Output Optimization by Lie Bracket Approximations

Dissertation zur Erlangung
des naturwissenschaftlichen Doktorgrades
der Julius-Maximilians-Universität Würzburg

vorgelegt von

Raik Suttner

Würzburg, 2020

Abstract

In this dissertation, we develop and analyze novel optimizing feedback laws for control-affine systems with real-valued state-dependent output (or objective) functions. Given a control-affine system, our goal is to derive an output-feedback law that asymptotically stabilizes the closed-loop system around states at which the output function attains a minimum value. The control strategy has to be designed in such a way that an implementation only requires real-time measurements of the output value. Additional information, like the current system state or the gradient vector of the output function, is not assumed to be known. A method that meets all these criteria is called an *extremum seeking control law*. We follow a recently established approach to extremum seeking control, which is based on approximations of Lie brackets. For this purpose, the measured output is modulated by suitable highly oscillatory signals and is then fed back into the system. Averaging techniques for control-affine systems with highly oscillatory inputs reveal that the closed-loop system is driven, at least approximately, into the directions of certain Lie brackets. A suitable design of the control law ensures that these Lie brackets point into descent directions of the output function. Under suitable assumptions, this method leads to the effect that minima of the output function are practically uniformly asymptotically stable for the closed-loop system. The present document extends and improves this approach in various ways.

One of the novelties is a control strategy that does not only lead to practical asymptotic stability, but in fact to asymptotic and even exponential stability. In this context, we focus on the application of distance-based formation control in autonomous multi-agent system in which only distance measurements are available. This means that the target formations as well as the sensed variables are determined by distances. We propose a fully distributed control law, which only involves distance measurements for each individual agent to stabilize a desired formation shape, while a storage of measured data is not required. The approach is applicable to point agents in the Euclidean space of arbitrary (but finite) dimension. Under the assumption of infinitesimal rigidity of the target formations, we show that the proposed control law induces local uniform asymptotic (and even exponential) stability. A similar statement is also derived for nonholonomic unicycle agents with all-to-all communication. We also show how the findings can be used to solve extremum seeking control problems.

Another contribution is an extremum seeking control law with an adaptive dither signal. We present an output-feedback law that steers a fully actuated control-affine system with general drift vector field to a minimum of the output function. A key novelty of the approach is an adaptive choice of the frequency parameter. In this way, the task of determining a sufficiently large frequency parameter becomes obsolete. The adaptive choice of the frequency parameter also prevents finite escape times in the presence of a drift. The proposed control law does not only lead to convergence into a neighborhood of a minimum, but leads to exact convergence. For the case of an output function with a global minimum and no other critical point, we prove global convergence.

Finally, we present an extremum seeking control law for a class of nonholonomic systems. A detailed averaging analysis reveals that the closed-loop system is driven approximately into descent directions of the output function along Lie brackets of the control vector

fields. Those descent directions also originate from an approximation of suitably chosen Lie brackets. This requires a two-fold approximation of Lie brackets on different time scales. The proposed method can lead to practical asymptotic stability even if the control vector fields do not span the entire tangent space. It suffices instead that the tangent space is spanned by the elements in the Lie algebra generated by the control vector fields. This novel feature extends extremum seeking by Lie bracket approximations from the class of fully actuated systems to a larger class of nonholonomic systems.

Acknowledgments

First of all, I would like to express my sincere gratitude to Prof. Dr. Uwe Helmke who initiated my interest in control theory when I attended his lectures as a student. His wide-ranging research activities put me in touch with extremum seeking control and Lie bracket approximations during my Master's project. I hope that the content of this dissertation resembles the ideas and plans that he had in mind.

I am also grateful to Prof. Dr. Sergey Dashkovskiy for accepting me as his Ph.D. student. In particular, I would like to thank him for his patience and his support.

There are many other people whom I would like to thank for their scientific contributions to this thesis. I am grateful to Dr. Christian Lageman, Dr. Zhiyong Sun, and Prof. Dr. Brian D.O. Anderson for their advice on formation shape control and the results in Chapter 2 of this dissertation. Moreover, I would like to thank Prof. Dr. Achim Ilchmann for his ideas during the Elgersburg Workshop, which have led to the results in Chapter 3. In the context of Chapter 4, I would like to thank Dr. Victoria Grushkovskaya and Prof. Dr. Alexander Zuyev for their inspiring work on extremum seeking control, Lie brackets, and nonholonomic systems. A special 'thanks' goes to Dr. Simon Michalowsky and Prof. Dr. Christian Ebenbauer for helpful discussions. I am grateful to Prof. Dr. Sergey Dashkovskiy and Dr. Jens Jordan for their valuable comments on my thesis. Last but not least, I would like to thank all present and former members of Chair II at the Institute of Mathematics in Würzburg that I met during my time as a Ph.D. student. It was a really nice and productive time.

Finally, I would like to thank my family and my friends for their moral support and their encouragement over the years.

Contents

Abstract i

Acknowledgments iii

1 Introduction 1

1.1 Some general remarks on extremum seeking control . . . 1
1.2 Lie brackets and averaging . . . 4
1.3 Lie brackets and extremum seeking control . . . 11
1.4 Outline . . . 18
1.5 Global definitions and notation for the entire document . . . 20

2 Distance-based formation control 23

2.1 Introduction and motivation . . . 23
2.2 Local definitions and notation for the chapter . . . 26
2.3 Infinitesimal rigidity and gradient estimates . . . 26
2.3.1 Infinitesimal rigidity . . . 26
2.3.2 Gradient estimates . . . 28
2.4 Formation control for point agents . . . 31
2.4.1 Problem description . . . 31
2.4.2 Control law and main statement . . . 32
2.4.3 Simulation examples . . . 36
2.5 Local asymptotic stability analysis for point agents . . . 38
2.5.1 Estimates for the Lie derivatives . . . 38
2.5.2 Averaging of the sinusoids . . . 44
2.5.3 Integral expansion . . . 45
2.5.4 Proofs of Theorems 2.12 and 2.13 . . . 48
2.6 Extension to formation control for unicycles with all-to-all communication . 50
2.6.1 Problem description . . . 50
2.6.2 Control law and main statement . . . 51
2.7 Local asymptotic stability analysis for unicycles . . . 55
2.7.1 Estimates for the Lie derivatives . . . 56
2.7.2 Averaging of the sinusoids . . . 57
2.7.3 Integral expansion . . . 61
2.7.4 Proof of Proposition 2.32 . . . 63
2.8 Extension to extremum seeking control . . . 63

3 Extremum seeking control with an adaptive dither signal 67

3.1 Introduction and motivation . . . 67
3.2 Local definitions and notation for the chapter . . . 69
3.3 Main result . . . 70
3.4 A numerical Example . . . 72

3.5 Convergence analysis 74
3.5.1 Estimates for the Lie derivatives 76
3.5.2 Averaging of the sinusoids 78
3.5.3 Integral expansion 81
3.5.4 Proof of Theorem 3.3 83

4 Extremum seeking control for a class of nonholonomic systems 85
4.1 Introduction and motivation 85
4.2 Local definitions and notation for the chapter 86
4.3 Problem statement and motivation of the proposed control strategy 88
4.4 Control law 90
4.4.1 Definition of the slow oscillations and the first-order hold 91
4.4.2 Definition of the fast oscillations 92
4.4.3 Control law and closed-loop system 94
4.5 Main results 94
4.6 Averaging of the fast oscillations 96
4.6.1 Iterated Lie derivatives and Lie brackets 96
4.6.2 Averaging of the sinusoids 98
4.6.3 Integral expansion 108
4.7 Averaging of the slow oscillations 110
4.7.1 Averaging of the sinusoids 111
4.7.2 Estimate for the output variations 112
4.7.3 Integral expansion 114
4.7.4 Proof of Theorem 4.7 116

Possible directions for future research 119

Bibliography 121

1 Introduction

1.1 Some general remarks on extremum seeking control

In many applications it is desired to stabilize a dynamical system about a reference trajectory or a set point that is optimal with respect to a certain performance criterion. This task is especially challenging if the optimization problem is complicated by uncertainties and a limited amount of available information. Take, for example, an anti-lock braking system in an automobile [6, 120]. In this case the objective is to maximize the friction force coefficient between the wheel and the ground in order to stop the vehicle as fast as possible while preventing the wheels from locking up. The braking system has to provide the optimal breaking torque. Clearly, if the breaking torque is too weak, then the friction force coefficient is below its optimal value. On the other hand, if the breaking torque is too strong, then the wheels start to lock and the friction force coefficient suffers due to dangerous slipping. Under suitable assumptions, the friction force coefficient can be characterized by the vehicle's current linear acceleration, which can be measured with an accelerometer. Based on these real-time measurements, the breaking system has to provide a breaking torque such that the linear acceleration attains an extreme value. Note that a functional dependence of the linear acceleration on the breaking torque is, in general, not known due to uncertainties like the road conditions and humidity. This situation requires a suitable feedback law that regulates the breaking torque in such a way that the measured performance function attains its (unknown) extreme value. Any solution to this real-time optimization problem can be considered as an example of an *extremum seeking control law*.

A mathematically rigid definition of *extremum seeking control* is difficult to state. In particular, it is not always possible to make a clear distinction between extremum seeking control and other optimization approaches. A common feature of extremum seeking control laws is the intention to steer a control system in such a way that a certain performance-evaluating function (or objective function) attains an extreme value. Methods have been proposed for discrete-time system [23, 42, 66] and continuous-time system [55, 44, 31]. In general, the objective function may depend on the time parameter, the system state, and the controls. It is usually assumed that the objective function only depends on the current values of the system state and the controls, but not on their prehistory. However, in some studies, such as [44], an additional cost functional takes past system states and control values into account. Moreover, it is frequently required that an extremum seeking control law only relies on real-time measurements of certain quantities and that it adapts to changing conditions. For this reason, extremum seeking control is associated with the fields of real-time optimization [6] and adaptive control [9, 10]. The description gets more difficult when it comes to the question what kind of information about the system and the performance function may be used to design and implement an extremum seeking control law. In the most ideal form, an extremum seeking control law would be a universally applicable control strategy that requires no other information than the current value of the performance function. It is clear that such a universal strategy does not exist. All known control laws rely on very specific assumptions on the system and its performance function.

The same is true for the results in the present document. There is no generally valid rule in the literature on what kind of assumptions may (or may not) be made. Sometimes, a method is referred to as an extremum seeking control law even if its implementation requires more information than just the current value of the performance function, such as partial information about the system state and the system dynamics [44] or the current gradient vector of the objective function [119]. As a general rule, one can say that the notion of extremum seeking control involves the desire to rely on a minimum amount of information about the system and the objective function. The goal is to provide real-time optimization for dynamic systems in the presence of uncertainties.

Over the past decades many different approaches to extremum seeking control have been proposed in the literature. Most of the studies are motivated by particular real-world problems. For example, one of the earliest papers [61] from the year 1922 addresses the electrification of railways by means of alternating currents of high frequency. Other examples are maximum power point tracking [63, 16], PID tuning [53], cam timing [89], electromechanical valve actuation [88]. Also, from a theoretical point of view, many directions have been pursued to design extremum seeking control. For example, there are methods based on sliding mode control [87, 118], parameter estimation techniques [44, 2, 43], or methods from numerical optimization [117, 52]. The existing strategies and applications are documented in various survey articles [12, 102, 115] and textbooks [6, 65, 98, 120]. Every approach requires certain assumptions on the control system and the performance function. A frequently studied extremum seeking control problem is described by the following continuous-time input-output model [6, 115]:

$$\dot{x} = f(x, u), \tag{1.1}$$

$$y = \psi(x). \tag{1.2}$$

It consists of a control system (1.1) and a performance-evaluating output (1.2). The right-hand side of (1.1) is assumed to be given by a vector field f that depends smoothly on the system state $x \in \mathbb{R}^n$ and a vector $u \in \mathbb{R}^m$ of input channels for a control law. In the following sections and chapters of this document, we will also impose the additional assumption that (1.1) has a control-affine structure. For the moment, however, we consider the more general nonlinear control system (1.1). The performance output (1.2) is assumed to be given by a smooth real-valued function ψ on the system state, which will be referred to as the *output function*. This contains the assumption that the performance of the system does not depend on the time parameter or the input vector. The assumption that ψ does not depend on the inputs will be important for the method that we apply in this document. Extensions to moderately time-dependent output functions are possible, see [39], but not further addressed here. To get a well-defined optimization problem, we assume that ψ attains a local (or even global) extreme value $y^* \in \mathbb{R}$ at a certain system state $x^* \in \mathbb{R}^n$; without loss of generality we always consider the case of a *minimum*. The current system state, the vector field f as well as the gradient of ψ are treated as unknown quantities. Moreover, the optimal state x^* and the optimal value y^* are not assumed to be known. Only real-time measurements of the current output value (1.2) are available. The ambitious goal is to find an output-feedback control law such that the unknown state x^* becomes asymptotically stable for the closed-loop system. Before we start to explain the method that is used in the present document, we briefly indicate one of the other existing approaches that can be used to obtain extremum seeking control for (1.1), (1.2). This will also highlight some of the conceptual differences of the method that is used here compared to other approaches.

A natural approach to derive extremum seeking control for (1.1), (1.2) is to determine descent directions of the output function. For this reason, many extremum seeking strategies involve suitable perturbation signals in order to extract gradient information from the response of the output signal [6]. The approach in the present document also belongs to this class of methods, but there is the following important difference. In contrast to the approach that we use here, the stability of many perturbation-based extremum seeking control schemes relies on the existence of a so-called steady-state input-output map [55, 90, 23, 44, 116, 118]. To be more precise, it is assumed that, for each constant input vector $u \in \mathbb{R}^m$, a certain point $l(u) \in \mathbb{R}^n$ is asymptotically stable for (1.1). This is usually referred to as a *steady-state assumption.* In particular, whenever the input u is kept (at least approximately) constant, then the system state will converge to the unknown point $l(u)$. The so-defined map $l\colon \mathbb{R}^m \to \mathbb{R}^n$ does not need to be known but should be at least sufficiently smooth. Moreover, to obtain a well-defined optimization problem, it is assumed that the composition $\psi \circ l$ attains a strict minimum value at some $u^* \in \mathbb{R}^m$. Note that $l(u^*)$ is not necessarily a point at which ψ attains a minimum value. Since, for each $u \in \mathbb{R}^m$, the point $l(u)$ is a steady state of the system, the function $\psi \circ l$ is called the *steady-state input-output map.* For an approximately constant input u, the output will converge to some approximately constant output value $\psi(l(u))$. Consequently, one can probe the response of the steady-state input-output map by inducing sufficiently slow variation of the input vector. This can be done in a systematic way by feeding in certain sinusoidal perturbation signals with sufficiently small amplitudes and low frequencies. As a result, the input u slowly moves into a descent direction of $\psi \circ l$. A suitable averaging analysis reveals [55, 116] that such an extremum seeking controller causes the input vector u to converge to some neighborhood of the optimal input vector u^*. The attracting neighborhood around u^* shrinks with decreasing amplitudes and frequencies of the perturbation signals. On the other hand, small amplitudes and frequencies also lead to the effect that the speed of convergence of u towards u^* decreases. Therefore, small amplitudes and frequencies lead to a trade off between accuracy and speed of convergence.

While the accuracy of the approach in the previous paragraph improves in the small-amplitude, low-frequency limit, the method that we study in the present document requires exactly the opposite limit. It also involves suitable perturbation signals for the inputs, but the amplitudes and frequencies need to be sufficiently large. The method does not rely on a steady-state assumption but can be applied to potentially unstable system. The strategy is to overpower unstable dynamics of the system and to force the system into a descent direction of the output function. Clearly, large amplitudes and high frequencies are not suitable for certain applications. However, the method does not lead to the trade off between accuracy and speed of convergence as in the previous paragraph. In fact, any prescribed accuracy and speed of convergence can be achieved by using sufficiently strong perturbations signals. An additional advantage of the method is that the underlying formalism for the analysis and the design of extremum seeking controllers is very general and rather easy to apply. We explain this method in the following two sections. The employed perturbation signals have the purpose to steer the control system into directions of certain Lie brackets. This property can be revealed by a suitable averaging analysis, which is indicated in Section 1.2. The Lie brackets can be chosen in such a way that they point into descent directions of the output function. Under suitable assumptions, this leads to the effect that the closed-loop system is asymptotically stable around states at which the output functions attains a minimum value. The design of such an extremum seeking control law is explained in Section 1.3. We also give references to related works and approaches.

The ideas and concepts in Sections 1.2 and 1.3 will be used in all subsequent chapters, which is outlined in Section 1.4. For later references, we collect basic definitions and notation at the end of the chapter in Section 1.5.

1.2 Lie brackets and averaging

In this section, we introduce the averaging method that is used in the entire document. The strategy in each of the subsequent Chapters 2-4 is to approximate the directions of certain Lie brackets in order to solve a particular optimization problem. For this reason, we now briefly recall the differential geometric definition of Lie brackets and explain how this is related to the averaging approach that we use to obtain approximations of Lie brackets. Suitable textbook reference on Lie brackets and their role in control theory are, for example, [3, 18, 50]. To provide an easy-to-understand introduction, the subsequent considerations are kept on a very elementary level. We only need the notion of flow maps, which are recalled in the next paragraph.

Let f be a smooth vector field on $\mathbb{R}^n$. Then, for each point x_0 of $\mathbb{R}^n$, there exists a unique maximal solution of the ordinary differential equation $\dot{x} = f(x)$ that passes through the initial value x_0 at initial time 0. This maximal solution is denoted by $t \mapsto \Phi_t^f(x_0)$. Since $t \mapsto \Phi_t^f(x_0)$ originates from integrating $\dot{x} = f(x)$, such a maximal solution is also called a *maximal integral curve of* f. The domain of $t \mapsto \Phi_t^f(x_0)$ is not necessarily equal to $\mathbb{R}$ but at least an open interval containing the initial time 0. Since f is assumed to be smooth, it is known from ordinary differential equations that the solutions of $\dot{x} = f(x)$ depend smoothly on the initial value x_0 and the time parameter t. Let $\mathcal{D}(f)$ denote the set of all $(t, x) \in \mathbb{R} \times \mathbb{R}^n$ at which $\Phi_t^f(x)$ exists. This set is an open subset of $\mathbb{R} \times \mathbb{R}^n$. The smooth map $\mathcal{D}(f) \to \mathbb{R}^n$ that assigns to each $(t, x) \in \mathcal{D}(f)$ the point $\Phi_t^f(x) \in \mathbb{R}^n$ is called the *flow of* f. Using the uniqueness of solutions, one can easily prove that, for each $x_0 \in \mathbb{R}^n$, there exists an open neighborhood U of x_0 in $\mathbb{R}^n$ and a sufficiently small $\varepsilon > 0$ such that, for each $t \in (-\varepsilon, \varepsilon)$, the map $U \to \mathbb{R}^n$, $x \mapsto \Phi_t^f(x)$ is a well-defined diffeomorphism onto its image. In the next paragraph we will use a suitable composition of flow maps to define the Lie bracket of two smooth vector fields.

For the rest of this section, let f_1 and f_2 be two smooth vector fields on $\mathbb{R}^n$. As in the previous paragraph, we denote their flows by Φ^{f_1} and Φ^{f_2}, respectively. For the moment, fix an arbitrary point x_0 of $\mathbb{R}^n$. Then, there exists some sufficiently small $\varepsilon > 0$ such that, for each $t \in (-\varepsilon, \varepsilon)$, the following construction can be carried out. First, we start at x_0 and follow the integral curve of f_1 that passes through x_0 at time 0 to the point $x_1 := \Phi_t^{f_1}(x_0)$. Next, we follow the integral curve of f_2 that passes through x_1 at time 0 to the point $x_2 := \Phi_t^{f_2}(x_1)$. Then, we repeat the procedure but along the reverse directions of $-f_1$ and $-f_2$; i.e., we go from x_2 along $-f_1$ to $x_3 := \Phi_t^{-f_1}(x_2)$, and finally from x_3 along $-f_2$ to $x_4 := \Phi_t^{-f_2}(x_3)$. In other words, x_4 is obtained by applying the composition of $\Phi_t^{f_1}$, $\Phi_t^{f_2}$, $\Phi_t^{-f_1}$, and $\Phi_t^{-f_2}$ to x_0. This composition can be applied to x_0 for each $t \in (-\varepsilon, \varepsilon)$, which leads to the curve $t \mapsto x_4(t)$ in $\mathbb{R}^n$. Since the flows of f_1 and f_2 are smooth, also $t \mapsto x_4(t)$ is smooth. Consequently, we obtain a tangent vector to $\mathbb{R}^n$ at x_0 if we take the derivative $\dot{x}_4(0)$ of $t \mapsto x_4(t)$ at $t = 0$. Note that the linearization of an integral curve $t \mapsto \Phi_t^f(x)$ reads $\Phi_t^j(x) = x + tf(x) + O(t^2)$, where $O(t^2)$ is the usual Landau symbol for remainders that tend to zero like t^2 as $t \to 0$. The particular combination of $\pm f_1$ and $\pm f_2$ in x_4 leads to a cancellation of the linear terms, and therefore $x_4(t) = x_0 + O(t^2)$. Thus, we have

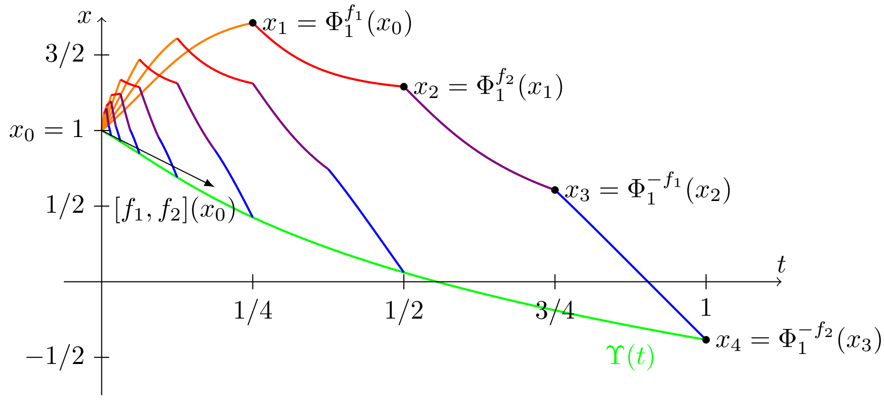


Figure 1.1: *Illustration of the curve* Υ *in* (1.3) *with initial value* $x_0 = 1$ *for the particular choice of the one-dimensional vector fields* f_1, f_2 *that are given by* (1.35)*. The flow maps* Φ^{f_1}, Φ^{f_2} *also appear in Figure 1.6 in the context of extremum seeking control.*

$\dot{x}_4(0) = 0$. This means that $t \mapsto x_4(t)$ runs through the point x_0 with zero velocity. To get a possibly nonzero velocity, we accelerate the curve around $t = 0$ by choosing a different parametrization. For this purpose, we define a continuous transformation $\sigma\colon \mathbb{R} \to \mathbb{R}$ of the time parameter t by $\sigma(t) := -\sqrt{|t|}$ for $t \leq 0$ and by $\sigma(t) := +\sqrt{|t|}$ for $t > 0$. Note that small changes of the time parameter t around 0 lead to infinitely large changes of the new time parameter $\sigma(t)$. If we consider $t \mapsto x_4(t)$ in the new time scale σ, then

$$\Upsilon(t) \; := \; \Big(\Phi^{-f_2}_{\sigma(t)} \circ \Phi^{-f_1}_{\sigma(t)} \circ \Phi^{f_2}_{\sigma(t)} \circ \Phi^{f_1}_{\sigma(t)}\Big)(x_0) \tag{1.3}$$

defines a continuous curve in $\mathbb{R}^n$ with domain $(-\varepsilon^2, \varepsilon^2)$ that passes through the point x_0 at time 0. The definition of the curve Υ is illustrated in Figure 1.1. Since $t \mapsto x_4(t)$ is a smooth map with vanishing derivative at $t = 0$, it follows that $\Upsilon = x_4 \circ \sigma$ is at least differentiable. Thus a tangent vector to $\mathbb{R}^n$ at x_0 is given by the derivative

$$[f_1, f_2](x_0) \; := \; \dot{\Upsilon}(0) \tag{1.4}$$

of Υ at 0, which is called the *Lie bracket* of f_1 and f_2 at x_0. The velocity vector of Υ at 0 is also shown in Figure 1.1. A direct computation of the Lie bracket from its definition in (1.4) is, in general, not possible since the flow maps Φ^{f_1} and Φ^{f_2} are usually not explicitly known. However, there is a much easier and direct formula to compute the Lie bracket; see equation (1.18) below. At this point, we could simply state this formula without proof by giving a reference to a textbook on differential geometry, such as [1]. Indeed, the formula can be easily proved using a suitable expansion of Υ around $t = 0$. However, in the following, we present an alternative proof in the next paragraphs. This procedure is certainly more labor intense than the traditional way to prove the Lie bracket formula, but, on the other hand, it provides an alternative interpretation of the Lie bracket in terms of dynamical systems. In particular, we will see that the Lie bracket (1.4) arises naturally from a suitable averaging analysis. Moreover, the subsequent procedure can be seen as the simplest possible example of the method that is used in the entire document. All results in the remaining chapters can be interpreted as generalizations of this strategy.

We continue to study the Lie bracket $[f_1, f_2](x_0)$ of the smooth vector fields f_1, f_2 at some point x_0 of $\mathbb{R}^n$. We have introduced the Lie bracket in (1.4) as the velocity vector of

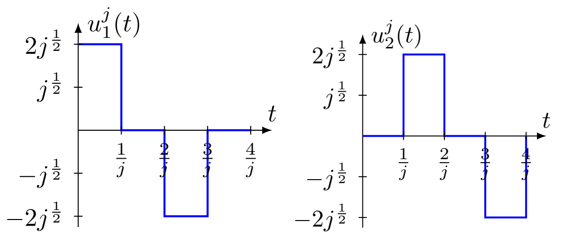


Figure 1.2: *Graphical definition of the $4/j$-periodic functions u_i^j in (1.5).*

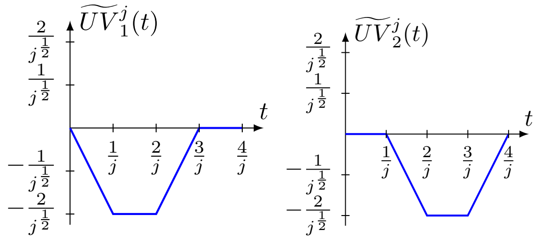


Figure 1.3: *Graphical definition of the $4/j$-periodic functions $\widetilde{UV}_i^j$ in (1.8).*

the curve Υ given by (1.3). Note that, for every smooth vector field f, every time t, and every point x, at which the flow $\Phi_t^f(x)$ exists, the scaling property $\Phi_t^f(x) = \Phi_{t/\lambda}^{\lambda f}(x)$ holds for every nonzero real number λ. Thus, if we evaluate Υ at $t = 4/j$ for some sufficiently large positive real number j, then we have

$$\Upsilon(4/j) \;=\; \Big(\Phi_{1/j}^{-\sqrt{4j}f_2} \circ \Phi_{1/j}^{-\sqrt{4j}f_1} \circ \Phi_{1/j}^{\sqrt{4j}f_2} \circ \Phi_{1/j}^{\sqrt{4j}f_1}\Big)(x_0).$$

To obtain the point $\Upsilon(4/j)$, we start at x_0 and then we move on time intervals of length $1/j$ along integral curves of the vector fields $\sqrt{4j}f_1$, $\sqrt{4j}f_2$, $-\sqrt{4j}f_1$, and $-\sqrt{4j}f_2$. This trajectory can also be generated by a dynamical system. For this purpose, we introduce the $4/j$-periodic time-varying functions u_1^j, u_2^j in Figure 1.2. Now consider the time-varying system

$$\dot{x} \;=\; u_1^j(t)\, f_1(x) + u_2^j(t)\, f_2(x) \tag{1.5}$$

on $\mathbb{R}^n$. For every $j > 0$, let γ^j be the maximal solution of (1.5) with $\gamma^j(0) = x_0$. It is easy to see that the particular choice of the functions u_1^j, u_2^j ensures that the vector fields f_1, f_2 on the right-hand side of (1.5) are "turned on and off" in such way that $\gamma^j(4/j) = \Upsilon(4/j)$ for sufficiently large $j > 0$. Thus, we obtain from (1.4) that the Lie bracket of f_1 and f_2 at x_0 is given by

$$[f_1, f_2](x_0) \;=\; \lim_{j\to\infty} \frac{\gamma^j(4/j) - x_0}{4/j - 0}. \tag{1.6}$$

This formula establishes a first connection between the Lie bracket (1.4) and the trajectories of the dynamical system (1.5). In the next two paragraphs, we present an averaging technique that can be applied to extract the behavior of system (1.5) in the large-amplitude, high-frequency limit $j \to \infty$. As an immediate consequence, this will lead us to a well-known formula for the Lie bracket.

For the moment, let γ be any solution of (1.5) for some $j > 0$, and fix arbitrary t_1, t_2 in the domain of γ. Let φ be a smooth real-valued function on $\mathbb{R}^n$. Using the fundamental theorem of calculus for the composition of γ and φ, we obtain that

$$\varphi(\gamma(t_2)) \;=\; \varphi(\gamma(t_1)) + \sum_{i=1,2} \int_{t_1}^{t_2} u_i^j(t)\, (f_i\varphi)(\gamma(t))\, \mathrm{d}t, \tag{1.7}$$

where $f_i\varphi$ denotes the *Lie derivative of* φ *along* f_i for $i = 1, 2$; i.e., for every $x \in \mathbb{R}$, we let $(f_i\varphi)(x)$ denote the real number that originates from applying the derivative of φ at x to the vector $f_i(x)$. In other words, we interpret the vector fields f_i as linear differential operators $\varphi \mapsto f_i\varphi$. Next, we use integration by parts in the above integral to average

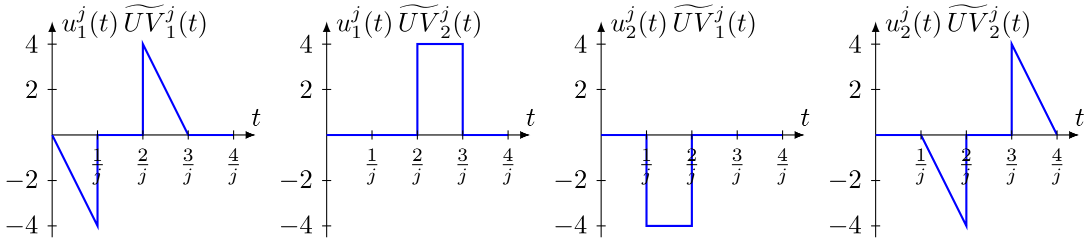


Figure 1.4: *Illustration of the* $4/j$*-periodic functions* $u_{i_1}^j\widetilde{UV}_{i_2}^j$ *in* (1.8b).

the $4/j$-periodic functions u_i^j. For this purpose, suitable $4/j$-periodic antiderivatives $\widetilde{UV}_i^j$ of the functions $-u_i^j$ are defined in Figure 1.3. The notation $\widetilde{UV}_i^j$ is taken from [68, 69]. Moreover, for all $i_1, i_2 \in \{1,2\}$, we denote the Lie derivative of the smooth function $f_{i_2}\varphi$ along the vector field f_{i_1} by $f_{i_1}(f_{i_2}\varphi)$. Since γ is a solution of (1.5), integration by parts in (1.7) leads to

$$\varphi(\gamma(t_2)) \;=\; \varphi(\gamma(t_1)) - \sum_{i=1,2}\Big[\widetilde{UV}_i^j(t)\,(f_i\varphi)(\gamma(t))\Big]_{t=t_1}^{t=t_2} \tag{1.8a}$$

$$+ \sum_{i_1,i_2=1,2}\int_{t_1}^{t_2} u_{i_1}^j(t)\,\widetilde{UV}_{i_2}^j(t)\,(f_{i_1}(f_{i_2}\varphi))(\gamma(t))\,\mathrm{d}t, \tag{1.8b}$$

where we use a notation of the form $[\alpha(t)]_{t=t_1}^{t=t_2}$ to denote the difference $\alpha(t_2)-\alpha(t_1)$. The integrals in (1.8b) contain the products $u_{i_1}^j\,\widetilde{UV}_{i_2}^j$ with $i_1, i_2 \in \{1,2\}$. It can be seen in Figure 1.4 that the products $u_1^j\,\widetilde{UV}_1^j$ and $u_2^j\,\widetilde{UV}_2^j$ are $4/j$-periodic functions with zero averages $v_{1,1} := 0$ and $v_{2,2} := 0$, respectively. The products $u_1^j\,\widetilde{UV}_2^j$ and $u_2^j\,\widetilde{UV}_1^j$ are also $4/j$-periodic functions but with nonzero averages $v_{1,2} := 1$ and $v_{2,1} := -1$, respectively. If we add and subtract the averages v_{i_1,i_2} to the products $u_{i_1}^j\,\widetilde{UV}_{i_2}^j$ in (1.8b), then we obtain

$$\varphi(\gamma(t_2)) \;=\; \varphi(\gamma(t_1)) - \sum_{i=1,2}\Big[\widetilde{UV}_i^j(t)\,(f_i\varphi)(\gamma(t))\Big]_{t=t_1}^{t=t_2} \tag{1.9a}$$

$$+ \sum_{i_1,i_2=1,2}\int_{t_1}^{t_2} v_{i_1,i_2}\,(f_{i_1}(f_{i_2}\varphi))(\gamma(t))\,\mathrm{d}t \tag{1.9b}$$

$$- \sum_{i_1,i_2=1,2}\int_{t_1}^{t_2} \big(v_{i_1,i_2} - u_{i_1}^j(t)\,\widetilde{UV}_{i_2}^j(t)\big)\,(f_{i_1}(f_{i_2}\varphi))(\gamma(t))\,\mathrm{d}t. \tag{1.9c}$$

The terms in (1.9b) represent the averaged contribution of (1.5), while the terms in (1.9a) and (1.9c) are remainders, which will be shown to become small with increasing j. To obtain an easy estimate for the contribution in (1.9c), we apply integration by parts a second time. For this purpose, suitable antiderivatives $\widetilde{UV}_{i_1,i_2}^j$ of the functions $(v_{i_1,i_2} - u_{i_1}^j\,\widetilde{UV}_{i_2}^j)$ are defined in Figure 1.5. Moreover, for all $i_1, i_2, i_3 \in \{1,2\}$, we denote the Lie derivative of the smooth function $f_{i_2}(f_{i_3}\varphi)$ along the vector field f_{i_1} by $f_{i_1}(f_{i_2}(f_{i_3}\varphi))$. Since γ is a solution of (1.5), integration by parts in (1.9c) leads to

$$\varphi(\gamma(t_2)) \;=\; \varphi(\gamma(t_1)) - \Big[(D_1^j\varphi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} \tag{1.10a}$$

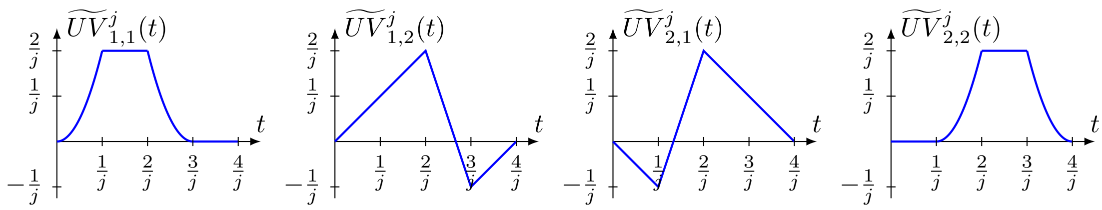


Figure 1.5: *Graphical definition of the* $4/j$*-periodic functions* $\widetilde{UV}^j_{i_1,i_2}$ *in* (1.10).

$$+\sum_{i_1,i_2=1,2}\int_{t_1}^{t_2} v_{i_1,i_2}\,(f_{i_1}(f_{i_2}\varphi))(\gamma(t))\,\mathrm{d}t+\int_{t_1}^{t_2}(D_2^j\varphi)(t,\gamma(t))\,\mathrm{d}t, \tag{1.10b}$$

where the time-varying functions $D_1^j\varphi, D_2^j\varphi\colon \mathbb{R}\times\mathbb{R}^n\to\mathbb{R}$ are defined by

$$(D_1^j\varphi)(t,x) := \sum_{i=1,2}\widetilde{UV}^j_i(t)\,(f_i\varphi)(x)+\sum_{i_1,i_2=1,2}\widetilde{UV}^j_{i_1,i_2}(t)\,(f_{i_1}(f_{i_2}\varphi))(x), \tag{1.11}$$

$$(D_2^j\varphi)(t,x) := \sum_{i_1,i_2,i_3=1,2} u^j_{i_1}(t)\,\widetilde{UV}^j_{i_2,i_3}(t)\,(f_{i_1}(f_{i_1}(f_{i_2}\varphi)))(x). \tag{1.12}$$

On a more abstract level, one can interpret (1.11) and (1.12) as the definitions of two time-varying differential operators D_1^j and D_2^j that act on smooth real-valued functions. The notation D_1^j and D_2^j and the interpretation as differential operators can also be found in [80].

