with

$$ \mathbf{Y}({}^{a}_{b}\mathbf{x}) = \mathbf{A}^{+}({}^{a}_{b}\mathbf{x}) = $$
$$ \begin{bmatrix} {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ {}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} & {}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} & {}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} & \frac{1}{2}{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} & \frac{1}{2}{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} \\ 0 & 0 & 0 & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{z}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{z}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{z}} & 0 & 0 & 0 & \frac{1}{2}{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} & \frac{1}{2}{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} & \frac{1}{2}{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} \\ 0 & 0 & 0 & \frac{1}{2}{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{y}} & \frac{1}{2}{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{y}} & \frac{1}{2}{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{y}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{x}}_{\mathsf{x}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{y}}_{\mathsf{x}} & -\frac{1}{2}{}^{a}_{b}R^{\mathsf{z}}_{\mathsf{x}} & 0 & 0 & 0 \end{bmatrix} \tag{A.183} $$

For the kinematics matrix $\mathbf{A}$ we have the corresponding commutation coefficients

$$ \mathsf{\Upsilon}^{r}_{pq} = \left(\mathsf{A}^{\sigma}_{p} \frac{\partial \mathsf{A}^{\rho}_{q}}{\partial \mathsf{x}^{\sigma}} - \mathsf{A}^{\sigma}_{q} \frac{\partial \mathsf{A}^{\rho}_{p}}{\partial \mathsf{x}^{\sigma}} \right) \mathsf{Y}^{r}_{\rho} = \left(\frac{\partial \mathsf{Y}^{r}_{\rho}}{\partial \mathsf{x}^{\sigma}} - \frac{\partial \mathsf{Y}^{r}_{\sigma}}{\partial \mathsf{x}^{\rho}} \right) \mathsf{A}^{\sigma}_{q} \mathsf{A}^{\rho}_{p}, \qquad p, q, r = 1, \dots, 6. \tag{A.184} $$

In (4.15) we have already seen that these turn out to be constant.

Body Jacobian. Assume we have appropriate configuration $\boldsymbol{x}$ and velocity coordinates $\boldsymbol{\xi}$ related by $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}$ and the body configuration coordinates ${}^a_b\mathsf{x} = {}^a_b\mathsf{x}(\boldsymbol{x})$ are parameterized by the system configuration coordinates $\boldsymbol{x}$. For the body velocities we have

$$
{}^a_b\boldsymbol{\xi} = \mathsf{A}^{+}({}^a_b\mathsf{x}(\boldsymbol{x}))\frac{\partial\, {}^a_b\mathsf{x}}{\partial \boldsymbol{x}}(\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi} \tag{A.185}
$$

Introducing the shorthand notation $\mathsf{A}^{+}({}^a_b\mathsf{x}(\boldsymbol{x})) = {}^a_b\mathsf{Y}(\boldsymbol{x})$ and omitting the dependencies, we can express the body Jacobian as

$$
{}^a_b\boldsymbol{J} = \frac{\partial\, {}^a_b\boldsymbol{\xi}}{\partial \boldsymbol{\xi}} = {}^a_b\mathsf{Y}\frac{\partial\, {}^a_b\mathsf{x}}{\partial \boldsymbol{x}}\boldsymbol{A} \tag{A.186}
$$

In index notation this is

$$
{}^a_bJ^p_i = {}^a_b\mathsf{Y}^p_\rho\, \partial_i\, {}^a_b\mathsf{x}^\rho, \tag{A.187}
$$

For the following commutation relation will be useful for the next paragraph

$$
\begin{aligned}
\partial_j\, {}^a_bJ^p_i - \partial_i\, {}^a_bJ^p_j \;&=\; \partial_j\, {}^a_b\mathsf{Y}^p_\rho\, \partial_i\, {}^a_b\mathsf{x}^\rho + {}^a_b\mathsf{Y}^p_\rho\, \partial_j\partial_i\, {}^a_b\mathsf{x}^\rho - \partial_i\, {}^a_b\mathsf{Y}^p_\rho\, \partial_j\, {}^a_b\mathsf{x}^\rho - {}^a_b\mathsf{Y}^p_\rho\, \partial_i\partial_j\, {}^a_b\mathsf{x}^\rho \\
&\overset{(3.34)}{=} \frac{\partial\, {}^a_b\mathsf{Y}^p_\rho}{\partial\, {}^a_b\mathsf{x}^\sigma}\partial_j\, {}^a_b\mathsf{x}^\sigma\, \partial_i\, {}^a_b\mathsf{x}^\rho - \frac{\partial\, {}^a_b\mathsf{Y}^p_\rho}{\partial\, {}^a_b\mathsf{x}^\sigma}\partial_i\, {}^a_b\mathsf{x}^\sigma\, \partial_j\, {}^a_b\mathsf{x}^\rho - {}^a_b\mathsf{Y}^p_\rho\, \gamma^s_{ij}\partial_s\, {}^a_b\mathsf{x}^\rho \\
&= \left(\frac{\partial\, {}^a_b\mathsf{Y}^p_\sigma}{\partial\, {}^a_b\mathsf{x}^\rho} - \frac{\partial\, {}^a_b\mathsf{Y}^p_\rho}{\partial\, {}^a_b\mathsf{x}^\sigma}\right)\partial_i\, {}^a_b\mathsf{x}^\sigma\, \partial_j\, {}^a_b\mathsf{x}^\rho - \gamma^s_{ij}\, {}^a_bJ^p_s \\
&\overset{(A.184)}{=} \Upsilon^p_{qr}\, {}^a_bJ^q_i\, {}^a_bJ^r_j - \gamma^s_{ij}\, {}^a_bJ^p_s.
\end{aligned} \tag{A.188}
$$

The last step uses the argument from (3.33) twice.

Connection coefficients. For the metric coefficients

$$
M_{ij} = \sum\nolimits_{a,b} {}^a_b\mathsf{M}_{pq}\, {}^a_bJ^p_i\, {}^a_bJ^q_j, \qquad {}^a_b\mathsf{M}_{pq} = {}^a_b\mathsf{M}_{qp} = const. \tag{A.189}
$$

we can express the connection coefficients as

$$
\begin{aligned}
\mathit{\Gamma}_{ijk} \;&=\; \tfrac{1}{2}\big(\partial_k M_{ij} + \partial_j M_{ik} - \partial_i M_{jk} + \gamma^s_{ij}M_{sk} + \gamma^s_{ik}M_{sj} - \gamma^s_{jk}M_{si}\big) \\
&\overset{(A.189)}{=} \tfrac{1}{2}\sum\nolimits_{a,b} {}^a_b\mathsf{M}_{pq}\big(\partial_k\, {}^a_bJ^p_i\, {}^a_bJ^q_j + {}^a_bJ^p_i\partial_k\, {}^a_bJ^q_j + \gamma^s_{ij}\, {}^a_bJ^p_s\, {}^a_bJ^q_k \\
&\qquad\qquad + \partial_j\, {}^a_bJ^p_i\, {}^a_bJ^q_k + {}^a_bJ^p_i\partial_j\, {}^a_bJ^q_k + \gamma^s_{ik}\, {}^a_bJ^p_s\, {}^a_bJ^q_j \\
&\qquad\qquad - \partial_i\, {}^a_bJ^p_j\, {}^a_bJ^q_k - {}^a_bJ^p_j\partial_i\, {}^a_bJ^q_k - \gamma^s_{jk}\, {}^a_bJ^p_s\, {}^a_bJ^q_i\big) \\
&\overset{(A.188)}{=} \tfrac{1}{2}\sum\nolimits_{a,b} {}^a_b\mathsf{M}_{pq}\big(\partial_k\, {}^a_bJ^p_i\, {}^a_bJ^q_j + {}^a_bJ^p_i\partial_k\, {}^a_bJ^q_j + \big(\Upsilon^p_{hr}\, {}^a_bJ^h_i\, {}^a_bJ^r_j - \partial_j\, {}^a_bJ^p_i + \partial_i\, {}^a_bJ^p_j\big)\, {}^a_bJ^q_k \\
&\qquad\qquad + \partial_j\, {}^a_bJ^p_i\, {}^a_bJ^q_k + {}^a_bJ^p_i\partial_j\, {}^a_bJ^q_k + \big(\Upsilon^p_{hr}\, {}^a_bJ^h_i\, {}^a_bJ^r_k - \partial_k\, {}^a_bJ^p_i + \partial_i\, {}^a_bJ^p_k\big)\, {}^a_bJ^q_j \\
&\qquad\qquad - \partial_i\, {}^a_bJ^p_j\, {}^a_bJ^q_k - {}^a_bJ^p_j\partial_i\, {}^a_bJ^q_k - \big(\Upsilon^p_{hr}\, {}^a_bJ^h_j\, {}^a_bJ^r_k - \partial_k\, {}^a_bJ^p_j + \partial_j\, {}^a_bJ^p_k\big)\, {}^a_bJ^q_i\big) \\
&= \tfrac{1}{2}\sum\nolimits_{a,b} {}^a_b\mathsf{M}_{pq}\big(2\, {}^a_bJ^p_i\partial_k\, {}^a_bJ^q_j + \Upsilon^p_{hr}\, {}^a_bJ^h_i\, {}^a_bJ^r_j\, {}^a_bJ^q_k + \Upsilon^p_{hr}\, {}^a_bJ^h_i\, {}^a_bJ^r_k\, {}^a_bJ^q_j - \Upsilon^p_{hr}\, {}^a_bJ^h_j\, {}^a_bJ^r_k\, {}^a_bJ^q_i\big) \\
&= \sum\nolimits_{a,b} {}^a_bJ^p_i\big({}^a_b\mathsf{M}_{pq}\partial_k\, {}^a_bJ^q_j + \underbrace{\tfrac{1}{2}\big(\Upsilon^h_{pq}\, {}^a_b\mathsf{M}_{hr} + \Upsilon^h_{pr}\, {}^a_b\mathsf{M}_{hq} - \Upsilon^h_{qr}\, {}^a_b\mathsf{M}_{hp}\big)}_{{}^a_b\Gamma_{pqr}}\, {}^a_bJ^q_j\, {}^a_bJ^r_k\big)
\end{aligned} \tag{A.190}
$$