Now we derive estimates for the remainders (1.11) and (1.12) in the integral expansion (1.10). From Figures 1.2, 1.3 and 1.5, we obtain that there exists a positive constant a such that

$$\big|u^j_i(t)\big| \le a\,j^{\frac{1}{2}}, \qquad \big|\widetilde{UV}^j_i(t)\big| \le a\,j^{-\frac{1}{2}}, \qquad \big|\widetilde{UV}^j_{i_1,i_2}(t)\big| \le a\,j^{-1} \tag{1.13}$$

for every $j>0$, all $i,i_1,i_2\in\{1,2\}$, and every $t\in\mathbb{R}$. Moreover, since the vector fields f_1 and f_2 are assumed to be smooth, for every smooth real-valued function φ on $\mathbb{R}^n$ and every compact subset K of $\mathbb{R}^n$, there exists a positive constant b such that

$$\big|(f_{i_1}\varphi)(x)\big| \le b, \qquad \big|(f_{i_1}(f_{i_2}\varphi))(x)\big| \le b, \qquad \big|(f_{i_1}(f_{i_2}(f_{i_3}\varphi)))(x)\big| \le b \tag{1.14}$$

for all $i_1,i_2,i_3\in\{1,2\}$ and every $x\in K$. Using the definitions in (1.11) and (1.12), it follows that, for every smooth real-valued function φ on $\mathbb{R}^n$ and every compact subset K of $\mathbb{R}^n$, there exist positive constants c_1,c_2 such that

$$\big|(D_1^j\varphi)(t,x)\big| \le c_1\,j^{-\frac{1}{2}} \quad\text{and}\quad \big|(D_2^j\varphi)(t,x)\big| \le c_2\,j^{-\frac{1}{2}} \tag{1.15}$$

for every $j>0$, every $t\in\mathbb{R}$, and every $x\in K$. The above estimates ensure that the remainders in (1.10) converge locally uniformly to zero as j tends to infinity. If we interpret D_1^j and D_2^j as time-varying differential operators, then estimates (1.15) are closely related to the concept of "DO-convergence" in [80], where *DO* abbreviates *differential operator*.

Now we return to our initial objective to derive a simple formula for the Lie bracket (1.4) via averaging. From Figures 1.3 and 1.5, we obtain that the functions $\widetilde{UV}^j_i$ and $\widetilde{UV}^j_{i_1,i_2}$ vanish at integer multiples of $4/j$. By (1.11), this implies $(D_1^j\varphi)(0,x)=0$ and $(D_1^j\varphi)(4/j,x)=0$

for every $j > 0$, every smooth real-valued function φ on $\mathbb{R}^n$, and every $x \in \mathbb{R}^n$. As in (1.6), let γ^j be the maximal solution of (1.5) with $\gamma^j(0) = x_0$. Using (1.10) with $t_1 = 0$ and $t_2 = 4/j$ as well as (1.15), we conclude that

$$\lim_{j\to\infty} \frac{\varphi(\gamma^j(4/j)) - \varphi(x_0)}{4/j - 0} \;=\; \sum_{i_1,i_2=1,2} v_{i_1,i_2}\,(f_{i_1}(f_{i_2}\varphi))(x_0) \tag{1.16}$$

for every smooth real-valued function φ on $\mathbb{R}^n$. Recall that the averaged coefficients v_{i_1,i_2} in (1.16) are given by

$$v_{1,1} \;=\; 0, \qquad v_{1,2} \;=\; 1, \qquad v_{2,1} \;=\; -1, \qquad v_{2,2} \;=\; 0. \tag{1.17}$$

Moreover, note that (1.16) is in particular true if φ is any of the component functions of the identity map on $\mathbb{R}^n$. Therefore, we obtain from (1.6), (1.16), and (1.17) the well-known formula

$$[f_1, f_2](x_0) \;=\; \mathrm{D}f_2(x_0)f_1(x_0) - \mathrm{D}f_1(x_0)f_2(x_0) \tag{1.18}$$

for the Lie bracket of f_1 and f_2, where $\mathrm{D}f_i$ denotes the derivative of f_i for $i = 1, 2$. In particular, we have shown that the Lie bracket arises naturally as the averaged vector field of the right-hand side of (1.5) in the large-amplitude, high-frequency limit. Thus, in the limit $j \to \infty$, we may consider

$$\dot{x} \;=\; f^\infty(x) \;:=\; [f_1, f_2](x) \tag{1.19}$$

as the averaged system of (1.5). Using (1.10), (1.15), (1.18), and the Gronwall inequality, it is now easy to prove the following approximation result (see, e.g., Proposition 8.3 in [68]).

Proposition 1.1. *For every compact subset K of $\mathbb{R}^n$ and every time span $T > 0$, there exist $c, j_0 > 0$ such that, for every initial time $t_0 \in \mathbb{R}$ and every maximal solution γ^∞ of* (1.19)*, the following implication holds: if γ^∞ exists on $[t_0, t_0 + T]$ with $\gamma^\infty(t) \in K$ for every $t \in [t_0, t_0 + T]$, then, for every $j \geq j_0$, also the maximal solution γ^j of* (1.5) *with initial condition $\gamma^j(t_0) = \gamma^\infty(t_0)$ exists on $[t_0, t_0 + T]$ and the estimate*

$$\|\gamma^\infty(t) - \gamma^j(t)\| \;\leq\; c\, j^{-\frac{1}{2}}$$

holds for every $t \in [t_0, t_0 + T]$, where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$.

Note that we have shown much more than just formula (1.18) for the Lie bracket. Proposition 1.1 states that the trajectories of (1.5) approximate the trajectories of (1.19) locally with increasing parameter value j. This approximation property can be extended to a much more general situation and can be explained by a well-established averaging theory. The notation and the approach that we have used in the preceding paragraphs is taken from one of the mile stones [69] of this averaging theory. A short review of some of the known results is given in the remaining paragraphs of this section.

One of the earliest results on the connection between Lie brackets and averaging of dynamical systems can be found in [57]. Therein, the authors consider a time-varying system of the form

$$\dot{x} \;=\; \sum_{i=1}^{m} u_i^j(t)\, f_i(x) \tag{1.20}$$

on $\mathbb{R}^n$, where the f_i are smooth vector fields and the time-varying functions u_i^j are given by

$$u_i^j(t) \;:=\; \sqrt{2j}\,\cos(jt+\vartheta_i), \tag{1.21}$$

where j is a positive real parameter and the phase shifts ϑ_i are arbitrary real numbers. Note that if $m = 2$, $\vartheta_1 = 0$, and $\vartheta_2 = \pi/2$, then the sinusoids $u_1^j(t) = \sqrt{2j}\cos(jt)$ and $u_2^j(t) = \sqrt{2j}\sin(jt)$ can be interpreted as smoothed versions of the rectangular shaped functions in Figure 1.2. It is natural to expect that a similar approximation property holds as in Proposition 1.1. Indeed, it is shown in [57] that the trajectories of (1.20) with the u_i^j as in (1.21) approximate the trajectories of the averaged system

$$\dot{x} \;=\; \sum_{i<k} \sin(\vartheta_k - \vartheta_i)\,[f_i, f_k](x) \tag{1.22}$$

with increasing parameter value j. Again, if $m = 2$, $\vartheta_1 = 0$, and $\vartheta_2 = \pi/2$, then (1.22) coincides with (1.19). In particular, this shows that the geometric shape of the functions u_i^j provides a certain degree of freedom. It is also possible to obtain the same result with square waves or triangular waves (see [99]). A further extension of the theory was done in [56]. Instead of (1.21), the functions u_i^j are now assumed to be of the form

$$u_i^j(t) \;:=\; \sum_{k<i} \lambda_{i,k}\,\sqrt{2j}\sin(j\omega_{k,i}t+\vartheta_{i,k}) + \sum_{k>i} \lambda_{i,k}\,\sqrt{2j}\sin(j\omega_{i,k}t+\vartheta_{i,k}), \tag{1.23}$$

where the amplitudes $\lambda_{i,k}$ and phase shifts $\vartheta_{i,k}$ are real numbers, and the frequency coefficients $\omega_{i,k}$, $i < k$, are pairwise distinct positive real numbers. It turns out that, in the limit $j \to \infty$, only sinusoids with the same frequency coefficients give rise to Lie brackets in the averaged system. Conversely, sinusoids with different frequency coefficients do not resonate and therefore do not contribute to the averaged system. It is shown in [56] that the trajectories of (1.20) with the u_i^j as in (1.23) approximate the trajectories of the averaged system

$$\dot{x} \;=\; \sum_{i<k} \lambda_{i,k}\,\lambda_{k,i}\,\sin(\vartheta_k - \vartheta_i)\,[f_i, f_k](x) \tag{1.24}$$

with increasing parameter value j. We will use a similar choice of sinusoids as in (1.23) in Chapters 2 and 3. The content of Chapter 4 requires an even more general approach, which is indicated in the next paragraph.

The above methods for approximations of Lie brackets can be extended even further. It is not only possible to approximate Lie brackets of pairs of vector fields as in (1.24), but also iterated Lie brackets of arbitrary order. This was mainly done in the papers [58, 107, 68, 69] and in the Ph.D. thesis [67] by Wensheng Liu. To approximate iterated Lie brackets, the functions u_i^j in (1.20) have to satisfy certain averaging conditions in the limit $j \to \infty$. To be more precise, it is required that certain iterated integrals of the u_i^j converge uniformly to zero as j tends to infinity. For example, if we consider the rectangular waves in Figure 1.2, then the iterated integrals are represented by the functions $\widetilde{UV}_i^j$ and $\widetilde{UV}_{i_1,i_2}^j$ in Figures 1.3 and 1.5, respectively. We have seen in (1.13) that the iterated integrals converge uniformly to zero as j tends to infinity. General definitions of suitable iterated integrals can be found in [58, 69] and also in Subsection 4.6.2 of the present document. If the functions u_i^j satisfy those averaging conditions in the limit $j \to \infty$, then one can prove that the trajectories of (1.20) approximate the trajectories of an averaged system of the form

$$\dot{x} \;=\; \sum_{k=1}^{r} \sum_{i_1,\ldots,i_k=1}^{m} v_{i_1,\ldots,i_k}\,[f_{i_1}, \cdots [f_{i_{k-1}}, f_{i_k}] \cdots](x), \tag{1.25}$$

where the $v_{i_1,\ldots,i_k}$ are certain real numbers that depend on the choice of the u_i^j. A system of the form (1.25) is also called an *extended system* of (1.20) because its right-hand side does not only contain the initial vector fields $f_1, \ldots, f_m$ but also their iterated Lie brackets. It is shown in [69] that the approximation of an extended system can be explained on a purely algebraic level in terms of the so-called *Chen-Fliess series* determined by the u_i^j. Note that Chen-Fliess series techniques have been widely used in control theory, e.g., by Chen [22], Fliess [36], and Sussmann [105]. Reference [68] provides an explicit recipe on how to choose the function u_i^j in (1.20) such that the extended system (1.25) contains prescribed coefficients $v_{i_1,\ldots,i_k}$. We will use the results from [68] in Chapter 4 to design extremum seeking control for nonholonomic systems. It is worth to mention that the approximation property also holds if (1.20) contains an additional drift vector field and if the control vector fields display a moderate time-dependence. We will encounter such a situation in Chapter 3.

1.3 Lie brackets and extremum seeking control

In the previous section, we have seen that directions of Lie brackets can be approximated by a dynamical system with a suitable highly oscillatory right-hand side. Now we explain how this approximation property can be used for the purpose of extremum seeking control. First, we describe the idea by a very simple toy example. Then, we indicate the general approach and give an overview on some of the existing results on extremum seeking control by Lie bracket approximations.

As an introductory example, we consider the one-dimensional single-input single-output system

$$\dot{x} = u, \tag{1.26}$$
$$y = \psi(x), \tag{1.27}$$

where u is a real-valued input channel, x is the scalar system state, and y is a real-valued output channel. We suppose that the output is given by a smooth real-valued function ψ on the system space $\mathbb{R}$. We also assume that the output can be measured constantly while the system state x and an analytic expression for the function ψ are not known. Our goal is to find an output-feedback control law that asymptotically stabilizes the closed-loop system around states at which ψ attains a minimum value. It is clear that for the simple toy model (1.26), (1.27) the problem could be solved by an easier method than what we describe in the following. However, in order to explain the underlying principles of extremum seeking control by Lie bracket approximations, this toy model serves as a good prototype for more general problems.

To solve above extremum seeking control problem, we return to system (1.5). Recall that the periodically time-varying functions u_1^j, u_2^j on the right-side of (1.5) are defined in Figure 1.2. The parameter j determines the amplitudes and frequencies of u_1^j, u_2^j. The right-hand side of (1.5) also involves the not further specified smooth vector fields f_1, f_2 on $\mathbb{R}^n$. With respect to our one-dimension toy example (1.26), (1.27), we will make a particular choice of the vector fields f_1, f_2 in dimension $n = 1$ below. It is already known from Proposition 1.1 that the trajectories of the highly oscillatory system (1.5) approximate the trajectories of the averaged system (1.19) for sufficiently large values of j. The averaged system is driven into the direction of the Lie bracket $[f_1, f_2]$. For our objective to steer the control system (1.26) towards a state at which the output (1.27) attains a minimum value, it is certainly helpful to know descent directions of the output function ψ. This information is provided by the negative gradient of ψ, which is denoted by $-\nabla\psi$. Note that, in our one-dimensional toy model, the gradient of ψ is just a real-valued function. However, we do not have direct access to the gradient of ψ. We can only measure the value of ψ at the current system state and this system state is also not known. To circumvent the problem, we use the approximation property from Proposition 1.1. We will choose the vector fields f_1, f_2 in such a way that their Lie bracket $[f_1, f_2]$ coincides with the negative gradient of ψ. Then, the averaged system (1.19) is constantly driven into a descent direction of ψ, which in turn implies that the same is also true (at least approximately) for the oscillatory system (1.5). To carry out this plan, we first need to know how to choose f_1, f_2, and secondly, we also need a time-varying output-feedback control law for (1.26) such that the closed-loop system coincides with (1.5).

The above idea to obtain extremum seeking control by Lie bracket approximations can be realized as follows. We introduce two smooth design functions $h_1, h_2 \colon \mathbb{R} \to \mathbb{R}$, which are not rather specified at the moment. In order to obtain a closed-loop system of the form (1.5), we choose the parameter-dependent, time-varying output-feedback control law

$$u \;=\; u_1^j(t)\, h_1(y) + u_2^j(t)\, h_2(y) \tag{1.28}$$

for (1.26), where u_1^j, u_2^j are given by Figure 1.2, and y denotes the output (1.27). Note that an implementation of (1.28) does not require any other information than real-time measurements of the output signal. Moreover, the closed-loop system can be written in the form (1.5) if we define the two vector fields f_1, f_2 on $\mathbb{R}$ by

$$f_1(x) \;:=\; h_1(\psi(x)), \tag{1.29a}$$

$$f_2(x) \;:=\; h_2(\psi(x)). \tag{1.29b}$$

Using equation (1.18), we can compute the Lie bracket of f_1 and f_2 at any $x \in \mathbb{R}$, which leads to

$$[f_1, f_2](x) \;=\; [h_1, h_2](\psi(x))\, \nabla\psi(x), \tag{1.30}$$

where the function $[h_1, h_2] \colon \mathbb{R} \to \mathbb{R}$ is defined by

$$[h_1, h_2](y) \;:=\; h_2'(y)\, h_1(y) - h_1'(y)\, h_2(y). \tag{1.31}$$

Thus, if we want to ensure that $[f_1, f_2] = -\nabla\psi$, then we have to choose the design functions h_1, h_2 in such a way that $[h_1, h_2]$ is identically equal to -1. There are infinitely many different ways to satisfy this property. For example, we can define

$$h_1(y) \;:=\; y, \qquad h_2(y) \;:=\; 1, \qquad \text{or} \tag{1.32}$$

$$h_1(y) \;:=\; \sin(y), \qquad h_2(y) \;:=\; \cos(y), \qquad \text{or} \tag{1.33}$$

$$h_1(y) \;:=\; \mathrm{e}^{y}/\sqrt{2}, \qquad h_2(y) \;:=\; \mathrm{e}^{-y}/\sqrt{2}. \tag{1.34}$$

Definitions (1.32), (1.33), and (1.34) appeared for the first time in [31], [96], and [41], respectively. We will discuss the choice of h_1, h_2 again when we study a more general problem. For our introductory example, we choose the design functions as in (1.33). Definition (1.33) of h_1, h_2 is also used with certain variations in the subsequent chapters. In contrast to (1.32) and (1.34), the definition in (1.33) ensures that control law (1.28) leads to bounded input signals even if y attains large values, which might be preferable for practical implementations.

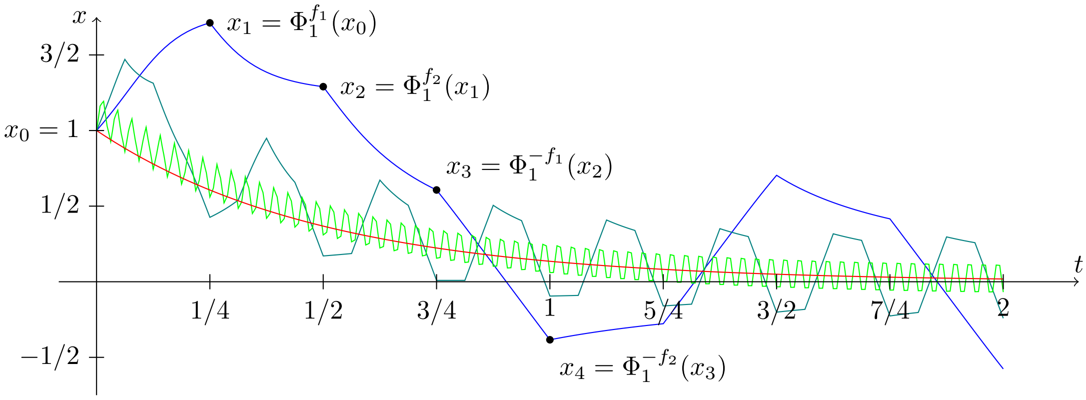


Figure 1.6: *The trajectory of control system* (1.26) *with output* $y = x^2$ *under the* j*-dependent control law* (1.28) *with initial condition* $x(0) = 1$ *is drawn in blue for* $j = 4$*, in cyan for* $j = 16$*, and in green for* $j = 128$*. In the limit* $j \to \infty$*, the trajectories of the closed-loop system converge locally uniformly to the trajectories of* (1.36)*. The trajectory of* (1.36) *with initial condition* $x(0) = 1$ *is drawn in red.*

We conclude the introductory example by discussing the behavior of the closed-loop system in some more detail. For the sake of simplicity, we suppose that the output function ψ is simply given by $\psi(x) := x^2$. Then, the optimal point x^* is the origin. Following the construction of control law (1.28), the closed loop-system (1.5) consists of the time-varying functions u_1^j, u_2^j in Figure 1.2 and, according to (1.29) and (1.33), contains the vector fields f_1, f_2 given by

$$f_1(x) \;=\; \cos(x^2), \qquad (1.35a)$$
$$f_2(x) \;=\; \sin(x^2). \qquad (1.35b)$$

The averaged system (1.19) simply reads

$$\dot{x} \;=\; f^\infty(x) \;=\; -\nabla\psi(x) \;=\; -2x. \qquad (1.36)$$

It is clear that the optimal point $x^* = 0$ is asymptotically stable for (1.36). Using the approximation property in Proposition 1.1, this implies (see [78] for a proof) that x^* is practically uniformly asymptotically stable for the closed-loop system, where the word *uniformly* means that the stability property is uniform with respect to the time parameter, and the word *practically* indicates the dependence on the parameter j. Figure 1.6 shows how the solutions of the closed-loop system approximate the solutions of the averaged system with increasing value of the parameter j. Note that the trajectories of the closed-loop system only converge into a certain neighborhood of x^*. Such a behavior of solutions is usually studied in the context of practical stability theory; see, e.g., [60]. However, the approximation property improves with increasing parameter j. Therefore, the attracting neighborhood of x^* can be made arbitrary small by choosing j sufficiently large. Since x^* is in fact globally asymptotically stable for (1.36), one can also ensure that the domain of attraction for the closed-loop system is an arbitrary large compact neighborhood of x^* by choosing j sufficiently large. Consequently, the output-feedback law (1.28) is indeed an extremum seeking control law, at least for the toy example (1.26), (1.27). The practical stability result can be also derived directly from the integral equation (1.10) for the trajectories of (1.5), which becomes

$$\psi(\gamma(t_2)) \;=\; \psi(\gamma(t_1)) - \Big[(D_1^j\psi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} (f^\infty\psi)(\gamma(t))\,\mathrm{d}t + \int_{t_1}^{t_2} (D_2^j\psi)(t,\gamma(t))\,\mathrm{d}t \tag{1.37}$$

for $\varphi = \psi$, where the Lie derivative of ψ along f^∞ is given by

$$(f^\infty\psi)(x) \;=\; -4\,x^2. \tag{1.38}$$

We know from (1.15) that the remainders $D_1^j\psi, D_2^j\psi$ converge uniformly to zero as j tends to infinity. Thus, outside x^* the negative averaged term (1.38) dominates the right-hand side of (1.37). This leads to a decay of the value of ψ along trajectories of (1.5). In the next paragraphs we will see that this strategy can be also applied to more complex control systems.

The idea of using Lie bracket approximations for the purpose of extremum seeking control appeared for the first time in the Master Thesis [26] by Hans-Bernd Dürr. Since then, the approach has been extended into various directions. An overview on some of the existing literature is given at the end of this section. The results in the present document also contribute to this field of research. To explain the method for a more general situation than for the toy example (1.26), (1.27), we consider a control-affine system of the form

$$\dot{x} \;=\; \sum_{i=1}^{m} u_k\, g_k(x), \tag{1.39}$$

where the u_k are real-valued input channels for a control law, and the g_k are smooth control vector fields on $\mathbb{R}^n$. All of the control systems that we study in the present document are assumed to have a control-affine structure (possibly with an additional drift vector field, which is omitted in (1.39) for the sake of simplicity). However, it is worth to mention that the Lie bracket approach can be also extended to certain systems that are not affine in control; see [100]. As in the introductory example, we assume that the only information about the current system state is provided by a real-valued output channel

$$y \;=\; \psi(x), \tag{1.40}$$

where ψ is a smooth real-valued function on the state space $\mathbb{R}^n$, called the *output function*. Again, the goal is to derive an output-feedback control law for (1.39) that asymptotically stabilizes the closed-loop system around states at which the output (1.40) attains a minimum value. It turns out that we can use basically the same strategy as for the toy model (1.26), (1.27), which is described in the following paragraphs.

To obtain an extremum seeking control law for the more general problem (1.39), (1.40), we return to the design function h_1, h_2 for the toy model (1.26), (1.27). Recall that the functions h_1, h_2 have to be chosen in such a way that the Lie bracket vector in (1.30) points into a descent direction of the output function. This can be easily extended to (1.39), (1.40) as follows. For every $k \in \{1, \ldots, m\}$, define two smooth vector fields $f_{(k,1)}, f_{(k,2)}$ on $\mathbb{R}^n$ by

$$f_{(k,1)}(x) \;:=\; h_1(\psi(x))\, g_k(x), \tag{1.41a}$$

$$f_{(k,2)}(x) \;:=\; h_2(\psi(x))\, g_k(x). \tag{1.41b}$$

Note that (1.41) reduces in dimension $n = 1$ to (1.29) if $g_k(x) = 1$ as in (1.26). As in (1.30), we compute the Lie bracket of $f_{(k,1)}$ and $f_{(k,2)}$ using (1.18). For every $k \in \{1, \ldots, m\}$, we obtain

$$[f_{(k,1)}, f_{(k,2)}](x) \;=\; [h_1, h_2](\psi(x))\,(g_k\psi)(x)\, g_k(x), \tag{1.42}$$

where the function $[h_1, h_2]$ is defined by (1.31), and $g_k\psi$ denotes the Lie derivative of ψ along g_k (i.e., $(g_k\psi)(x)$ is the derivative of ψ at x applied to $g_k(x)$; cf. Section 1.2). Note that the product $(g_k\psi)g_k$ of the real-valued function $g_k\psi$ and the vector field g_k on the right-hand side of (1.42) is again a vector field on $\mathbb{R}^n$. If we take the Lie derivative of ψ along $(g_k\psi)g_k$, then we obtain the nonnegative function $(g_k\psi)^2$ on $\mathbb{R}^n$. For the purpose of extremum seeking control, we are interested in descent directions of ψ. To ensure that the Lie bracket vector in (1.42) points into a descent directions of ψ, we need that the factor $[h_1, h_2](\psi(x))$ is always negative; i.e., we have to choose the design functions h_1, h_2 such that

$$[h_1, h_2](y) \; < \; 0 \tag{1.43}$$

for every $y \in \mathbb{R}$. For example, we can define h_1, h_2 by (1.32), (1.33), or (1.34). A general study on suitable choices of h_1, h_2 can be found in [41]. Next, we explain how the Lie brackets in (1.42) can be approximated by time-varying output feedback.

As explained in the previous paragraph, we are interested in the directions of the Lie brackets in (1.42) because they contain valuable information about descent directions of ψ. For the toy model (1.26), (1.27), this can be done by choosing a control law of the form (1.28) with suitable highly oscillatory functions u_1^j, u_2^j. Then, the closed-loop system (1.5) approximates the averaged system (1.19) with increasing parameter value j. We know from Section 1.2 that this approximation property is not restricted to a single pair of vector fields but can be extended to several pairs of vector fields by choosing suitable highly oscillatory functions with distinct frequency coefficients as, for example, in (1.23). For each $k \in \{1, \ldots, m\}$, we choose two suitable time-varying functions $u_{(k,1)}^j, u_{(k,2)}^j$ to approximate the Lie bracket of $f_{(k,1)}, f_{(k,2)}$. For example, one can choose the sinusoids

$$u_{(k,1)}^j(t) \; := \; \sqrt{2\, j\, \omega_k}\; \cos(j\, \omega_k\, t), \tag{1.44a}$$

$$u_{(k,2)}^j(t) \; := \; \sqrt{2\, j\, \omega_k}\; \sin(j\, \omega_k\, t), \tag{1.44b}$$

where $\omega_1, \ldots, \omega_k$ are pairwise distinct positive real frequency coefficients. Differently shaped highly oscillatory functions can be found, for example, in [99, 111]. For each $k \in \{1, \ldots, m\}$, we propose the parameter-dependent, time-varying output-feedback control law

$$u_k \; = \; u_{(k,1)}^j(t)\, h_1(y) + u_{(k,2)}^j(t)\, h_2(y) \tag{1.45}$$

for (1.39), where y denotes the output (1.40). Then, the closed-loop system reads

$$\dot{x} \; = \; \sum_{k=1}^{m} \big(u_{(k,1)}^j(t)\, f_{(k,1)}(x) + u_{(k,2)}^j(t)\, f_{(k,1)}(x) \big), \tag{1.46}$$

which is of the form (1.20). The same averaging methods as in Section 1.2 shows that the trajectories of (1.46) approximate the trajectories of the averaged system

$$\dot{x} \; = \; \sum_{k=1}^{m} [f_{(k,1)}, f_{(k,2)}](x) \tag{1.47}$$

with increasing frequency parameter j. Let $\dot{\psi}$ denote the derivative of the output function ψ along solutions of (1.47). Using (1.42) for the Lie brackets in (1.47), we obtain that

$$\dot{\psi}(x) \; = \; [h_1, h_2](\psi(x)) \sum_{k=1}^{m} (g_k\psi)(x)^2.$$

Recall that the design functions h_1, h_2 are chosen to satisfy (1.43). Consequently, we obtain the estimate

$$\dot{\psi}(x) \ \leq \ -\sum_{k=1}^{m}(g_k\psi)(x)^2 \ \leq \ 0$$

for the derivative of ψ along solutions of (1.47). Under suitable assumptions on the control vector fields g_k and the output function ψ, one can use ψ as a Lyapunov function and prove that a minimum point $x^* \in \mathbb{R}^n$ of ψ is asymptotically stable for (1.47). Because of the approximation property, this in turn implies that x^* is practically uniformly asymptotically stable for (1.46) as explained earlier for the toy model (1.26). Therefore, the proposed output feedback (1.45) has the desired properties of an extremum seeking control law for (1.39).

The first journal paper that introduces the above approach to extremum seeking control by Lie bracket approximations is [31]. This paper provides local and semi-global practical stability results for a more general situation than what we discussed in the preceding paragraphs. As mentioned earlier, the approach can be also extended to time-varying control-affine systems with a possible drift. Also the choice of the highly-oscillatory functions can be relaxed from sinusoids as in (1.44) to a larger class of time-varying functions that can be characterized by suitable averaging conditions. All these extensions are addressed in [31]. In some earlier conference paper, the method is applied to various problems, such as distributed positioning of autonomous mobile sensors [33], source seeking [32], distance-based synchronization [30], and obstacle avoidance [29]. Inspired by the results in [31], also other research groups started to investigate Lie bracket approach to extremum seeking control. For example, in [95] the output function plays the role of a control Lyapunov function for the purpose of practical stabilization. This allows practical stabilization of linear time-varying system without explicit knowledge of the system's matrices or the system state [93], and applications to the problem of tracking [94]. In the present document, we will also consider the output function as a Lyapunov function to prove stability properties of closed-loop systems. A real-world implementation of a Lie bracket-based extremum seeking algorithm, which optimizes the rise time of the output voltage of a high voltage converter modulator, is documented in [92]. As explained earlier, an extremum seeking control law of the form (1.45) provides a certain degree of freedom in the choice of the highly-oscillatory functions and the design functions h_1, h_2. For example, non-sinusoidal oscillations are applied in [99, 111]. The smooth design functions in (1.32) and (1.33) were introduced in [31] and [96], respectively. A first non-smooth definition of h_1, h_2 appears in [97]. Under certain additional assumptions, such non-smooth design functions do not only lead to practical stability but to asymptotic stability [41, 113] or even exponential stability [111]. We will study such control laws in Chapter 2. A completely different choice of the oscillatory inputs and the design functions is proposed in [109]. The approach in [109] provides the first extremum seeking control law with an adaptive frequency parameter. While all other studies involve the uncertainty of a sufficiently large value of j in (1.45), the method in [109] chooses j adaptively and leads to guaranteed convergence to an optimal state. We will study this control law in Chapter 3.

There are many other extensions and applications of the Lie bracket approach to extremum seeking control from [31]. For example, extensions to control systems on submanifolds of the Euclidean space are studied in [34, 75, 76]. We will see throughout the document that the Lie bracket approach allows a coordinate-free description of extremum seeking control on arbitrary smooth manifolds. A frequently studied application of the

Lie bracket approach is the problem of source seeking, both theoretically [91] and experimentally [40]. Other applications include iterative learning [21], numerical optimization methods [35], and distributed optimization over graphs [72, 73], where the latter requires an approximation of iterated Lie brackets as in (1.25). We will use approximations of iterated Lie brackets in Chapter 4 to derive extremum seeking control for nonholonomic systems. Lie brackets of higher order are also approximated in [59] for the purpose of Newton-based extremum seeking. Interestingly, it is also possible to establish a connection between the Lie bracket averaging theory and singular perturbation theory, which is used to analyze classical extremum controllers as in [55] under a steady state assumption; see [27, 28]. A suitable rescaling of the time parameter reveals that the large-amplitude, high-frequency sinusoids in (1.44) are related to the small-amplitude, low-frequency sinusoids in other extremum seeking control schemes.

All of the above references consider first-order control system; i.e., the first derivative of the system state is controlled directly through the input channels. In the present document, we also restrict our studies to such first-order (or kinematic) models. Extensions of the Lie bracket approach to second-order control systems, like mechanical systems, or higher-order control systems can be found in [70, 71]. However, this results into even larger amplitudes and frequencies of the employed oscillatory signals than for first-order control systems, which might be undesirable for practical implementations. However, for some applications to second-order mechanical system, there is a suitable alternative. Instead of approximating Lie brackets, one can also approximate so-called *symmetric products* of vector fields by using the class of vibrational signals from [17]. Recent studies [108, 114, 112] show that an approximation of symmetric products can be used to design extremum seeking control for mechanical systems, which causes smaller amplitudes and frequencies than the Lie bracket approach.

Finally, it is worth to mention that the Lie bracket approach to extremum seeking control can be viewed as a particular case of a stabilization method that was already pursued in earlier references than [31]. Instead of using output feedback of the form (1.45), it is also possible to apply a state feedback control law of the form

$$u_k \;=\; \sum_{\nu} u^j_{(k,\nu)}(t)\, H_{(k,\nu)}(x) \tag{1.48}$$

to a control-affine system of the form (1.39), where the $u^j_{(k,\nu)}$ are suitable highly oscillatory inputs, and the $H_{(k,\nu)}$ are smooth real-valued design functions on the state space. Note that (1.48) reduces to (1.45) if the summation index ν is restricted to $\nu \in \{1,2\}$ and if $H_{(k,\nu)}(x) = h_\nu(\psi(x))$. If we apply (1.48) to (1.39), then the closed-loop system is of the form (1.20), where the vector fields f_i in (1.20) originate from the products of the functions $H_{(k,\nu)}$ and the control vector fields g_k in (1.39). By choosing suitable highly oscillatory inputs $u^j_{(k,\nu)}$, we can induce that the trajectories of the closed-loop system approximate the trajectories of an extended system of the form (1.25). This in turn implies that if the extended system is asymptotically stable, then the closed-loop system is at least practically asymptotically stable. Thus, a control-affine system of the form (1.39) can be stabilized around a given point x^* if we can find suitable highly oscillatory inputs $u^j_{(k,\nu)}$ and design functions $H_{(k,\nu)}$ in (1.25) such that x^* is asymptotically stable for the extended system of the closed-loop system. Such a procedure is described, for example, in [68]. The particular notion of practical asymptotic stability, which is used in this context, can be traced back to [77, 78]. For example, the same arguments as in [78] to conclude practical asymptotic stability for the closed-loop system from asymptotic stability for the averaged system are also used in [31] in the context of extremum seeking control. Stabilizing state feedback by Lie bracket approximation is intensively studied for homogeneous systems [80, 81, 79]. For homogeneous systems, the approach does not only lead to practical asymptotic stability, but in fact to asymptotic stability or even exponential stability. In particular, the approach in [80] to design state feedback for nonholonomic control systems is closely related to the extremum seeking control law in Chapter 4. A more recent study on exponential stabilization of nonholonomic systems by means of Lie bracket approximations can be also found in [121]. All of these stabilization methods are based on the Lie bracket averaging theory from [58, 68, 69]; cf. Section 1.2. This is in particular true for the findings and the results in the present document.