Appendix B

On nonholonomic systems

Nonholonomic systems are characterized by having constraints expressed in terms of the *velocity* $\dot{\boldsymbol{x}}$. We restrict to linear constraints, i.e.

$$N^a_\alpha(\boldsymbol{x})\,\dot{x}^\alpha = 0, \qquad a = 1 \ldots \mu. \tag{B.1}$$

since practical examples of nonlinear nonholonomic constraints seem to not exist, see [Hamel, 1949, p. 499] or [Neimark and Fufaev, 1972, ch. IV].

Nonholonomy. So far (B.1) should rather be called a kinematic constraint. It is nonholonomic if it is *nonintegrable*, i.e. there exists no function $\psi(\boldsymbol{x})$ such that $\frac{\mathrm{d}}{\mathrm{d}t}\psi = \frac{\partial\psi}{\partial\boldsymbol{x}}\dot{\boldsymbol{x}} = \boldsymbol{N}\dot{\boldsymbol{x}}$. The reverse would imply

$$N^a_\alpha = \frac{\partial\psi^a}{\partial x^\alpha} \qquad \Rightarrow \qquad \frac{\partial N^a_\alpha}{\partial x^\beta} = \frac{\partial^2\psi^a}{\partial x^\beta \partial x^\alpha} = \frac{\partial N^a_\beta}{\partial x^\alpha} \tag{B.2}$$

If this condition is fulfilled, we call (B.1) a holonomic kinematic constraint and the only difference to a geometric constraint (as treated in (**??**)) is a missing initial condition.

Some trouble. It is well documented in the dedicated literature (e.g. [Hamel, 1935], [Bloch, 2003, p. 208], probably first mentioned in [Korteweg, 1899]) that in the case of nonholonomic constraints, the Lagrange-d'Alembert principle (**??**) and the calculus of variations (**??**) lead to *different* results. In the context of the Lagrange-d'Alembert principle, the *virtual displacements* $\delta\boldsymbol{x}$ have to be compatible with system constraints which implies

$$N^a_\alpha\,\delta x^\alpha = 0. \tag{B.3}$$

This results in the same dynamical equations as derived with a Newton-Eulerian approach.

On the other hand to get the solution to an associated variational problem the *varied path* $t \mapsto \bar{\boldsymbol{x}}(t) = \boldsymbol{x}(t) + \varepsilon\boldsymbol{\chi}(t)$ has to obey the constraint, i.e.

$$\begin{aligned} 0 = N^a_\alpha(\bar{\boldsymbol{x}})\dot{\bar{x}}^{[}\alpha] &= \underbrace{N^a_\alpha \dot{x}^\alpha}_{0} + N^a_\alpha \underbrace{\varepsilon\dot{\chi}^{[}\alpha]}_{\delta\dot{x}^\alpha} + \frac{\partial N^a_\alpha}{\partial x^\beta}\underbrace{\varepsilon\chi^\beta}_{\delta x^\beta}\,\dot{x}^\alpha + \mathcal{O}(\varepsilon^2), \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(N^a_\alpha \delta x^\alpha\right) + \left(\frac{\partial N^a_\alpha}{\partial x^\beta} - \frac{\partial N^a_\beta}{\partial x^\alpha}\right)\dot{x}^\alpha \delta x^\beta + \mathcal{O}(\varepsilon^2) \end{aligned} \tag{B.4}$$

which is obviously a different constraint on the *variation* $\delta\boldsymbol{x}$. This branch is called *vakonomic mechanics*, see e.g. [Arnold et al., 2002, sec. 1.4], and leads to different equations of motion unless the constraint is holonomic, i.e. the term in the brackets in (B.4) vanishes. See also [Hamel, 1935] for a deeper analysis.

There is consensus that the "correct" equations for physics, i.e. the ones that are backed by experiments (see [Lewis and Murray, 1995]), are the ones derived with Lagrange-d'Alembert principle with the constraint (B.3). Hamilton's principle can still be used when the constraints (B.4) that arise form the calculus of variations are replaced by the "correct" constraints $N^a_\alpha\,\delta x^\alpha = 0$ for the variation $\delta\boldsymbol{x}$, see [Bloch, 2003, sec. 5.2] The vakonomic approach however, is interesting in its own right and has application for optimal control, e.g. [Bloch, 2003].

Equations of motion with Lagrange multipliers Assume we have a given holonomic mechanical system we have chosen possibly redundant coordinates $\boldsymbol{x}$ and a kinematic relation $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{\xi}$. Now we like to add the nonholonomic constraint (B.1) that can be expressed in terms of the velocity coordinates $\boldsymbol{\xi}$ as

$$N^a_\alpha\,\dot{x}^\alpha = 0 \qquad \Leftrightarrow \qquad \underbrace{N^a_\alpha A^\alpha_i}_{Z^a_i}\,\dot{\xi}^i = 0, \qquad a = 1\ldots\mu \tag{B.5}$$

Using the concept of the Lagrange multipliers from (**??**) we can write the kinetic equation as

$$f^{\mathrm{M}}_i + Z^a_i\boldsymbol{\lambda}[a] = f^{\mathrm{E}}_i - \partial_i\mathcal{V}, \qquad i = 1\ldots n. \tag{B.6}$$

Expressing the inertia forces $\boldsymbol{f}^{\mathrm{M}}$ by the kinetic energy (**??**) and introducing the Lagrangian $\mathcal{L} = \mathcal{T} - \mathcal{V}$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\xi^i} + \gamma^k_{ij}\xi^j\frac{\partial\mathcal{L}}{\partial\xi^k} - A^\alpha_i\frac{\partial\mathcal{L}}{\partial x^\alpha} + Z^a_i\boldsymbol{\lambda}[a] = f^{\mathrm{E}}_i, \qquad i = 1\ldots n. \tag{B.7}$$

This form for the special case that $\boldsymbol{x} = \boldsymbol{q}$ and $\boldsymbol{\xi} = \dot{\boldsymbol{q}}$, so γ, is widely used for nonholonomic systems and can be found in e.g. [Lurie, 2002, eq. 7.1.6] or [Goldstein, 1951, eq. 2.29]. However, we could have expressed the inertia forces $f^{\mathrm{M}}_i = \partial\mathcal{S}/\partial\dot{\xi}^i$ by the acceleration energy (??) or directly $f^{\mathrm{M}}_i = M_{ij}\dot{\xi}^i + \mathit{\Gamma}_{ijk}\xi^j\xi^k$ by the inertia matrix just as well.

Equations of motion without Lagrange multipliers Usually we do not want to calculate the Lagrange multipliers explicitly. Instead we like to use the kinematic constraint to reduce the order of the resulting equations of motion. The crucial idea[1] of Appell [Appell, 1900] and Hamel [Hamel, 1904] is to chose the velocity coordinates $\boldsymbol{\xi}$ such that the last μ coincide with the kinematic constraint[2], i.e. $\boldsymbol{\xi}[n-\mu+a] = N^a_\alpha\dot{x}^\alpha, a = 1\ldots\mu$.

[1] Actually Hamel and Appell derived their formulations directly from Lagrange-d'Alembert's principle without the use of Lagrange multipliers, but the main idea remains the same.

[2] This seemed to be their motivation for the use of velocity coordinates in the first place.

Using the structure from (**??**) this is

$$\underbrace{\begin{bmatrix} \tilde{Y} \\ \boldsymbol{N} \\ \boldsymbol{\Phi} \end{bmatrix}}_{\boldsymbol{Y}_\square} \dot{\boldsymbol{x}} = \begin{bmatrix} \tilde{\xi} \\ \eta \\ 0 \end{bmatrix}, \quad \operatorname{rank} \boldsymbol{Y}_\square = \nu. \qquad \Leftrightarrow \qquad \dot{\boldsymbol{x}} = [\underbrace{\tilde{A}\ \Upsilon}_{A}\ \Psi\,] \begin{bmatrix} \tilde{\xi} \\ \eta \\ 0 \end{bmatrix} \tag{B.8}$$

With this choice we can express the nonholonomic constraint simply by $\eta = 0$ and we have $\boldsymbol{Z} = [\,0\ \mathbf{I}_\mu\,]$. Consequently the last μ components of the force balance (B.7) define the reaction forces $\boldsymbol{\lambda}[1] \ldots \boldsymbol{\lambda}[\mu]$ while the first $n - \mu$ are independent of them.

Combining these considerations and expressing the generalized inertia force $\boldsymbol{f}^{\mathrm{M}}$ in terms of the *acceleration energy* $\mathcal{S}$ from (??) we find *Appell's equation*

$$\frac{\partial \mathcal{S}}{\partial \dot{\xi}^i}\Big|_{\boldsymbol{\xi}[n-\mu+1]\ldots\boldsymbol{\xi}[n]=0} + \partial_i \mathcal{V} = f_i^{\mathrm{E}}, \qquad i = 1 \ldots n - \mu. \tag{B.9}$$

Expressing the generalized inertia force $\boldsymbol{f}^{\mathrm{M}}$ in terms of the *kinetic energy* $\mathcal{T}$ from (??) and introducing the Lagrangian $\mathcal{L} = \mathcal{T} - \mathcal{V}$ we find *Hamel's equation*

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \xi^i} + \gamma_{ij}^k \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} - A_i^\alpha \frac{\partial \mathcal{L}}{\partial x^\alpha} \right)\Big|_{\boldsymbol{\xi}[n-\mu+1]\ldots\boldsymbol{\xi}[n]=0} = f_i^{\mathrm{E}}, \qquad i = 1 \ldots n - \mu. \tag{B.10}$$

Note that the dummy index k still runs over $1 \ldots n$ so we still need to differentiate $\mathcal{L}$ w.r.t. the velocity coordinates $\boldsymbol{\xi}[n - \mu + 1] \ldots \boldsymbol{\xi}[n]$ that will be set to zero *afterwards*.

We like to investigate this a bit further: First introduce the *constraint Lagrangian* $\tilde{\mathcal{L}}$:

$$\tilde{\mathcal{L}} = \mathcal{L}\big|_{\boldsymbol{\xi}[n-\mu+1]\ldots\boldsymbol{\xi}[n]=0}. \tag{B.11}$$

Then recall the splitting of the velocity coordinates $\xi = [\tilde{\xi}^\top \eta^\top]^\top$ from (B.8) and split the commutation coefficients in the same way:

$$\gamma_{ij}^k = \left(\frac{\partial \tilde{Y}_\alpha^k}{\partial x^\beta} - \frac{\partial \tilde{Y}_\beta^k}{\partial x^\alpha} \right) \tilde{A}_i^\alpha \tilde{A}_j^\beta = \tilde{\gamma}_{ij}^k, \qquad i, j, k = 1 \ldots n - \mu \tag{B.12a}$$

$$\gamma_{ij}^{n-\mu+a} = \left(\frac{\partial N_\alpha^a}{\partial x^\beta} - \frac{\partial N_\beta^a}{\partial x^\alpha} \right) \tilde{A}_i^\alpha \tilde{A}_j^\beta = \hat{\gamma}_{ij}^a, \qquad i, j = 1 \ldots n - \mu, \quad a = 1, \ldots, \mu \tag{B.12b}$$

Then we can write (B.10) as

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \tilde{\mathcal{L}}}{\partial \tilde{\xi}^i} + \tilde{\gamma}_{ij}^k \tilde{\xi}^j \frac{\partial \tilde{\mathcal{L}}}{\partial \xi^k} + \underbrace{\hat{\gamma}_{ij}^a \tilde{\xi}^j \frac{\partial \mathcal{L}}{\partial \eta^a}\Big|_{\eta=0}}_{c_i^{\mathrm{N}}} - \tilde{A}_i^\alpha \frac{\partial \tilde{\mathcal{L}}}{\partial x^\alpha} = f_i^{\mathrm{E}}, \qquad i = 1 \ldots n - \mu. \tag{B.13}$$

Note that if the kinematic constraint (B.1) is holonomic then $\hat{\gamma} = 0$ and consequently $\boldsymbol{c}^{\mathrm{N}} = 0$.

Discussion

EoM with minimal coordinates. Assume that for a special choice of minimal coordinates $\boldsymbol{q} = [q_{\mathsf{F}}^\top, q_{\mathsf{C}}^\top]^\top$ the kinematic constraint (B.1) can be transformed to

$$\dot{q}_{\mathsf{C}}^a = -W_i^a \dot{q}_{\mathsf{F}}^i, \qquad a = 1 \dots \mu. \tag{B.14}$$

Choose the velocity coordinates $\tilde{\xi} = \dot{q}_{\mathsf{F}}$ and $\eta = \dot{q}_{\mathsf{C}} + W\dot{q}_{\mathsf{F}}$ then we find that $\tilde{\gamma} = 0$ and

$$\hat{\gamma}_{ij}^a = \frac{\partial W_i^a}{\partial q_{\mathsf{F}}^j} - \frac{\partial W_j^a}{\partial q_{\mathsf{F}}^i} + \frac{\partial W_j^a}{\partial q_{\mathsf{C}}^b} W_i^b - \frac{\partial W_i^a}{\partial q_{\mathsf{C}}^b} W_j^b, \qquad a = 1, \dots, \mu \quad i, j = 1, \dots, n - \mu. \tag{B.15}$$

The kinetic equation (B.13) takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \tilde{\mathcal{L}}}{\partial \dot{q}_{\mathsf{F}}^i} - \frac{\partial \tilde{\mathcal{L}}}{\partial q_{\mathsf{F}}^i} + \underbrace{\hat{\gamma}_{ij}^a \dot{q}_{\mathsf{F}}^j \frac{\partial \mathcal{L}}{\partial \dot{q}_{\mathsf{C}}^a}|_{\dot{q}_{\mathsf{C}} = -S\dot{q}_{\mathsf{F}}}}_{c_i^{\mathrm{N}}} + W_i^a \frac{\partial \tilde{\mathcal{L}}}{\partial q_{\mathsf{C}}^a} = f_i^{\mathrm{E}}, \quad i = 1, \dots, n - \mu. \tag{B.16}$$

The combination of (B.14), (B.15) and (B.16) are called *the Lagrange-d'Alembert equations of motion* in [Bloch et al., 1996, Def. 2.1].

Example 14. A rolling rigid body Consider a rigid body, e.g. the ellipsoid from (??), rolling on a plane. We use the position $\boldsymbol{r}$ of a body fixed point and the orientation $\boldsymbol{R}$ as redundant configuration coordinates $\boldsymbol{x} \cong (\boldsymbol{r}, \boldsymbol{R})$. The condition of rolling without slipping can be formulated as

$$\dot{\boldsymbol{r}} + \dot{\boldsymbol{R}} p = 0 \qquad \cong \qquad \eta = \boldsymbol{N}(\boldsymbol{x})\dot{\boldsymbol{x}} = 0 \tag{B.17}$$

where p are the coordinates of the contact point w.r.t. the body fixed frame which depends on the current configuration $p = p(\boldsymbol{x})$.

As velocity coordinates we choose $\boldsymbol{\xi} = \boldsymbol{\omega}$ and obtain the kinematic relations

$$\boldsymbol{\omega} = (\boldsymbol{R}^\top \dot{\boldsymbol{R}})^\vee \qquad \Rightarrow \qquad \dot{\boldsymbol{r}} = \boldsymbol{R}\widehat{p}\boldsymbol{\omega}, \ \dot{\boldsymbol{R}} = \boldsymbol{R}\widehat{\boldsymbol{\omega}} \qquad \cong \qquad \dot{\boldsymbol{x}} = \tilde{A}(\boldsymbol{x})\boldsymbol{\xi} \tag{B.18}$$

The kinetic energy of the rigid body in terms of the chosen coordinates is

$$\mathcal{T} = \tfrac{1}{2}m\|\dot{\boldsymbol{r}}\|^2 + \tfrac{1}{2}\boldsymbol{\omega}^\top \Theta \boldsymbol{\omega} = \underbrace{\tfrac{1}{2}\boldsymbol{\xi}^\top \overbrace{(\Theta + m\widehat{p}^\top \widehat{p})}^{\tilde{\boldsymbol{M}}} \boldsymbol{\xi}}_{\tilde{\mathcal{T}}} - m(R\widehat{\boldsymbol{\xi}}p)^\top \eta + \tfrac{1}{2}m\|\eta\|^2 \tag{B.19}$$

potential energy from gravity

$$\mathcal{V} = -m\boldsymbol{a}_{\mathsf{G}}^\top \boldsymbol{r} \tag{B.20}$$

With these ingredients we can evaluate (B.10) to obtain the kinetic equation. Using $\partial_i p \xi^i = \dot{p}$ we can write them in a matrix vector form