1.4 Outline

The considerations in the preceding Sections 1.2 and 1.3 have shown that the Lie bracket approach provides a general tool to design extremum seeking control. In many cases, stability properties of the closed-loop system can be proved in a systematic way. Indeed, the procedures in the subsequent chapters follow basically the same pattern as for the toy example (1.26), (1.27) in Section 1.3. The main steps are summarized in Figure 1.7. To allow an easy comparison, the table in Figure 1.7 lists the most relevant equations, definitions, and statements in each of the chapters.

The subsequent Section 1.5 summarizes basic definitions and notations that are used in all later chapters. The rest of the document is organized as follows.

Chapter 2 addresses an optimization problem which is usually not associated with extremum seeking control, namely formation shape control. In this case the control system is a team of autonomous point agents with the common goal to establish a certain formation shape. The formation shape is defined by prescribed distances between the agents. The standard way to solve this problem is a gradient-based control law. For this purpose, each agent is assigned with a suitable local potential function. An implementation of the control law assumes that each agent knows the gradient direction of its local potential function. A computation of the gradient requires measurements of relative positions. However, this means that much more variables need to be sensed (relative positions) than the variables that are actively controlled (relative distances). It is therefore natural to ask whether the formation shape control problem can be also solved if only distance measurements are available. We will see in Chapter 2 that this can be done by using ideas from extremum seeking control. Measurements of inter-agent distances provide already enough information so that each agent can compute the current value of its local potential function. Therefore, we can apply the Lie bracket approach to extract the gradient directions from the current values of the local potential functions. In this case, the Lie bracket approach does not only lead to practical stability but to the novel feature of asymptotic stability and even exponential stability. Additionally, we extend the distance-only formation control law for point agents to a team of nonholonomic unicycles under the assumption of all-to-all communication. As for point agents, we show that the proposed control law induces exponential stability. We also explain how this method can be extended to other optimization problems. We improve the existing extremum seeking methods, which only lead to practical asymptotic stability, by presenting the first extremum seeking control law that can lead to exponential stability.

A common feature of all the existing Lie bracket-based extremum seeking control laws

In each chapter, the optimization problem is described by

(1) a control-affine system

on a certain state manifold together with

(2) a smooth real-valued output function.

The time-varying output-feedback control law is composed of

(3) highly oscillatory functions

and also suitable

(4) design functions

to modulate the measured output signal. These two components determine

(5) the output-feedback control law.

Then, we analyze

(6) the closed-loop system.

This requires a suitable averaging analysis in order to extract

(7) the averaged system.

The closed-loop system can be interpreted as a control-affine system under highly oscillatory open-loop controls. The first task is to derive

(8) estimates for iterated Lie derivatives

along the control vector fields. Then, a suitable averaging analysis for the highly oscillatory functions leads to

(9) estimates for sinusoidal remainders

and an

(10) extraction of the averaged coefficients.

This allows us to derive an

(11) integral expansion for the trajectories of the closed-loop system.

The integral expansion consists of an averaged term and remainder terms. We derive

(12) estimates for the averaged term

as well as

(13) estimates for the remainder terms.

Finally, we use the above preparations to prove the

(14) main stability result

for the closed-loop system.

	Toy example	Section 2.5	Section 2.7	Section 3.5	Section 4.6	Section 4.7
(1)	(1.26)	(2.5), (2.19)	(2.40), (2.43)	(3.8), (3.21)	(4.5)	(4.13)
(2)	(1.27)	(2.9)	(2.44)	(3.9)	(4.17)	(4.6)
(3)	Figure 1.2	(2.8)	(2.45)	(3.25)	Subsection 4.4.2	(4.11)
(4)	(1.33)	(2.7)	(2.7)	(3.17)	not applicable	(4.9)
(5)	(1.28)	(2.11), (2.12)	(2.48), (2.49)	(3.10), (3.26)	(4.21)	(4.14)
(6)	(1.5)	(2.21)	(2.58)	(3.29)	(4.15)	(4.12)
(7)	(1.19)	(2.24), (2.25)	(2.55)	(3.28)	(4.71)	(4.27)
(8)	(1.14)	Lemma 2.19	Lemma 2.33	Lemma 3.9	Lemma 4.11	not applicable
(9)	(1.13)	Lemma 2.22	Lemma 2.36	Lemma 3.10	Lemma 4.13	Lemma 4.17
(10)	(1.17)	Lemma 2.23	Lemma 2.37	Lemma 3.11	Lemma 4.14	Lemma 4.18
(11)	(1.37)	Proposition 2.24	Proposition 2.39	Proposition 3.12	Proposition 4.15	Proposition 4.23
(12)	(1.38)	Proposition 2.25	Proposition 2.40	Proposition 3.13	not applicable	Proposition 4.24
(13)	(1.15)	Proposition 2.26	Proposition 2.41	Proposition 3.14	Proposition 4.16	Proposition 4.25
(14)	page 13	Theorems 2.12, 2.13	Theorem 2.30	Theorem 3.3	Theorem 4.7	

Figure 1.7: *Overview on the repeating structures in the document. The numbers (1)-(14) in the table refer to the list in the text on the left-hand side.*

is that they require the choice of a sufficiently large frequency parameter for the employed perturbation signals. Otherwise no stability property can be guaranteed and even finite escape times can occur. This is certainly a problem in practical implementations. All of the existing studies do not provide explicit information on how large the frequency parameter has to be chosen for a successful implementation. The proposed method in Chapter 3 provides the first solution to this problem. The idea is to choose the amplitudes and frequencies of the oscillatory signals in an adaptive fashion. An increase of the measured output value automatically leads to larger amplitudes and frequencies. This way, the control law itself chooses a suitable frequency parameter. Under suitable assumptions, the proposed control strategy does not only lead to convergence of the system state into a neighborhood of the optimal state but to exact convergence. Moreover, the control law has the ability to compensate the influence of an arbitrary drift vector field. In particular, finite escape times cannot occur.

As explained in Section 1.3, the Lie bracket approach to extremum seeking control gives access to descent directions of the output function along the control vector fields of a system. A proof of (practical) asymptotic stability for the closed-loop system usually requires that the averaged Lie bracket system is asymptotically stable. For this reason, many studies implicitly assume that the linear span of the control vector field contains a proper descent direction of the output function. The same is true for the results in Chapters 2 and 3. In general, however, this condition requires that the control system is fully actuated; i.e., the control vector fields span the entire tangent space. Otherwise, the averaged system might have undesired equilibrium points at which all Lie derivatives of the output function along the control vector fields vanish. In this case, the existing results cannot guarantee stability properties of the closed-loop system. This problem motivates the investigations on extremum seeking control for nonholonomic system in Chapter 4. There are many examples of control-affine systems which are not fully actuated, but have at least the property that the Lie brackets of their control vector fields span the entire tangent space (also known as the *Lie algebra rank condition*). This feature can be used to design an extremum seeking control law that leads to the same practical asymptotic stability results as for fully actuated systems. The idea is to induce a two-fold approximation of Lie brackets. In the first step, we approximate Lie brackets of the control vector fields. For a suitable class of nonholonomic systems, this gives access to all directions of the tangent space. In the second step, we use the ideas from Section 1.3 and approximate descent directions of the output function along Lie brackets of the control vector fields. Under standard assumptions on the output function, this approach leads to practical asymptotic stability for the closed-loop system.

1.5 Global definitions and notation for the entire document

By a *smooth manifold* we mean a second-countable Hausdorff space endowed with a real finite-dimensional smooth structure; see [62]. The word smooth always means of class C^∞. The word *function* will be only used for maps whose codomain is the set of real numbers. The notion of a smooth manifold allows the definitions of basic objects like *tangent spaces*[1],

[1] Let $C^\infty(M)$ denote the algebra of smooth functions on a smooth manifold M. For the objectives in the present document it is convenient to treat a *tangent vector* to M at a point x of M as a *derivation* on $C^\infty(M)$ at x; i.e., a linear function v_x on $C^\infty(M)$ such that $v_x(\varphi\psi) = (v_x\varphi)\psi(x) + \varphi(x)(v_x\psi)$ for all $\varphi, \psi \in C^\infty(M)$

vector fields[2], *smooth maps*[3], and so on, which are not recalled here. Instead we refer the reader to standard textbooks like [1], [18], and [62].

Let ψ be a function on a smooth manifold M. For every real number y, we denote the *fiber of ψ over y* by $\psi^{-1}(y)$; i.e., the set of all $x \in M$ with $\psi(x) = y$. Let r be either a real number or the symbol $+\infty$. We denote the *r-sublevel set of ψ* by $\psi^{-1}(\leq r)$; i.e., the (possibly empty) set of all $x \in M$ with $\psi(x) \leq r$. Let $x^* \in M$. We define a set $\psi^{-1}(\leq r, x^*)$ as follows. If x^* is not contained in $\psi^{-1}(\leq r)$, then $\psi^{-1}(\leq r, x^*)$ denotes the empty set, and otherwise, it denotes the connected component of $\psi^{-1}(\leq r)$ containing x^*. We say that *ψ attains a local minimum at x^** if there exists a neighborhood V of x^* in M such that $\psi(x^*) \leq \psi(x)$ for every $x \in V$.

Let v_x be a tangent vector at some point x of M. If there exists a real number, denoted by $v_x\psi$, such that, for every smooth curve γ from an open interval around 0 into M with $\gamma(0) = x$ and $\dot{\gamma}(0) = v_x$, the derivative of $\psi \circ \gamma$ exists at 0 and coincides with $v_x\psi$, then $v_x\psi$ is called the *directional derivative of ψ along v_x*. Let f be a vector field on M. If the directional derivative of ψ along the tangent vector $f(x)$ exists, then we denote it by $(f\psi)(x) := f(x)\psi$ and call it the *Lie derivative of ψ along f at x*. If the Lie derivative of ψ along f exists at every point of M, then the resulting function $f\psi$ on M is called the *Lie derivative of ψ along f*. Suppose that ψ is differentiable at x. Then, the *derivative of ψ at x* is the linear function $\mathrm{D}\psi(x)$ on the tangent space to M at x such that $\mathrm{D}\psi(x)v_x = v_x\psi$ for every tangent vector v_x at x. Moreover, the Lie derivative of f along ψ at x is then given by the well-known formula

$$(f\psi)(x) \;=\; \mathrm{D}\psi(x)f(x).$$

If $\mathrm{D}\psi(x) = 0$, then x is said to be a *critical point of ψ*, and otherwise it is said to be a *regular point of ψ*. Let g be another vector field on M. Suppose that, for every smooth function φ on M, both the Lie derivative of $g\varphi$ along f and the Lie derivative of $f\varphi$ along g exist at x. Then, there exists a unique tangent vector $[f, g](x)$ to M at x, called the *Lie bracket of f and g at x* such that

$$[f, g](x)\varphi \;=\; (f(g\varphi))(x) - (g(f\varphi))(x)$$

for every smooth function φ on M. For example, if f, g are locally Lipschitz continuous and if f, g vanish at x, then their Lie bracket exists at x and vanishes there. If f, g are smooth, then one can show that the above algebraic definition of the Lie bracket coincides with the geometric definition of the Lie bracket in Section 1.2.

Suppose that ψ is smooth and that $x \in M$ is a critical point of ψ. Then, $([f, g]\psi)(x) = 0$ and therefore $(f(g\psi))(x) = (g(f\psi))(x)$ for all smooth vector fields f, g on M. It is now easy to show that there exists a unique symmetric bilinear function $\mathrm{D}^2\psi(x)$ of the tangent space to M at x, called the *second derivative of ψ at x*, such that

$$\mathrm{D}^2\psi(x)(f(x), g(x)) \;=\; (f(g\psi))(x) \;=\; (g(f\psi))(x)$$

for all smooth vector fields f, g on M; see, e.g., [74]. The second derivative of ψ at x is said to be *positive definite* if $\mathrm{D}^2\psi(x)(v_x, v_x) > 0$ for every nonzero tangent vector v_x to M at x. The following statements are known from real analysis; see, e.g., [18].

[2] For the objectives in the present document it is convenient to treat a *vector field* on a smooth manifold M as a *derivation* on the algebra $C^\infty(M)$ of smooth functions on M; i.e., a linear map f from $C^\infty(M)$ into the algebra of functions on M such that $f(\varphi\psi) = (f\varphi)\psi + \varphi(f\psi)$ for all $\varphi, \psi \in C^\infty(M)$.

[3] For instance, a vector field f on a smooth manifold M is *smooth* if and only if, for every smooth function φ on M, an application of the derivation f to φ results into a smooth function $f\varphi$ on M.

Remark 1.2. Let ψ be a smooth function on a smooth manifold M. Suppose that ψ attains a local minimum value $y^* \in \mathbb{R}$ at some $x^* \in M$ and that the second derivative of ψ at x^* is positive definite. Then the following statements hold.

(i) The function ψ attains a *strict local minimum at* x^*; i.e., there exists a neighborhood V of x^* in M such that $\psi(x) > y^*$ for every $x \in V$ with $x \neq x^*$.

(ii) There exists a neighborhood V of x^* in M such that every $x \in V$ with $x \neq x^*$ is a regular point of ψ.

(iii) For every neighborhood V of x^*, there exists $\tilde{y} > y^*$ such that $\psi^{-1}(\leq \tilde{y}, x^*)$ is a compact subset of V. $\Diamond$

If a is function on an interval I and if $t_1, t_2 \in I$, then we use the notation

$$\big[a(t)\big]_{t=t_1}^{t=t_2} \;:=\; a(t_2) - a(t_1).$$

Additionally, if a is locally integrable[4], then we use the standard convention

$$\int_{t_1}^{t_2} a(t)\,\mathrm{d}t \;:=\; -\int_{t_2}^{t_1} a(t)\,\mathrm{d}t$$

for the integral if $t_1 > t_2$. If A, B are locally absolutely continuous functions[5] on an interval I, and if $t_1, t_2 \in I$, then, using the above notation, integration by parts can be written as

$$\int_{t_1}^{t_2} A(t)\,\dot{B}(t)\,\mathrm{d}t \;=\; \big[A(t)\,B(t)\big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \dot{A}(t)\,B(t)\,\mathrm{d}t.$$

For multiple later references, we state the well-known trigonometric identity

$$\sin(\alpha+\beta) \;=\; \cos(\alpha)\,\sin(\beta) + \sin(\alpha)\,\cos(\beta) \tag{1.49}$$

for all $\alpha, \beta \in \mathbb{R}$.

[4] Every integral in the present document is meant as the standard Lebesgue integral. Recall that a function a on an interval I is said to be *locally integrable* if, for all real numbers $t_1 \leq t_2$ in I, the integral of a over the compact interval $[t_1, t_2]$, denoted by $\int_{t_1}^{t_2} a(t)\,\mathrm{d}t$, exists as a real number.

[5] Recall that a function A on an interval I is said to be locally absolutely continuous if its derivative $\dot{A}$ exists almost everywhere on I as a locally integrable function and $\int_{t_1}^{t_2} \dot{A}(t)\,\mathrm{d}t = \big[A(t)\big]_{t=t_1}^{t=t_2}$ for all $t_1, t_2 \in I$.

2 Distance-based formation control

The content of this chapter is an extended version of [111] and [113].

2.1 Introduction and motivation

Distance-based formation control is an extensively studied subject in the field of autonomous multi-agent systems. The wish to achieve and maintain prescribed distances among autonomous agents in a distributed way arises in various applications such as leader-follower systems or in the context of formation shape control [86]. This task becomes especially difficult if the agents can measure only distances to other members of the team but not their relative positions.

In this chapter, we focus on the model of kinematic points in the Euclidean space of arbitrary dimension. The interaction topology is described by an undirected graph, where each node represents one of the agents. When we connect the current positions of the agents by line segments according to the edges of the graph, we obtain a graph in the Euclidean space, which is also referred to as a formation. We study the problem of distance-based formation control; i.e., the target formations are defined by distances. To be more precise, a target formation is reached if, for each edge of the graph, the distance between the corresponding pair of agents is equal to a desired value. These distances are the actively controlled variables. The aim is to find a distributed control law that steers the agents into one of the target formations. The agents have to accomplish this goal without any shared information like a global coordinate system or a common clock to synchronize their motion.

A well-established approach to solve the above problem is a gradient descent control law [54, 24, 85, 82, 103]. For this purpose, every agent is assigned with a local potential function. These functions penalize deviations of the distances to the prescribed values. Each local potential function is defined in such a way that it attains its global minimum value if and only if the distances to the neighbors are equal to the desired values. Thus, a target formation is reached if all agents have minimized the values of their local potential functions. To reach the minimum, every agent follows the negative gradient direction of its local potential function. It is shown in [54, 24, 85] that this approach can lead to local asymptotic stability with respect to the set of desired states. In fact, by imposing suitable rigidity assumptions on the target formations, one can prove local exponential stability; see, e.g., [82, 103].

An implementation of the gradient descent control law requires that all agents should be able to measure the *relative positions* to their neighbors in the underlying graph. It is clear that relative positions contain much more information than distances. In other words, the amount of sensed variables exceeds the amount of actively controlled variables. It is therefore natural to ask whether distance-based formation control is still possible even if the sensed variables coincide with the controlled variables. This means that each agent can only use its own real-time distance measurements to steer itself into a target formation. We also remark that distance sensing and measurement has emerged as a mature technique through the development of many low-cost, high precision sensors, such as ultrasonic sensors or laser scanners (see, e.g., the survey in [46]). Therefore, it motivates us to explore feasible solutions to formation control with distance-only measurement, which also finds significant applications in relevant areas, e.g., multi-robotic coordination in practice.

To our best knowledge, there are just a few studies on formation control by distance-only measurements. The idea in [4] is to compute relative positions directly from distance measurements. However, in order to do so, the agents need more information than just the distances to their neighbors in the underlying graph. It is shown in [4] that if the target frameworks are rigid, and if each agent also has access to the distances to its two-hop neighbors, then they can compute the relative positions by means of a Cholesky factorization of a suitable matrix, which is obtained from distance measurements. Since this factorization is only unique up to an orthogonal transformation, each agent also has to harmonize these relative positions with its individual coordinate system. This requires a certain ability to sense bearing. Thus, it is not sufficient to sense only the actively controlled distances.

Another approach is presented in [19]. In contrast to the above strategy, it suffices that each agent measures the distances to its neighbors in the underlying graph. The multi-agent system is divided into subgroups. Following a prescribed schedule, only one of these subgroups is active at a time while the other agents remain at their positions. This requires that the agents share a common clock. It is assumed that the agents of the currently active group have the ability to first localize the resting neighbors of the team by means of distance measurements, and then move into the best possible position. Note that the strategy requires that each agent can map and memorize its own motion within its own local coordinate system. For a minimally rigid graph in the plane, this algorithm leads locally to the desired convergence. However, a generalization to higher dimensions is limited, since the strategy requires a minimally rigid graph that can be constructed by means of a so-called Henneberg sequence [5], which is, in general, possible only in two dimensions.

A recent attempt to control formation shapes by distance-only measurements can be found in [51]. In this case, the agents perform suitable circular motions with commensurate frequencies. Using collected data from distance measurements during a prescribed time interval, each agent can extract relative positions and relative velocities of its neighbors by means of Fourier analysis. As in [19], the approach in [51] relies on the assumption that the agents share a precise common clock to synchronize their motions. The proposed strategy leads to convergence if certain control parameters are chosen sufficiently small. However, only existence of these parameters can be ensured but there is no explicit rule how to obtain them. Moreover, the control law only induces convergence to the set of desired formations but not convergence to a single static formation. In general, a common drift of the multi-agent system remains. An extension to higher dimensions is not obvious, since the extraction of relative positions and velocities relies on the geometry of the plane.

A common feature of all of the above strategies is that the agents should be able to compute or infer relative positions from distance measurements. In this chapter, we use a different approach. To explain the idea, we return to the gradient descent control law. In this case, each agent tries to minimize its own local potential function by moving into the negative gradient direction. A computation of the gradient requires measurements of relative positions. However, the value of each local potential function can be computed from individual distance measurements, and is therefore accessible to every agent. This leads to the question of whether an agent can find the minimum of its local potential function when only the values of the function are available. To solve this problem, we use an approach that was recently introduced in the context of extremum seeking control, see, e.g., [31, 34, 27, 95, 97, 96]. The reader is referred to Chapter 1 for an introduction to this method. By feeding in suitable sinusoidal perturbations, we induce that the agents are driven, at least approximately, into descent directions of their local potential functions. On average, this leads to a decay of all local potential functions, and therefore convergence to a target formation. The proposed control law for each agent needs no other information than the current value of the local potential function. Under the assumption that the target formations are infinitesimally rigid (see Section 2.2 for the definition), we can ensure local uniform asymptotic stability. Our control strategy is fully distributed, and can be applied to point agents in any finite dimension.

An earlier attempt to apply Lie bracket approximations to the problem of formation shape control can be found in [111]. The control law therein requires a permanent all-to-all communication between the agents for an exchange of distance information. The control law in this chapter is based on individual distance measurements and works without any exchange of measured data. Moreover, the results in [111] contain an unknown frequency parameter for the sinusoidal perturbations. It is assumed that the frequency parameter is chosen sufficiently large; otherwise convergence to a desired formation cannot be guaranteed. The results in the above paper provides only the existence of a sufficiently large frequency parameter, but there is no explicit rule on how to find that value. The control law in this chapter can lead to local uniform asymptotic stability even if the frequency parameter is chosen arbitrarily small. We discuss the influence of the frequency parameter on the performance of our control law in the main part.

The idea of using Lie bracket approximations to extract directional information from distance measurements can also be found in several other studies. The range of applications includes, among others, multi-agent source seeking [32], synchronization [30], and obstacle avoidance [29]. A common feature of the above papers and the content of this chapter is that the desired states are characterized by minima of (artificial) potential functions. Another similarity is that a purely distance-based control law is derived by using Lie bracket approximations in order to get access to the direction of steepest decent. However, the above studies only guarantee practical asymptotic stability, and depend on the unknown frequency parameter that we mentioned in the previous paragraph. Our results for formation shape control are stronger because they can lead to local asymptotic stability without the dependence on the frequency parameter. Thus, our findings might also be of interest to the above fields of applications.

The chapter is organized as follows. In Section 2.2, we introduce basic definitions and notations, which we use throughout the chapter. As indicated above, our approach involves the notion of infinitesimal rigidity, which is recalled in Section 2.3. We also derive suitable estimates for the derivatives of the potential functions in this section. The distance-only control law and the main stability results are presented in Section 2.4, which are supported by numerical simulations in the same section. A detailed analysis of the closed-loop system and the proofs of the main theorems are carried out in Section 2.5. In addition to formation control for point agents, we show in Section 2.6 that the Lie bracket approach can be extended to nonholonomic unicycles under the additional assumption of all-to-all communication among the agents. Again, we present a stability result under the proposed control strategy and then prove the theorem in Section 2.7. Finally, in Section 2.8, we indicate how the approach can be extended for the purpose of extremum seeking control. If the minimum value of the output function is known, then basically the same control strategy as for point agents can be applied to a more general type of control-affine system. Under suitable assumptions, this approach leads to asymptotic (and even exponential) stability for the closed-loop system.

2.2 Local definitions and notation for the chapter

To emphasize the difference between points and tangent vectors in the notation, we use the notion of a Euclidean space $P \cong \mathbb{R}^n$ of (finite and nonzero) dimension n with underlying translation space $T \cong \mathbb{R}^n$ (see [84]). The elements of P are called *points* and the elements of T are called *translations*. The Euclidean inner product on T is denoted by $\langle\langle \cdot, \cdot \rangle\rangle$ and the induced norm is denoted by $\| \cdot \|$. Let ϕ be a map from an open subset U of P into $\mathbb{R}^m$. If ϕ is k-times differentiable at some point x of U, then we denote its kth derivative at x by $\mathrm{D}^k\phi(x)$, which is a k-linear map of T into $\mathbb{R}^m$. The induced operator norm of $\mathrm{D}^k\phi(x)$ is denoted by $\| \mathrm{D}^k\phi(x)\|$. As usual, for any subset S of U, we say that ϕ is Lipschitz continuous on S if there exists a Lipschitz constant $L > 0$ such that $\|\phi(x_2) - \phi(x_1)\| \leq L\|x_2 - x_1\|$ for all x_1, x_2 in S, where we use the same symbol for the norm on $\mathbb{R}^m$ and the norm on T. If each point of U has a neighborhood on which ϕ is Lipschitz continuous, then ϕ is said to be *locally Lipschitz continuous*. Equivalently, ϕ is locally Lipschitz continuous if and only if ϕ is Lipschitz continuous on every compact subset of U. If a function φ on U is differentiable at some point x of U, then the *gradient of* φ *at* x is the unique translation vector $\nabla\varphi(x) \in T$ that satisfies $\mathrm{D}\varphi(x)(v) = \langle\langle \nabla\varphi(x), v \rangle\rangle$ for every $v \in T$. For later references, we collect the following statements on nonnegative smooth functions.

Lemma 2.1. *Let φ be a nonnegative smooth function on an open subset U of a Euclidean space P. Then the following statements hold.*

(a) The square root of φ is locally Lipschitz continuous.

(b) For every compact subset K of U, there exists $c_1 > 0$ such that $\|\nabla\varphi(x)\|^2 \leq c_1\, \varphi(x)$ for every $x \in K$.

(c) Suppose that, for some $x^* \in U$, we have $\varphi(x^*) = 0$ and the second derivative of φ at x^* is positive definite. Then, there exist $c_0 > 0$ and a neighborhood V of x^* in U such that $\|\nabla\varphi(x)\|^2 \geq c_0\, \varphi(x)$ for every $x \in V$.

Proof. Statement (a) is a particular case of the more general result that every nonnegative definite matrix-valued smooth map has a locally Lipschitz continuous square root. The proof can be found in [37]. A proof of statements (b) and (c) for nonnegative smooth functions on Riemannian manifolds can be found in [18]. □

Since we restrict our considerations to a Euclidean space P with translation space T, each tangent space to P can be identified with T. Therefore, a vector field on P can be simply considered as a map from P into T. For example, the gradient of a differentiable function on P is a vector field on P. If f, g are vector fields on P and if g is differentiable at some $x \in P$, then the ("covariant") derivative $\mathrm{D}g(x)f(x)$ of g with respect to f at x is denoted by $\nabla_f g(x)$.

2.3 Infinitesimal rigidity and gradient estimates

2.3.1 Infinitesimal rigidity

In this subsection, we recall several definitions and statements from [7, 8].

An (*undirected*) *graph* $G = (V, E)$ is a set $V = \{1, \ldots, N\}$ together with a nonempty set E of two-element subsets[1] of V. Each element of V is referred to as a *vertex* of G and

[1] Note that self-loops are excluded by requiring that E consists of two-element subsets of V.

each element of E is called an *edge* of G. As an abbreviation, we denote an edge $\{i_1, i_2\} \in E$ simply by $i_1 i_2$. A *framework* $G(p)$ in P consists of a graph G with N vertices and a point

$$p \;=\; (p_1, \ldots, p_N) \;\in\; P \times \cdots \times P \;=:\; P^N.$$

Note that for a framework $G(p)$ in P, we may have $p_{i_1} = p_{i_2}$ for $i_1 \neq i_2$.

Consider a graph $G = (V, E)$ with N vertices and M edges, that is, $V = \{1, \ldots, N\}$, and E has M elements. Order the M edges of G in some way and define the *edge map* $e_G \colon P^N \to \mathbb{R}^M$ of G by

$$e_G(p) \;:=\; (\ldots, \|p_{i_2} - p_{i_1}\|^2, \ldots)_{i_1 i_2 \in E}$$

for every $p = (p_1, \ldots, p_N) \in P^N$. Thus, the value of e_G at any $(p_1, \ldots, p_N) \in P^N$ is a vector that collects the squared distances $\|p_{i_2} - p_{i_1}\|^2$ for all edges $i_1 i_2 \in E$. A point $p \in P^N$ is said to be a *regular point* of e_G if the rank of $\mathrm{D}e_G$ attains its global maximum value at p. For later references, we state the following result from [7], which is an easy consequence of the Inverse Function Theorem.

Proposition 2.2. *Let G be a graph with N vertices and M edges. If $p \in P^N$ is a regular point of e_G, then there exists an open neighborhood U of p in P^N such that the image of U under e_G is a smooth submanifold of $\mathbb{R}^M$ of dimension* $\operatorname{rank} \mathrm{D}e_G(p)$.

The *complete graph with N vertices* is the graph with N vertices that has each two-element subset of $\{1, \ldots, N\}$ as an edge.

Definition 2.3. Let G be a graph with N vertices, let C be the complete graph with N vertices, and let $p \in P^N$. The framework $G(p)$ in P is said to be *rigid* if there exists a neighborhood U of p in P^N such that

$$e_G^{-1}(e_G(p)) \cap U \;=\; e_C^{-1}(e_C(p)) \cap U,$$

where $e_G^{-1}(e_G(p))$ denotes the fiber of e_G over $e_G(p)$ and $e_C^{-1}(e_C(p))$ denotes the fiber of e_C over $e_C(p)$. $\Diamond$

Thus, a framework $G(p)$ is rigid if and only if, whenever q sufficiently close to p with $\|q_{i_2} - q_{i_1}\| = \|p_{i_2} - p_{i_1}\|$ for every edge $i_1 i_2$ of G, we have in fact $\|q_{i_2} - q_{i_1}\| = \|p_{i_2} - p_{i_1}\|$ for all vertices i_1, i_2 of G. Another result from [7] is the following.

Proposition 2.4. *Let C be the complete graph with N vertices. For every $p \in P^N$, the set $e_C^{-1}(e_C(p))$ is a smooth submanifold of P^N.*

The manifold $e_C^{-1}(e_C(p))$ is actually analytic and one can derive an explicit formula for its dimension; see again [7]. As in [8], we use the manifold structure of $e_C^{-1}(e_C(p))$ to introduce the notion of infinitesimal rigidity.

Definition 2.5. A framework $G(p)$ in P is said to be *infinitesimally rigid* if the tangent space to the smooth submanifold $e_C^{-1}(e_C(p))$ at p coincides with the kernel of $\mathrm{D}e_G(p)$. $\Diamond$

To make the notion of infinitesimal rigidity more intuitive, we recall a geometric interpretation from [38]. For this purpose, we consider *smooth isometric deformations* of a given framework $G(p)$; i.e., smooth curves from an open time interval around 0 into the set $e_G^{-1}(e_G(p))$ passing through p at time 0. By definition, each such curve $\gamma = (\gamma_1, \ldots, \gamma_N)$

preserves the squared distances $\|\gamma_{i_2}(t) - \gamma_{i_1}(t)\|^2$ for all edges $i_1 i_2$ of G, and we have $e_G(\gamma(t)) = e_G(p)$ for every t in the domain of γ. By the chain rule, this implies that the velocity vector $\dot{\gamma}(0)$ of γ at time 0 is an element of the kernel of $\mathrm{D}e_G(p)$ (which is termed *rigidity matrix* in the literature of graph rigidity; see, e.g., [8]). This explains why vectors in the kernel of $\mathrm{D}e_G(p)$ are referred to as *infinitesimal isometric perturbations* of $G(p)$. On the other hand, for the complete graph C, the tangent space to the smooth manifold $e_C^{-1}(e_C(p))$ at p consists of the velocities of all smooth curves in $e_C^{-1}(e_C(p))$ passing through p. By definition, the curves in $e_C^{-1}(e_C(p))$ preserve the squared distances for all vertices of G. Thus, infinitesimal rigidity of $G(p)$ means that, for every smooth curve γ of the form $\gamma(t) = p+tv$ with v being an infinitesimal isometric perturbations of $G(p)$, changes of the squared distances $\|\gamma_{i_2}(t) - \gamma_{i_1}(t)\|^2$ are not detectable around $t = 0$ in *first-order* terms for all vertices i_1, i_2 of G.

For our purposes, it is more convenient to characterize the notion of infinitesimal rigidity by the following result from [8].

Theorem 2.6. *A framework $G(p)$ in P is infinitesimally rigid if and only if p is a regular point of e_G and if $G(p)$ is rigid.*

It follows that the notions of rigidity and infinitesimal rigidity coincide at regular points of the edge map. Finally, we note that it is also possible to characterize infinitesimal rigidity of $G(p)$ in P by means of an explicit formula for $\operatorname{rank} \mathrm{D}e_G(p)$; see again [8].

2.3.2 Gradient estimates

In this subsection, $G = (V, E)$ is a graph with N vertices and M edges. Let $e_G\colon P^N \to \mathbb{R}^M$ be the edge map of G. For each edge $i_1 i_2 \in E$, let $d_{i_1 i_2}$ be a nonnegative real number. Define $d := (d_{i_1 i_2}^2)_{i_1 i_2 \in E} \in \mathbb{R}^M$, where the components of d are ordered in the same way as the components of e_G. Define a nonnegative smooth function $\psi_{G,d}$ on P^N by

$$\psi_{G,d}(p) \;:=\; \frac{1}{4}\,\|e_G(p) - d\|^2 \;=\; \frac{1}{4}\sum_{i_1 i_2 \in E} \big(\|p_{i_2} - p_{i_1}\|^2 - d_{i_1 i_2}^2\big)^2 \tag{2.1}$$

for every $p = (p_1, \ldots, p_N) \in P^N$. This type of function will appear again in the subsequent sections as local and global potential function of a system of N agents in P. Our aim is to derive boundedness properties for the gradient of $\psi_{G,d}$. As in Section 1.5, we denote the r-sublevel set of $\psi_{G,d}$ by $\psi_{G,d}^{-1}(\leq r)$.

Proposition 2.7. *For the function $\psi_{G,d}$ on P^N in* (2.1)*, the following statements hold.*

(a) For every $r > 0$, the square root of $\psi_{G,d}$ is Lipschitz continuous on $\psi_{G,d}^{-1}(\leq r)$.

(b) For every $r > 0$, there exists $c_1 > 0$ such that

$$\|\nabla\psi_{G,d}(p)\|^2 \;\leq\; c_1\,\psi_{G,d}(p) \tag{2.2}$$

for every $p \in \psi_{G,d}^{-1}(\leq r)$.

(c) For every $r > 0$ and every integer $k \geq 2$, there exists $c_2 > 0$ such that

$$\|\,\mathrm{D}^k\psi_{G,d}(p)\| \;\leq\; c_2 \tag{2.3}$$

for every $p \in \psi_{G,d}^{-1}(\leq r)$.

(d) *Suppose that, for each* $p \in \psi_{G,d}^{-1}(0)$, *the framework* $G(p)$ *is infinitesimally rigid. Then, there exist* $r_0, c_0 > 0$ *such that*

$$\|\nabla\psi_{G,d}(p)\|^2 \;\geq\; c_0\,\psi_{G,d}(p) \tag{2.4}$$

for every $p \in \psi_{G,d}^{-1}(\leq r_0)$.

Proof. For the proof of Proposition 2.7, we need some additional facts from differential geometry, which can be found in [62]. An *isometry of* P is a map $\alpha\colon P \to P$ such that $\|\alpha(x_2) - \alpha(x_1)\| = \|x_2 - x_1\|$ for all $x_1, x_2 \in P$. It is known that the set $\mathrm{E}(n)$ of all isometries of P forms a Lie group, called the *Euclidean group*. For each $\alpha \in \mathrm{E}(n)$, we define $\alpha^N\colon P^N \to P^N$ by $\alpha^N(p) := (\alpha(p_1), \ldots, \alpha(p_N))$ for every $p = (p_1, \ldots, p_N) \in P^N$. It is known that the map $\mathrm{E}(n) \times P^N \to P^N$, $(\alpha, p) \mapsto \alpha^N(p)$ is a smooth group action of $\mathrm{E}(n)$ on P^N. For every subset S of P^N, we let $S^{\mathrm{E}(n)}$ denote the set of all $\alpha^N(p)$ with $p \in S$ and $\alpha \in \mathrm{E}(n)$. In particular, for a single point $p \in P^N$, the set $\{p\}^{\mathrm{E}(n)}$ is called the *orbit* of p under the action of $E(n)$. The set $P^N/\,\mathrm{E}(n)$ of all orbits endowed with the quotient topology is called the *orbit space*. Note that $\psi_{G,d}$ is *invariant* under the action of $\mathrm{E}(n)$; i.e., we have $\psi_{G,d} \circ \alpha^N = \psi_{G,d}$ for every $\alpha \in \mathrm{E}(n)$. It is easy to check that every sublevel set of $\psi_{G,d}$ can be reduced to a compact set by isometries; i.e., for every $r > 0$, there exists a compact subset K of P^N such that $\psi_{G,d}^{-1}(\leq r) = K^{\mathrm{E}(n)}$.