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \tilde{\mathcal{T}}}{\partial \tilde{\xi}^i} + \tilde{\gamma}_{ij}^k \tilde{\xi}^j \frac{\partial \tilde{\mathcal{T}}}{\partial \xi^k} - \tilde{A}_i^\alpha \frac{\partial \tilde{\mathcal{T}}}{\partial x^\alpha}\right]_{i=1\dots3} = \tilde{\boldsymbol{M}}\dot{\boldsymbol{\xi}} + \big(\underbrace{m(\dot{\widehat{p}}^\top \widehat{p} + \widehat{p}^\top \dot{\widehat{p}})}_{\dot{\tilde{\boldsymbol{M}}}} + \widehat{\boldsymbol{\xi}}\tilde{\boldsymbol{M}}\big)\boldsymbol{\xi} \tag{B.21}$$

$$\boldsymbol{c}^{\mathrm{N}} = \left[\hat{\gamma}_{ij}^a \tilde{\xi}^j \frac{\partial \mathcal{T}}{\partial \eta^a}\Big|_{\eta=0}\right]_{i=1\dots3} = m\,\dot{\widehat{p}}\,\widehat{p}\,\boldsymbol{\xi} \tag{B.22}$$

$$\boldsymbol{f}^{\mathrm{V}} = \boldsymbol{\nabla}\mathcal{V} = m\widehat{p}^\top \boldsymbol{R}^\top \boldsymbol{a}_{\mathsf{G}} \tag{B.23}$$

Rolling ellipsoid: Constraint $\psi(p)$ for the ellipsoid surface (with $D = \mathrm{diag}(d_1^2, d_2^2, d_3^2)$) and the tangent condition yield

$$\psi(p) = p^\top D^{-1} p - 1 = 0, \quad \nabla\psi(p) \times R^\top e_3 = 0 \quad \Rightarrow \quad p = \frac{DR^\top e_3}{\sqrt{e_3^\top R D R^\top e_3}}. \tag{B.24}$$

Newton-Euler

$$m\ddot{r} = F^R + F^A, \qquad \Theta\dot{\omega} + \widehat{\omega}\Theta\omega = \widehat{p}R^\top F^R + \tau^A \tag{B.25}$$

The condition of rolling without slipping implies

$$\dot{r} + R\widehat{\omega}p = 0 \quad \Rightarrow \quad \ddot{r} = R(\widehat{p}\dot{\omega} - \widehat{\omega}^2 p - \widehat{\omega}\dot{p}) \tag{B.26}$$

where p are the coordinates of the contact point w.r.t. the body fixed frame. We can eliminate $\ddot{r}$ and the reaction force F^R from (B.25) to obtain (cf. [Borisov and Mamaev, 2002])

$$\tilde{\Theta}\dot{\omega} + \widehat{\omega}\tilde{\Theta}\omega + m\widehat{p}^\top \dot{\widehat{p}}\,\omega = \tau^A - \widehat{p}R^\top F^A \tag{B.27}$$

where $\tilde{\Theta} = \Theta + m\widehat{p}^\top \widehat{p}$.

B.1 Hamilton's principle

Hamilton's canonical equations. Define the *generalized momentum* $\boldsymbol{p}$ as

$$p_i = \frac{\partial \mathcal{L}}{\partial \xi^i}, \quad i = 1, \dots, n.. \tag{B.28}$$

and assume that this relation can be inverted to express the velocity $\boldsymbol{\xi} = \zeta(\boldsymbol{x}, \boldsymbol{p}, t)$ in terms of the momentum. Then define the *Hamiltonian* $\mathcal{H}$ as

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{p}, t) = \Big[p_i \xi^i - \mathcal{L}(\boldsymbol{x}, \boldsymbol{\xi}, t)\Big]_{\boldsymbol{\xi} = \zeta(\boldsymbol{x}, \boldsymbol{p}, t)} = p_i \zeta^i(\boldsymbol{x}, \boldsymbol{p}, t) - \mathcal{L}(\boldsymbol{x}, \zeta(\boldsymbol{x}, \boldsymbol{p}, t), t). \tag{B.29}$$

The definitions (B.28) and (B.29) describe the *Legendre transformation* $(\mathcal{L}, \boldsymbol{\xi}) \to (\mathcal{H}, \boldsymbol{p})$, see [Lanczos, 1986, ch. VI.1] or [Arnold, 1989, sec. 14] for some geometric background. Note that the configuration coordinates $\boldsymbol{x}$ and the time t do not participate in the transformation.

Evaluation of the differentials of (B.29), we get the relations

$$\partial_j \mathcal{H} = p_i \partial_j \zeta^i - \partial_j \mathcal{L} - \frac{\partial \mathcal{L}}{\partial \xi^i} \partial_j \zeta^i = -\partial_j \mathcal{L} \tag{B.30a}$$

$$\frac{\partial \mathcal{H}}{\partial p_j} = \zeta^j + p_i \frac{\partial \zeta^i}{\partial p_j} - \frac{\partial \mathcal{L}}{\partial \xi^i} \frac{\partial \zeta^i}{\partial p_j} = \xi^j \tag{B.30b}$$

$$\frac{\partial \mathcal{H}}{\partial t} = p_i \frac{\partial \zeta^i}{\partial t} - \frac{\partial \mathcal{L}}{\partial \xi^i} \frac{\partial \zeta^i}{\partial t} - \frac{\partial \mathcal{L}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}, \tag{B.30c}$$

With this we can express the kinematic equation (**??**) and Lagrange's equation (**??**) in terms of the generalized momentum $\boldsymbol{p}$ and the Hamiltonian $\mathcal{H}$.

$$\dot{x}^\alpha = A_i^\alpha \frac{\partial \mathcal{H}}{\partial p_i}, \qquad \dot{p}_i = f_i^{\mathrm{E}} - \gamma_{ij}^k \frac{\partial \mathcal{H}}{\partial p_j} p_k - A_i^\alpha \frac{\partial \mathcal{H}}{\partial x^\alpha} \tag{B.31}$$

In matrix notation they can be combined as

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{p} \end{bmatrix} = \begin{bmatrix} 0 & \boldsymbol{A} \\ -\boldsymbol{A}^\top & G \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{H}}{\partial \boldsymbol{x}} \\ \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}} \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{f}^{\mathrm{E}} \end{bmatrix} \tag{B.32}$$

with the skew symmetric matrix $G_{ij}(\boldsymbol{x}, \boldsymbol{p}) = -\gamma_{ij}^k(\boldsymbol{x}) p_k = -G_{ji}(\boldsymbol{x}, \boldsymbol{p})$. For the special case of minimal configuration coordinates $\boldsymbol{q}$ and velocity coordinates $\boldsymbol{\xi} = \dot{\boldsymbol{q}}$ we have $\boldsymbol{A} = \mathbf{I}_n$ and $G = 0$ and (B.32) is called *Hamilton's canonical equations*.

A general conservation law. The time derivative of the Hamiltonian along the motion is

$$\begin{aligned} \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} &= \frac{\partial \mathcal{H}}{\partial x^\alpha} \dot{x}^\alpha + \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i + \frac{\partial \mathcal{H}}{\partial t} \\ &= \underbrace{\frac{\partial \mathcal{H}}{\partial x^\alpha} A_i^\alpha \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial p_i} A_i^\alpha \frac{\partial \mathcal{H}}{\partial x^\alpha}}_{0} + p_k \underbrace{\gamma_{ij}^k \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial \mathcal{H}}{\partial p_j}}_{0} + \frac{\partial \mathcal{H}}{\partial p_i} f_i^{\mathrm{E}} + \frac{\partial \mathcal{H}}{\partial t}. \end{aligned} \tag{B.33}$$

In terms of the original coordinates this is

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \xi^i f_i^{\mathrm{E}} - \frac{\partial \mathcal{L}}{\partial t}. \tag{B.34}$$

This is the well known conservation law for the Hamiltonian (e.g. [Lanczos, 1986, ch. VI.6]). The remarkable aspect of the conservation law (and for the Legendre transformation) is that there is no particular assumption on the structure of the Lagrangian $\mathcal{L}$.

However for our case of systems of mechanical particles the Lagrangian is $\mathcal{L} = \frac{1}{2} \boldsymbol{\xi}^\top \boldsymbol{M} \boldsymbol{\xi} - \mathcal{V}$, so the generalized momentum is $\boldsymbol{p} = \boldsymbol{M} \boldsymbol{\xi}$ and the Hamiltonian is the total energy $\mathcal{H} = \frac{1}{2} \boldsymbol{\xi}^\top \boldsymbol{M} \boldsymbol{\xi} + \mathcal{V}$.

Conclusion. The redundant configuration coordinates and velocity coordinates do behave well in the context of Hamiltonian mechanics, tanking into account the considerations from **??**. As before, there is no new physical insight, but the formulations allow a more sophisticated parameterization that might be useful for other fields of physics or optimal control.

Example. Let the Lagrangian have the form

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\xi}, t) = \tfrac{1}{2} M_{ij}(\boldsymbol{x}, t) \xi^j \xi^i + b_i(\boldsymbol{x}, t) \xi^i + \mathcal{L}_0(\boldsymbol{x}, t) \tag{B.35}$$

The Euler-Lagrange equation evaluates to

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \xi^i} + \gamma_{ij}^k \xi^j \frac{\partial \mathcal{L}}{\partial \xi^k} - \partial_i \mathcal{L} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(M_{ij}\xi^j + b_i\big) + \gamma_{ij}^k \xi^j \big(M_{kj}\xi^j + b_k\big) - \tfrac{1}{2}\partial_i M_{kj}\xi^j \xi^k - \partial_i b_i \xi^i - \partial_i \mathcal{L}_0 \\
&= M_{ij}\dot{\xi}^j + \partial_k M_{ij}\xi^k\xi^j + \frac{\partial M_{ij}}{\partial t}\xi^j + \frac{\partial b_i}{\partial t} + \gamma_{ij}^k \xi^j \big(M_{kj}\xi^j + b_k\big) - \tfrac{1}{2}\partial_i M_{kj}\xi^j\xi^k - \partial_i \mathcal{L}_0 \\
&= M_{ij}\dot{\xi}^j + \varGamma_{ijk}\xi^k\xi^j + \frac{\partial M_{ij}}{\partial t}\xi^j + \frac{\partial b_i}{\partial t} + \gamma_{ij}^k \xi^j b_k - \partial_i \mathcal{L}_0
\end{aligned} \tag{B.36}$$

generalized momentum

$$p_i = \frac{\partial \mathcal{L}}{\partial \xi^i} = M_{ij}\xi^i + b_i \qquad \Leftrightarrow \qquad \xi^i = M^{ij}(p_j - b_j) \tag{B.37}$$

Hamiltonian

$$\mathcal{H} = \xi^i(M_{ij}\xi^i + b_i) - \mathcal{L} = \tfrac{1}{2}M_{ij}\xi^j\xi^i - \mathcal{L}_0 \tag{B.38}$$

B.2 On error coordinates

Error coordinates. Introduce (possibly redundant) error coordinates $\boldsymbol{e} \in \mathbb{R}^{\nu_e}$ as

$$\boldsymbol{e} = \boldsymbol{\chi}_{\mathrm{E}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}), \qquad \boldsymbol{\phi}_{\mathrm{E}}(\boldsymbol{e}) = 0. \tag{B.39}$$

and require that this relation is invertible with $\boldsymbol{x} = \boldsymbol{\chi}(\boldsymbol{e}, \boldsymbol{x}_{\mathsf{R}})$, i.e. $\boldsymbol{\chi}(\boldsymbol{\chi}_{\mathrm{E}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}), \boldsymbol{x}_{\mathsf{R}}) = \boldsymbol{x}\,\forall\,\boldsymbol{x} \in \mathbb{X}$. The inverse function theorem now implies that the differential $\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}} = \frac{\partial \boldsymbol{\chi}_{\mathrm{E}}}{\partial \boldsymbol{x}}\boldsymbol{A}$ has full rank: $\operatorname{rank}(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}}) = \dim \mathbb{X} = n$.

Let $\boldsymbol{\varPhi}_{\mathrm{E}}$ be the linear independent rows of $\partial\boldsymbol{\phi}_{\mathrm{E}}/\partial\boldsymbol{e}$. Then the derivative of the geometric constraint $\dot{\boldsymbol{\phi}}_{\mathrm{E}} = 0$ implies

$$\boldsymbol{\varPhi}_{\mathrm{E}}\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}} = 0, \qquad \boldsymbol{\varPhi}_{\mathrm{E}}\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}} = 0, \tag{B.40}$$

Since the matrices $\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}}$ and $\boldsymbol{\varPhi}_{\mathrm{E}}$ have full rank, their pseudo-inverses are

$$(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+ = \big((\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^\top(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})\big)^{-1}(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^\top, \qquad (\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}}) = \mathbf{I}_n \tag{B.41}$$

$$\boldsymbol{\varPhi}_{\mathrm{E}}^+ = \boldsymbol{\varPhi}_{\mathrm{E}}^\top\big(\boldsymbol{\varPhi}_{\mathrm{E}}\boldsymbol{\varPhi}_{\mathrm{E}}^\top\big)^{-1}, \qquad \boldsymbol{\varPhi}_{\mathrm{E}}\boldsymbol{\varPhi}_{\mathrm{E}}^+ = \mathbf{I}_{\nu_e - n}. \tag{B.42}$$

Furthermore, due to the orthogonality $\boldsymbol{\varPhi}_{\mathrm{E}}\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}} = 0$ we have

$$(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+ + \boldsymbol{\varPhi}_{\mathrm{E}}^+\boldsymbol{\varPhi}_{\mathrm{E}} = \mathbf{I}_{\nu_e}. \tag{B.43}$$

Error potential. We require that the potential $\bar{\mathcal{V}}$ can be expressed as a function $\bar{\mathcal{V}}_{\mathrm{E}}$ of the error coordinates $\boldsymbol{e}$ alone, i.e.

$$\bar{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}}) = \bar{\mathcal{V}}_{\mathrm{E}}(\boldsymbol{\chi}_{\mathrm{E}}(\boldsymbol{x}, \boldsymbol{x}_{\mathsf{R}})). \tag{B.44}$$

Now the requirement (5.43) for the transport map $\boldsymbol{Q}$ can be written as

$$\begin{aligned}
\boldsymbol{\nabla}_{\mathsf{R}}\bar{\mathcal{V}} + \boldsymbol{Q}^\top \boldsymbol{\nabla}\bar{\mathcal{V}} &= \big(\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}} + \boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}}\boldsymbol{Q}\big)^\top \frac{\partial \bar{\mathcal{V}}_{\mathrm{E}}}{\partial \boldsymbol{e}} \\
&= \Big(\underbrace{\big((\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+ + \boldsymbol{\Phi}_{\mathrm{E}}^+\boldsymbol{\Phi}_{\mathrm{E}}\big)}_{\mathbf{I}_{\nu_e}} \boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}} + \boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}}\boldsymbol{Q}\Big)^\top \frac{\partial \bar{\mathcal{V}}_{\mathrm{E}}}{\partial \boldsymbol{e}} \\
&= \big((\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+(\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}}) + \boldsymbol{Q}\big)^\top (\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^\top \frac{\partial \bar{\mathcal{V}}_{\mathrm{E}}}{\partial \boldsymbol{e}} + \boldsymbol{\Phi}_{\mathrm{E}}^+ \underbrace{\boldsymbol{\Phi}_{\mathrm{E}}\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}}}_{0} \frac{\partial \bar{\mathcal{V}}_{\mathrm{E}}}{\partial \boldsymbol{e}} = 0
\end{aligned} \tag{B.45}$$

which has the simple solution

$$\boldsymbol{Q} = -(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+(\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}}). \tag{B.46}$$

Error kinematics. With the same approach as above we can derive a kinematic relation between the error coordinates $\boldsymbol{e}$ and the error velocity $\boldsymbol{\xi}_{\mathrm{E}} = \boldsymbol{\xi} - \boldsymbol{Q}\boldsymbol{\xi}_{\mathsf{R}}$ as

$$\begin{aligned}
\dot{\boldsymbol{e}} &= (\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})\boldsymbol{\xi} + (\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}})\boldsymbol{\xi}_{\mathsf{R}} \\
&= (\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})\boldsymbol{\xi} + \underbrace{\big((\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})(\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+ + \boldsymbol{\Phi}_{\mathrm{E}}^+\boldsymbol{\Phi}_{\mathrm{E}}\big)}_{\mathbf{I}_{\nu_e}} \boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}}\boldsymbol{\xi}_{\mathsf{R}} \\
&= (\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}}) \underbrace{\big(\boldsymbol{\xi} + (\boldsymbol{\nabla}\boldsymbol{\chi}_{\mathrm{E}})^+(\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}})\boldsymbol{\xi}_{\mathsf{R}}\big)}_{\boldsymbol{\xi}_{\mathrm{E}}} + \boldsymbol{\Phi}_{\mathrm{E}}^+ \underbrace{\boldsymbol{\Phi}_{\mathrm{E}}(\boldsymbol{\nabla}_{\mathsf{R}}\boldsymbol{\chi}_{\mathrm{E}})}_{0} \boldsymbol{\xi}_{\mathsf{R}}
\end{aligned} \tag{B.47}$$