To prove parts (a), (b), and (c), fix an arbitrary $r > 0$. Then, there exists a compact subset K of P^N such that $\psi_{G,d}^{-1}(\leq r) = K^{\mathrm{E}(n)}$. Suppose for the sake of contradiction that $\psi_{G,d}$ is not Lipschitz continuous on $\psi_{G,d}^{-1}(\leq r)$. Then, using the invariance of $\psi_{G,d}$ under the action of $\mathrm{E}(n)$ and the compactness of K, it follows that there exist sequences of points p^j, q^j in $\psi_{G,d}^{-1}(\leq r)$ with $p^j \neq q^j$ that converge to a common point $p^\infty \in K$ and such that $|\psi_{G,d}^{1/2}(q^j) - \psi_{G,d}^{1/2}(p^j)|/\|q^j - p^j\|$ tends to infinity as $j \to \infty$. However, this would contradict Lemma 2.1 (a), which states that the square root of $\psi_{G,d}$ is at least locally Lipschitz continuous. Next, we prove part (b). By Lemma 2.1 (b), there exists $c_1 > 0$ such that (2.2) holds for every $p \in K$. Note that the derivative of any $\alpha \in \mathrm{E}(n)$ is identically equal to an orthogonal map of T and therefore leaves the norm on T invariant. By the chain rule, we obtain that $\|(\nabla\psi_{G,d}) \circ \alpha^N\| = \|\nabla\psi_{G,d}\|$ for every $\alpha \in \mathrm{E}(n)$, which implies that (2.2) holds in fact for every $p \in K^{\mathrm{E}(n)}$. Let $k \geq 2$ be an integer. Since $\psi_{G,d}$ is smooth, there exists $c_2 > 0$ such that (2.3) holds for every $p \in K$. As for the gradient, it follows from the invariance of $\psi_{G,d}$ under the action of $\mathrm{E}(n)$, the chain rule, and the invariance of the norm under orthogonal transformations that (2.3) holds for every $p \in K^{\mathrm{E}(n)}$.

It remains to prove part (d). For the rest of the proof, we suppose that $G(p)$ is infinitesimally rigid for every $p \in \psi_{G,d}^{-1}(0)$. Note that $\psi_{G,d}^{-1}(0) = e_G^{-1}(d)$, where e_G is the edge map from Subsection 2.3.1. For the moment, fix an arbitrary $q \in e_G^{-1}(d)$. We will show that there exist a neighborhood W of q in P^N and some constant $c_0 > 0$ such that (2.4) holds for every $p \in W$. By Proposition 2.2 and Theorem 2.6, there exists an open neighborhood U of q in P^N such that the image $e_G(U)$ of U under e_G is a smooth submanifold of $\mathbb{R}^M$ of dimension $k := \operatorname{rank} \mathrm{D}e_G(q)$. After possibly shrinking U around q, we can find a parametrization $\phi\colon V \to e_G(U)$ for the entire manifold $e_G(U)$. Then, $\bar{e}_G := (\phi^{-1} \circ e_G)|_U\colon U \to V$ is a smooth map with $\operatorname{rank} \mathrm{D}\bar{e}_G(q) = k$. Define a smooth function g_d on V by $g_d(x) := \|\phi(x) - d\|^2/4$ for every $x \in V$. Then, the restriction of $\psi_{G,d}$ to U equals $g_d \circ \bar{e}_G$, and by the chain rule, we obtain

$$\nabla\psi_{G,d}(p) \;=\; \mathrm{D}\bar{e}_G(p)^\top \nabla g_d(\bar{e}_G(p))$$

for every $p \in U$, where $\mathrm{D}\bar{e}_G(p)^\top : \mathbb{R}^k \to T^N$ denotes the adjoint[2] of $\mathrm{D}\bar{e}_G(p) : T^N \to \mathbb{R}^k$ with respect to the inner products on $\mathbb{R}^k$ and T^N. Since $p \mapsto \mathrm{D}\bar{e}_G(p)^\top$ is continuous and has maximum rank k at q, there exist[3] a neighborhood W of q in U and a constant $c_0' > 0$ such that $\| \mathrm{D}\bar{e}_G(p)^\top v\| \geq c_0'\|v\|$ for every $p \in W$ and every $v \in \mathbb{R}^k$. In particular, this implies

$$\|\nabla\psi_{G,d}(p)\| \;\geq\; c_0' \, \|\nabla g_d(\bar{e}_G(p))\|$$

for every $p \in W$. Using $\phi(z) = d$ at $z := \bar{e}_G(q) \in V$, a direct computation shows that $\mathrm{D}^2 g_d(z)(v,v) = \|\mathrm{D}\phi(z)v\|^2/2$ for every $v \in \mathbb{R}^k$. Since $\operatorname{rank}\mathrm{D}\phi(z) = k$, it follows that the second derivative of g_d at z is positive definite. Because of Lemma 2.1 (c), we can shrink W sufficiently around q and find some $c_0'' > 0$ such that

$$\|\nabla g_d(\bar{e}_G(p))\|^2 \;\geq\; c_0'' \, g_d(\bar{e}_G(p)) \;=\; c_0'' \, \psi_{G,d}(p)$$

for every $p \in W$. Thus, (2.4) holds for every $p \in W$ with $c_0 := (c_0')^2 \, c_0''$.

Let $\pi\colon P^N \to P^N/\,\mathrm{E}(n)$ be the projection onto the orbit space. Let C be the complete graph with N vertices. For the edge map e_c of C, it is known (see [84]) that $e_C^{-1}(e_C(p)) = \{p\}^{E(n)}$ for every $p \in P^N$. Note that both e_C and e_G are continuous, and also invariant under the action of $\mathrm{E}(n)$; i.e., we have $e_C \circ \alpha^N = e_C$ and $e_G \circ \alpha^N = e_G$ for every $\alpha \in \mathrm{E}(n)$. Thus, there exist unique continuous maps $\tilde{e}_C, \tilde{e}_G\colon P^N/\,\mathrm{E}(n) \to \mathbb{R}^M$ such that $e_C = \tilde{e}_C \circ \pi$ and $e_G = \tilde{e}_G \circ \pi$ (see [62]). The assumption of rigidity means in the orbit space that, for every orbit $\tilde{p} \in \tilde{e}_G^{-1}(d)$, there exists a neighborhood $\tilde{U}$ of $\tilde{p}$ in $P^N/\,\mathrm{E}(n)$ such that $\tilde{e}_G^{-1}(d) \cap \tilde{U} = \tilde{e}_C^{-1}(\tilde{e}_C(\tilde{p})) \cap \tilde{U}$. Since $\tilde{e}_G^{-1}(d)$ is compact, and since $\tilde{e}_C^{-1}(\tilde{e}_C(\tilde{p})) = \{\tilde{p}\}$, it follows that $\tilde{e}_G^{-1}(d)$ only consists of finitely many orbits. Thus, there exists a finite set $F \subseteq e_G^{-1}(d)$ such that $e_G^{-1}(d) = F^{\mathrm{E}(n)}$. Since F is finite, we obtain from the previous paragraph that there exist a neighborhood W of F in P^N and some constant $c_0 > 0$ such that (2.4) holds for every $p \in W$. Since both $\psi_{G,d}$ and $\|\nabla\psi_{G,d}\|$ are invariant under the action of $\mathrm{E}(n)$, we conclude that (2.4) holds for every $p \in W^{\mathrm{E}(n)}$. The proof is complete, if we can show that there exists $r_0 > 0$ such that $\psi_{G,d}^{-1}(\leq r_0) \subseteq W^{\mathrm{E}(n)}$. Since $\psi_{G,d}\colon P^N \to \mathbb{R}$ is continuous and invariant under the action of $\mathrm{E}(n)$, there exists a unique continuous function $\tilde{\psi}_{G,d}$ on $P^N/\,\mathrm{E}(n)$ such that $\psi_{G,d} = \tilde{\psi}_{G,d} \circ \pi$. Since the projection map π is open (see [62]), the set $\tilde{W} := \pi(W)$ is a neighborhood of $\tilde{P} := \pi(P) = \tilde{\psi}_{G,d}^{-1}(0)$ in $P^N/\,\mathrm{E}(n)$. Since $\tilde{\psi}_{G,d}$ is continuous and has compact sublevel sets, there exists a sufficiently small $r_0 > 0$ such that $\tilde{\psi}_{G,d}^{-1}(\leq r_0) \subseteq \tilde{W}$. Thus, $\psi_{G,d}^{-1}(\leq r_0) \subseteq W^{\mathrm{E}(n)}$, which completes the proof. □

Remark 2.8. In general, the noncompact set $\psi_{G,d}^{-1}(0)$ of global minima of $\psi_{G,d}$ might have a complicated structure. However, the proof of Proposition 2.7 reveals that under the assumption of infinitesimal rigidity, the set $\psi_{G,d}^{-1}(0)$ is simply the union of orbits of finitely many points in P^N under action of the Euclidean group. It therefore suffices to consider $\psi_{G,d}$ in a small neighborhood of a single point of each orbit. A similar strategy is also applied in several other studies on formation shape control (see, e.g., [48, 82]). The assumption of infinitesimal rigidity allows us to derive the lower bound (2.4) for the gradient of $\psi_{G,d}$ on a noncompact sublevel set. This estimate will play an important role in the proofs of our main results. ◊

[2]By the *adjoint* of $\mathrm{D}\bar{e}_G(p)\colon T^N \to \mathbb{R}^k$, we mean the unique linear map $\mathrm{D}\bar{e}_G(p)^\top : \mathbb{R}^k \to T^N$ that satisfies the property $\langle\!\langle \mathrm{D}\bar{e}_G(p)v, w\rangle\!\rangle = \langle\!\langle v, \mathrm{D}\bar{e}_G(p)^\top w\rangle\!\rangle$ for every $v \in T^N$ and every $w \in \mathbb{R}^k$, where $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ denotes the inner product on T^N and on $\mathbb{R}^k$.

[3]Note that the dimension k is less than or equal to the dimension of T^N.

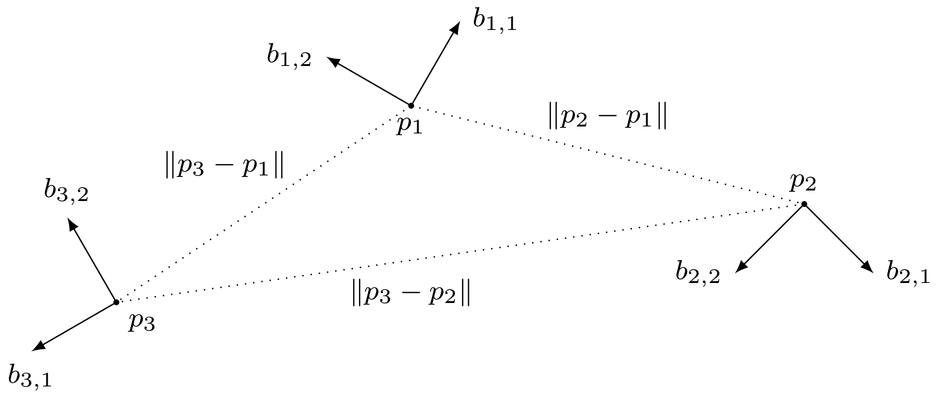


Figure 2.1: *A system of $N = 3$ point agents in $n = 2$ dimensions. Their current distances $\|p_{i_2} - p_{i_1}\|$ are indicated by dotted lines. The agents do not share information about a global coordinate system. Instead, each agent navigates with respect to its individual body frame, which is defined by the orthonormal velocity directions $b_{i,k}$.*

2.4 Formation control for point agents

2.4.1 Problem description

We consider a system of N point agents in the n-dimensional Euclidean space P. For each $i \in \{1, \dots, N\}$, let $b_{i,1}, \dots, b_{i,n}$ be an orthonormal basis of the underlying translation space T. We assume that the motion of agent $i \in \{1, \dots, N\}$ is determined by the kinematic equation

$$\dot{p}_i \;=\; \sum_{k=1}^{n} u_{i,k}\, b_{i,k}, \tag{2.5}$$

where each $u_{i,k}$ is a real-valued input channel to control the velocity into direction $b_{i,k}$. The situation is depicted in Figure 2.1. It is worth to mention that the directions $b_{i,k}$ do not need to be known for an implementation of the control law that is presented in the next subsection.

Suppose that the agents are equipped with very primitive sensors so that they can only measure distances to certain other members of the team. These measurements are described by an (undirected) graph $G = (V, E)$; see Subsection 2.3.1 for the definition. If there is an edge $i_1 i_2 \in E$ between agents $i_1, i_2 \in V$, then it means that agent i_1 can measure the Euclidean distance $\|p_{i_2} - p_{i_1}\|$ to agent i_2 and vice versa. Note that the agents cannot measure relative positions $p_{i_2} - p_{i_1}$ but only distances. For each edge $i_1 i_2 \in E$, let $d_{i_1 i_2} \geq 0$ be a nonnegative real number, which is the *desired distance* between agents i_1 and i_2. We assume that these distances are *realizable* in P; i.e., there exists $p = (p_1, \dots, p_N) \in P^N$ such that $\|p_{i_2} - p_{i_1}\| = d_{i_1 i_2}$ for every $i_1 i_2 \in E$. We are interested in a distributed and distance-only control law that steers the multi-agent system into such a target formation. The control law that we propose in Subsection 2.4.2 requires only distance measurements and can be implemented directly in each agent's local coordinate frame, which is independent of any global coordinate frame.

We remark that, throughout the chapter, we consider an undirected graph for modeling a multi-agent formation system, as it is commonly assumed in the literature on multi-agent coordination control (see the surveys [86, 20]). This assumption is motivated by various application scenarios. In practice, agents are often equipped with homogeneous sensors that have the same sensing ability, e.g., same sensing ranges for range sensors. Therefore, it is justifiable to assume bidirectional sensing (described by an undirected graph) in modeling a multi-agent system. Undirected graphs also enable a gradient-based control law for stabilizing formation shapes, which may not be possible for general directed graphs. Extensions of the current results to directed graphs will be a content of future research.

2.4.2 Control law and main statement

The control law will be composed of the following constituents.

(1) For each $i \in \{1, \ldots, N\}$, define a local potential function ψ_i on P^N by

$$\psi_i(p) \; := \; \frac{1}{4} \sum_{i' \in V \colon ii' \in E} \big(\|p_{i'} - p_i\|^2 - d_{ii'}^2\big)^2 \tag{2.6}$$

for every $p = (p_1, \ldots, p_N) \in P^N$ with distances $d_{ii'} \geq 0$ as in Subsection 2.4.1.

(2) Let a be a smooth and bounded function on $\mathbb{R}$ such that $a(0) = 0$ and $a'(0) \neq 0$. Define two functions h_1, h_2 on $\mathbb{R}$ by $h_1(y) := h_2(y) := 0$ for $y \leq 0$ and by

$$h_1(y) \; := \; a(y) \sin(\log y), \tag{2.7a}$$

$$h_2(y) \; := \; a(y) \cos(\log y) \tag{2.7b}$$

for $y > 0$.

(3) Choose nN pairwise distinct positive real frequency coefficients $\omega_{i,k}$ for $i \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, n\}$. Moreover, for every $j > 0$, every $i \in \{1, \ldots, N\}$, and every $k \in \{1, \ldots, n\}$, define two sinusoids $u^j_{(i,k,1)}, u^j_{(i,k,2)} \colon \mathbb{R} \to \mathbb{R}$ by

$$u^j_{(i,k,1)}(t) \; := \; \sqrt{2\, j\, \omega_{i,k}}\, \cos(j\, \omega_{i,k} t + \varphi_{i,k}), \tag{2.8a}$$

$$u^j_{(i,k,2)}(t) \; := \; \sqrt{2\, j\, \omega_{i,k}}\, \sin(j\, \omega_{i,k} t + \varphi_{i,k}) \tag{2.8b}$$

with arbitrary shifts $\varphi_{i,k} \in \mathbb{R}$.

Remark 2.9. We briefly give some preliminary comments on the above functions without going into details here.

(1) Note that agent i only needs to measure the distances $\|p_{i'} - p_i\|$ to its neighbors $i' \in V$ with $ii' \in E$ in order to compute the current value of its local potential function ψ_i. In particular, the requirements of a distributed distance-only control law are met if each agent only uses the current value

$$y_i \; := \; \psi_i(p) \tag{2.9}$$

of its local potential function (2.6) in the feedback loop. In the context of extremum seeking control, one can interpret (2.9) as an output channel for agent i. A computation of the current value of the local potential function corresponds to a measurement of (2.9). For this reason, the local potential function ψ_i plays the role of an output function for agent i. We are interested in a control law that asymptotically stabilizes the agents at states at which the output functions attain their minimum value 0. Moreover, note that, for each $i \in \{1, \ldots, N\}$, the local potential function ψ_i is of the form (2.1) if we consider the graph that originates from G by keeping only the edges to the neighbors of the vertex i in G.

(2) An admissible choice of the function a in (2.7) is, for example, given by $a := \tanh$. The particular combination of sine and cosine in (2.7) leads to the Lie bracket

$$[h_1, h_2](y) \;:=\; h_2'(y)h_1(y) - h_1'(y)h_2(y) \;=\; -a(y)^2/y \tag{2.10}$$

for every $y > 0$. This Lie bracket can be extended to a locally Lipschitz continuous function on $\mathbb{R}$ if we let $[h_1, h_2](y) := 0$ for every $y \leq 0$.

(3) The choice of pairwise distinct frequency coefficients $\omega_{i,k}$ for the sinusoids $u^j_{(i,k,\nu)}$ in (2.8) has the purpose to excite certain Lie brackets of vector fields, which are directly linked to the Lie bracket of h_1, h_2 in (2.10). This fact is revealed by a suitable averaging analysis in Section 2.5. ◊

Given real numbers $\kappa \geq 1/2$ and $j > 0$, we propose the control law

$$u_{i,k} \;=\; u^j_{(i,k,1)}(t)\, h_1(y_i^\kappa) + u^j_{(i,k,2)}(t)\, h_2(y_i^\kappa) \tag{2.11}$$

for every $i \in \{1, \ldots, N\}$ and every $k \in \{1, \ldots, n\}$, where y_i^κ denotes the κth power of the current value (2.9) of the local potential function ψ_i. Whenever $y_i > 0$, we can write control law (2.11) also as

$$u_{i,k} \;=\; \sqrt{2\, j\, \omega_{i,k}}\; a(y_i^\kappa)\, \sin(j\, \omega_{i,k}\, t + \varphi_{i,k} + \log y_i^\kappa), \tag{2.12}$$

where we have used the trigonometric identity (1.49).

Remark 2.10. An implementation of control law (2.11) requires that each agent knows the desired inter-agent distances to its neighbors, and its own pairwise distinct frequencies (and arbitrary shifts). Such information can be embedded into the memory of each agent prior to an implementation of the control law. Also, each agent needs to measure the current inter-agent distances (in contrast to relative positions, as assumed in most papers on formation shape control) relative to its neighbors in order to compute the value of its local potential (2.6). The setting of such a control scenario is common in most distributed control laws, which is acknowledged by the term 'centralized design, distributed implementation', which does not contradict with the principle of distributed control (see, e.g., the surveys [20, 86]). Therefore, the proposed control law is fully distributed.

It is also important to note that we allow arbitrary phase shifts $\varphi_{i,k}$ in the sinusoids (2.8). The phase shifts for one agent are not assumed to be known to the other members of the team. In particular, this means that the control law (2.11) requires no time synchronization among the agents. Moreover, since we merely assume that the frequency coefficients $\omega_{i,k}$ are pairwise distinct, it is not necessary that the sinusoids have a common period. ◊

It is shown later in Lemma 2.18 (a) that, for every $i \in \{1, \ldots, N\}$ and every $\nu \in \{1, 2\}$, the function $h_\nu \circ \psi_i^\kappa$ is locally Lipschitz continuous. It therefore follows from standard theorems for ordinary differential equations that system (2.5) under the control law (2.11) has a unique maximal solution for any initial condition. These solutions do not have a finite escape time because (2.11) is bounded. In summary, we have the following result.

Proposition 2.11. *For any initial condition, system* (2.5) *under control law* (2.11) *has a unique global solution, which we call a* trajectory *of* (2.5) *under* (2.11).

To state our main result, we introduce the global potential function $\psi = \psi_{G,d}$ on P^N with $\psi_{G,d}$ as in (2.1); i.e., we define

$$\psi(p) \;:=\; \frac{1}{4} \sum_{i_1 i_2 \in E} \big(\|p_{i_2} - p_{i_1}\|^2 - d_{i_1 i_2}^2\big)^2. \tag{2.13}$$

Note that the fiber of ψ over 0,

$$\psi^{-1}(0) \;=\; \{(p_1, \ldots, p_N) \in P^N \mid \forall i_1 i_2 \in E\colon \|p_{i_2} - p_{i_1}\| = d_{i_1 i_2}\}, \tag{2.14}$$

is the *set of desired formations*. Since we assume that the distances $d_{i_1 i_2}$ are realizable in P, the set (2.14) is not empty.

Theorem 2.12. *Suppose that, for every point p of* (2.14)*, the framework $G(p)$ is infinitesimally rigid. Let $\kappa = 1/2$. Then, there exist $\mu, r > 0$ such that, for every $\lambda > 1$, there exists $j_0 > 0$ such that, for every $j \geq j_0$, every $t_0 \in \mathbb{R}$, and every $p_0 \in \psi^{-1}(\leq r)$, the trajectory p of system* (2.5) *under control law* (2.11) *with initial condition $p(t_0) = p_0$ has the following two properties: $p(t)$ converges to some point of* (2.14) *as $t \to \infty$, and the estimate*

$$\psi(p(t)) \;\leq\; \lambda\, \psi(p_0)\, \mathrm{e}^{-\mu(t-t_0)} \tag{2.15}$$

holds for every $t \geq t_0$.

Theorem 2.12 states that, under the assumption of infinitesimal rigidity, control law (2.11) with $\kappa = 1/2$ leads to local exponential stability if the frequency parameter j is sufficiently large. If we increase the exponent κ in (2.11) to a value $> 1/2$, then this has the following two effects. On the one hand, an increase of κ reduces the speed of convergence to a desired state, and therefore we only get asymptotic stability instead of exponential stability. On the other hand, this also leads to the effect that the quality of approximation of the averaged system (which is presented below) improves the closer the agents are to a desired formation. This phenomenon allows us to circumvent the assumption that the frequency parameter j is chosen sufficiently large. The averaging analysis in Section 2.5 will reveal why an increase of κ changes the speed of convergence and the quality of approximation. We will discuss this in more detail at the end of the averaging analysis in Remark 2.27. For $\kappa > 1/2$, the following statement holds.

Theorem 2.13. *Suppose that, for every point p of* (2.14)*, the framework $G(p)$ is infinitesimally rigid. Let $\kappa > 1/2$ and let $j \geq 1$. Then, there exists $\mu > 0$ such that, for every $\lambda > 1$, there exists $r > 0$ such that, for every $t_0 \in \mathbb{R}$ and every $p_0 \in \psi^{-1}(\leq r)$, the trajectory p of system* (2.5) *under control law* (2.11) *with initial condition $p(t_0) = p_0$ has the following two properties: $p(t)$ converges to some point of* (2.14) *as $t \to \infty$, and the estimate*

$$\psi(p(t)) \;\leq\; \frac{\lambda\, \psi(p_0)}{\big(1 + (2\kappa - 1)\, \psi(p_0)^{2\kappa-1}\, \mu\, (t - t_0)\big)^{\frac{1}{2\kappa-1}}} \tag{2.16}$$

holds for every $t \geq t_0$.

Detailed proofs of Theorems 2.12 and 2.13 are presented in Section 2.5. At this point, we only indicate the reason why the set (2.14) becomes locally uniformly asymptotically stable for system (2.5) under control law (2.11). Note that the closed-loop system is an ordinary

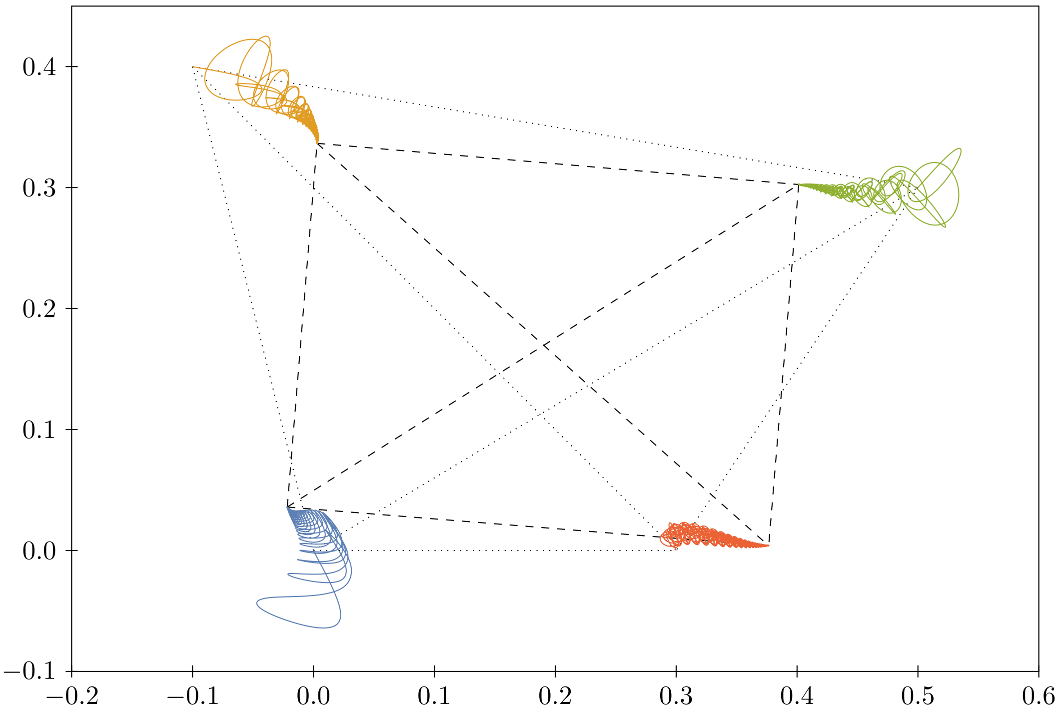


Figure 2.2: *Simulation on stabilization control of a four-agent rectangular formation shape. We denote the positions by $p_i = (x_i, y_i)$ for $i = 1, \ldots, 4$. The initial formation is indicated by dotted lines, and the finial formation is indicated by dashed lines.*

differential equation in the product space P^N, which consists of the coupled differential equations

$$\dot{p}_i \;=\; \sum_{k=1}^{n}\sum_{\nu=1}^{2} u^j_{(i,k,\nu)}(t)\, h_\nu(\psi_i^\kappa(p))\, b_{i,k} \tag{2.17}$$

on P for $i = 1, \ldots, N$. One can interpret the right-hand side of (2.17) as a time-varying linear combination of the state dependent maps $p \mapsto h_\nu(\psi_i^\kappa(p))\, b_{i,k}$ with highly oscillatory functions $u^j_{(i,k,\nu)}$. When we consider the closed-loop system in the product space, each of the maps $p \mapsto h_\nu(\psi_i^\kappa(p))\, b_{i,k}$ defines a vector field $f_{(i,k,\nu)}$ on P^N. The analysis in Section 2.5 will show that the trajectories of (2.17) are driven into directions of certain Lie brackets of the vector fields $f_{(i,k,\nu)}$. To be more precise, the particular choice of the sinusoids $u^j_{(i,k,\nu)}$ with pairwise distinct frequencies $\omega_{i,k}$ causes the trajectories of (2.17) to follow Lie brackets of the form $[f_{(i,k,1)}, f_{(i,k,2)}]$. The ordinary differential equation on P^N with the sum of all Lie brackets $\frac{1}{2}[f_{(i,k,1)}, f_{(i,k,2)}]$ on the right-hand side is referred to as the corresponding *Lie bracket system*; cf. Section 1.2. A direct computation shows that the Lie bracket system is given by the coupled differential equations

$$\dot{p}_i \;=\; -\kappa\, h(\psi_i^\kappa(p))\, \psi_i(p)^{2\kappa-1}\, \nabla_{p_i}\psi(p) \tag{2.18}$$

on P for $i = 1, \ldots, N$, where $\nabla_{p_i}\psi\colon P^N \to T$ is the gradient of the global potential function ψ with respect to the ith position variable, and $h(\psi_i^\kappa(p))$ denotes a certain positive factor for $\psi_i(p) > 0$ sufficiently close to 0. Thus, in a neighborhood of (2.14), the system state of (2.18) is constantly driven into a descent direction of ψ. The assumption of infinitesimal rigidity ensures that the decay of ψ along trajectories of (2.18) is sufficiently

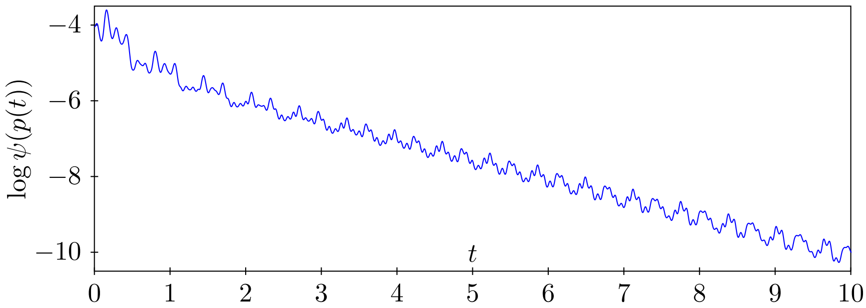


Figure 2.3: *Exponential decay of the global potential function* (2.13) *for the multi-agent system in Figure 2.2 on the time interval* $[0, 10]$.

fast. Since the trajectories of (2.17) approximate the behavior of (2.18) in a neighborhood of (2.14), this in turn implies that also the value of ψ along trajectories of (2.5) under (2.11) decays on average.

Remark 2.14. We emphasize that both Theorem 2.12 and Theorem 2.13 are local but not global stability results. Both theorems ensure convergence to a desired formation for initial points from a certain neighborhood of the set (2.14) if j is sufficiently large. The size of the domain of attraction $\psi^{-1}(\leq r)$ is characterized by the sublevel $r > 0$. The value of r depends on the choice of the frequency parameter j. An increase of j leads to an increase of r. An upper bound for r is naturally given by the domain of attraction of the averaged system (2.18). Note that a gradient-based control law can lead to undesired equilibria at critical points of the potential function. Therefore we cannot expect global asymptotic stability for (2.18), and also not semi-global uniform asymptotic stability for (2.5) under (2.11). ♢

2.4.3 Simulation examples

In this subsection, we provide two simulation results to demonstrate the behavior of (2.5) under (2.11). We consider a rectangular formation shape in two dimensions and a double tetrahedron formation shape in three dimensions. One can check that the corresponding frameworks are infinitesimally rigid by means of the *rank condition* in [8] for the derivative of the edge map. The same formations are also considered in [103] for system (2.5) under the well-established negative gradient control law. Note that in contrast to the method in this chapter, *relative position measurements* are required in [103] to stabilize the desired formation shapes.

Our first example is a system of $N = 4$ point agents in the Euclidean space of dimension $n = 2$. For each $i \in \{1, \ldots, N\}$, the orthonormal velocity vectors of agent i in (2.5) are given by $b_{i,1} = (\cos\phi_i, \sin\phi_i)^\top$ and $b_{i,2} = (-\sin\phi_i, \cos\phi_i)^\top$, where $\phi_i = i\pi/3$.[4] We let G be the complete graph of N nodes. This means that each agent can measure the distances to all other members of the team. The common goal of the agents is to reach a rectangular formation with desired distances $d_{12} = d_{34} = 0.3$, $d_{23} = d_{14} = 0.4$, and $d_{13} = d_{24} = 0.5$. The initial conditions are given by $p_1(0) = (0.0, 0.0)$, $p_2(0) = (-0.1, 0.4)$, $p_3(0) = (0.5, 0.3)$, and $p_4(0) = (0.3, 0.0)$. The amplitude a of the functions h_ν in (2.7) is chosen as $a := \tanh$.

[4] To distinguish *points* and *tangent vectors* in the notation we write points as row vectors and tangent vectors as column vectors.

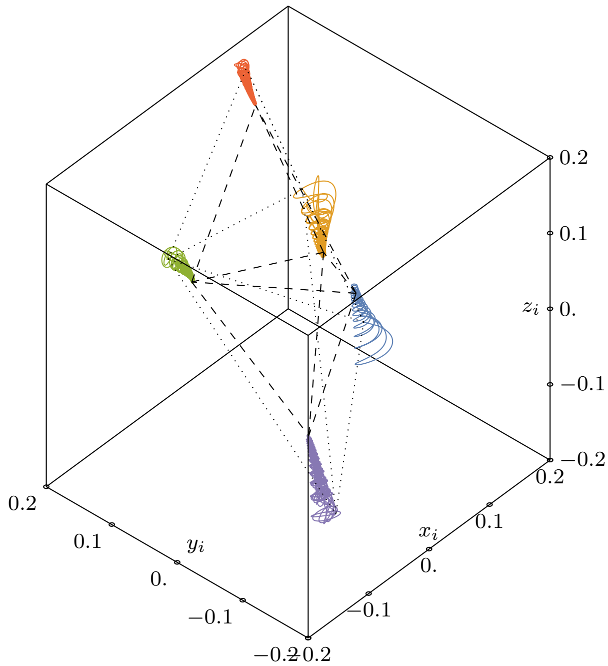


Figure 2.4: *Simulation on stabilization control of a double tetrahedron formation. We denote the positions by $p_i = (x_i, y_i, z_i)$ for $i = 1, \ldots, 5$. The initial formation is indicated by dotted lines, and the finial formation is indicated by dashed lines.*

The frequency coefficients $\omega_{i,k}$ for the sinusoids $u^j_{(i,k,\nu)}$ in (2.8) are chosen as the pairwise distinct integers $\omega_{i,k} = (i-1)\,n + k$ for $i = 1, \ldots, N$ and $k = 1, \ldots, n$. For the sake of simplicity, the phase shifts $\varphi_{i,k}$ are all set equal to zero. To obtain local exponential stability as in Theorem 2.12, the exponent κ in (2.11) is chosen to be $\kappa = 1/2$. It turns out that the initial positions are not in the domain of attraction if we choose $j = 1$. As indicated in Remark 2.14, the domain of attraction becomes larger when we increase j. The trajectories for $j = 10$ are shown in Figure 2.2. An exponential decay of the global potential function can be observed in Figure 2.3.

In the second example, we consider a system of $N = 5$ point agents in the Euclidean space of dimension $n = 3$. For each $i \in \{1, \ldots, N\}$, the orthonormal velocity vectors of agent i in (2.5) are given by $b_{i,1} = (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i)^\top$, $b_{i,2} = (-\sin\phi_i, \cos\phi_i, 0)^\top$, and $b_{i,3} = (-\cos\theta_i \cos\phi_i, -\cos\theta_i, \sin\theta_i)^\top$, where $\phi_i = i\pi/3$ and $\theta_i = i\pi/6$. We let G be the graph that originates from the complete graph of N nodes by removing the edge between the nodes 4 and 5. The common goal of the agents is to reach a formation shape of a double tetrahedron with desired distances $d_{i_1 i_2} = 0.2$ for every edge $i_1 i_2$ of G. The initial conditions are given by $p_1(0) = (0, -0.1, 0.05)$, $p_2(0) = (0.18, 0.16, -0.01)$, $p_3(0) = (-0.02, 0.18, 0.005)$, $p_4(0) = (0.12, 0.19, 0.17)$ and $p_5(0) = (-0.1, -0.15, -0.12)$. The functions h_ν, the frequency coefficients $\omega_{i,k}$, the phase shifts $\varphi_{i,k}$, and the exponent κ are chosen as in the first example. Again, the initial positions are not within the domain of attraction of (2.5) under (2.11) for $j = 1$. However, for $j = 10$, one can see in Figure 2.4 that the trajectories converge to the desired formation shape.

One may interpret the oscillatory trajectories in the simulations as follows. Each agent constantly explores how small changes of its current position influences the value of its local potential function ψ_i. This way an agent obtains gradient information. On average this leads to a decay of all local potential functions. Sufficiently fast oscillations are necessary in our approach to ensure that every agent can explore its neighborhood properly. If the value of ψ_i is small, then the terms $\sin(\log \psi_i^\kappa)$ and $\cos(\log \psi_i^\kappa)$ in (2.11) induce sufficiently fast oscillations. When ψ_i is not small, then an increase of the global frequency parameter j can compensate the lack of oscillations. It is clear that the energy effort to implement (2.11) is much larger than for a gradient-based control law. This is in some sense the price that we have to pay when we reduce the amount of utilized information from the gradient of ψ_i to the values of ψ_i.

2.5 Local asymptotic stability analysis for point agents

The aim of this section is to prove Theorems 2.12 and 2.13. In the first step, we rewrite system (2.5) under control law (2.11) as a control-affine system under open-loop controls. For this purpose, we have to introduce a suitable notation. Recall that, for every $i \in \{1, \ldots, n\}$, the velocity directions $b_{i,1}, \ldots, b_{i,n} \in T$ in (2.5) are assumed to be an orthonormal basis of T. For every $i \in \{1, \ldots, N\}$ and every $k \in \{1, \ldots, n\}$, let $g_{i,k}(p) \in T^N$ be the vector with $b_{i,k} \in T$ as its kth component and all other $(N-1)$ components are equal to $0 \in T$. Then, we can write the multi-agent control system (2.5) equivalently as the control-affine system

$$\dot{p} \;=\; \sum_{i=1}^{N} \sum_{k=1}^{n} u_{i,k}\, g_{i,k}(p) \tag{2.19}$$

on P^N. It is clear that the vectors $g_{i,k}(p)$ form an orthonormal basis of T^N at any $p \in P^N$. Moreover, it follows directly from the definitions that, for every $i \in \{1, \ldots, N\}$ and every $k \in \{1, \ldots, n\}$, the Lie derivatives of the local potential function ψ_i and global potential function ψ along $g_{i,k}$ coincide; i.e., $g_{i,k}\psi = g_{i,k}\psi_i$. As an abbreviation, we define an indexing set Λ to be the set of all triples (i, k, ν) with $i \in \{1, \ldots, N\}$, $k \in \{1, \ldots, n\}$, and $\nu \in \{1, 2\}$. For each $\ell = (i, k, \nu) \in \Lambda$, define a vector field f_ℓ on P^N by

$$f_\ell(p) \;:=\; h_\nu(\psi_i^\kappa(p))\, g_{i,k}(p). \tag{2.20}$$

When we insert (2.11) into (2.5), the closed-loop system can be written as

$$\dot{p} \;=\; f^j(t, p) \;:=\; \sum_{\ell \in \Lambda} u_\ell^j(t)\, f_\ell(p), \tag{2.21}$$

which may be interpreted as a control-affine system with control vector fields f_ℓ under open-loop controls u_ℓ^j.

2.5.1 Estimates for the Lie derivatives

In this subsection, we derive suitable boundedness properties of (iterated) Lie derivatives of the global potential function ψ along the control vector fields f_ℓ in (2.20). These boundedness properties will ensure later in Subsection 2.5.4 that certain remainder terms become small when the agents are close to the set (2.14) of target formations.

For later references, we collect the following properties of the functions h_ν in (2.7), which easily follow from their definitions.

Lemma 2.15. *For every $\nu \in \{1, 2\}$, the following statements hold:*

(a) h_ν *is smooth on* $(0,\infty)$*,*

(b) h_ν *is locally Lipschitz continuous on* $\mathbb{R}$*,*

(c) $\limsup_{y\downarrow 0} |h''_\nu(y)y| < \infty$.

Let W_1, W_2 be subsets of a Euclidean space, and let W be a subset of the (possibly empty) intersection of W_1, W_2. Let b be a nonnegative function on W_1. For the sake of convenience, we introduce the following terminology. We say that a function α on W_2 is *bounded by a multiple of* b *on* W if there exists $c > 0$ such that $|\alpha(x)| \leq c\,b(x)$ for every $x \in W$. We say that a vector field f on W_2 is *bounded by a multiple of* b *on* W if there exists $c > 0$ such that $\|f(x)\| \leq c\,b(x)$ for every $x \in W$.

Note that the vector fields f_ℓ in (2.20) are, in general, not differentiable at every point of P^N. However, we will show that the f_ℓ are at least locally Lipschitz continuous. For this purpose, it turns out to be convenient to use the notion of a *pointwise Lipschitz constant.* In non-smooth calculus [47], this quantity provides an upper bound for the difference quotient of a function. It is known from [25] that local Lipschitz continuity can be characterized as follows.

Lemma 2.16. *A function* α *on an open subset* U *of a Euclidean space is locally Lipschitz continuous if and only if, for each* $x \in U$*, the* pointwise Lipschitz constant

$$\mathrm{L}(\alpha)(x) \;:=\; \limsup_{x\neq x'\to x} \frac{|\alpha(x') - \alpha(x)|}{\|x' - x\|}$$

of α *at* x *exists as a nonnegative real number*[5] *and if, for each compact subset* K *of* U*, the function* $\mathrm{L}(\alpha)$ *is bounded by a constant on* K*. The same statement holds for a vector field* f *on* U *with respect to the* pointwise Lipschitz constant

$$\mathrm{L}(f)(x) \;:=\; \limsup_{x\neq x'\to x} \frac{\|f(x') - f(x)\|}{\|x' - x\|}$$

of f *at* x*.*

The following rules help in estimating the pointwise Lipschitz constant for a function constructed from other functions with known pointwise Lipschitz constants.

Lemma 2.17. *Let* $\alpha, \alpha_1, \alpha_2$ *be locally Lipschitz continuous functions on an open subset* U *of a Euclidean space and let* β *be a locally Lipschitz continuous function on* $\mathbb{R}$*. Then, the following inequalities hold on* U*:*

(a) $\mathrm{L}(\alpha_1 + \alpha_2) \leq \mathrm{L}(\alpha_1) + \mathrm{L}(\alpha_2)$ *(sum rule),*

(b) $\mathrm{L}(\alpha_1 \cdot \alpha_2) \leq \mathrm{L}(\alpha_1)\cdot|\alpha_2| + |\alpha_1|\cdot\mathrm{L}(\alpha_2)$ *(product rule),*

(c) $\mathrm{L}(\beta\circ\alpha) \leq (\mathrm{L}(\beta)\circ\alpha)\cdot\mathrm{L}(\alpha)$ *(chain rule).*

Moreover, if α *is differentiable at some* $x \in U$*, then*

$$\mathrm{L}(\alpha)(x) \;=\; \|\nabla\alpha(x)\|.$$

[5] By its definition, the *upper limit* $\limsup_{x\neq x'\to x} \frac{|\alpha(x')-\alpha(x)|}{\|x'-x\|} = \lim_{r\downarrow 0} \sup\left\{ \frac{|\alpha(x')-\alpha(x)|}{\|x'-x\|} \;\middle|\; x \neq x' \in U\colon \|x'-x\| \leq r \right\}$ is either a nonnegative real number or the symbol $+\infty$.

We leave a verification of the above rules to the reader. Next, we use Lemmas 2.16 and 2.17 to prove the following technical result.

Lemma 2.18. *Let $\ell = (i, k, \nu) \in \Lambda$, let $\kappa \geq 1/2$, and let $r > 0$.*

(a) The function $h_\nu \circ \psi_i^\kappa$ is locally Lipschitz continuous and the following boundedness properties hold:

(i) $h_\nu \circ \psi_i^\kappa$ is bounded by a multiple of ψ_i^κ on $\psi_i^{-1}(\leq r)$;

(ii) $\mathrm{L}(h_\nu \circ \psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$.

(b) The Lie derivative $f_\ell \psi$ of ψ along f_ℓ is differentiable with locally Lipschitz continuous derivative and the following boundedness properties hold:

(i) $f_\ell \psi$ is bounded by a multiple of $\psi_i^{\kappa+1/2}$ on $\psi_i^{-1}(\leq r)$;

(ii) $\nabla(f_\ell \psi)$ is bounded by a multiple of ψ_i^κ on $\psi_i^{-1}(\leq r)$;

(iii) $\mathrm{L}(\nabla(f_\ell \psi))$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$.

(c) If $\kappa > 1/2$, then $h_\nu \circ \psi_i^\kappa$ is differentiable and $\nabla(h_\nu \circ \psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$.

Proof. It follows from Lemma 2.15 (b) and $h_\nu(0) = 0$ that h_ν is bounded by a multiple of the identity $y \mapsto y$ on $[0, r^\kappa]$, which implies part (i) of statement (a). We already know from Proposition 2.7 (a) that the square root of ψ_i is Lipschitz continuous on its sublevel sets. Therefore, $\mathrm{L}(\psi_i^{1/2})$ is bounded by constant on $\psi_i^{-1}(\leq r)$. Note that ψ_i^κ is the composition of the functions $\psi_i^{1/2}$ and $y \mapsto y^{2\kappa}$. By the chain rule in Lemma 2.17, we conclude that $\mathrm{L}(\psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$. Because of Lemma 2.15 (b), $\mathrm{L}(h_\nu)$ is bounded by a constant on $[0, r^\kappa]$. Again, by the chain rule, it follows that $\mathrm{L}(h_\nu \circ \psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$. In particular, $h_\nu \circ \psi_i^\kappa$ is locally Lipschitz continuous. If $\kappa > 1/2$, then part (i) of statement (a) implies that $h_\nu \circ \psi_i^\kappa$ is also differentiable at every $p \in \psi_i^{-1}(0)$ with $\nabla(h_\nu \circ \psi_i^\kappa)(p) = 0$. It now follows from part (ii) of statement (a) and Lemma 2.17 that $\nabla(h_\nu \circ \psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$. This completes the proofs of statements (a) and (c). Next, we prove statement (b).

It follows from Proposition 2.7 (b) that $g_{i,k}\psi = g_{i,k}\psi_i$ is bounded by a multiple of $\psi_i^{1/2}$ on $\psi_i^{-1}(\leq r)$. Using Lemma 2.18 (a), we conclude that $f_\ell \psi = (h_\nu \circ \psi_i^\kappa)(g_{i,k}\psi)$ is bounded by a multiple of $\psi_i^{\kappa+1/2}$ on $\psi_i^{-1}(\leq r)$. Since $\kappa \geq 1/2$, this in turn implies that $f_\ell \psi$ is differentiable at every $p \in \psi_i^{-1}(0)$ with $\nabla(f_\ell \psi)(p) = 0$. Outside of $\psi_i^{-1}(0)$, $f_\ell \psi$ is a composition of differentiable functions and therefore we may compute

$$\nabla(f_\ell \psi) \;=\; (h'_\nu \circ \psi_i^\kappa)\,(g_{i,k}\psi)\,\nabla\psi_i^\kappa + (h_\nu \circ \psi_i^\kappa)\,\nabla(g_{i,k}\psi).$$

It follows from Lemma 2.15 (b) that h'_ν is bounded by a constant on $(0, r^\kappa]$. Thus, $h'_\nu \circ \psi_i^\kappa$ is also bounded by a constant on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$. We already know that $g_{i,k}\psi$ is bounded by a multiple of $\psi_i^{1/2}$ on $\psi_i^{-1}(\leq r)$. Moreover, it follows from Proposition 2.7 (b) that $\nabla\psi_i^\kappa$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$. By Lemma 2.18 (a), $h_\nu \circ \psi_i^\kappa$ is bounded by a multiple of ψ_i^κ on $\psi_i^{-1}(\leq r)$. Finally, it follows from Proposition 2.7 (c) that $\nabla(g_{i,k}\psi) = \nabla(g_{i,k}\psi_i)$ is bounded by a constant on $\psi_i^{-1}(\leq r)$. Consequently, $\nabla(f_\ell \psi)$ is bounded by a multiple of ψ_i^κ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$. Since $\nabla(f_\ell \psi) = 0$ on $\psi_i^{-1}(0)$, this implies part (ii) of statement (b).

It is left to prove part (iii) of statement (b). Because of previous part (ii), there exists $c > 0$ such that

$$\|\nabla(f_\ell\psi)(p) - \nabla(f_\ell\psi)(p^*)\| \;\leq\; c\,|\psi_i^\kappa(p) - \psi_i^\kappa(p^*)| \tag{2.22}$$

for every $p^* \in \psi^{-1}(0)$ and every $p \in \psi_i^{-1}(\leq r)$. We have already argued that ψ_i^κ is locally Lipschitz continuous and that $\mathrm{L}(\psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$. Because of (2.22), we conclude that $\mathrm{L}(\nabla(f_\ell\psi))$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi^{-1}(0)$. Outside of $\psi_i^{-1}(0)$, we know that $\nabla(f_\ell\psi)$ is composition of differentiable maps and therefore, using Lemma 2.17, we obtain

$$\begin{aligned}\mathrm{L}(\nabla(f_\ell\psi)) \;\leq\; &\mathrm{L}(h_\nu' \circ \psi_i^\kappa)\,|g_{i,k}\psi|\,\|\nabla\psi_i^\kappa\| + |h_\nu' \circ \psi_i^\kappa|\,\mathrm{L}(g_{i,k}\psi)\,\|\nabla\psi_i^\kappa\| \\ &+ |h_\nu' \circ \psi_i^\kappa|\,|g_{i,k}\psi|\,\mathrm{L}(\nabla\psi_i^\kappa) + \mathrm{L}(h_\nu \circ \psi_i^\kappa)\,\|\nabla(g_{i,k}\psi)\| + |h_\nu \circ \psi_i^\kappa|\,\mathrm{L}(\nabla(g_{i,k}\psi))\end{aligned}$$

outside of $\psi_i^{-1}(0)$. It follows from Lemma 2.15 (c) that h_ν'' is bounded by a multiple of $y \mapsto y^{-1}$ on $(0, r^\kappa]$. We already know that $\mathrm{L}(\psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$. By the chain rule, we conclude that $\mathrm{L}(h_\nu' \circ \psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{-1/2}$ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$. It follows from Proposition 2.7 (c) that $\mathrm{L}(g_{i,k}\psi)$ and $\mathrm{L}(\nabla(g_{i,k}\psi))$ are bounded by constants on $\psi_i^{-1}(\leq r)$ and that $\mathrm{L}(\nabla\psi_i^\kappa)$ is bounded by a multiple of $\psi_i^{\kappa-1}$ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$. For the remaining constituents on the right-hand side of the above estimate for $\mathrm{L}(\nabla(f_\ell\psi))$, we have already derived suitable upper bounds. This allows us to conclude that $\mathrm{L}(\nabla(f_\ell\psi))$ is bounded by a multiple of $\psi_i^{\kappa-1/2}$ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$, which completes the proof. □

Note that, for every $i \in \{1, \ldots, N\}$, we have $\psi_i \leq \psi$ on P^N. This implies that $\psi^{-1}(\leq r)$ is a subset of $\psi_i^{-1}(\leq r)$ for every $r > 0$ and every $i \in \{1, \ldots, N\}$. In the next step, we use Lemma 2.18 to derive the following result.

Lemma 2.19. *Let $\ell_m = (i_m, k_m, \nu_m) \in \Lambda$ for $m = 1, 2$, let $\kappa \geq 1/2$, and let $r > 0$.*

(a) *(i)* *f_{ℓ_1} is locally Lipschitz continuous and bounded by a multiple of ψ^κ on $\psi^{-1}(\leq r)$.*

(ii) *$\mathrm{L}(f_{\ell_1})$ is bounded by a multiple of $\psi^{\kappa-1/2}$ on $\psi^{-1}(\leq r)$.*

(b) *(i)* *$f_{\ell_1}\psi$ is differentiable, has a locally Lipschitz continuous gradient, and is bounded by a multiple of $\psi^{\kappa+1/2}$ on $\psi^{-1}(\leq r)$.*

(ii) *$f_{\ell_2}(f_{\ell_1}\psi)$ is locally Lipschitz continuous and bounded by a multiple of $\psi^{2\kappa}$ on $\psi^{-1}(\leq r)$.*

(iii) *$\mathrm{L}(f_{\ell_2}(f_{\ell_1}\psi))$ is bounded by a multiple of $\psi^{2\kappa-1/2}$ on $\psi^{-1}(\leq r)$.*

(c) *If $\kappa > 1/2$, then f_{ℓ_1} is differentiable and $\nabla_{f_{\ell_2}} f_{\ell_1}$ is bounded by a multiple of $\psi^{2\kappa-1/2}$ on $\psi^{-1}(\leq r)$.*

Proof. By definition, we have $f_{\ell_1} = (h_{\nu_1} \circ \psi_{i_1}^\kappa) g_{i_1,k_1}$. Since g_{i_1,k_1} is constant, statement (a) is an immediate consequence of Lemma 2.18 (a). Moreover, part (i) of statement (b) follows immediately from Lemma 2.18 (b). Since $f_{\ell_1}\psi$ is differentiable, we may compute

$$f_{\ell_2}(f_{\ell_1}\psi) \;=\; \langle\!\langle \nabla(f_{\ell_1}\psi), f_{\ell_2} \rangle\!\rangle.$$

Using Lemma 2.18 (b) and Lemma 2.19 (a), we obtain part (ii) of statement (b). We obtain from Lemma 2.18 (b) that $\nabla(f_{\ell_1}\psi)$ is bounded by a multiple of ψ^κ on $\psi^{-1}(\leq r)$

and that $\mathrm{L}(\nabla(f_{\ell_1}\psi))$ is bounded by a multiple of $\psi^{\kappa-1/2}$ on $\psi^{-1}(\leq r)$. We also know from statement (a) that f_{ℓ_2} is bounded by a multiple of ψ^{κ} on $\psi^{-1}(\leq r)$ and that $\mathrm{L}(f_{\ell_2})$ is bounded by a multiple of $\psi^{\kappa-1/2}$ on $\psi^{-1}(\leq r)$. Using the sum and the product rule in Lemma 2.17, this implies the remaining part (iii) of statement (b). If $\kappa > 1/2$, then we know from Lemma 2.18 (c) that $(h_{\nu_1} \circ \psi_{i_1}^{\kappa})$ is differentiable and therefore also f_{ℓ_1} is differentiable and we may compute

$$\nabla_{f_{\ell_2}} f_{\ell_1} \;=\; \langle\!\langle \nabla(h_{\nu_1} \circ \psi_{i_1}^{\kappa}), f_{\ell_2} \rangle\!\rangle \, g_{i_1,k_1}.$$

Finally, it follows from Lemma 2.18 (c) and part (i) of Lemma 2.19 (a) that the inner product of $\nabla(h_{\nu_1} \circ \psi_{i_1}^{\kappa})$ and f_{ℓ_2} is bounded by a multiple of $\psi^{2\kappa-1/2}$, which completes the proof. □

Lemma 2.20. *For every, $i \in \{1, \ldots, N\}$ and every $k \in \{1, \ldots, n\}$, the Lie bracket of $f_{(i,k,1)}$ and $f_{(i,k,2)}$ exists as a locally Lipschitz continuous vector field on P^N with*

$$[f_{(i,k,1)}, f_{(i,k,2)}] \;=\; -\kappa\, (h \circ \psi_i^{\kappa})\, \psi_i^{2\kappa-1}\, (g_{i,k}\psi_i)\, g_{i,k}, \tag{2.23}$$

where $h\colon \mathbb{R} \to \mathbb{R}$ is defined by

$$h(y) \;:=\; \begin{cases} a'(0)^2 & \text{for } y \leq 0, \\ a(y)^2/y^2 & \text{for } y > 0. \end{cases}$$

Proof. Since the vector fields $f_{(i,k,1)}$ and $f_{(i,k,2)}$ are locally Lipschitz continuous and since both vector fields vanish on $\psi_i^{-1}(0)$, also their Lie bracket vanishes on $\psi_i^{-1}(0)$. Since also $(g_{i,k}\psi_i)$ vanishes on $\psi_i^{-1}(0)$, we conclude that (2.23) holds on $\psi_i^{-1}(0)$. A direct computation, using (2.10), shows that (2.23) also holds outside $\psi_i^{-1}(0)$.

It is left to show that $[f_{(i,k,1)}, f_{(i,k,2)}]$ is locally Lipschitz continuous. By Lemma 2.16, this follows if we can show that the pointwise Lipschitz constant of $[f_{(i,k,1)}, f_{(i,k,2)}]$ is bounded by a constant on each sublevel set of ψ_i. Fix an arbitrary $r > 0$. Since a is assumed to be smooth with $a(0) = 0$, we may conclude that h is bounded by a constant on $[0, r^{\kappa}]$. We also conclude from Proposition 2.7 (b) that $g_{i,k}\psi_i$ is bounded by a multiple of $\psi_i^{1/2}$ on $\psi_i^{-1}(\leq r)$. Because of (2.23), we obtain that $[f_{(i,k,1)}, f_{(i,k,2)}]$ is bounded by a multiple of $\psi_i^{2\kappa-1/2}$ on $\psi_i^{-1}(\leq r)$. Thus, there exists $c > 0$ such that

$$\|[f_{(i,k,1)}, f_{(i,k,2)}](p) - [f_{(i,k,1)}, f_{(i,k,2)}](p^*)\| \;\leq\; c\,|\psi_i^{2\kappa-1/2}(p) - \psi_i^{2\kappa-1/2}(p^*)|$$

for every $p^* \in \psi^{-1}(0)$ and every $p \in \psi_i^{-1}(\leq r)$. Since $\kappa \geq 1/2$, we conclude from Proposition 2.7 (b) and Lemma 2.17 that $\mathrm{L}(\psi_i^{2\kappa-1/2})$ is bounded by a multiple of $\psi_i^{2\kappa-1}$ on $\psi_i^{-1}(\leq r)$. Therefore, $\mathrm{L}([f_{(i,k,1)}, f_{(i,k,2)}])$ is bounded by a multiple of $\psi_i^{2\kappa-1}$ on $\psi_i^{-1}(0)$. Outside of $\psi_i^{-1}(0)$, we conclude from (2.23) and Lemma 2.17 that

$$\begin{aligned} \mathrm{L}([f_{(i,k,1)}, f_{(i,k,2)}]) \;\leq\;& \kappa\, |h' \circ \psi_i^{\kappa}|\, \|\nabla(\psi_i^{\kappa})\|\, \psi_i^{2\kappa-1}\, |g_{i,k}\psi_i| \\ &+ \kappa\, |h \circ \psi_i^{\kappa}|\, \|\nabla\psi_i^{2\kappa-1}\|\, |g_{i,k}\psi_i| + \kappa\, |h \circ \psi_i^{\kappa}|\, \psi_i^{2\kappa-1}\, \|\nabla(g_{i,k}\psi_i)\|, \end{aligned}$$

where we have used that $g_{i,k}$ is identically equal to a constant vector of length 1. We already know that h is bounded by a constant on $(0, r^{\kappa}]$. Since a is assumed to be smooth with $a(0) = 0$, it is easy to check that h' is bounded by a multiple of $y \mapsto y^{-1}$ on $(0, r^{\kappa}]$. It follows from Proposition 2.7 (b) that $\nabla\psi_i^{\kappa}$ and $\nabla\psi_i^{2\kappa-1}$ are bounded by a multiples of $\psi_i^{\kappa-1/2}$

and $\psi_i^{2\kappa-3/2}$ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$, respectively, and that $g_{i,k}\psi_i$ is bounded by a multiple of $\psi_i^{1/2}$ on $\psi_i^{-1}(\leq r)$. We also conclude from Proposition 2.7 (c) that $\nabla(g_{i,k}\psi_i)$ is bounded by a constant on $\psi_i^{-1}(\leq r)$. Consequently, $\mathrm{L}([f_{(i,k,1)}, f_{(i,k,2)}])$ is bounded by a multiple of $\psi_i^{2\kappa-1}$ on $\psi_i^{-1}(\leq r) \setminus \psi_i^{-1}(0)$. Thus, the same is true on $\psi_i^{-1}(\leq r)$, and, since $\kappa \geq 1/2$, it follows from Lemma 2.16 that $[f_{(i,k,1)}, f_{(i,k,2)}]$ is locally Lipschitz continuous. □

Because of Lemma 2.20, a well-defined locally Lipschitz continuous vector field f^∞ on P^N is given by

$$f^\infty \ := \ \sum_{i=1}^{N}\sum_{k=1}^{n}[f_{(i,k,1)}, f_{(i,k,2)}]. \tag{2.24}$$

Using (2.23) and $g_{i,k}\psi_i = g_{i,k}\psi$, we can write f^∞ also as

$$f^\infty \ = \ -\kappa\sum_{i=1}^{N}(h\circ\psi_i^\kappa)\,\psi_i^{2\kappa-1}\sum_{k=1}^{n}(g_{i,k}\psi)\,g_{i,k} \tag{2.25}$$

with the function h defined in Lemma 2.20. Using that the vector fields $g_{i,k}$ form an orthonormal frame of T^N, it is now easy to see that the differential equation $\dot{p} = f^\infty(p)$ on P^N coincides with the N coupled differential equations (2.18) on P. As indicated earlier, in a neighborhood of the set (2.14), the system state of (2.18) is constantly driven into a descent direction of ψ. We make this statement more precise by providing an estimate for the Lie derivative of ψ along f^∞:

Lemma 2.21. *There exist $c, r > 0$ such that*

$$(f^\infty\psi)(p) \ \leq \ -c\,\|\nabla\psi(p)\|^{4\kappa}$$

for every $p \in \psi^{-1}(\leq r)$.

Proof. Since a is assumed to be smooth with $a(0) = 0$ and $a'(0) \neq 0$, there exist $c_h, r > 0$ such that $h(y) \geq c_h y$ for every $y \in [0, r^\kappa]$. Because of (2.25), this implies

$$f^\infty\psi \ \leq \ -c_h\sum_{i=1}^{N}\psi_i^{2\kappa-1}\sum_{k=1}^{n}(g_{i,k}\psi)^2$$

on $\psi^{-1}(\leq r)$. We obtain from Proposition 2.7 (b) that, for every $i \in \{1, \ldots, N\}$, there exists $c_i > 0$ such that, for every $k \in \{1, \ldots, n\}$, we have

$$\psi_i \ \geq \ c_i\,\|\nabla\psi_i\|^2 \ \geq \ c_i\,(g_{i,k}\psi_i)^2 \ = \ c_i\,(g_{i,k}\psi)^2$$

on $\psi^{-1}(\leq r)$. Thus, there exists $\tilde{c} > 0$ such that

$$f^\infty\psi \ \leq \ -\tilde{c}\sum_{i=1}^{N}\sum_{k=1}^{n}(g_{i,k}\psi)^{4\kappa} \tag{2.26}$$

on $\psi^{-1}(\leq r)$. For each $\alpha \geq 1$ and each $v \in \mathbb{R}^{nN}$, the α-norm of v is defined as usual to be the αth root of the sum of the αth powers of the components of v. Note that the sum on the right-hand side of (2.26) is equal to the (4κ)th power of the 4κ-norm of the vector in $\mathbb{R}^{nN}$ with components $g_{i,k}\psi$. On the other hand, we have $\|\nabla\psi\|^2 = \sum_{i=1}^{N}\sum_{k=1}^{n}(g_{i,k}\psi)^2$ since the vector fields $g_{i,k}$ form an orthonormal frame of T^N. Using that all norms on $\mathbb{R}^{nN}$ are equivalent, we obtain the asserted estimate. □

2.5.2 Averaging of the sinusoids

The next step in the analysis of the closed-loop system (2.21) addresses the sinusoids u_ℓ^j therein. Instead of the differential equation (2.21), it is more convenient to consider the corresponding integral equation. Repeated integration by parts on the right-hand side of this integral equation will reveal that the functions u_ℓ^j give rise to an averaged vector field, which is given by the sum (2.24) of Lie brackets of the f_ℓ. A much more general treatment of this averaging procedure can be found in [58, 68, 69]. In the following, we introduce the notation from [68, 69]. For every $\ell = (i,k,\nu) \in \Lambda$, define two complex-valued constants $\eta_{\pm\omega_{i,k},\ell}$ as follows. If $\nu = 1$, let $\eta_{\pm\omega_{i,k},\ell} := \sqrt{2\omega_{i,k}}\,\mathrm{e}^{\pm\mathrm{i}\varphi_{i,k}}/2$, and otherwise, i.e., if $\nu = 2$, let $\eta_{\pm\omega_{i,k},\ell} := \pm\sqrt{2\omega_{i,k}}\,\mathrm{e}^{\pm\mathrm{i}\varphi_{i,k}}/(2\mathrm{i})$, where i denotes the imaginary unit and e denotes Euler's number. Moreover, let $\Omega(\ell) := \{\pm\omega_{i,k}\}$.

Let $\ell \in \Lambda$. Using the above notation, we can write $u_\ell^j(t)$ in (2.8) as

$$u_\ell^j(t) \;=\; j^{\frac{1}{2}} \sum_{\omega\in\Omega(\ell)} \eta_{\omega,\ell}\,\mathrm{e}^{\mathrm{i}j\omega t} \tag{2.27}$$

for every $t \in \mathbb{R}$. When we integrate $-u_\ell^j$, we get

$$-\int_{t_1}^{t_2} u_\ell^j(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_\ell^j(t)\Big]_{t=t_1}^{t=t_2}, \tag{2.28}$$

where

$$\widetilde{UV}_\ell^j(t) \;:=\; -j^{-\frac{1}{2}} \sum_{\omega\in\Omega(\ell)} \frac{\eta_{\omega,\ell}}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}j\omega t}. \tag{2.29}$$

Let $\ell_1, \ell_2 \in \Lambda$. When we multiply $u_{\ell_1}^j(t)$ by $\widetilde{UV}_{\ell_2}^j(t)$, we get

$$u_{\ell_1}^j(t)\,\widetilde{UV}_{\ell_2}^j(t) \;=\; v_{\ell_1,\ell_2} - \widetilde{uv}_{\ell_1,\ell_2}^j(t), \tag{2.30}$$

where

$$v_{\ell_1,\ell_2} \;:=\; -\sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2=0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}\,\omega_2}, \tag{2.31}$$

$$\widetilde{uv}_{\ell_1,\ell_2}^j(t) \;:=\; \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2\neq 0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}\,\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \tag{2.32}$$

for every $t \in \mathbb{R}$. When we integrate $\widetilde{uv}_{\ell_1,\ell_2}^j$, we get

$$\int_{t_1}^{t_2} \widetilde{uv}_{\ell_1,\ell_2}^j(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_{\ell_1,\ell_2}^j(t)\Big]_{t=t_1}^{t=t_2}, \tag{2.33}$$

where

$$\widetilde{UV}_{\ell_1,\ell_2}^j(t) \;:=\; j^{-1} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2\neq 0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}^2\,\omega_2(\omega_2+\omega_1)}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t}. \tag{2.34}$$

It is easy to see that the functions in (2.27), (2.29) and (2.34) satisfy the following estimates.

Lemma 2.22. *There exists $c > 0$ such that*

$$\begin{aligned} \left|u_\ell^j(t)\right| &\leq c\, j^{\frac{1}{2}}, \\ \left|\widetilde{UV}_\ell^j(t)\right| &\leq c\, j^{-\frac{1}{2}}, \\ \left|\widetilde{UV}_{\ell_1,\ell_2}^j(t)\right| &\leq c\, j^{-1} \end{aligned}$$

for every $j \geq 1$, all $\ell, \ell_1, \ell_2 \in \Lambda$, and every $t \in \mathbb{R}$.

This means that the functions $\widetilde{UV}_\ell^j$ and $\widetilde{UV}_{\ell_1,\ell_2}^j$ converge uniformly to 0 as the global frequency parameter j tends to ∞. Moreover, a direct computation shows that the v_{ℓ_1,ℓ_2} in (2.31) are given as follows.

Lemma 2.23. *For all $\ell_1 = (i_1, k_1, \nu_1), \ell_2 = (i_2, k_2, \nu_2) \in \Lambda$, we have*

$$v_{\ell_1,\ell_2} \;=\; \begin{cases} +1 & \text{if } (i_1,k_1) = (i_2,k_2) \text{ and } \nu_1 = 1 \text{ and } \nu_2 = 2, \\ -1 & \text{if } (i_1,k_1) = (i_2,k_2) \text{ and } \nu_1 = 2 \text{ and } \nu_2 = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Because of Lemma 2.23, we have

$$\sum_{\ell_1,\ell_2\in\Lambda} v_{\ell_1,\ell_2}\, f_{\ell_1}(f_{\ell_2}\varphi) \;=\; \sum_{i=1}^{N}\sum_{k=1}^{n}[f_{(i,k,1)}, f_{(i,k,2)}]\varphi \;=\; f^\infty\varphi \tag{2.35}$$

for every smooth function φ on P^N, where the vector field f^∞ on P^N is given by (2.24).