B.3 A possible generalization of the rigid body energies

Note that any symmetric, positive matrix $\mathbf{K} \in \mathbb{SYM}^+(6)$ can be decomposed into

$$\mathbf{K} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \\ \mathbf{H} & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} \mathbf{K}_{\mathrm{r}} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{\mathrm{R}} \end{bmatrix} \begin{bmatrix} \mathbf{I}_3 & \mathbf{H}^\top \\ \mathbf{0} & \mathbf{I}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{\mathrm{r}} & \mathbf{K}_{\mathrm{r}}\mathbf{H}^\top \\ \mathbf{H}\mathbf{K}_{\mathrm{r}} & \mathbf{K}_{\mathrm{R}} + \mathbf{H}\mathbf{K}_{\mathrm{r}}\mathbf{H}^\top \end{bmatrix} \tag{B.48}$$

where $\mathbf{K}_{\mathrm{r}}, \mathbf{K}_{\mathrm{R}} \in \mathbb{SYM}^+(3)$ and $\mathbf{H} \in \mathbb{R}^{3\times 3}$. For variables $\mathbf{x}_i \in \mathbb{R}^3$ and $\mathbf{Y}_i \in \mathbb{R}^{3\times 3}$ collected in the square matrix $\boldsymbol{\Xi}_i = \left[\begin{smallmatrix} \mathbf{Y}_i & \mathbf{x}_i \\ \mathbf{0} & 0 \end{smallmatrix}\right]$ define the inner product as

$$\langle \boldsymbol{\Xi}_1, \boldsymbol{\Xi}_2 \rangle_{\mathbf{K}} = \tilde{\mathbf{x}}_1^\top \mathbf{K}_{\mathrm{r}} \tilde{\mathbf{x}}_2 + \mathrm{tr}(\mathbf{Y}_1 \,\mathrm{Vee}(\mathbf{K}_{\mathrm{R}}) \mathbf{Y}_2^\top), \tag{B.49}$$

where $\tilde{\mathbf{x}}_i = \mathbf{x}_i + \frac{1}{2}\mathbf{Y}_i \,\mathrm{vee2}(\mathbf{H}) + \frac{1}{4}(\mathbf{H} + \mathbf{H}^\top)\,\mathrm{vee2}(\mathbf{Y}_i)$. To prove that this is indeed an inner product one can use the same argument as in subsection 4.2.2 and note that positive definiteness of $\mathbf{K}_{\mathrm{R}}$ implies positive definiteness of $\mathrm{Vee}(\mathbf{K}_{\mathrm{R}})$. For the special parameters $\mathbf{K}_{\mathrm{r}} = k\mathbf{I}_3$ and $\mathbf{H} = \mathrm{wed}(\boldsymbol{h}), \boldsymbol{h} \in \mathbb{R}^3$ this inner product coincides with the definition of (A.8). For the special arguments $\mathbf{Y}_i = \mathrm{wed}(\mathbf{y}_i), \mathbf{y}_i \in \mathbb{R}^3$ we have the following simplifications $\tilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{H}^\top \mathbf{y}_i$ and $\mathrm{tr}(\mathrm{wed}(\mathbf{y}_1)\,\mathrm{Vee}(\mathbf{K}_{\mathrm{R}})\,\mathrm{wed}(\mathbf{y}_2)^\top) = \mathbf{y}_1^\top \mathbf{K}_{\mathrm{R}} \mathbf{y}_2$. So, for the case $\boldsymbol{\Xi}_i = \mathrm{wed}(\boldsymbol{\xi}_i), \boldsymbol{\xi}_i = [\mathbf{x}_i^\top, \mathbf{y}_i^\top]^\top$ the inner product (B.49) simplifies to

$$\langle \mathrm{wed}(\boldsymbol{\xi}_1), \mathrm{wed}(\boldsymbol{\xi}_2) \rangle_{\mathbf{K}} = \boldsymbol{\xi}_1^\top \mathbf{K} \boldsymbol{\xi}_2. \tag{B.50}$$

For any inner product we may define an induced norm and metric as

$$\|\boldsymbol{\Xi}\|_{\mathbf{K}} = \sqrt{\langle \boldsymbol{\Xi}, \boldsymbol{\Xi} \rangle_{\mathbf{K}}}, \qquad d_{\mathbf{K}}(\boldsymbol{\Xi}_1, \boldsymbol{\Xi}_2) = \|\boldsymbol{\Xi}_1 - \boldsymbol{\Xi}_2\|_{\mathbf{K}}. \tag{B.51}$$

Note that the same inner product, norm and metric can also be defined for elements of the form $\boldsymbol{\Xi}_i = \left[\begin{smallmatrix} \mathbf{Y}_i & \mathbf{x}_i \\ \mathbf{0} & 1 \end{smallmatrix}\right]$.

With this we may define the following energies as the square of the metrics

$$\mathcal{V} = \tfrac{1}{2} d^2_{\mathbf{K}}(\mathbf{G}, \mathbf{G}_{\mathsf{R}}), \qquad \mathbf{K} \in \mathbb{SYM}^+(6), \tag{B.52a}$$
$$\mathcal{R} = \tfrac{1}{2} d^2_{\mathbf{D}}(\dot{\mathbf{G}}, \dot{\mathbf{G}}_{\mathsf{R}}), \qquad \mathbf{D} \in \mathbb{SYM}^+(6), \tag{B.52b}$$
$$\mathcal{T} = \tfrac{1}{2} d^2_{\mathbf{M}}(\ddot{\mathbf{G}}, \ddot{\mathbf{G}}_{\mathsf{R}}), \qquad \mathbf{M} \in \mathbb{SYM}^+(6). \tag{B.52c}$$

For the special parameters $\mathbf{K}_{\mathrm{r}} = k\mathbf{I}_3$ and $\mathbf{H} = \mathrm{wed}(\boldsymbol{h}), \boldsymbol{h} \in \mathbb{R}^3$ this coincides with the energies defined in (5.11). Alternatively, using $\mathbf{G}_{\mathrm{E}} = \mathbf{G}_{\mathsf{R}}^{-1}\mathbf{G}$, we may define

$$\mathcal{V} = \tfrac{1}{2} d^2_{\mathbf{K}}(\mathbf{G}_{\mathrm{E}}, \mathbf{I}_4), \quad \mathbf{K} \in \mathbb{SYM}^+(6), \tag{B.53a}$$
$$\mathcal{R} = \tfrac{1}{2} d^2_{\mathbf{D}}(\dot{\mathbf{G}}_{\mathrm{E}}, \mathbf{0}), \quad \mathbf{D} \in \mathbb{SYM}^+(6), \tag{B.53b}$$
$$\mathcal{T} = \tfrac{1}{2} d^2_{\mathbf{M}}(\ddot{\mathbf{G}}_{\mathrm{E}}, \mathbf{0}), \quad \mathbf{M} \in \mathbb{SYM}^+(6), \tag{B.53c}$$

For the special parameters mentioned above, this coincides with the energies defined in (5.20).

The inner product and its induced norm and metric can be regarded as a generalization of the inner product (with 10 parameters) proposed in subsection 4.2.2 in the sense that it incorporates all 21 independent coefficients of the symmetric matrix $\mathbf{K} \in \mathbb{SYM}^+(6)$. This results in more tuning parameters for the control design. On the downside, the translation invariance (A.15) and the physical interpretation of the parameters are lost for the general case.

Bibliography

[Abraham and Marsden, 1978] Abraham, R. and Marsden, J. E. (1978). *Foundations of Mechanics*. Westview Press.

[Appell, 1900] Appell, P. E. (1900). Sur une forme générale des équations de la dynamique. *J. Reine Angew. Math.*, 121:310–319.

[Arnold, 1989] Arnold, V. I. (1989). *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer, New York.

[Arnold et al., 2002] Arnold, V. I., Kozlov, V. V., and Neishtadt, A. I. (2002). *Mathematical Aspects of Classical and Celestial Mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer.

[Bloch, 2003] Bloch, A. M. (2003). *Nonholonomic Mechanics and Control*. Springer.

[Bloch et al., 1996] Bloch, A. M., Krishnaprasad, P. S., Marsden, J. E., and Murray, R. M. (1996). Nonholonomic Mechanical Systems with Symmetry. *Archive for Rational Mechanics and Analysis*, 136(1):21–99.

[Bloch et al., 2000] Bloch, A. M., Leonard, N. E., and Marsden, J. E. (2000). Controlled Lagrangians and the stabilization of mechanical systems. I. The first matching theorem. *IEEE Trans. on Automatic Control*, 45(12):2253–2270.

[Boltzmann, 1897] Boltzmann, L. (1897). *Vorlesungen ueber die Pincipe der Mechanik*. Johan Ambrosius Barth.

[Boltzmann, 1902] Boltzmann, L. (1902). Über die Form der Lagrangeschen Gleichungen für nichtholonome, generalisierte Koordinaten. In *Wissenschaftliche Abhandlungen*, volume 3, pages 682–692. Johann Ambrosius Barth.

[Boothby, 1986] Boothby, W. M. (1986). *An introduction to differentiable manifolds and Riemannian geometry*. Academic Press.

[Borisov and Mamaev, 2002] Borisov, A. V. and Mamaev, I. S. (2002). The rolling motion of a rigid body on a plane and sphere: Hierarchy of dynamics. *Regular and Chaotic Dynamics*, 7(2):177–200.

[Bremer, 2008] Bremer, H. (2008). *Elastic Multibody Dynamics: A Direct Ritz Approach*. Springer.

[Bullo and Lewis, 2004] Bullo, F. and Lewis, A. (2004). *Geometric Control of Mechanical Systems*. Texts in Applied Mathematics. Springer.

[Bullo and Murray, 1999] Bullo, F. and Murray, R. M. (1999). Tracking for fully actuated mechanical systems: A geometric framework. *Automatica*, 35(1):17–34.

[Courant and Hilbert, 1924] Courant, R. and Hilbert, D. (1924). *Methoden der Mathematischen Physik*. Julius Springer.

[Desloge, 1987] Desloge, E. A. (1987). Relationship between Kane's equations and the Gibbs-Appell equations. *Journal of Guidance, Control, and Dynamics*, 10(1):120–122.

[Dontchev and Rockafellar, 2010] Dontchev, A. L. and Rockafellar, R. T. (2010). *Implicit Functions and Solution Mappings*. Springer, 2nd edition.