2.5.3 Integral expansion

For the moment, fix an arbitrary $j > 0$, let $\gamma\colon \mathbb{R} \to P^N$ be a trajectory of (2.21), and let $t_1, t_2 \in \mathbb{R}$. The fundamental theorem of calculus applied to the composition of γ and ψ implies that

$$\psi(\gamma(t_2)) \;=\; \psi(\gamma(t_1)) + \sum_{\ell\in\Lambda}\int_{t_1}^{t_2} u_\ell^j(t)\,(f_\ell\psi)(\gamma(t))\,\mathrm{d}t.$$

We know from Lemma 2.19 (b) that each of the Lie derivatives $f_\ell\psi$ is differentiable. Thus, we may apply integration by parts, which leads to

$$\begin{aligned} \psi(\gamma(t_2)) \;=\; & \psi(\gamma(t_1)) - \sum_{\ell\in\Lambda}\Big[\widetilde{UV}_\ell^j(t)\,(f_\ell\psi)(\gamma(t))\Big]_{t=t_1}^{t=t_2} \\ & + \sum_{\ell_1,\ell_2\in\Lambda}\int_{t_1}^{t_2} u_{\ell_1}^j(t)\,\widetilde{UV}_{\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\psi))(\gamma(t))\,\mathrm{d}t, \end{aligned}$$

where we have used (2.28) and that γ is a solution of (2.21). We know from Lemma 2.19 (b) that each of the Lie derivatives $f_{\ell_1}(f_{\ell_2}\psi)$ is locally Lipschitz continuous. Since γ is continuously differentiable, the composition of γ and $f_{\ell_1}(f_{\ell_2}\psi)$ is locally Lipschitz continuous, and therefore, in particular, locally absolutely continuous. Consequently, $f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma$ is differentiable almost everywhere on $\mathbb{R}$ with a locally integrable derivative, which will be denoted by $(f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma)^\cdot$. This justifies that we may apply again integration by parts, which leads to

$$\psi(\gamma(t_2)) \;=\; \psi(\gamma(t_1)) - \Big[(D_1^j\psi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \sum_{\ell_1,\ell_2\in\Lambda}\int_{t_1}^{t_2} v_{\ell_1,\ell_2}\,(f_{\ell_1}(f_{\ell_2}\psi))(\gamma(t))\,\mathrm{d}t$$

$$+ \sum_{\ell_1,\ell_2\in\Lambda} \int_{t_1}^{t_2} \widetilde{UV}^j_{\ell_1,\ell_2}(t)\,(f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma)\dot{}(t)\,\mathrm{d}t,$$

where we have used first (2.30) and then (2.33) as well as that γ is a solution of (2.21). The function $D^j_1\psi$ on $\mathbb{R}\times P^N$ is defined by

$$(D^j_1\psi)(t,p) \;:=\; \sum_{\ell_1\in\Lambda} \widetilde{UV}^j_{\ell_1}(t)\,(f_{\ell_1}\psi)(p) + \sum_{\ell_1,\ell_2\in\Lambda} \widetilde{UV}^j_{\ell_2,\ell_1}(t)\,(f_{\ell_2}(f_{\ell_1}\psi))(p). \tag{2.36}$$

If $f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma$ is differentiable at some $t\in\mathbb{R}$, then the local Lipschitz continuity of $f_{\ell_1}(f_{\ell_2}\psi)$ implies that

$$(f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma)\dot{}(t) \;=\; \lim_{s\to 0}\frac{1}{s}\Big((f_{\ell_1}(f_{\ell_2}\psi))\big(\gamma(t)+s\dot\gamma(t)\big) - (f_{\ell_1}(f_{\ell_2}\psi))(\gamma(t))\Big).$$

Let f^j be the time-varying vector field on P^N defined by (2.21). Then $\dot\gamma(t) = f^j(t,\gamma(t))$ for every $t\in\mathbb{R}$. The above equation, for the derivative of $f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma$ motivates us to define the (fixed-time) Lie derivative of $f_{\ell_1}(f_{\ell_2}\psi)$ along f^j by

$$(f^j(f_{\ell_1}(f_{\ell_2}\psi)))(t,p) \;:=\; \lim_{s\to 0}\frac{1}{s}\Big((f_{\ell_1}(f_{\ell_2}\psi))\big(p+sf^j(t,p)\big) - (f_{\ell_1}(f_{\ell_2}\psi))(p)\Big) \tag{2.37}$$

for every $t\in\mathbb{R}$ and every $p\in P^N$ at which the limit on the right-hand side of (2.37) exists. Then, obviously, we have

$$(f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma)\dot{}(t) \;=\; (f^j(f_{\ell_1}(f_{\ell_2}\psi)))(t,\gamma(t))$$

for almost every $t\in\mathbb{R}$. Therefore, we define

$$(D^j_2\psi)(t,p) \;:=\; \sum_{\ell_1,\ell_2\in\Lambda} \widetilde{UV}^j_{\ell_1,\ell_2}(t)\,(f^j(f_{\ell_1}(f_{\ell_2}\psi)))(t,p) \tag{2.38}$$

for every $t\in\mathbb{R}$ and every $p\in P^N$ at which the finitely many (fixed-time) Lie derivatives $f^j(f_{\ell_1}(f_{\ell_2}\psi))$ exist. Consequently, we have

$$\sum_{\ell_1,\ell_2\in\Lambda} \int_{t_1}^{t_2} \widetilde{UV}^j_{\ell_1,\ell_2}(t)\,(f_{\ell_1}(f_{\ell_2}\psi)\circ\gamma)\dot{}(t)\,\mathrm{d}t \;=\; \int_{t_1}^{t_2} (D^j_2\psi)(t,\gamma(t))\,\mathrm{d}t.$$

Because of (2.35), we have derived the following integral expansion for the propagation of ψ along trajectories of (2.21); cf. equation (1.37) in Chapter 1.

Proposition 2.24. *For every $j>0$, every trajectory $\gamma\colon\mathbb{R}\to P^N$ of* (2.21)*, and all $t_1,t_2\in\mathbb{R}$, we have*

$$\psi(\gamma(t_2)) \;=\; \psi(\gamma(t_1)) - \Big[(D^j_1\psi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2}(f^\infty\psi)(\gamma(t))\,\mathrm{d}t + \int_{t_1}^{t_2}(D^j_2\psi)(t,\gamma(t))\,\mathrm{d}t,$$

where $f^\infty\psi$, $D^j_1\psi$, and $D^j_2\psi$ are given by (2.35)*,* (2.36)*, and* (2.38)*, respectively.*

Proposition 2.24 implies that the propagation of ψ along trajectories of (2.21) is dominated by the averaged term $f^\infty\psi$ if the contributions of the remainder terms $D^j_1\psi$ and $D^j_2\psi$ are small compared to $f^\infty\psi$. A lower bound for the magnitude of $f^\infty\psi$ is given in the following result.

Proposition 2.25. *Suppose that, for every point p of (2.14), the framework $G(p)$ is infinitesimally rigid. Then, there exist $r_0, c_0 > 0$ such that*

$$(f^\infty\psi)(p) \;\leq\; -c_0\,\psi(p)^{2\kappa}$$

for every $p \in \psi^{-1}(\leq r_0)$.

Proof. The result is an immediate consequence of Lemma 2.21 and Proposition 2.7 (d). □

For the proofs of Theorems 2.12 and 2.13, we also need the subsequent estimates for the remainders $D_1^j\psi$ and $D_2^j\psi$ in Proposition 2.24.

Proposition 2.26. *(a) For every $r > 0$, there exists $c_1 > 0$ such that*

$$|D_1^j\psi(t,p)| \;\leq\; c_1\,j^{-\frac{1}{2}}\,\psi(p)^{\kappa+1/2}$$

for every $j \geq 1$, every $t \in \mathbb{R}$, and every $p \in \psi^{-1}(\leq r)$.

(b) For every $r > 0$, there exists $c_2 > 0$ such that

$$|D_2^j\psi(t,p)| \;\leq\; c_2\,j^{-\frac{1}{2}}\,\psi(p)^{3\kappa-1/2}$$

for every $j \geq 1$, every $t \in \mathbb{R}$, and every $p \in \psi^{-1}(\leq r)$ at which (2.38) exists.

Proof. We already know estimates for the Lie derivatives $f_{\ell_1}\psi$, $f_{\ell_1}(f_{\ell_2}\psi)$ and the sinusoids $\widetilde{UV}_{\ell_1}^j$, $\widetilde{UV}_{\ell_1,\ell_2}^j$ from Lemma 2.19 (b) and Lemma 2.22, respectively. If we apply those estimates to the constituents of $D_1^j\psi$ in (2.36), then we immediately obtain part (a) of Proposition 2.26. Moreover, from Lemma 2.22 and (2.38), we derive that there exists $c > 0$ such that

$$|D_2^j\psi(t,p)| \;\leq\; c\,j^{-1}\sum_{\ell_1,\ell_2\in\Lambda}\big|(f^j(f_{\ell_1}(f_{\ell_2}\psi)))(t,p)\big|$$

for every $j \geq 1$, every $t \in \mathbb{R}$, and every $p \in P^N$ at which (2.38) exists. We know from Lemma 2.19 (b) that each of the functions $f_{\ell_1}(f_{\ell_2}\psi)$ is locally Lipschitz continuous. Using Lemma 2.17, it is easy to check that, for all $\ell_1, \ell_2 \in \Lambda$, every $j \geq 1$, every $t \in \mathbb{R}$, and every $p \in P^N$ at which (2.37) exists, we have

$$|(f^j(f_{\ell_1}(f_{\ell_2}\psi)))(t,p)| \;\leq\; \mathrm{L}(f_{\ell_1}(f_{\ell_2}\psi))(p)\,\|f^j(t,p)\|,$$

where $\mathrm{L}(f_{\ell_1}(f_{\ell_2}\psi))(p)$ is the pointwise Lipschitz constant of $f_{\ell_1}(f_{\ell_2}\psi)$ at p from Lemma 2.16. We know from Lemma 2.19 (b) that $\mathrm{L}(f_{\ell_1}(f_{\ell_2}\psi))$ is bounded by a multiple of $\psi^{2\kappa-1/2}$ on $\psi^{-1}(\leq r)$. Using the definition of f^j in (2.21), Lemma 2.22, and Lemma 2.19 (a), it is easy to see that there exists $c' > 0$ such that $\|f^j(t,p)\| \leq c'\,j^{1/2}\,\psi^\kappa$ for every $j \geq 1$, every $t \in \mathbb{R}$, and every $p \in \psi^{-1}(\leq r)$. This implies statement (b) and completes the proof. □

Remark 2.27. The above Propositions 2.25 and 2.26 give a detailed description on how the frequency parameter j and the exponent κ influence the remainders $D_1^j\psi$ and $D_2^j\psi$ compared to the averaged term $f^\infty\psi$ in the integral expansion of Proposition 2.24. To induce a decay of the potential function ψ along solutions of the closed-loop system, we need that the negative contribution of $f^\infty\psi$ is sufficiently large compared to the contributions of $D_1^j\psi$ and $D_2^j\psi$. We distinguish the following two cases, which correspond to Theorem 2.12 and Theorem 2.13, respectively.

Suppose that $\kappa = 1/2$. Then, the exponents of $\psi(p)$ in the estimates of Propositions 2.25 and 2.26 are all equal to 1. This means that contribution of both the averaged term $f^\infty\psi$ and the remainders $D_1^j\psi$ and $D_2^j\psi$ vanish with the same speed as the value of ψ tends to the optimal value 0. To ensure that the contribution of $f^\infty\psi$ is much larger than the contribution of $D_1^j\psi$ and $D_2^j\psi$, we have to choose a sufficiently large value of j. If we apply the estimates in Propositions 2.25 and 2.26 to the integral expansion in Proposition 2.24 and take the limit $j \to \infty$, then we obtain an estimate of the form

$$\psi(\gamma(t_2)) \;\leq\; \psi(\gamma(t_1)) - \int_{t_1}^{t_2} c_0\,\psi(\gamma(t))\,\mathrm{d}t.$$

This indicates why Theorem 2.12 ensures an exponential decay of ψ under the assumption of a sufficiently large frequency parameter j. A more precise argument is given in Subsection 2.5.4.

Suppose that $\kappa > 1/2$. Then, the exponents of $\psi(p)$ in the estimates of Propositions 2.25 and 2.26 are all greater than 1. Moreover, the exponent $3\kappa - 1/2$ of $\psi(p)$ in the estimate for $D_2^j\psi$ is greater than the exponent 2κ in the estimate for $f^\infty\psi$. Thus, even if j is not large, the contribution of the $D_2^j\psi$ becomes arbitrary small compared to $f^\infty\psi$ when ψ is sufficiently close to the optimal value 0. This explains why Theorem 2.13 ensures a decay of ψ for an arbitrary value of j, but under the assumption that the initial value of ψ is already sufficiently close to 0. If we increase the frequency parameter j, then the remainders become smaller and therefore a decay of ψ also occurs for larger initial values of ψ. If we apply the estimates in Propositions 2.25 and 2.26 to the integral expansion in Proposition 2.24 and take the limit $j \to \infty$, then we obtain an estimate of the form

$$\psi(\gamma(t_2)) \;\leq\; \psi(\gamma(t_1)) - \int_{t_1}^{t_2} c_0\,\psi(\gamma(t))^{2\kappa}\,\mathrm{d}t,$$

where $2\kappa > 1$. This indicates why Theorem 2.13 does not ensure an exponential decay of ψ but only a power law decay. A more precise argument is given below. ♢

2.5.4 Proofs of Theorems 2.12 and 2.13

Suppose that, for every point p of (2.14), the framework $G(p)$ is infinitesimally rigid. Then, there exist $r_0, c_0 > 0$ as in Proposition 2.25. For this sublevel r_0, there exist $c_1, c_2 > 0$ as in Proposition 2.26 (a) and (b). From Proposition 2.24, we conclude that, for every $j \geq 1$, every trajectory $\gamma\colon \mathbb{R} \to P^N$ of (2.21), and all $t_1 < t_2$ in $\mathbb{R}$, the following implication holds: if $\gamma(t) \in \psi^{-1}(\leq r_0)$ for every $t \in [t_1, t_2]$, then

$$\psi(\gamma(t_2)) \;\leq\; \psi(\gamma(t_1)) + c_1\,j^{-\frac{1}{2}}\,\psi(\gamma(t_2))^{\kappa+1/2} + c_1\,j^{-\frac{1}{2}}\,\psi(\gamma(t_1))^{\kappa+1/2} \tag{2.39a}$$

$$- \int_{t_1}^{t_2} \big(c_0 - c_2\,j^{-\frac{1}{2}}\,\psi(\gamma(t))^{\kappa-1/2}\big)\,\psi(\gamma(t))^{2\kappa}\,\mathrm{d}t. \tag{2.39b}$$

Next, we distinguish two cases.

Theorem 2.12

Suppose that $\kappa = 1/2$. If $j > 1/c_1^2$, then inequality (2.39) can be written as

$$\psi(\gamma(t_2)) \;\leq\; \frac{1 + c_1 j^{-\frac{1}{2}}}{1 - c_1 j^{-\frac{1}{2}}}\psi(\gamma(t_1)) - \frac{c_0 - c_2\,j^{-\frac{1}{2}}}{1 - c_1 j^{-\frac{1}{2}}}\int_{t_1}^{t_2} \psi(\gamma(t))\,\mathrm{d}t.$$

Let $r := r_0/2$, $\mu \in (0, c_0/2)$ and $\lambda \in (1,2)$. It is now clear that we can find a sufficiently large $j_0 > 0$ such that, for every $j \geq j_0$, every trajectory $\gamma\colon \mathbb{R} \to P^N$ of (2.21), and all $t_1 < t_2$ in $\mathbb{R}$, the following implication holds: if $\gamma(t) \in \psi^{-1}(\leq 2r)$ for every $t \in [t_1, t_2]$, then

$$\psi(\gamma(t_2)) \;\leq\; \lambda\, \psi(\gamma(t_1)) - \lambda\, \mu \int_{t_1}^{t_2} \psi(\gamma(t))\, \mathrm{d}t.$$

Now a standard comparison argument for integral inequalities implies that inequality (2.15) holds.[6] Moreover, the exponential decay of ψ along γ implies that $\gamma(t)$ converges to some element of $\psi^{-1}(0)$ as $t \to \infty$.

Theorem 2.13

Suppose that $\kappa > 1/2$ and that $j \geq 1$. If $\psi(\gamma(t_2))^{\kappa-1/2} < j^{1/2}/c_1$, then inequality (2.39) can be written as

$$\psi(\gamma(t_2)) \;\leq\; \frac{1 + c_1 j^{-\frac{1}{2}} \psi(\gamma(t_1))^{\kappa-1/2}}{1 - c_1 j^{-\frac{1}{2}} \psi(\gamma(t_2))^{\kappa-1/2}} \psi(\gamma(t_1)) - \int_{t_1}^{t_2} \frac{c_0 - c_2 j^{-\frac{1}{2}} \psi(\gamma(t))^{\kappa-1/2}}{1 - c_1 j^{-\frac{1}{2}} \psi(\gamma(t_2))^{\kappa-1/2}} \psi(\gamma(t))^{2\kappa}\, \mathrm{d}t.$$

Let $\mu \in (0, c_0/2)$ and $\lambda \in (1,2)$. It is now clear that we can find a sufficiently small $r > 0$ such that, for every trajectory $\gamma\colon \mathbb{R} \to P^N$ of (2.21), and all $t_1 < t_2$ in $\mathbb{R}$, the following implication holds: if $\gamma(t) \in \psi^{-1}(\leq 2r)$ for every $t \in [t_1, t_2]$, then

$$\psi(\gamma(t_2)) \;\leq\; \lambda\, \psi(\gamma(t_1)) - \lambda\, \mu \int_{t_1}^{t_2} \psi(\gamma(t))^{2\kappa}\, \mathrm{d}t.$$

Now a standard comparison argument for integral inequalities implies that inequality (2.16) holds.[7] It is left to prove that the trajectories of (2.21) with initial values in $\psi^{-1}(\leq r)$ converge to some point of (2.14). For this purpose, fix a trajectory γ of (2.21) with $\psi(\gamma(t_0)) \leq r$ for some $t_0 \in \mathbb{R}$. We already know from (2.16) that $\psi(\gamma(t)) \leq 2r$ for every $t \geq t_0$. By integrating (2.21), we obtain that

$$\gamma(t_2) - \gamma(t_1) \;=\; \sum_{\ell\in\Lambda} \int_{t_1}^{t_2} u_\ell^j(t)\, f_\ell(\gamma(t))\, \mathrm{d}t$$

for all $t_2 \geq t_1 \geq t_0$. Note that the above equation is understood as an equation on the translation space T^N of P^N. We know from Lemma 2.19 (c) that each of the vector fields f_ℓ is differentiable. This justifies that we may apply integration by parts in the above integral. Using (2.28) and that γ is a solution of (2.21), we obtain

$$\begin{aligned} \gamma(t_2) - \gamma(t_1) \;=\; & -\sum_{\ell\in\Lambda} \widetilde{UV}_\ell^j(t_2)\, f_\ell(\gamma(t_2)) + \sum_{\ell\in\Lambda} \widetilde{UV}_\ell^j(t_1)\, f_\ell(\gamma(t_1)) \\ & + \sum_{\ell_1,\ell_2\in\Lambda} \int_{t_1}^{t_2} u_{\ell_1}^j(t) \widetilde{UV}_{\ell_2}^j(t)\, \nabla_{f_{\ell_1}} f_{\ell_2}(\gamma(t))\, \mathrm{d}t \end{aligned}$$

[6] To see this, use the fact that $y(t) := \psi(\gamma(t_0))\, \mathrm{e}^{-\mu(t-t_0)}$ defines a function $y\colon [t_0,\infty) \to \mathbb{R}$ that satisfies the integral equation $y(t_2) = y(t_1) - \mu \int_{t_1}^{t_2} y(t)\, \mathrm{d}t$ for all $t_2 \geq t_1 \geq t_0$.

[7] To see this, use the fact that $y(t) := \psi(\gamma(t_0)) \left(1 + (2\kappa-1)\psi(\gamma(t_0))^{2\kappa-1}\, \mu\, (t-t_0)\right)^{-\frac{1}{2\kappa-1}}$ defines a function $y\colon [t_0,\infty) \to \mathbb{R}$ that satisfies the integral equation $y(t_2) = y(t_1) - \mu \int_{t_1}^{t_2} y(t)^{2\kappa}\, \mathrm{d}t$ for all $t_2 \geq t_1 \geq t_0$.

for all $t_2 \geq t_1 \geq t_0$. It now follows from Lemmas 2.19 and 2.22 that there exist $c_1', c_2' > 0$ such that

$$\|\gamma(t_2) - \gamma(t_1)\| \;\leq\; c_1'\,\psi(\gamma(t_2))^{\kappa} + c_1'\,\psi(\gamma(t_1))^{\kappa} + \int_{t_1}^{t_2} c_2'\,\psi(\gamma(t))^{2\kappa-1/2}\,\mathrm{d}t$$

for all $t_2 \geq t_1 \geq t_0$. Now we apply estimate (2.16) and obtain that

$$\|\gamma(t_2) - \gamma(t_1)\| \;\leq\; \frac{2\,c_1'\,\lambda_0^{\kappa}}{\big(1+\mu_0\,(t_1-t_0)\big)^{\frac{\kappa}{2\kappa-1}}} + \int_{t_1}^{t_2} \frac{c_2'\,\lambda_0^{2\kappa-1/2}}{\big(1+\mu_0\,(t-t_0)\big)^{\frac{2\kappa-1/2}{2\kappa-1}}}\,\mathrm{d}t$$

for all $t_2 \geq t_1 \geq t_0$, where $\lambda_0 := \lambda\,\psi(\gamma(t_0))$ and $\mu_0 := (2\kappa-1)\psi(\gamma(t_0))^{2\kappa-1}\,\mu$. Note that for the exponent of the denominator in the above integral, we have $\frac{2\kappa-1/2}{2\kappa-1} > 1$. This allows us to conclude that, for every $\varepsilon > 0$, there exists $T > t_0$ such that $\big\|\gamma(t_2) - \gamma(t_1)\big\| \leq \varepsilon$ for all $t_2 \geq t_1 \geq T$. It follows that $\gamma(t)$ converges to some $p^* \in P^N$ as $t \to \infty$. Since $\psi(\gamma(t)) \to 0$ as $t \to \infty$, we conclude that p^* is an element of (2.14).

2.6 Extension to formation control for unicycles with all-to-all communication

In this section, we propose an extension of the control strategy in Section 2.4 for point agents to the case of nonholonomic unicycles.

2.6.1 Problem description

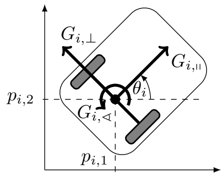
We consider a system of N unicycles in the Euclidean plane. In local coordinates with, a vector $p_i \in \mathbb{R}^2$ for the position and an angle $\theta_i \in \mathbb{R}$ for the orientation, the kinematic equations for the ith unicycle read

$$\dot{p}_i \;=\; u_{i,\shortparallel}\begin{pmatrix}\cos\theta_i\\ \sin\theta_i\end{pmatrix}, \tag{2.40a}$$

$$\dot{\theta}_i \;=\; u_{i,\sphericalangle}, \tag{2.40b}$$

where $u_{i,\shortparallel}$ and $u_{i,\sphericalangle}$ are real-valued input channels for the translational and the rotational velocity, respectively. As in Subsection 2.4.1, we suppose that the agents are equipped with sensors so that they can measure the distances to other members of the team according to an (undirected) graph $G = (V, E)$ of N vertices. For each edge $i_1 i_2 \in E$, let $d_{i_1 i_2} \geq 0$ be a nonnegative real number, which is the *desired distance* between agents i_1 and i_2. We assume again that these distances are realizable in the two-dimensional space. In contrast to the situation described in Subsection 2.4.1, we additionally require that the agents have the ability to exchange measured data so that each agent knows the values of all current distances $\|p_{i_2} - p_{i_1}\|$ with $i_1 i_2 \in E$ at any given time. This means that we require *all-to-all communication*. For this situation, we are interested in a distance-only control law that steers the multi-agent system into a target formation with $\|p_{i_2} - p_{i_1}\| = d_{i_1 i_2}$ for all $i_1 i_2 \in E$. Note that a target formation does not impose any constraints on the orientations of the unicycles. In particular, a target formation does not exclude time-varying orientations.

The entire multi-agent system of N unicycles is a system on the N-fold product $\mathrm{SE}(2)^N$ of the special Euclidean Group[8] SE(2) of dimension 2. For each $i \in \{1, \ldots, N\}$, let $G_{i,\parallel}$ and $G_{i,\triangleleft}$ denote the smooth vector fields on $\mathrm{SE}(2)^N$ that describe the direction of the translational and the directional velocity of the ith unicycle, respectively; i.e., in the local coordinates of (2.40), we have



$$G_{i,\parallel} \;=\; \cos(\theta_i)\,\frac{\partial}{\partial p_{i,1}} + \sin(\theta_i)\,\frac{\partial}{\partial p_{i,2}}, \tag{2.41}$$

$$G_{i,\triangleleft} \;=\; \frac{\partial}{\partial \theta_i}. \tag{2.42}$$

Now, the N equations in (2.40) can be combined to the control-affine system

$$\dot{x} \;=\; \sum_{i=1}^{N} \big(u_{i,\parallel}\, G_{i,\parallel}(x) + u_{i,\triangleleft}\, G_{i,\triangleleft}(x)\big) \tag{2.43}$$

on $\mathrm{SE}(2)^N$. For the rest of this section, we proceed in the coordinate-free description (2.43).

2.6.2 Control law and main statement

The control law for (2.43) will be composed of the following constituents.

(1) Let $\psi\colon P^N \to \mathbb{R}$ be the global potential function defined by (2.13). Since we assume all-to-all communication, each agent knows the current value

$$y \;=\; \psi(p) \tag{2.44}$$

of ψ at any given time.

(2) Let $a\colon \mathbb{R} \to \mathbb{R}$ be a smooth function and bounded such that $a(0) = 0$ and $a'(0) \neq 0$. Let the functions h_1, h_2 on $\mathbb{R}$ be defined as in (2.7).

(3) Let $n_1, n_2, \ldots$ denote the prime numbers in increasing order; i.e., $n_1 = 2, n_2 = 3, \ldots$. For every $i \in \{1, \ldots, N\}$ and every $j > 0$, define three sinusoids $u^j_{(i,1)}, u^j_{(i,2)}, u^j_{(i,3)}$ by

$$u^j_{(i,1)}(t) \;:=\; (j\,\omega_{(i,1)})^{\frac{3}{4}}\, \cos(j\,\omega_{(i,1)}t + \varphi_i), \tag{2.45a}$$

$$u^j_{(i,2)}(t) \;:=\; (j\,\omega_{(i,2)})^{\frac{3}{4}}\, \sin(j\,\omega_{(i,2)}t + \varphi_i), \tag{2.45b}$$

$$u^j_{(i,3)}(t) \;:=\; 2^{13/8}\,(j\,\omega_{(i,3)})^{\frac{3}{4}}\, \cos(j\,\omega_{(i,3)}t + \varphi_i) \tag{2.45c}$$

with arbitrary shifts $\varphi_i \in \mathbb{R}$ and frequency coefficients

$$\omega_{(i,1)} \;:=\; 3\sqrt{n_{i+1}} + 2\sqrt{2n_{i+1}}, \tag{2.46a}$$

$$\omega_{(i,2)} \;:=\; \sqrt{n_{i+1}}, \tag{2.46b}$$

$$\omega_{(i,3)} \;:=\; 2\sqrt{n_{i+1}} + \sqrt{2n_{i+1}}. \tag{2.46c}$$

Remark 2.28. We briefly give some preliminary comments on the above functions without going into details here.

[8] For the objectives in this section it is convenient to treat the three-dimensional smooth manifold SE(2) as the Cartesian product of the two-dimensional Euclidean space (to describe the position) and the unit circle (to describe the orientation).

(1) In contrast to Subsection 2.4.2, we do not use local potential functions but the global potential function ψ for the distance-only control law. A computation of the current value of ψ by the agents requires all-to-all communication. The reason for using a global potential function instead of local potential functions is of rather "technical nature". The stability analysis of the proposed control method in Section 2.7 requires an extension of Lemma 2.19 to iterated Lie derivatives of ψ along more than just two vector fields. However, due to non-smoothness, those higher-order Lie derivatives do not necessarily exist if the vector fields depend on the values of the local potential functions. Lemma 2.19 (b) already indicates that iterated Lie derivatives of ψ along two vector fields are locally Lipschitz continuous but, in general, not differentiable. It turns out that this problem can be circumvented by using the global potential function for each agent. A suitable stability analysis for a fully distributed control law with local potential functions is left to future research.

(2) To obtain a global potential function on the state manifold $\mathrm{SE}(2)^N$, we introduce the canonical projection

$$\pi\colon\ \mathrm{SE}(2)^N \to P^N$$

of $\mathrm{SE}(2)^N$ onto P^N, where $P \cong \mathbb{R}^2$ is the underlying Euclidean space of dimension 2. Now we can define the global potential function

$$\Psi \ := \ \psi \circ \pi\colon\ \mathrm{SE}(2)^N \to \mathbb{R}. \tag{2.47}$$

It follows from Lemma 2.18 (a) that, for each $\nu \in \{1, 2\}$, the composition

$$h_\nu \circ \Psi^{1/2}\colon\ \mathrm{SE}(2)^N \to \mathbb{R}$$

is locally Lipschitz continuous, where $\Psi^{1/2}$ denotes the square root of Ψ.

(3) As for point agents, the choice of the sinusoids has the purpose to approximate certain Lie brackets. While the approach for point agents is based on an approximation of Lie brackets of two vector fields, the approach for unicycles will involve approximations of (iterated) Lie brackets of four vector fields. This strategy is further explained in the text after Theorem 2.30. ◊

Given a real number $j > 0$, we propose the control law

$$u_{i,\parallel} \ = \ u^j_{(i,1)}(t)\, h_1(y^{1/2}) + u^j_{(i,2)}(t)\, h_2(y^{1/2}), \tag{2.48a}$$

$$u_{i,\triangleleft} \ = \ u^j_{(i,3)}(t) \tag{2.48b}$$

for every $i \in \{1, \ldots, N\}$, where $y^{1/2}$ denotes the square root of the value y as in (2.44). Note that a computation of $y = \Psi(\pi(x))$ only requires measurements of inter-agent distances at the current system state $x \in \mathrm{SE}(2)^N$. Whenever $y > 0$, we can write control law (2.48) also as

$$u_{i,k} \ = \ \sqrt{2\, j\, \omega_{(i,k)}}\; a(y^{1/2})\, \sin(j\, \omega_{(i,k)}\, t + \varphi_i + \log y^{1/2}), \tag{2.49}$$

where we have used the trigonometric identity (1.49). Because of Remark 2.28 and the boundedness of h_1, h_2, we obtain the following result from standard existence and uniqueness properties for ordinary differential equations.

Proposition 2.29. *For any initial condition, system* (2.40) *under control law* (2.48) *has a unique global solution, which we call a* trajectory *of* (2.40) *under* (2.48)*.*

For unicycles, the set of desired states is the preimage of the set (2.14) for point agents under the projection map π; i.e., the set $\Psi^{-1}(0)$. Similar to Theorem 2.12, we will prove exponential stability under the assumption of infinitesimal rigidity:

Theorem 2.30. *Suppose that, for every point p of (2.14), the framework $G(p)$ is infinitesimally rigid. Then, there exist $\mu, r > 0$ such that, for every $\lambda > 1$, there exists $j_0 > 0$ such that, for every $j \geq j_0$, every $t_0 \in \mathbb{R}$, and every $x_0 \in \Psi^{-1}(\leq r)$, the trajectory x of system (2.40) under control law (2.48) with initial condition $x(t_0) = x_0$ has the following two properties: $\pi(x(t))$ converges to some point of (2.14) as $t \to \infty$, and the estimate*

$$\Psi(x(t)) \;\leq\; \lambda\, \Psi(x_0)\, \mathrm{e}^{-\mu(t-t_0)} \tag{2.50}$$

holds for every $t \geq t_0$.

At this point, we only indicate why control law (2.48) leads to local exponential stability. A detailed proof of Theorem 2.30 is given in Section 2.7. Recall that the team of unicycles is described by the control system (2.43) on the state manifold $\mathrm{SE}(2)^N$ with the control vector fields $G_{i,\shortparallel}$, $G_{i,\triangleleft}$ given by (2.41), (2.42) for every $i \in \{1, \ldots, N\}$. Each of the $G_{i,\triangleleft}$ describes the rotation of a unicycle around its axis. Such rotations do not influence the value of the global potential function Ψ. This means that the Lie derivative of Ψ along each of the $G_{i,\triangleleft}$ is identically equal to zero. Changes of Ψ are induced by motions along the current alignment of the unicycles, which are described by the vector fields $G_{i,\shortparallel}$. However, the vector fields $G_{i,\shortparallel}$ do not give immediate access to all directions on the underlying space P^N. In particular, the steepest descent direction of Ψ at some $x \in \mathrm{SE}(2)^N$ is not necessarily in the span of the vectors $G_{1,\shortparallel}(x), \ldots, G_{N,\shortparallel}(x)$. Therefore, a gradient-based strategy as in Section 2.4 for point agents cannot be applied directly. An approximation of the steepest descent direction of Ψ would require that the unicycles can also move perpendicular to their current alignments. These directions are described by the Lie brackets

$$G_{i,\perp} \;:=\; [G_{i,\triangleleft}, G_{i,\shortparallel}] \;=\; -\sin(\theta_i)\,\frac{\partial}{\partial p_{i,1}} + \cos(\theta_i)\,\frac{\partial}{\partial p_{i,2}} \tag{2.51}$$

for $i = 1, \ldots, N$, where the expression on the right-hand side uses the same local coordinates as in (2.41) and (2.42). The averaging analysis in Section 2.7 will reveal that the particular choice of the sinusoids in (2.45) allows the unicycles to move (approximately) along the directions of the iterated Lie brackets

$$\big[(h_1 \circ \Psi^{1/2})G_{i,\shortparallel}, \big[G_{i,\triangleleft}, \big[(h_2 \circ \Psi^{1/2})G_{i,\shortparallel}, G_{i,\triangleleft}\big]\big]\big] \;=\; -\frac{1}{2}(h \circ \Psi^{1/2})\,(G_{i,\shortparallel}\Psi)\,G_{i,\shortparallel}, \tag{2.52}$$

$$\big[\big[G_{i,\triangleleft}, (h_1 \circ \Psi^{1/2})G_{i,\shortparallel}\big], \big[G_{i,\triangleleft}, (h_2 \circ \Psi^{1/2})G_{i,\shortparallel}\big]\big] \;=\; -\frac{1}{2}(h \circ \Psi^{1/2})\,(G_{i,\perp}\Psi)\,G_{i,\perp} \tag{2.53}$$

for $i = 1, \ldots, N$, where the function h on $\mathbb{R}$ is defined in Lemma 2.20. To be more precise, one can prove that, in the limit $j \to \infty$, the trajectories of the closed-loop system approximate the trajectories of the averaged system

$$\dot{x} \;=\; F^{\infty}(x), \tag{2.54}$$

where the averaged vector field

$$F^{\infty} \;:=\; -\frac{1}{2}(h \circ \Psi^{1/2}) \sum_{i=1}^{N} \big((G_{i,\shortparallel}\Psi)\,G_{i,\shortparallel} + (G_{i,\perp}\Psi)\,G_{i,\perp}\big) \tag{2.55}$$

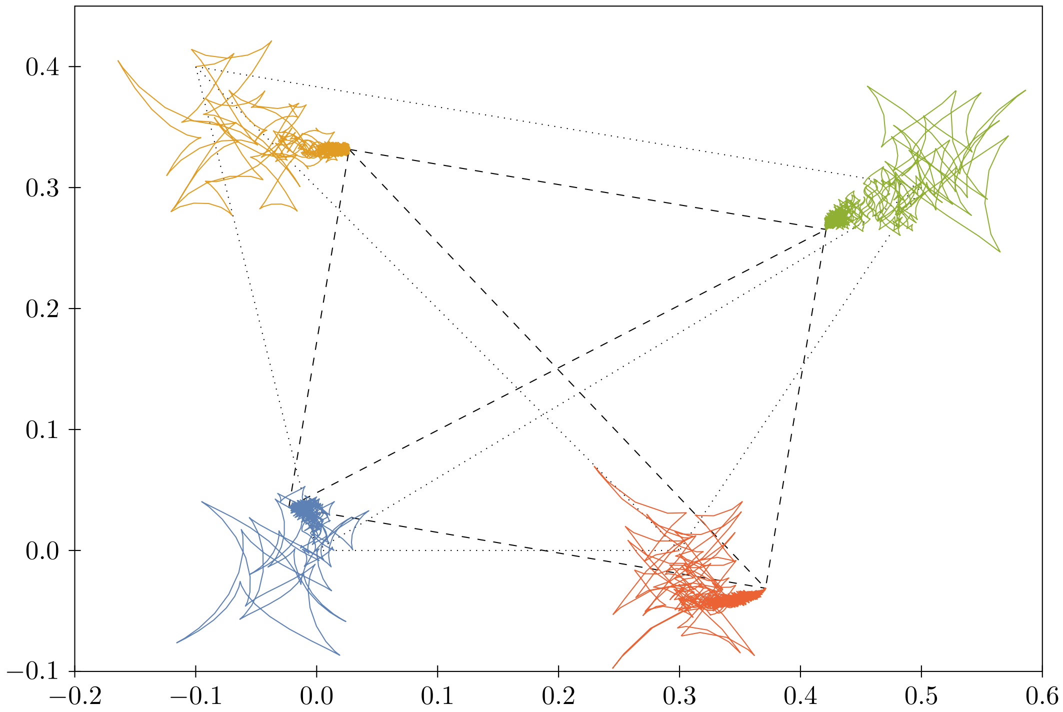


Figure 2.5: *Simulation result for a team of four unicycle agents under control law* (2.48). *The initial positions and the desired distances are the same as for the team of point agents in Figure 2.2.*

on $\mathrm{SE}(2)^N$ is the sum of the Lie brackets in (2.52) and (2.53). It is easy to check that, in local coordinates, system (2.54) reads

$$\dot{p}_i = -\frac{1}{2} h(\psi^{1/2}(p))\, \nabla_{p_i}\psi(p), \tag{2.56a}$$

$$\dot{\theta}_i = 0 \tag{2.56b}$$

for $i = 1, \ldots, N$, where $\nabla_{p_i}\psi$ is coordinate representation of the gradient of ψ with respect to the ith position variable. Thus, in the limit $j \to \infty$, the nonholonomic unicycles approximate the behavior of fully actuated kinematic points under a gradient-based control law. For this reason, we get the same local exponential stability for unicycles as in Section 2.4 for point agents.

Remark 2.31. It is also possible to derive a control law that leads to asymptotic stability without the dependence on the frequency parameter j as in Theorem 2.13 for point agents. For this purpose, the exponent 1/2 in control law (2.48) has to be replaced by an exponent $\kappa > 1/2$. In this case, the proof of asymptotic stability is a suitable combination of the arguments in Sections 2.5 and 2.7. ♢

Before we begin with the stability analysis, we present some numerical results. We consider a team of $N = 4$ nonholonomic unicycles with all-to-all communication. The communication graph G, the desired distances $d_{ii'}$, and the initial positions $p_i(0)$ are the same as in Subsection 2.4.3 for the team of non-communicating point agents. The initial orientations are given by $\theta_i(0) = \phi_i$ with the same angles ϕ_i as in Subsection 2.4.3 for

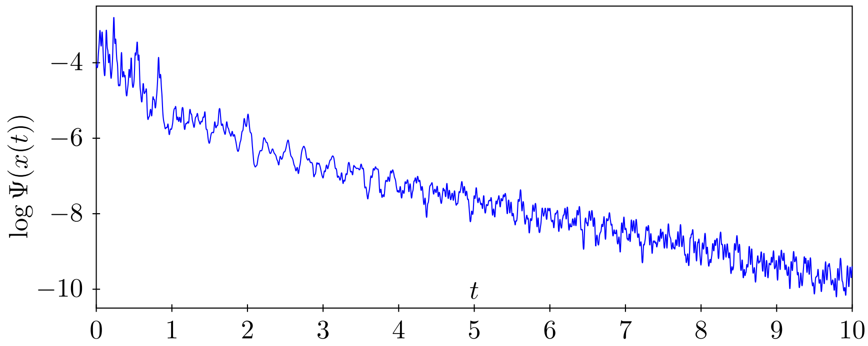


Figure 2.6: *Exponential decay of the global potential function* (2.47) *for the multi-agent system in Figure 2.5 on the time interval* $[0, 10]$.

$i = 1, \ldots, N$. The simulation results are generated for the frequency parameter $j = 10$. It can be seen in Figures 2.5 and 2.6 that the unicycles converge exponentially fast to a desired formation. The speed of convergence is approximately the same as for point agents in Figures 2.2 and 2.3.

2.7 Local asymptotic stability analysis for unicycles

For the moment, fix an arbitrary trajectory $\gamma\colon \mathbb{R} \to \mathrm{SE}(2)^N$ of (2.40) under (2.48) for some initial value $\gamma(t_0) \in \mathrm{SE}(2)^N$ at some initial time $t_0 \in \mathbb{R}$. If $\gamma(t_0)$ is an element of $\Psi^{-1}(0)$, then (2.48a) and the uniqueness of solutions imply that $\pi \circ \gamma$ is identically equal to some point of (2.14). Thus, there is nothing to show for initial values in $\Psi^{-1}(0)$. Conversely, if $\gamma(t_0)$ is outside $\Psi^{-1}(0)$, then γ will never enter $\Psi^{-1}(0)$ at any time. For this reason, it suffices to study the closed-loop system on the open submanifold

$$M \;:=\; \{x \in \mathrm{SE}(2)^N \mid \Psi(x) > 0\}$$

of $\mathrm{SE}(2)^N$. For the rest of the section, we restrict Ψ to the submanifold M and denote the restriction again by the symbol Ψ. Thus, in particular, for any $r > 0$, the sublevel set $\Psi^{-1}(\leq r)$ only contains the points $x \in M$ with $\Psi(x) \leq r$.

For every $i \in \{1, \ldots, N\}$, define three smooth vector fields $F_{(i,1)}, F_{(i,2)}, F_{(i,3)}$ on M by

$$F_{(i,1)}(x) \;:=\; h_1(\Psi^{1/2}(x))\, G_{i,\shortparallel}(x), \tag{2.57a}$$

$$F_{(i,2)}(x) \;:=\; h_2(\Psi^{1/2}(x))\, G_{i,\shortparallel}(x), \tag{2.57b}$$

$$F_{(i,3)}(x) \;:=\; G_{i,\triangleleft}(x). \tag{2.57c}$$

As an abbreviation, we define an indexing set Λ to be the set of all pairs (i, ν) with $i \in \{1, \ldots, N\}$ and $\nu \in \{1, 2, 3\}$. It is now easy to see that the restriction of system (2.40) under the control law (2.48) to the submanifold M can be written as the closed-loop system

$$\dot{x} \;=\; \sum_{\ell \in \Lambda} u_\ell^j(t)\, F_\ell(x) \tag{2.58}$$

on M, which may be interpreted as a control-affine system with control vector fields F_ℓ under open-loop controls u_ℓ^j. For every $r > 0$, let $\Psi^{-1}(\leq r)$ denote the set of all $x \in M$ with $\Psi(x) \leq r$. The statement of Theorem 2.30 follows immediately if we can show the subsequent result.

Proposition 2.32. *Suppose that, for every point p of (2.14), the framework $G(p)$ is infinitesimally rigid. Then, there exist $\mu, r > 0$ such that, for every $\lambda > 1$, there exists $j_0 > 0$ such that, for every $j \geq j_0$, every $t_0 \in \mathbb{R}$, every $x_0 \in \Psi^{-1}(\leq r)$, and every trajectory $\gamma\colon \mathbb{R} \to M$ of (2.58) with initial condition $\gamma(t_0) = x_0$, we have*

$$\Psi(\gamma(t)) \;\leq\; \lambda\, \Psi(x_0)\, \mathrm{e}^{-\mu(t-t_0)}$$

holds for every $t \geq t_0$.

2.7.1 Estimates for the Lie derivatives

Since the vector fields F_ℓ in (2.57) are smooth, for every positive integer k, all $\ell_1, \cdots, \ell_k \in \Lambda$, and every smooth function φ on M, we may define the iterated Lie derivative

$$F_{\ell_1,\ldots,\ell_k}\varphi \;:=\; F_{\ell_1}(\cdots(F_{\ell_k}\varphi)\cdots)$$

of φ along $F_{\ell_1}, \ldots, F_{\ell_k}$, which is again a smooth function on M. As in Subsection 2.5.1, for a subset S of M, a function φ on M, and a nonnegative function b on M, we say that φ is bounded by a multiple of b on S if there exists some positive constant c such that $|\varphi(x)| \leq c\, b(x)$ for every $x \in S$.

Lemma 2.33. *For every $r > 0$, every $k \in \{1, \ldots, 5\}$, and all $\ell_1, \ldots, \ell_k \in \Lambda$, the iterated Lie derivative $F_{\ell_1,\ldots,\ell_k}\Psi$ is bounded by a multiple of Ψ on $\Psi^{-1}(\leq r)$.*

Proof. We only indicate the proof for $k = 1$ and $k = 2$. A similar strategy, which requires considerably more computational effort, can be applied for $k = 3, 4, 5$.

Let $r > 0$ and $\ell_1 = (i_1, \nu_1), \ell_2 = (i_2, \nu_2) \in \Lambda$. If $\nu_2 = 3$, then $F_{\ell_2}\Psi$ is identically equal to zero. It therefore suffices to consider the case in which $\nu_2 \in \{1, 2\}$. Then,

$$F_{\ell_2}\Psi \;=\; (h_{\nu_2} \circ \Psi^{1/2})\,(G_{i_2,\shortparallel}\Psi).$$

If follows from Lemma 2.15 (b) and $h_{\nu_2}(0) = 0$ that h_{ν_2} is bounded by a multiple of the identity $y \mapsto y$ on $[0, r^{1/2}]$. Therefore, $h_{\nu_2} \circ \Psi^{1/2}$ is bounded by a multiple of $\Psi^{1/2}$ on $\Psi^{-1}(\leq r)$. It follows from Proposition 2.7 (b) that $G_{i_2,\shortparallel}\Psi$ is bounded by a multiple of $\Psi^{1/2}$ on $\Psi^{-1}(\leq r)$. Thus, $F_{\ell_2}\Psi$ is bounded by a multiple of Ψ on $\Psi^{-1}(\leq r)$.

For the Lie derivative of $F_{\ell_2}\Psi$ along F_{ℓ_1}, we distinguish three cases. Again, we only consider the nontrivial case with $\nu_2 \in \{1, 2\}$. First, suppose that $\nu_1 = 3$ and $i_1 = i_2$. Then, we have

$$F_{\ell_1}(F_{\ell_2}\Psi) \;=\; (h_{\nu_2} \circ \Psi^{1/2})\,(G_{i_2,\triangleleft}(G_{i_2,\shortparallel}\Psi)).$$

A direct computation leads to $G_{i_2,\triangleleft}(G_{i_2,\shortparallel}\Psi) = G_{i_2,\perp}\Psi$. It follows from Proposition 2.7 (b) that $G_{i_2,\perp}\Psi$ is bounded by a multiple of $\Psi^{1/2}$ on $\Psi^{-1}(\leq r)$. Therefore, $F_{\ell_1}(F_{\ell_2}\Psi)$ is bounded by a multiple of Ψ on $\Psi^{-1}(\leq r)$. Second, suppose that $\nu_1 = 3$ and $i_1 \neq i_2$. Then, $F_{\ell_1}(F_{\ell_2}\Psi)$ is identically equal to zero. It remains to consider the case in which $\nu_1 \neq 3$. Then,

$$\begin{aligned} F_{\ell_1}(F_{\ell_2}\Psi) \;=\;& (h_{\nu_1} \circ \Psi^{1/2})\,(h'_{\nu_2} \circ \Psi^{1/2})\,(G_{i_1,\shortparallel}\Psi^{1/2})\,(G_{i_2,\shortparallel}\Psi) \\ &+ (h_{\nu_1} \circ \Psi^{1/2})\,(h_{\nu_2} \circ \Psi^{1/2})\,(G_{i_1,\shortparallel}(G_{i_2,\shortparallel}\Psi)). \end{aligned}$$

We already know that $h_{\nu_1} \circ \Psi^{1/2}$, $h_{\nu_2} \circ \Psi^{1/2}$, and $G_{i_2,\shortparallel}\Psi$ are bounded by multiples of $\Psi^{1/2}$ on $\Psi^{-1}(\leq r)$. By the chain rule, we conclude that $G_{i_1,\shortparallel}\Psi^{1/2}$ is bounded by a constant on $\Psi^{-1}(\leq r)$. It follows from Lemma 2.15 (b) that $h'_{\nu_2} \circ \Psi^{1/2}$ is also bounded by a constant on

$\Psi^{-1}(\leq r)$. Finally, we obtain from Proposition 2.7 (c) that $G_{i_1,\shortparallel}(G_{i_2,\shortparallel}\Psi)$ is again bounded by a constant on $\Psi^{-1}(\leq r)$. We conclude that $F_{\ell_1}(F_{\ell_2}\Psi)$ is bounded by a multiple of Ψ on $\Psi^{-1}(\leq r)$.

For $k = 3, 4, 5$, the asserted boundedness properties can be derived from explicit computations of $F_{\ell_1,\ldots,\ell_k}\Psi$ as above. This requires suitable estimates for derivatives of higher order. For instance, it is easy to check that, for every $\nu \in \{1, 2\}$ and every integer $k \geq 2$, the kth derivative of h_ν is bounded by a multiple of $y \mapsto y^{k-1}$ on $(0, r^{1/2}]$. Moreover, boundedness statement for iterated Lie derivatives of Ψ and $\Psi^{1/2}$ along more than one vector fields can be derived from Proposition 2.7 (c) and the chain rule. □

2.7.2 Averaging of the sinusoids

An averaging analysis for the sinusoids u^j_ℓ in (2.45) requires the following two technical lemmas on integer linear combinations of the frequency coefficients ω_ℓ, $\ell \in \Lambda$, in (2.46).

Lemma 2.34. *Let $n_1, n_2, \ldots$ denote the prime numbers in increasing order; i.e., $n_1 = 2, n_2 = 3, \ldots$. Then, the real numbers $\sqrt{n_2}, \sqrt{2n_2}, \sqrt{n_3}, \sqrt{2n_3}, \ldots$ are linearly independent over the ring of integers.*

Proof. The statement is an immediate consequence of the main result in [11]. □

Lemma 2.35. *The frequency coefficients ω_ℓ in (2.46) with ℓ ranging over Λ are pairwise distinct. Moreover, for all $\omega_1, \ldots, \omega_4 \in \{\pm\omega_\ell \mid \ell \in \Lambda\}$, the following statements hold:*

(1) we always have $\omega_1 \neq 0$;

(2) if $\omega_1 + \omega_2 = 0$, then there exists $\ell \in \Lambda$ such that $\{\omega_1, \omega_2\} = \{\pm\omega_\ell\}$;

(3) we always have $\omega_1 + \omega_2 + \omega_3 \neq 0$;

(4) if $\omega_1 + \cdots + \omega_4 = 0$, then there exists a permutation σ of $1, \ldots, 4$ such that either

(i) $\omega_{\sigma(1)} + \omega_{\sigma(2)} = 0$ and $\omega_{\sigma(3)} + \omega_{\sigma(4)} = 0$, or

(ii) there exists $i \in \{1, \ldots, N\}$ and $s \in \{\pm 1\}$ such that $\omega_{\sigma(1)} = s\omega_{(i,1)}$, $\omega_{\sigma(2)} = s\omega_{(i,2)}$, $\omega_{\sigma(3)} = \omega_{\sigma(4)} = -s\omega_{(i,3)}$.

Proof. For the proof, let Γ denote the set of the six integer pairs $\pm(3, 2)$, $\pm(1, 0)$, and $\pm(2, 1)$. Note that, by Lemma 2.34, for each $k \in \{1, \ldots, 4\}$, there exist unique $i_k \in \{1, \ldots, N\}$ and $(x_k, y_k) \in \Gamma$ such that

$$\omega_k \;=\; x_k\sqrt{n_{i_k+1}} + y_k\sqrt{2n_{i_k+1}}. \tag{2.59}$$

Statement (1) is obvious. Statement (2) follows immediately from the unique representation (2.59) and Lemma 2.34. To prove statement (3), suppose for the sake of contradiction that $\omega_1 + \omega_2 + \omega_3 = 0$. Then, Lemma 2.34 implies that the numbers i_1, i_2, i_3 are all equal and that the sum of the vectors $(x_1, y_1), (x_2, y_3), (x_3, y_3)$ is equal to $(0, 0)$. However, this is a contradiction since one can easily check that the sum of three vectors from the set Γ is never equal to zero. Finally, to prove statement (4), suppose that $\omega_1 + \cdots + \omega_4 = 0$. It is easy to see that cases (i) and (ii) cannot be satisfied simultaneously. If the numbers $i_1, \ldots, i_4$ in (2.59) are not equal to some $i \in \{1, \ldots, N\}$, then Lemma 2.34 and statements (1) and (3) imply that the canceling comes from pairs of the ω_k as in case (i). It is left to consider the case in which the numbers $i_1, \ldots, i_4$ in (2.59) are all equal to some $i \in \{1, \ldots, N\}$. Then, because of Lemma 2.34, the sum of the vectors $(x_1, y_1), \ldots, (x_4, y_4)$ is equal to $(0, 0)$. It is now easy to check that there are only the following two cases in which the sum of four vectors from the set Γ is equal to zero: either, each of the (x_k, y_k) is canceled by its negative, which corresponds again to case (i), or, there exists $s \in \{\pm 1\}$ such that the (x_k, y_k) are a permutation of the vectors $s(3,2)$, $s(1,0)$, $-s(2,1)$, $-s(2,1)$ which corresponds to case (ii). □

In the following, we introduce the notation from [68, 69]. For every $\ell \in \Lambda$, define $\Omega(\ell) := \{\pm\omega_\ell\}$ with $\omega_\ell > 0$ as in (2.46). It follows from Lemma 2.35 that the sets $\Omega(\ell)$ are pairwise disjoint. For every $i \in \{1, \ldots, N\}$, we define the six complex-valued constants

$$\eta_{\pm\omega_{(i,1)}} := \frac{1}{2}\,\omega_{(i,1)}^{\frac{3}{4}}\,\mathrm{e}^{\pm \mathrm{i}\varphi_i}, \qquad \eta_{\pm\omega_{(i,2)}} := \pm\frac{1}{2\mathrm{i}}\,\omega_{(i,2)}^{\frac{3}{4}}\,\mathrm{e}^{\pm \mathrm{i}\varphi_i}, \qquad \eta_{\pm\omega_{(i,3)}} := \frac{2^{13/8}}{2}\,\omega_{(i,3)}^{\frac{3}{4}}\,\mathrm{e}^{\pm \mathrm{i}\varphi_i},$$

where i denotes the imaginary unit and e denotes Euler's number.

Let $\ell \in \Lambda$. Using the above notation, we can write $u_\ell^j(t)$ in (2.45) as

$$u_\ell^j(t) = j^{\frac{3}{4}} \sum_{\omega\in\Omega(\ell)} \eta_\omega\,\mathrm{e}^{\mathrm{i}j\omega t} \tag{2.60}$$

for every $t \in \mathbb{R}$. Note that $\omega \neq 0$ for every $\omega \in \Omega(\ell)$ by Lemma 2.35 (1). Thus, when we integrate $-u_\ell^j$, we get

$$-\int_{t_1}^{t_2} u_\ell^j(t)\,\mathrm{d}t = \Big[\widetilde{UV}_\ell^j(t)\Big]_{t=t_1}^{t=t_2}, \tag{2.61}$$

where

$$\widetilde{UV}_\ell^j(t) := -j^{-\frac{1}{4}} \sum_{\omega\in\Omega(\ell)} \frac{\eta_\omega}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}j\omega t}. \tag{2.62}$$

Let $\ell_1, \ell_2 \in \Lambda$. When we multiply $u_{\ell_1}^j(t)$ by $\widetilde{UV}_{\ell_2}^j(t)$, we get

$$u_{\ell_1}^j(t)\widetilde{UV}_{\ell_2}^j(t) = -\widetilde{uv}_{\ell_1,\ell_2}^j(t), \tag{2.63}$$

where

$$\widetilde{uv}_{\ell_1,\ell_2}^j(t) := j^{\frac{2}{4}} \sum_{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)} \frac{\eta_{\omega_1}\,\eta_{\omega_2}}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \tag{2.64}$$

for every $t \in \mathbb{R}$. Note that, for each pair (ω_1, ω_2) in $\Omega(\ell_1) \times \Omega(\ell_2)$, also $(-\omega_1, -\omega_2)$ is in $\Omega(\ell_1) \times \Omega(\ell_2)$. By definition of the η_ω, we then have

$$\frac{\eta_{\omega_1}\,\eta_{-\omega_1}}{\mathrm{i}(-\omega_1)} + \frac{\eta_{-\omega_1}\,\eta_{\omega_1}}{\mathrm{i}\omega_1} = 0.$$

Because of this pairwise canceling, the summation in (2.64) reduces to the elements of the set

$$\Omega(\ell_1,\ell_2) := \{(\omega_1,\omega_2) \in \Omega(\ell_1)\times\Omega(\ell_2) \mid \omega_1+\omega_2 \neq 0\}.$$

Thus, when we integrate $\widetilde{uv}_{\ell_1,\ell_2}^j$, we get

$$\int_{t_1}^{t_2} \widetilde{uv}_{\ell_1,\ell_2}^j(t)\,\mathrm{d}t = \Big[\widetilde{UV}_{\ell_1,\ell_2}^j(t)\Big]_{t=t_1}^{t=t_2}, \tag{2.65}$$

where

$$\widetilde{UV}_{\ell_1,\ell_2}^j(t) := j^{-\frac{2}{4}} \sum_{(\omega_1,\omega_2)\in\Omega(\ell_1,\ell_2)} \frac{\eta_{\omega_1}\,\eta_{\omega_2}}{\mathrm{i}^2\omega_2(\omega_1+\omega_2)}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t}. \tag{2.66}$$

Let $\ell_1, \ell_2, \ell_3 \in \Lambda$. When we multiply $u^j_{\ell_1}(t)$ by $\widetilde{UV}^j_{\ell_2,\ell_3}(t)$, we obtain

$$u^j_{\ell_1}(t)\widetilde{UV}^j_{\ell_2,\ell_3}(t) \;=\; -\widetilde{uv}^j_{\ell_1,\ell_2,\ell_3}(t), \tag{2.67}$$

where

$$\widetilde{uv}^j_{\ell_1,\ell_2,\ell_3}(t) \;:=\; j^{\frac{1}{4}} \sum_{(\omega_1,\omega_2,\omega_3)\in\Omega(\ell_1,\ell_2,\ell_3)} \frac{\eta_{\omega_1}\,\eta_{\omega_2}\,\eta_{\omega_3}}{\mathrm{i}^2\omega_2(\omega_1+\omega_2)}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2+\omega_3)t} \tag{2.68}$$

and

$$\Omega(\ell_1,\ell_2,\ell_3) \;:=\; \{(\omega_1,\omega_2,\omega_3) \mid \omega_1 \in \Omega(\ell_1), (\omega_2,\omega_3) \in \Omega(\ell_2,\ell_3)\}.$$

Because of Lemma 2.35 (3), we have $\omega_1+\omega_2+\omega_3 \neq 0$ for every $(\omega_1,\omega_2,\omega_3) \in \Omega(\ell_1,\ell_2,\ell_3)$. Thus, when we integrate $\widetilde{uv}^j_{\ell_1,\ell_2,\ell_3}$, we get

$$\int_{t_1}^{t_2} \widetilde{uv}^j_{\ell_1,\ell_2,\ell_3}(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}^j_{\ell_1,\ell_2,\ell_3}(t)\Big]_{t=t_1}^{t=t_2}, \tag{2.69}$$

where

$$\widetilde{UV}^j_{\ell_1,\ell_2,\ell_3}(t) \;:=\; -j^{-\frac{3}{4}} \sum_{\hat\omega\in\Omega(\ell_1,\ell_2,\ell_3)} \frac{\eta_{\hat\omega}}{\mathrm{i}^3\Pi(\hat\omega)}\,\mathrm{e}^{\mathrm{i}j\Sigma(\hat\omega)t} \tag{2.70}$$

with the abbreviations

$$\begin{aligned} \eta_{\hat\omega} &:= \eta_{\omega_1}\,\eta_{\omega_2}\,\eta_{\omega_3}, \\ \Pi(\hat\omega) &:= \omega_3(\omega_2+\omega_3)(\omega_1+\omega_2+\omega_3), \\ \Sigma(\hat\omega) &:= \omega_1+\omega_2+\omega_3 \end{aligned}$$

for every $\hat\omega := (\omega_1,\omega_2,\omega_3) \in \Omega(\ell_1,\ell_2,\ell_3)$.

Let $\ell_1, \ell_2, \ell_3, \ell_4 \in \Lambda$. When we multiply $u^j_{\ell_1}(t)$ by $\widetilde{UV}^j_{\ell_2,\ell_3,\ell_4}(t)$, we obtain

$$u^j_{\ell_1}(t)\widetilde{UV}^j_{\ell_2,\ell_3,\ell_4}(t) \;=\; v_{\ell_1,\ell_2,\ell_3,\ell_4} - \widetilde{uv}^j_{\ell_1,\ell_2,\ell_3,\ell_4}(t), \tag{2.71}$$

where

$$v_{\ell_1,\ell_2,\ell_3,\ell_4} \;:=\; -\sum_{\substack{(\omega_1,\hat\omega)\in\Omega(\ell_1)\times\Omega(\ell_2,\ell_3,\ell_4) \\ \omega_1+\Sigma(\hat\omega)=0}} \frac{\eta_{\omega_1}\eta_{\hat\omega}}{\mathrm{i}^3\Pi(\hat\omega)}, \tag{2.72}$$

$$\widetilde{uv}^j_{\ell_1,\ell_2,\ell_3,\ell_4}(t) \;:=\; \sum_{(\omega_1,\hat\omega)\in\Omega(\ell_1)\times\Omega(\ell_2,\ell_3,\ell_4)} \frac{\eta_{\omega_1}\eta_{\hat\omega}}{\mathrm{i}^3\Pi(\hat\omega)}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat\omega))t}. \tag{2.73}$$

When we integrate $\widetilde{uv}^j_{\ell_1,\ell_2,\ell_3,\ell_4}$, we get

$$\int_{t_1}^{t_2} \widetilde{uv}^j_{\ell_1,\ell_2,\ell_3,\ell_4}(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}^j_{\ell_1,\ell_2,\ell_3,\ell_4}(t)\Big]_{t=t_1}^{t=t_2}, \tag{2.74}$$

where

$$\widetilde{UV}^j_{\ell_1,\ell_2,\ell_3,\ell_4}(t) \;:=\; j^{-1} \sum_{\substack{(\omega_1,\hat\omega)\in\Omega(\ell_1)\times\Omega(\ell_2,\ell_3,\ell_4) \\ \omega_1+\Sigma(\hat\omega)\neq 0}} \frac{\eta_{\omega_1}\eta_{\hat\omega}}{\mathrm{i}^4\Pi(\hat\omega)(\omega_1+\Sigma(\hat\omega))}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat\omega))t}. \tag{2.75}$$

The functions in (2.60), (2.62), (2.66), (2.70) and (2.75) satisfy the following estimates.

Lemma 2.36. *There exists $c > 0$ such that*

$$\begin{aligned} \left|u_\ell^j(t)\right| &\leq c\, j^{\frac{3}{4}}, \\ \left|\widetilde{UV}^j_{\ell_1,\ldots,\ell_k}(t)\right| &\leq c\, j^{-\frac{k}{4}} \end{aligned}$$

for every $j \geq 1$, every $k \in \{1,\ldots,4\}$, all $\ell, \ell_1, \ldots, \ell_k \in \Lambda$, and every $t \in \mathbb{R}$.

The coefficients $v_{\ell_1,\ell_2,\ell_3,\ell_4}$ in (2.72) can be computed explicitly as follows.

Lemma 2.37. *For all $\ell_1,\ldots,\ell_4 \in \Lambda$, we have $v_{\ell_1,\ldots,\ell_4} = 0$ except for the coefficients*

$$\begin{aligned} v_{(1,2,3,3)_i} &= -\frac{1}{2} + \frac{1}{\sqrt{2}}, & v_{(1,3,2,3)_i} &= +2 - \sqrt{2}, \\ v_{(1,3,3,2)_i} &= -2, & v_{(2,1,3,3)_i} &= +\frac{1}{2} + \frac{1}{\sqrt{2}}, \\ v_{(2,3,1,3)_i} &= -2 - \sqrt{2}, & v_{(2,3,3,1)_i} &= +2, \\ v_{(3,1,2,3)_i} &= -1, & v_{(3,1,3,2)_i} &= +2 + \sqrt{2}, \\ v_{(3,2,1,3)_i} &= +1, & v_{(3,2,3,1)_i} &= -2 + \sqrt{2}, \\ v_{(3,3,1,2)_i} &= -\frac{1}{2} - \frac{1}{\sqrt{2}}, & v_{(3,3,2,1)_i} &= +\frac{1}{2} - \frac{1}{\sqrt{2}} \end{aligned}$$

for every $i \in \{1,\ldots,N\}$, where we have used the abbreviation

$$(\nu_1,\nu_2,\nu_3,\nu_4)_i \;:=\; (i,\nu_1),(i,\nu_2),(i,\nu_3),(i,\nu_4)$$

for all $\nu_1,\nu_2,\nu_3,\nu_4 \in \{1,2,3\}$.

Proof. Let $\ell_1,\ldots,\ell_4 \in \Lambda$. Suppose that $\omega_1 \in \Omega(\ell_1)$ and $\hat{\omega} := (\omega_2,\omega_3,\omega_4) \in \Omega(\ell_2,\ell_3,\ell_4)$ satisfy $\omega_1 + \cdots + \omega_4 = 0$. Then, there exists a permutation of $1,\ldots,4$ such that we are either in the situation of case (i) or in the situation of case (ii) of Lemma 2.35 (4). Suppose that we are in the situation of case (i). Note that also $-\omega_1 \in \Omega(\ell_1)$ and $-\hat{\omega} \in \Omega(\ell_2,\ell_3,\ell_4)$ satisfy $-\omega_1 + \Sigma(-\hat{\omega}) = 0$. Then, by definition of the coefficients η_ω, we have $\eta_{\omega_1}\eta_{\hat{\omega}} = \eta_{-\omega_1}\eta_{-\hat{\omega}}$. Moreover, we have $\Pi(\hat{\omega}) = -\Pi(-\hat{\omega})$, and therefore

$$\frac{\eta_{\omega_1}\eta_{\hat{\omega}}}{\mathrm{i}^3\Pi(\hat{\omega})} + \frac{\eta_{-\omega_1}\eta_{-\hat{\omega}}}{\mathrm{i}^3\Pi(-\hat{\omega})} \;=\; 0.$$

Thus, a necessary condition for a nonvanishing contribution by $(\omega_1,\hat{\omega})$ is that we are in the situation of case (ii) in Lemma 2.35 (4). This implies that there exist $i \in \{1,\ldots,4\}$ such that $(\ell_1,\ell_2,\ell_3,\ell_4) = (\nu_1,\nu_2,\nu_3,\nu_4)_i$, where $(\nu_1,\nu_2,\nu_3,\nu_4)$ is one of the 12 permutations of $(1,2,3,3)$. A direct computations shows that those 12 coefficients coincide with the ones in the list of Lemma 2.37. □

Remark 2.38. There are at least two different ways to identify the vector fields F_ℓ in (2.57) and the coefficients $v_{\ell_1},\ldots_{\ell_4}$ in Lemma 2.37 with the vector field F^∞ in (2.55). The first option is a long but direct computation that leads to

$$\sum_{\ell_1,\ldots,\ell_4 \in \Lambda} v_{\ell_1,\ldots,\ell_4} F_{\ell_1,\ldots,\ell_4}\varphi \;=\; F^\infty\varphi \tag{2.76}$$

for every smooth function φ on M, where, by a slight abuse of notation, the symbol F^∞ denotes restriction of the vector field in (2.55) from $\mathrm{SE}(2)^N$ to M. The second option is less labor intense but requires a result from [69] that is also used later in Lemma 4.14. In the terminology of [69], the estimates in Lemma 2.36 imply that the functions u_i^j satisfy the property of so-called *GD(4)-convergence* with respect to the coefficients $v_{\ell_1,\ldots,\ell_4}$ in Lemma 2.37. The same algebraic argument as in [69], see also Lemma 4.14, implies that

$$\sum_{\ell_1,\ldots,\ell_4\in\Lambda} v_{\ell_1,\ldots,\ell_4}\, F_{\ell_1,\ldots,\ell_4}\varphi \;=\; \frac{1}{4}\sum_{\ell_1,\ldots,\ell_4\in\Lambda} v_{\ell_1,\ldots,\ell_4}\,[F_{\ell_1,\ldots,\ell_4}]\varphi$$

for every smooth function φ on M, where $[F_{\ell_1,\ldots,\ell_4}]$ abbreviates the iterated Lie bracket

$$[F_{\ell_1,\ldots,\ell_4}] \;=\; [F_{\ell_1},[F_{\ell_2},[F_{\ell_3},F_{\ell_4}]]]$$

for all $\ell_1,\ldots,\ell_4\in\Lambda$. The Lie brackets $[F_{\ell_1,\ldots,\ell_4}]$ are easier to compute than the Lie derivatives $F_{\ell_1,\ldots,\ell_4}\varphi$. This leads to

$$\begin{aligned}
[F_{(1,2,3,3)_i}] &= -[F_{(2,1,3,3)_i}] = 0,\\
[F_{(1,3,3,2)_i}] &= -[F_{(1,3,2,3)_i}] = -(h\circ\Psi^{1/2})\,(G_{i,\shortparallel}\Psi)G_{i,\shortparallel},\\
[F_{(2,3,1,3)_i}] &= -[F_{(2,3,3,1)_i}] = -(h\circ\Psi^{1/2})(G_{i,\shortparallel}\Psi)G_{i,\shortparallel},\\
[F_{(3,1,2,3)_i}] &= -[F_{(3,1,3,2)_i}] = +(h\circ\Psi^{1/2})(G_{i,\shortparallel}\Psi)G_{i,\shortparallel} - (h\circ\Psi^{1/2})(G_{i,\perp}\Psi)G_{i,\perp},\\
[F_{(3,2,1,3)_i}] &= -[F_{(3,2,3,1)_i}] = -(h\circ\Psi^{1/2})(G_{i,\shortparallel}\Psi)G_{i,\shortparallel} + (h\circ\Psi^{1/2})(G_{i,\perp}\Psi)G_{i,\perp},\\
[F_{(3,3,1,2)_i}] &= -[F_{(3,3,2,1)_i}] = -2\,(h\circ\Psi^{1/2})(G_{i,\shortparallel}\Psi)G_{i,\shortparallel} + 2\,(h\circ\Psi^{1/2})(G_{i,\perp}\Psi)G_{i,\perp}
\end{aligned}$$

on M for every $i\in\{1,\ldots,N\}$ in the notation of Lemma 2.37. Now, we easily obtain from Lemma 2.37 and (2.55) that

$$\frac{1}{4}\sum_{\ell_1,\ldots,\ell_4\in\Lambda} v_{\ell_1,\ldots,\ell_4}\,[F_{\ell_1,\ldots,\ell_4}] \;=\; F^\infty$$

holds on M. ◊

Using (2.55) and (2.76) with $\varphi=\Psi$, we obtain that

$$(F^\infty\Psi)(x) \;=\; -\frac{1}{2}\,h(\Psi^{1/2}(x))\sum_{i=1}^N\big((G_{i,\shortparallel}\Psi)(x)^2-(G_{i,\perp}\Psi)(x)^2\big)$$

for every $x\in M$. Since the vector fields $G_{i,\shortparallel}$, $G_{i,\perp}$ in (2.41), (2.51) generate an orthonormal frame of the underlying space P^N, and because of the definition of Ψ in (2.47), this implies

$$(F^\infty\Psi)(x) \;=\; -\frac{1}{2}\,h((\psi^{1/2}\circ\pi)(x))\,\|((\nabla\psi)\circ\pi)(x)\|^2 \tag{2.77}$$

for every $x\in M$.