[Fliess et al., 1999] Fliess, M., Levine, J., Martin, P., and Rouchon, P. (1999). A Lie-Bäcklund approach to equivalence and flatness of nonlinear systems. *IEEE Trans. Automat. Contr.*, 44(5):922–937.

[Fliess et al., 1995] Fliess, M., L´evine, J., and Rouchon, P. M. M. (1995). Flatness and defect of non-linear systems: introductory theory and examples. *International Journal of Control*, 61(6):1327–1361.

[Frankel, 1997] Frankel, T. (1997). *The Geometry of Physics*. Cambridge University Press.

[Galilei, 1638] Galilei, G. (1638). *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*.

[Gauß, 1829] Gauß, C. F. (1829). Über ein neues allgemeines Grundgesetz der Mechanik. *Journal für die reine und angewandte Mathematik*.

[Gibbs, 1879] Gibbs, J. W. (1879). On the fundamental formulae of dynamics. *Am. J. Math.*, 2(1):49–64.

[Goldstein, 1951] Goldstein, H. (1951). *Classical Mechanics*. Addison-Wesley.

[Gould et al., 2001] Gould, N., Hribar, M., and Nocedal, J. (2001). On the solution of equality constrained quadratic programming problems arising in optimization. *SIAM Journal on Scientific Computing*, 23(4):1376–1395.

[Grafarend and Kühnel, 2011] Grafarend, E. W. and Kühnel, W. (2011). A minimal atlas for the rotation group SO(3). *GEM-International Journal on Geomathematics*, 2(1):113–122.

[Güler, 2010] Güler, O. (2010). *Foundations of Optimization*, volume 258 of *Graduate Texts in Mathematics*. Springer.

[Hall, 2015] Hall, B. (2015). *Lie Groups, Lie Algebras, and Representations*, volume 222 of *Graduate Texts in Mathematics*. Springer, 2nd edition.

[Hamel, 1904] Hamel, G. (1904). Die Lagrange-Eulerschen Gleichungen der Mechanik. *Z. für Mathematik u. Physik*, 50:1–57.

[Hamel, 1935] Hamel, G. (1935). Das Hamiltonsche Prinzip bei nichtholonomen Systemen. *Mathematische Annalen*, 111:94–97.

[Hamel, 1949] Hamel, G. (1949). *Theoretische Mechanik*. Springer.

[Hauser et al., 1992] Hauser, J., Sastry, S., and Meyer, G. (1992). Nonlinear control design for slightly non-minimum phase systems: Application to V/STOL aircraft. *Automatica*, 28(4):664–679.

[Hooke, 1678] Hooke, R. (1678). *Lectures De Potentia Restitutiva, or of Spring. Explaining the Power of Springing Bodies.* Printed for John Martyn Printer to the Royal Society.

[Horn, 1987] Horn, B. K. P. (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, 4(4):629–642.

[Horn and Johnson, 1985] Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press.

[Kane and Levinson, 1985] Kane, T. R. and Levinson, D. A. (1985). *Dynamics*. McGraw Hill.

[Koditschek, 1989] Koditschek, D. E. (1989). The Application of Total Energy as a Lyapunov Function for Mechanical Control Systems. In Marsden, J. E., Krishnaprasad, P. S., and Simo, J. C., editors, *Dynamics and Control of Multibody Systems*, volume 97 of *Contemporary Mathematics*, pages 131–157. American Mathematical Society.

[Konz and Rudolph, 2013] Konz, M. and Rudolph, J. (2013). Quadrotor tracking control based on a moving frame. In *Proc. 9th IFAC Symposium on Nonlinear Control Systems*, pages 80–85.

[Konz and Rudolph, 2016] Konz, M. and Rudolph, J. (2016). Beispiele für einen direkten Zugang zu einer globalen, energiebasierten Modellbildung und Regelung von Starrkörpersystemen. *at - Automatisierungstechnik*, 64(2):96–109.

[Korteweg, 1899] Korteweg, D. J. (1899). Über eine ziemlich verbreitete unrichtige Behandlungsweise eines Problems der rollenden Bewegung, über die Theorie dieser Bewegung, und ins besondere über kleine rollende Schwingungen um eine Gleichgewichtslage. *Nieuw Archiefvoor Wiskunde*, 4:130–155.

[Lanczos, 1986] Lanczos, C. (1986). *The Variational Principles of Mechanics*. Dover Publications, 4 edition.

[Landau and Lifshitz, 1960] Landau, L. D. and Lifshitz, E. M. (1960). *Mechanics*, volume 1 of *Course of Theoretical Physics*. Pergamon Press.

[Lee, 2003] Lee, J. M. (2003). *Introduction to Smooth Manifolds*. Springer.

[Lee et al., 2010] Lee, T., Leok, M., and McClamroch, H. (2010). Control of complex maneuvers for a quadrotor UAV using geometric methods on SE(3). *ArXiv e-prints*. http://arxiv.org/abs/1003.2005v4.

[Lesser, 1992] Lesser, M. (1992). A geometric interpretation of kane's equations. *Proceedings of the Royal Society of London*, 436(1896):69–87.

[Lewis and Murray, 1995] Lewis, A. and Murray, R. M. (1995). Variational principles in constrained systems: theory and experiments. *Int. J. Nonlinear Mech.*, 30:793–815.

[Lewis, 1996] Lewis, A. D. (1996). The geometry of the Gibbs–Appell equations and Gauss's principle of least constraint. *Reports on Mathematical Physics*, 38(1):11–28.

[Luenberger and Ye, 2015] Luenberger, D. G. and Ye, Y. (2015). *Linear and Nonlinear Programming*. Springer, 4th edition.

[Lurie, 2002] Lurie, A. I. (2002). *Analytical Mechanics*. Springer.

[Markley, 1988] Markley, F. L. (1988). Attitude determination using vector observations and the singular value decomposition. *Journal of the Astronautical Sciences*, 38:245–258.

[Marsden and Ratiu, 1998] Marsden, J. E. and Ratiu, T. S. (1998). *Introduction to Mechanics and Symmetry*. Springer, Berlin Heidelberg, 2nd edition.

[Martin et al., 1997] Martin, P., Murray, R. M., and Rouchon, P. (1997). Flat systems, equivalence and trajectory generation. In *Mini-course ECC'97*.

[Misner et al., 1973] Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation*. W. H. Freeman and Company.

[Murray et al., 1994] Murray, R. M., Li, Z., and Sastry, S. S. (1994). *A Mathematical Introduction to Robotic Manipulation*. CRC Press.

[Murray et al., 1995] Murray, R. M., Rathinam, M., and Sluis, W. (1995). Differential flatness of mechanical control systems: A catalog of prototype systems. In *Proceedings of the ASME International Congressand Exposition*.

[Neimark and Fufaev, 1972] Neimark, J. I. and Fufaev, N. A. (1972). *Dynamics of Nonholonomic Systems*. Amer. Mathem. Society.

[Nocedal and Wright, 2006] Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer, 2nd edition.

[Olfati-Saber, 2001] Olfati-Saber, R. (2001). *Nonlinear Control of Underactuated Mechanical Systems with Applications to Robotics and Aerospace Vehicles*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

[Papastavridis, 2002] Papastavridis, J. G. (2002). *Analytical Mechanics*. World Scientific.

[Penrose, 1955] Penrose, R. (1955). A generalized inverse for matrices. *Mathematical proceedings of the Cambridge philosophical society*, 51(3):406–413.

[Pfeiffer and Glocker, 1996] Pfeiffer, F. and Glocker, C. (1996). *Multibody Dynamics with unilateral contacts*. Wiley-VCH.

[Poincaré, 1885] Poincaré, H. (1885). Sur les courbes définies par les équations différentielles. *Journal de Mathématiques Pures et Appliquées*, 4:167–244.

[Päsler, 1968] Päsler, M. (1968). *Prinzipe der Mechanik*. Walter de Gruyter & Co.

[Rathinam and Murray, 1998] Rathinam, M. and Murray, R. M. (1998). Configuration flatness of lagrangian systems underactuated by one control. *SIAM Journal of Control and Optimization*, 36(1):164–179.

[Rayleigh, 1877] Rayleigh, J. W. S. B. (1877). *The Theory of Sound*. Macmillan and Co.

[Roberson and Schwertassek, 1988] Roberson, R. E. and Schwertassek, R. (1988). *Dynamics of Multibody Systems*. Springer.

[Schlacher and Schöberl, 2007] Schlacher, K. and Schöberl, M. (2007). Construction of flat outputs by reduction and elimination. *CD Proceedings IFAC Symposium on Nonlinear Control Systems*, pages 666–671.

[Shabana, 2005] Shabana, A. A. (2005). *Dynamics of Multibody Systems*. Cambridge University Press, 3 edition.

[Slotine and Li, 1991] Slotine, J.-J. and Li, W. (1991). *Applied Nonlinear Control*. Springer.

[Spivak, 1999] Spivak, M. (1999). *A Comprehensive Introduction to Differential Geometry*. Publish or Perish.

[Spong et al., 2006] Spong, M. W., Hutchinson, S., and Vidyasagar, M. (2006). *Robot Modeling and Control*. Wiley.

[Szabó, 1956] Szabó, I. (1956). *Höhere Technische Mechanik*. Springer.

[Udwadia and Kalaba, 2002] Udwadia, F. E. and Kalaba, R. E. (2002). What is the general form of the explicit equations of motion for constrained mechanical systems? *Journal of Applied Mechanics*, 69:335–339.

[Wahba, 1965] Wahba, G. (1965). A least squares estimate of satellite attitude. *SIAM Review*, 7(3):409.

[Wittenburg, 2008] Wittenburg, J. (2008). *Dynamics of Multibody Systems*. Springer, 2 edition.

Appendix C

Templates

C.1 Math fonts

Latin alphabet in math mode

default	$A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z$ $a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z$
mathrm	$\mathrm{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$ $\mathrm{a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z}$
mathsf	$\mathsf{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$ $\mathsf{a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z}$
mathtt	$\mathtt{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$ $\mathtt{a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z}$
boldsymbol	$\boldsymbol{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$ $\boldsymbol{a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z}$
mathbf	$\mathbf{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$ $\mathbf{a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z}$
mathfrak	$\mathfrak{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$ $\mathfrak{a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z}$
mathcal	$\mathcal{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$
mathbb	$\mathbb{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z}$

Greek alphabet in math mode

default			Γ	Δ				Θ			Λ			Ξ		Π		Σ		Υ	Φ		Ψ	Ω
	α	β	γ	δ	ϵ	ζ	η	θ	ι	κ	λ	μ	ν	ξ	o	π	ρ	σ	τ	υ	ϕ	χ	ψ	ω
var	A	B	$\varGamma$	$\varDelta$	E	Z	H	$\varTheta$	I	K	$\varLambda$	M	N	$\varXi$	O	$\varPi$	P	$\varSigma$	T	$\varUpsilon$	$\varPhi$	X	$\varPsi$	$\varOmega$
					ε			ϑ								ϖ	ϱ	ς			φ			
boldsymbol			$\boldsymbol{\Gamma}$	$\boldsymbol{\Delta}$				$\boldsymbol{\Theta}$			$\boldsymbol{\Lambda}$			$\boldsymbol{\Xi}$		$\boldsymbol{\Pi}$		$\boldsymbol{\Sigma}$		$\boldsymbol{\Upsilon}$	$\boldsymbol{\Phi}$		$\boldsymbol{\Psi}$	$\boldsymbol{\Omega}$
	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\delta}$	$\boldsymbol{\epsilon}$	$\boldsymbol{\zeta}$	$\boldsymbol{\eta}$	$\boldsymbol{\theta}$	$\boldsymbol{\iota}$	$\boldsymbol{\kappa}$	$\boldsymbol{\lambda}$	$\boldsymbol{\mu}$	$\boldsymbol{\nu}$	$\boldsymbol{\xi}$	$\boldsymbol{o}$	$\boldsymbol{\pi}$	$\boldsymbol{\rho}$	$\boldsymbol{\sigma}$	$\boldsymbol{\tau}$	$\boldsymbol{\upsilon}$	$\boldsymbol{\phi}$	$\boldsymbol{\chi}$	$\boldsymbol{\psi}$	$\boldsymbol{\omega}$
var	$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{\varGamma}$	$\boldsymbol{\varDelta}$	$\boldsymbol{E}$	$\boldsymbol{Z}$	$\boldsymbol{H}$	$\boldsymbol{\varTheta}$	$\boldsymbol{I}$	$\boldsymbol{K}$	$\boldsymbol{\varLambda}$	$\boldsymbol{M}$	$\boldsymbol{N}$	$\boldsymbol{\varXi}$	$\boldsymbol{O}$	$\boldsymbol{\varPi}$	$\boldsymbol{P}$	$\boldsymbol{\varSigma}$	$\boldsymbol{T}$	$\boldsymbol{\varUpsilon}$	$\boldsymbol{\varPhi}$	$\boldsymbol{X}$	$\boldsymbol{\varPsi}$	$\boldsymbol{\varOmega}$
					$\boldsymbol{\varepsilon}$			$\boldsymbol{\vartheta}$								$\boldsymbol{\varpi}$	$\boldsymbol{\varrho}$	$\boldsymbol{\varsigma}$			$\boldsymbol{\varphi}$			
mathsf	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Μ	Ν	Ξ	Ο	Π	Ρ	Σ	Τ	ϒ	Φ	Χ	Ψ	Ω
	α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο	π	ρ	σ	τ	υ	φ	χ	ψ	ω
mathbf	$\mathbf{A}$	$\mathbf{B}$	$\mathbf{\Gamma}$	$\mathbf{\Delta}$	$\mathbf{E}$	$\mathbf{Z}$	$\mathbf{H}$	$\mathbf{\Theta}$	$\mathbf{I}$	$\mathbf{K}$	$\mathbf{\Lambda}$	$\mathbf{M}$	$\mathbf{N}$	$\mathbf{\Xi}$	$\mathbf{O}$	$\mathbf{\Pi}$	$\mathbf{P}$	$\mathbf{\Sigma}$	$\mathbf{T}$	$\mathbf{\Upsilon}$	$\mathbf{\Phi}$	$\mathbf{X}$	$\mathbf{\Psi}$	$\mathbf{\Omega}$

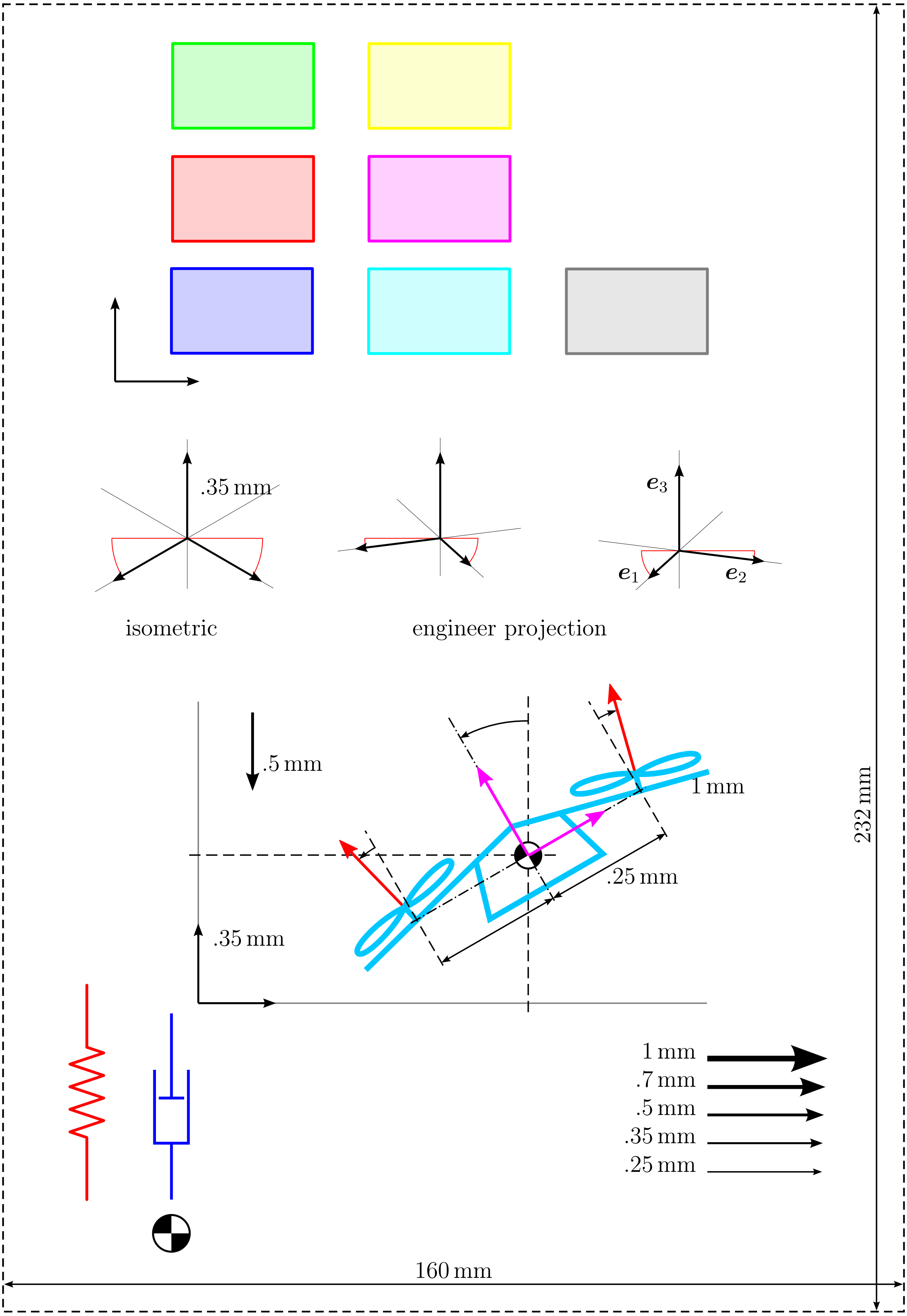


Figure C.1: Inkscape figure template