2.7.3 Integral expansion

For the moment, fix an arbitrary $j>0$, let $\gamma\colon\mathbb{R}\to M$ be a trajectory of (2.58), and let $t_1,t_2\in\mathbb{R}$. The fundamental theorem of calculus applied to the composition of γ and Ψ implies that

$$\Psi(\gamma(t_2)) \;=\; \Psi(\gamma(t_1)) + \sum_{\ell\in\Lambda}\int_{t_1}^{t_2} u_\ell^j(t)\,(F_\ell\Psi)(\gamma(t))\,\mathrm{d}t.$$

We apply integration by parts, which leads to

$$\begin{aligned}\Psi(\gamma(t_2)) &= \Psi(\gamma(t_1)) - \sum_{\ell\in\Lambda}\Big[\widetilde{UV}{}^j_\ell(t)\,(F_\ell\Psi)(\gamma(t))\Big]_{t=t_1}^{t=t_2} \\ &\quad + \sum_{\ell_1,\ell_2\in\Lambda}\int_{t_1}^{t_2} u^j_{\ell_1}(t)\,\widetilde{UV}{}^j_{\ell_2}(t)\,(F_{\ell_1}(F_{\ell_2}\Psi))(\gamma(t))\,\mathrm{d}t,\end{aligned}$$

where we have used first (2.63) and then (2.65) as well as that γ is a solution of (2.58). Next, we apply integration by parts three more times. In this context, we also use (2.63), (2.65), (2.67), (2.69), (2.71) and (2.74) as well as that γ is a solution of (2.58). This leads to

$$\begin{aligned}\Psi(\gamma(t_2)) &= \Psi(\gamma(t_1)) - \Big[(D^j_1\Psi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2}(D^j_2\Psi)(t,\gamma(t))\,\mathrm{d}t \\ &\quad + \sum_{\ell_1,\ldots,\ell_4\in\Lambda}\int_{t_1}^{t_2} v_{\ell_1,\ldots,\ell_4}\,(F_{\ell_1,\ldots,\ell_4}\Psi)(\gamma(t))\,\mathrm{d}t,\end{aligned}$$

where the functions $D^j_1\Psi$, $D^j_2\Psi$ on $\mathbb{R}\times M$ are defined by

$$(D^j_1\Psi)(t,x) := \sum_{k=1}^{4}\sum_{\ell_1,\ldots,\ell_k\in\Lambda}\widetilde{UV}{}^j_{\ell_1,\ldots,\ell_k}(t)\,(F_{\ell_1,\ldots,\ell_k}\Psi)(x), \tag{2.78}$$

$$(D^j_2\Psi)(t,x) := \sum_{\ell_1,\ldots,\ell_5\in\Lambda} u^j_{\ell_1}(t)\,\widetilde{UV}{}^j_{\ell_2,\ldots,\ell_5}(t)\,(F_{\ell_1,\ldots,\ell_5}\Psi)(x). \tag{2.79}$$

Thus, we are basically in the same situation as in Proposition 2.24 but with different functions $D^j_1\Psi$ and $D^j_2\Psi$. Because of (2.76) with $\varphi = \Psi$, we have derived the following integral expansion for the propagation of Ψ along trajectories of (2.58).

Proposition 2.39. *For every $j > 0$, every trajectory $\gamma\colon\mathbb{R}\to M$ of* (2.58)*, and all $t_1, t_2\in\mathbb{R}$, we have*

$$\Psi(\gamma(t_2)) = \Psi(\gamma(t_1)) - \Big[(D^j_1\Psi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2}(F^\infty\Psi)(\gamma(t))\,\mathrm{d}t + \int_{t_1}^{t_2}(D^j_2\Psi)(t,\gamma(t))\,\mathrm{d}t,$$

where $F^\infty\Psi$, $D^j_1\Psi$, and $D^j_2\Psi$ are given by (2.77)*,* (2.78)*, and* (2.79)*, respectively.*

As in Proposition 2.25, we need an estimate for the Lie derivative of Ψ along F^∞ under the assumption of infinitesimal rigidity.

Proposition 2.40. *Suppose that, for every point p of* (2.14)*, the framework $G(p)$ is infinitesimally rigid. Then, there exist $r_0, c_0 > 0$ such that*

$$(F^\infty\Psi)(x) \le -c_0\,\Psi(x)$$

for every $x\in\Psi^{-1}(\le r_0)$.

Proof. Since the potential function ψ in (2.13) is of the form (2.1), we know from Proposition 2.7 (d) that there exist $c_\psi, r_0 > 0$ such that $\|\nabla\psi(p)\|^2 \ge c_\psi\psi(p)$ for every $p\in\psi^{-1}(\le r_0)$. Recall that the function h is defined in Lemma 2.20. Since a is assumed to be smooth with $a(0) = 0$ and $a'(0)\ne 0$, there exist $c_h > 0$ such that $h(y)\ge c_h y$ for every $y\in[0, r_0^{1/2}]$. The asserted estimate with $c_0 := c_h c_y/2$ now follows from the definition of Ψ in (2.47) and equation (2.77). □

The remainders $D_1^j\Psi$ and $D_2^j\Psi$ in Proposition 2.39 satisfy the following estimates.

Proposition 2.41. *For every $r > 0$, there exist $c_1, c_2 > 0$ such that*

$$\begin{aligned} \left|(D_1^j\Psi)(t,x)\right| &\leq j^{-1/4}\, c_1\, \Psi(x), \\ \left|(D_2^j\Psi)(t,x)\right| &\leq j^{-1/4}\, c_2\, \Psi(x) \end{aligned}$$

for every $t \in \mathbb{R}$ and every $x \in \Psi^{-1}(\leq r)$.

Proof. From Lemma 2.33 and Lemma 2.36, we already know estimates for the (iterated) Lie derivatives $F_{\ell_1,\ldots,\ell_k}\Psi$ and the sinusoids u_ℓ^j, $\widetilde{UV}^j_{\ell_1,\ldots,\ell_k}$, respectively. The asserted estimates therefore follow immediately from the definitions of $D_1^j\Psi$ and $D_2^j\Psi$ in (2.78) and (2.79), respectively. □

2.7.4 Proof of Proposition 2.32

Suppose that, for every point p of (2.14), the framework $G(p)$ is infinitesimally rigid. Then, there exist $r_0, c_0 > 0$ as in Proposition 2.40. For this sublevel r_0, there exist $c_1, c_2 > 0$ as in Proposition 2.41. From Proposition 2.39, we conclude that, for every $j \geq 1$, every trajectory $\gamma\colon \mathbb{R} \to P^N$ of (2.58), and all $t_1 < t_2$ in $\mathbb{R}$, the following implication holds: if $\gamma(t) \in \Psi^{-1}(\leq r_0)$ for every $t \in [t_1, t_2]$, then

$$\Psi(\gamma(t_2)) \;\leq\; \Psi(\gamma(t_1)) + c_1\, j^{-\frac{1}{4}}\, \Psi(\gamma(t_2)) + c_1\, j^{-\frac{1}{4}}\, \Psi(\gamma(t_1)) - \int_{t_1}^{t_2} \big(c_0 - c_2\, j^{-\frac{1}{4}}\big)\, \Psi(\gamma(t))\, \mathrm{d}t.$$

Thus, we are in the same situation as in the proof of Theorem 2.12 in Subsection 2.5.4, and the same argument as therein proves Proposition 2.32.

2.8 Extension to extremum seeking control

So far, we have restricted our considerations to the problem of formation control. However, control law (2.11) for point agents can be also useful in the context of extremum seeking control. To see this, we return to the control-affine system (1.39) with output (1.40) in Section 1.3. This means, we consider

$$\dot{x} \;=\; \sum_{i=1}^{m} u_k\, g_k(x), \tag{2.80}$$

$$y \;=\; \psi(x), \tag{2.81}$$

where $u_1, \ldots, u_m$ are real-valued input channels for a control law, $g_1, \ldots, g_m$ are smooth vector fields on a smooth manifold M, and the output channel y is given by a smooth function ψ on M. Note that (2.80) reduces to the kinematic equation (2.5) of a single point agent if the control vector fields g_k form an orthonormal basis of the Euclidean space. In the same way, the output function ψ can be considered as the individual potential function of an agent. We already know from Section 1.3 that an extremum seeking control law for (2.80) is given by (1.45). Note that the extremum seeking control law (1.45) is almost the same as the formation control law (2.11) for point agents. The only notational difference is that (1.45) contains the design functions $y \mapsto h_\nu(y)$ while (2.11) contains the functions $y \mapsto h_\nu(y^\kappa)$ with some additional exponent $\kappa \geq 1/2$. Clearly, this is just a matter of notation: It is, of course, also possible to write $h_\nu(y)$ instead of $h_\nu(y^\kappa)$ in (2.11) if we replace y by y^κ on the right-hand side in the definition of the h_ν in (2.7). Consequently, control laws (1.45) and (2.11) have the same structure. It is therefore natural to ask whether it makes sense to apply the formation control law (2.11) to the more general input-output system (2.80), (2.81). This is indeed a suitable control strategy if the following additional assumption is satisfied, which is, at least for some applications, quite restrictive.

Assumption 2.42. The output function ψ attains a *known* minimum value y^* at some (not necessarily known) point x^* of M. ♢

There are two reasons why we need Assumption 2.42. Firstly, the definition of h_1, h_2 in (2.7) involves the logarithm, which is only defined for positive values. Secondly, the functions h_1, h_2 are designed in such a way that their amplitude $a(y)$ in (2.7) vanishes if $y = 0$. These two properties are crucial to obtain the full notion of asymptotic stability and not only practical asymptotic stability as in Section 1.3. If we know the minimum value y^* as in Assumption 2.42, then it makes sense to consider the parameter-dependent time-varying output-feedback control law

$$u_k \;=\; u^j_{(k,1)}(t)\, h_1\big((y-y^*)^\kappa\big) + u^j_{(k,2)}(t)\, h_2\big((y-y^*)^\kappa\big), \tag{2.82}$$

where the sinusoids $u^j_{(k,1)}, u^j_{(k,2)}$ are defined in (1.44), the design functions h_1, h_2 are defined in (2.7), and the constant κ is $\geq 1/2$.

In the situation of Assumption 2.42, we can apply exactly the same local asymptotic stability analysis as in Section 2.5 to system (2.80) under control law (2.82). At the end, this analysis leads to the same statements as in Propositions 2.24 and 2.26, where the averaged vector field f^∞ on M is now given by

$$f^\infty(x) \;:=\; -\kappa\, h\big(\psi(x)^\kappa\big)\, \psi(x)^{2\kappa-1} \sum_{k=1}^{m} (g_k\psi)(x)\, g_k(x) \tag{2.83}$$

with the function h defined in Lemma 2.20. Note that the right-hand side of (2.83) coincides with the right-hand side of (2.18) if the g_k form an orthonormal basis of the Euclidean space. To apply the same argument as in the proofs of Theorems 2.12 and 2.13, we also need a suitable replacement for Proposition 2.25. This requires two additional assumptions. The first assumption concerns the control vector fields g_k. Note that point agents can be steered instantaneously into any direction of the space. To get the same for control system (2.80), we need the following property.

Assumption 2.43. The vectors $g_1(x^*), \ldots, g_m(x^*)$ span the tangent space to M at x^*. ♢

If Assumption 2.43 is satisfied, then, by continuity, for every x from a sufficiently small neighborhood of x^* in M, the vectors $g_1(x), \ldots, g_m(x)$ also span the tangent space at x. The second assumption concerns the output function ψ. In the context of formation control, the assumption of infinitesimal rigidity ensures that the gradient of the global potential function satisfies the estimate in Proposition 2.7 (d); cf. also Remark 2.8. It turns out that we can apply basically the same argument as in the proof of Proposition 2.7 (d) if we assume the following.

Assumption 2.44. The second derivative of ψ at x^* is positive definite. ♢

Recall from Section 1.5 that the existence of a well-defined second derivative of ψ at x^* requires that x^* is a critical point of ψ, which is ensured by Assumption 2.42. Using Remark 1.2, we obtain the following replacement for Proposition 2.25: If Assumptions 2.42-2.44 are satisfied, then there exist $y_0 > y^*$ and $c_0 > 0$ such that

$$(f^{\infty}\psi)(x) \;\leq\; -c_0\,\psi(x)^{2\kappa}$$

for every $x \in \psi^{-1}(\leq y_0, x^*)$. Now the same reasoning as in Subsection 2.5.4 leads to the following versions of Theorems 2.12 and 2.13 for the purpose of extremum seeking control.

Theorem 2.45. *Suppose that Assumptions 2.42-2.44 are satisfied. Let $\kappa = 1/2$. Then, there exist $\mu > 0$ and $y^+ > y^*$ such that, for every $\lambda > 1$, there exists $j_0 > 0$ such that, for every $j \geq j_0$, every $t_0 \in \mathbb{R}$, and every $x_0 \in \psi^{-1}(\leq y^+, x^*)$, the maximal solution x of system* (2.80) *under control law* (2.82) *with initial condition $x(t_0) = x_0$ has the following two properties: $x(t)$ converges to x^* as $t \to \infty$, and the estimate*

$$\psi(x(t)) \;\leq\; \lambda\,\psi(x_0)\,\mathrm{e}^{-\mu(t-t_0)}$$

holds for every $t \geq t_0$.

Theorem 2.46. *Suppose that Assumptions 2.42-2.44 are satisfied. Let $\kappa > 1/2$ and let $j \geq 1$. Then, there exists $\mu > 0$ such that, for every $\lambda > 1$, there exists $y^+ > y^*$ such that, for every $t_0 \in \mathbb{R}$ and every $x_0 \in \psi^{-1}(\leq y^+, x^*)$, the maximal solution x of system* (2.80) *under control law* (2.82) *with initial condition $x(t_0) = x_0$ has the following two properties: $x(t)$ converges to x^* as $t \to \infty$, and the estimate*

$$\psi(x(t)) \;\leq\; \frac{\lambda\,\psi(x_0)}{\big(1 + (2\kappa - 1)\psi(x_0)^{2\kappa-1}\,\mu\,(t - t_0)\big)^{\frac{1}{2\kappa-1}}}$$

holds for every $t \geq t_0$.

As explained in Remark 2.27, the magnitude of the sublevel $y^+ > y^*$ for the domain of attraction $\psi^{-1}(\leq y^+, x^*)$ in Theorems 2.45 and 2.46 depends on the choice of the frequency parameter j. An increase of j, will lead to the existence of a larger value of y^+. In general, however, we cannot expect that y^+ grows unbounded with increasing j. We can merely expect that y^+ tends to some upper bound $> y^*$ in the limit $j \to \infty$. This is due to the fact that Assumptions 2.42-2.44 only contain local conditions at the optimal point x^*. Therefore, we only get local stability results. Under suitable additional global assumptions on the control vector fields $g_1, \ldots, g_m$ and the output function ψ, it is also possible to extend Theorems 2.45 and 2.46 to semi-global stability results.

Example 2.47. To allow a visual comparison with the extremum seeking method in Figure 1.6, we provide numerical results for the following situation. As in Section 1.3, we consider control system (1.26) with output (1.27) given by $y = \psi(x) := x^2$. Then, Assumptions 2.42-2.44 are satisfied with $y^* = x^* = 0$. Thus, if we apply a control law of the form (2.82) with $\kappa = 1/2$, then Theorem 2.45 ensures that x^* locally uniformly exponentially stable for the closed loop system. As explained in Section 1.2, the choice of highly oscillatory inputs is not restricted to the sinusoids (1.44). It is also possible to employ the rectangular inputs u_1^j, u_2^j in Figure 1.2. To ensure that the averaged vector field (2.83)

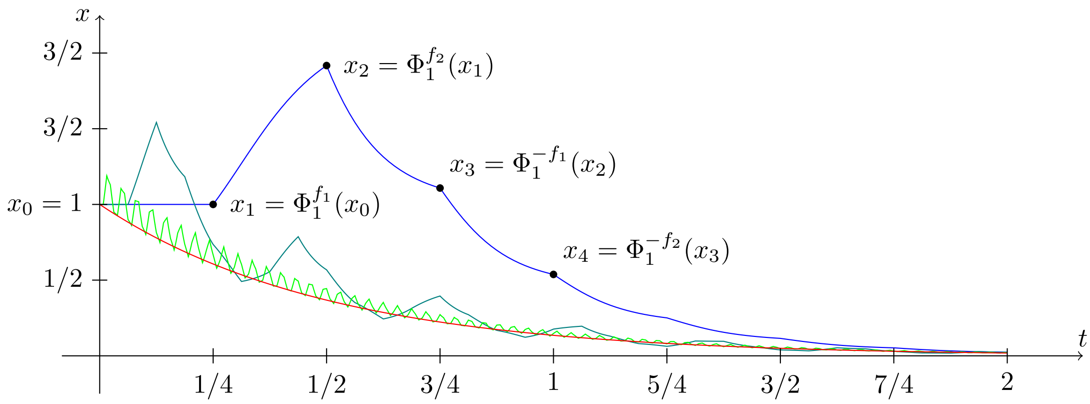


Figure 2.7: *Control system* (1.26) *with output* $y = x^2$ *under the* j*-dependent control law* (2.85) *with* u_1^j, u_2^j *as in Figure 1.2 and* h_1, h_2 *as in* (2.84)*. The trajectory of the closed-loop system with initial condition* $x(0) = 1$ *is drawn in blue for* $j = 4$*, in cyan for* $j = 16$*, and in green for* $j = 128$*. As in Figure 1.6, the flow maps* $\Phi_t^{\pm f_\nu}$ *are indicated for* $j = 4$*, where* $f_\nu(x) := h_\nu(|x|)$ *for* $\nu = 1, 2$*. In the limit* $j \to \infty$*, the trajectories of the closed-loop system converge locally uniformly to the trajectories of* (1.36)*. The trajectory of* (1.36) *with initial condition* $x(0) = 1$ *is drawn in red.*

coincides precisely with the right-hand side of (1.19), we choose slightly different design functions h_1, h_2 compared to (2.7). We define $h_1(y) := h_2(y) := 0$ for $y \leq 0$ and

$$h_1(y) \;:=\; y\,\sin(2\log y), \tag{2.84a}$$

$$h_2(y) \;:=\; y\,\cos(2\log y) \tag{2.84b}$$

for $y > 0$. This corresponds to (2.7) with amplitude a defined by $a(y) := y$. The additional factor 2 in the arguments of sine and cosine leads to an additional factor 2 on the right-hand side of (2.83). For the simple case of a single input channel u in control system (1.26), control law (2.82) reduces to

$$u \;=\; u_1^j(t)\,h_1(y^{1/2}) + u_2^j(t)\,h_2(y^{1/2}). \tag{2.85}$$

Then, the averaged system of (1.26) under control law (2.85) is given by (1.36). In contrast to the practical asymptotic stability in Figure 1.6, we can now observe exponential stability in Figure 2.7. ♢

An extremum seeking control law like (2.82) that can lead to exponential stability was proposed for the first time in [111]. Later versions of this approach can be found in [41] and [113]. A common assumption in all of these papers is that the minimum value y^* needs to be known to ensure exact convergence to the optimal state x^*. Moreover, the uncertainty in the choice of a sufficiently large parameter j for the dither signals causes difficulties in practical implementations. These questions are addressed in the next chapter.

3 Extremum seeking control with an adaptive dither signal

The content of this chapter is an extended version of [109].

3.1 Introduction and motivation

For the sake of simplicity, we restrict our introductory discussion to a control-affine system on $\mathbb{R}^n$ with a smooth drift vector field $g_0 \colon \mathbb{R}^n \to \mathbb{R}^n$, smooth control vector fields $g_1, \ldots, g_m \colon \mathbb{R}^n \to \mathbb{R}^n$, and a smooth output function $\psi \colon \mathbb{R}^n \to \mathbb{R}$. Thus, we consider the following control-affine system with output:

$$\dot{x} \;=\; g_0(x) + \sum_{k=1}^{m} u_k \, g_k(x), \tag{3.1}$$

$$y \;=\; \psi(x), \tag{3.2}$$

where $u_1, \ldots, u_m$ are real-valued input channels for a control law, and y is a real-valued output channel given by ψ. Again, to simplify the discussion, we assume for the moment that ψ is a quadratic function of the form $\psi(x) = y^* + \|x - x^*\|^2$, where $\|\cdot\|$ denotes the Euclidean norm. Our goal is to derive (time-varying) output-feedback that steers (3.1) to x^*. The optimal point $x^* \in \mathbb{R}^n$, the optimal value $y^* \in \mathbb{R}^n$, as well as the current system state of (3.1) and the vector fields $g_0, g_1, \ldots, g_m$ are treated as unknown quantities. Only real-time measurements of the output value (3.2) are available.

Note that the above extremum seeking control problem is more challenging than the one in Section 1.3 because, in contrast to (1.39), there is also a possibly nonvanishing drift involved in (3.1). For example, the drift could be of the form $g_0(x) = x - x^*$, which leads to the undesired effect that the system is driven into an ascent direction of ψ if the influence of the control vector fields is too weak. One possible way to overcome this problem, is to modify control law (1.45) in Section 1.3 as follows. Instead of using the j-dependent sinusoids $u^j_{(k,1)}, u^j_{(k,2)}$ in (1.44), we introduce an additional parameter $\lambda > 0$ and then, for every $k \in \{1, \ldots, m\}$, we define $u^{\lambda,j}_{(k,1)}, u^{\lambda,j}_{(k,2)} \colon \mathbb{R} \to \mathbb{R}$ by

$$u^{\lambda,j}_{(k,1)}(t) \;:=\; \sqrt{\lambda\, j\, \omega_k}\, \cos(j\, \omega_k\, t), \tag{3.3a}$$

$$u^{\lambda,j}_{(k,2)}(t) \;:=\; \sqrt{\lambda\, j\, \omega_k}\, \sin(j\, \omega_k\, t), \tag{3.3b}$$

where $\omega_1, \ldots \omega_k$ are again pairwise distinct positive real constants. Thus, the parameter λ provides an additional degree of freedom to enlarge the amplitudes of the sinusoids. If we use this slight modification of the sinusoids, then control law (1.45) becomes

$$u_k \;=\; u^{\lambda,j}_{(k,1)}(t)\, h_1(y) + u^{\lambda,j}_{(k,2)}(t)\, h_2(y) \tag{3.4}$$

for $k = 1, \ldots, m$, where y is the measured output signal (3.2) and the design functions h_1, h_2 are chosen as in Section 1.3. By applying control law (3.4) to control system (3.1), we obtain the closed loop system

$$\dot{x} \;=\; g_0(x) + \sum_{k=1}^{m} \big(u^{\lambda,j}_{(k,1)}(t)\, f_{(k,1)}(x) + u^{\lambda,j}_{(k,2)}(t)\, f_{(k,2)}(x)\big), \tag{3.5}$$

where the vector fields $f_{(k,1)}, f_{(k,2)}$ on $\mathbb{R}^n$ are defined by (1.41). If the drift g_0 vanishes and if $\lambda = 2$, then (3.5) reduces to (1.46). We already know that the trajectories of the closed-loop system (1.46) approximate the trajectories of the averaged system (1.47) if the parameter j is sufficiently large. A similar statement also holds for (3.5). Because of the additional drift g_0 and the parameter λ, the averaged system of (3.5) reads

$$\dot{x} \;=\; g_0(x) + \frac{\lambda}{2} \sum_{k=1}^{m} [f_{(k,1)}, f_{(k,2)}](x). \tag{3.6}$$

For the sake of simplicity, we assume for our further discussion that the design functions h_1, h_2 are given by (1.32), (1.33), or (1.34). Then, using (1.42), the averaged system can be also written as

$$\dot{x} \;=\; g_0(x) - \frac{\lambda}{2} \sum_{k=1}^{m} (g_k\psi)(x)\, g_k(x). \tag{3.7}$$

In the subsequent paragraphs, we explain how the parameter λ in (3.7) can be used to compensate the possibly negative influence of the drift g_0.

For the purpose of extremum seeking control, we are interested in the case in which the averaged system (3.7) is driven into a descent direction of ψ. Then the approximation property ensures that the closed-loop system (3.5) displays a similar behavior for sufficiently large j. To analyze the behavior of the averaged system, we consider the derivative $\dot{\psi}$ along solutions of (3.7), which is given by

$$\dot{\psi}(x) \;=\; (g_0\psi)(x) - \frac{\lambda}{2} \sum_{k=1}^{m} (g_k\psi)(x)^2.$$

Now assume that the vectors $g_1(x), \ldots, g_m(x)$ span $\mathbb{R}^n$ at every $x \in \mathbb{R}^n$. Recall that ψ is assumed to be of the form $\psi(x) = y^* + \|x - x^*\|^2$. For the moment, let K be the spherical shell of all $x \in \mathbb{R}^n$ with $r \leq \|x - x^*\| \leq R$ for certain radii $R > r > 0$. Since $g_0\psi$ is continuous and since K is compact, there exists $c > 0$ such that $(g_0\psi)(x) < c$ for every $x \in K$. Using the assumptions on $g_1, \ldots, g_m$ and ψ, a similar argument leads to the existence of some $d > 0$ such that $\sum_{k=1}^{m} (g_k\psi)(x)^2 > d$ for every $x \in K$. Thus, if $\lambda \geq \lambda_0 := 2c/d$, then $\dot{\psi}$ only takes negative values on K. In other words, for $\lambda \geq \lambda_0$, the solutions of (3.7) are driven into a descent direction of ψ within K. It is now easy to see that if λ is sufficiently large, then a neighborhood of x^* becomes locally asymptotically stable for (3.7). In the limit $\lambda \to \infty$, the attracting set shrinks to x^* and the domain of attraction expands to $\mathbb{R}^n$. To be more precise: For all compact neighborhoods $K^* \subseteq K_0 \subseteq \mathbb{R}^n$ of x^* and every $T > 0$, there exists a sufficiently large $\lambda_0 > 0$ such that, for every $\lambda \geq \lambda_0$ the solutions of (3.7) with initial values in K_0 enter K^* at latest after the time span T and then stay in K^*. Next, we return to the closed-loop system (3.5).

As argued in the previous paragraph, a large amplitude λ is needed to ensure desirable stability properties of the averaged system (3.7) with respect to the optimal point x^*.

On the other hand, the approximation of trajectories of (3.7) by of trajectories of (3.5) suffers with increasing λ. To compensate this effect, one has to increase the frequency parameter j sufficiently. This leads to the following result, which can be found, for example, in [95, 96, 76]: For all compact neighborhoods $K^* \subseteq K_0 \subseteq \mathbb{R}^n$ of x^* and every $T > 0$, there exist sufficiently large $\lambda > 0$, $j > 0$ (where the choice of j also depends on the choice of λ) such that the solutions of (3.5) with initial values in K_0 enter K^* at latest after the time span T and then stay in K^*. However, there is no known rule so far on how large λ, j have to be chosen for given K_0, K^*, T. Just the existence of certain lower bounds is ensured. Moreover, finite escape times can occur if λ, j are too small, or if the initial state is too far away from x^*. It is certainly difficult to implement (3.4) successfully without knowing suitable values for the parameters λ, j.

The intention of this chapter is to propose a solution to the above problems. The idea is to choose the amplitude λ and the frequency j in an adaptive way. This means that λ and j increase automatically so that the system state $x(t)$ of (3.5) converges to the desired state x^* as $t \to \infty$. We prevent finite escape times by introducing a dynamic funnel for the output value. In this way, the system state is always contained in a prescribed sublevel set of the output function ψ. Under the assumptions of our introductory discussion, we can ensure global convergence to the optimal state without the obstacle of unknown control parameters. We will see that this convergence result also holds in a more general situation than for the introductory example.

The rest of the chapter is organized as follows. In Section 3.2, we recall some basic definitions, which are related to time-varying control-affine systems with drift. The extremum seeking control law and the corresponding convergence result are presented in Section 3.3. We apply the main result to a nontrivial example in Section 3.4 and also provide numerical data. The proof of the convergence result is presented in Section 3.5.

3.2 Local definitions and notation for the chapter

We briefly introduce some of the terminology from [18, 101, 106], which is used throughout the chapter. Let M be a smooth manifold. A *time-varying function on M* is a function whose domain is $\mathbb{R} \times M$. A *Carathéodory function on M* is a time-varying function α on M with the property that, for each $t \in \mathbb{R}$, the function $x \mapsto \alpha(t, x)$ on M is continuous, and that, for each $x \in M$, the function $t \mapsto \alpha(t, x)$ on $\mathbb{R}$ is (Lebesgue) measurable. Let α be a time-varying function on M. We say that α is *uniformly bounded* on a subset V of M if there exists a constant $c > 0$ such that $|\alpha(t, x)| \leq c$ for every $t \in \mathbb{R}$ and every $x \in V$. We say that α is *locally uniformly bounded* if every point of M has a neighborhood V in M such that α is uniformly bounded on V. We say that α is *locally integrally Lipschitz continuous* if, for every $x \in M$, there exist a smooth chart (U, ϕ) for M around x and a positive locally integrable function L on $\mathbb{R}$ such that $|\alpha(t, x'') - \alpha(t, x')| \leq L(t)\, \|\phi(x'') - \phi(x')\|$ for every $t \in \mathbb{R}$ and all $x', x'' \in U$. We say that the time-varying function α on M is *smooth* if α is smooth as a function on the product manifold $\mathbb{R} \times M$. If α is smooth, then, for each $x \in M$, we denote the derivative of $s \mapsto \alpha(s, x)$ at $t \in \mathbb{R}$ by $(\partial_t \alpha)(t, x)$. This defines a smooth time-varying function $\partial_t \alpha$ on M, which we refer to as the *time derivative* of α. It is clear that any given function ψ on M can be considered as the time-varying function $(t, x) \mapsto \psi(x)$ on M.

A *time-varying vector field* on M is a map f with domain $\mathbb{R} \times M$ such that, for each $t \in \mathbb{R}$, the map $x \mapsto f(t, x)$ is a vector field on M. Suppose that α is a time-varying function on M, and f, g are time-varying vector fields on M. If, for each $t \in \mathbb{R}$, the function $\xi \mapsto \alpha(t, \xi)$ is smooth, then we denote the *fixed-time Lie derivative* of $\xi \mapsto \alpha(t, \xi)$ along the vector field $\xi \mapsto f(t, \xi)$ at $x \in M$ by $(f\alpha)(t, x)$. This defines another time-varying function $f\alpha$ on M. If, for each $t \in \mathbb{R}$, the vector fields $\xi \mapsto f(t, \xi)$, $\xi \mapsto g(t, \xi)$ are smooth, then we denote their *fixed-time Lie bracket* at $x \in M$ by $[f, g](t, x)$. This defines another time-varying vector field $[f, g]$ on M. We say that f is *locally uniformly bounded* if, for every smooth function φ on M, the time-varying function $f\varphi$ is locally uniformly bounded. In the same way, we define the notions of a *Carathéodory vector field*, *local integrally Lipschitz continuity*, and the property of being *smooth* for a time-varying vector field. It follows from Carathéodory's existence and uniqueness theorems for ordinary differential equations (see, e.g., [101]) that a locally uniformly bounded, locally integrally Lipschitz continuous Carathéodory vector field has a unique maximal integral curve for any initial condition. As for functions, we identify each vector field on M with its corresponding time-varying extension.

3.3 Main result

We consider a control-affine system

$$\dot{x} \;=\; g_0(t, x) + \sum_{k=1}^{m} u_k \, g_k(t, x) \tag{3.8}$$

with output

$$y \;=\; \psi(x) \tag{3.9}$$

on M under the following assumptions (in the terminology of Section 3.2).

Assumption 3.1. Suppose that

(1) M is a smooth manifold,

(2) ψ is a smooth function on M,

(3) (a) g_0 is a locally uniformly bounded and locally integrally Lipschitz continuous Carathéodory vector field on M,

(b) $g_1, \ldots, g_m$ are smooth time-varying vector fields on M

such that, for all $k_1, k_2, k_3 \in \{1, \ldots, m\}$, the time-varying functions $g_{k_1}\psi$, $\partial_t(g_{k_1}\psi)$, $g_0(g_{k_1}\psi)$, $g_{k_2}(g_{k_1}\psi)$, $\partial_t(g_{k_2}(g_{k_1}\psi))$, $g_0(g_{k_2}(g_{k_1}\psi))$, $g_{k_3}(g_{k_2}(g_{k_1}\psi))$ on M are locally uniformly bounded. ◊

Remark 3.2. If g_0 is a locally Lipschitz continuous vector field on M, and if $g_1, \ldots, g_m$ are smooth vector fields on M (all of them not time-varying), then property (3) in Assumption 3.1 is satisfied for every smooth function ψ on M. ◊

In (3.8) and (3.9), we call x the *system state*, t the *time parameter*, g_0 the *drift vector field*, $g_1, \ldots, g_m$ the *control vector fields*, and $u_1, \ldots, u_m$ the real-valued *input channels* for a control law. The *output function* ψ converts the current state into an *output value* y. In the context of extremum seeking control, only real-time measurements of y are available. We are interested in an output-feedback law for (3.8) so that y converges to a minimum value of ψ.

Let $\omega_1, \dots, \omega_m$ be pairwise distinct positive real constants. We propose the control law

$$u_k \;=\; \frac{\sqrt{\omega_k}}{(z-y)^2}\,\sin\Big(\omega_k\,J + \frac{1}{z-y}\Big), \qquad k = 1, \dots, m, \tag{3.10}$$

for system (3.8) with output (3.9), where the functions z, J are determined by the differential equations

$$\dot{z} \;=\; -(z-y), \tag{3.11}$$

$$\dot{J} \;=\; \frac{1}{(z-y)^5}. \tag{3.12}$$

The following results gives sufficient conditions under which (3.10)-(3.12) minimizes the output (3.9) of (3.8). We use the terminology and notation from Sections 1.5 and 3.2.

Theorem 3.3. *Suppose that Assumption 3.1 is satisfied. Suppose that ψ attains a local minimum value y^* at some point x^* of M. Assume that there exist y^+ with $y^* < y^+ \leq \infty$ and a continuous function b on $\psi^{-1}(\leq y^+, x^*)$ such that*

$$\sum_{k=1}^{m} (g_k\psi)(t,x)^2 \;\geq\; b(x) \;>\; 0 \tag{3.13}$$

for every $t \in \mathbb{R}$ and every $x \in \psi^{-1}(\leq y^+, x^*)$ with $x \neq x^*$. Assume that $\psi^{-1}(\leq \tilde{y}, x^*)$ is compact for every $\tilde{y} < y^+$. Then, for every $t_0 \in \mathbb{R}$, every $x_0 \in \psi^{-1}(\leq y^+, x^*)$, and all $z_0, J_0 \in \mathbb{R}$ with $\psi(x_0) < z_0 < y^+$, there exist unique solutions x, z, J of system (3.8), (3.9) *under control law* (3.10)-(3.12) *on the interval $[t_0, \infty)$ with initial values x_0, z_0, J_0 at t_0, and the following holds:*

(a) for every $t \geq t_0$, we have $y^* \leq y(t) = \psi(x(t)) < z(t)$,

(b) $t \mapsto z(t)$ is strictly decreasing with $z(t) \to y^*$ as $t \to \infty$.

In particular, statements (a) and (b) of Theorem 3.3 imply $y(t) \to y^*$ as $t \to \infty$.

Remark 3.4. Suppose that the vector fields $g_1, \dots, g_m$ are time-invariant. Moreover, suppose that there exists y^+ with $y^* < y^+ \leq \infty$ such that, for every $x \in \psi^{-1}(\leq y^+, x^*)$ with $x \neq x^*$, we have $(g_k\psi)(x) \neq 0$ for some $k \in \{1, \dots, m\}$. Then, the nonnegative smooth function $b := \sum_{k=1}^{m} (g_k\psi)^2$ on M satisfies (3.13) for every $x \in \psi^{-1}(\leq y^+, x^*)$ with $x \neq x^*$. $\Diamond$

Because of Remarks 1.2 and 3.4, a local version[1] of Theorem 3.3 can be stated as follows.

Corollary 3.5. *Suppose that Assumption 3.1 is satisfied and that the vector fields $g_1, \dots, g_m$ are time-invariant. Assume that*

1. *the function ψ attains a local minimum value y^* at some point x^* of M,*
2. *the second derivative of ψ at x^* is positive definite,*
3. *the vectors $g_1(x^*), \dots, g_m(x^*)$ span the tangent space to M at x^*.*

[1] By *local version*, we mean sufficient conditions to ensure the existence of some (possibly small) sublevel $y^+ > y^*$ as in Theorem 3.3.

Then, there exists $y^+ > y^*$ such that the conclusions of Theorem 3.3 hold.

Remark 3.6. Assume that we are in the situation of Theorem 3.3. Suppose that we are not interested in exact convergence to the optimal output value y^* but only in reaching a prescribed z^*-sublevel set of ψ for some $z^* \in (y^*, y^+)$. Then, one can modify the control law (3.10)-(3.12) as follows. As soon as the value of the function z reaches the desired value z^* one can replace (3.11) by $\dot{z} = 0$ to keep z at the constant value z^*. Then, the system state of (3.8) remains in $\psi^{-1}(\leq z^*, x^*)$, and the right-hand side of (3.12) remains bounded. We illustrate this modification by an example in the next section. ◊

Remark 3.7. Control law (3.10)-(3.12) bears a certain resemblance to the idea of *funnel control* (see, e.g., [45, 49]). By Theorem 3.3, the function z acts as an upper bound for the output y. The condition $\psi(x) = y < z$ defines a performance funnel for the system state x of (3.8). The boundary of the funnel is described by z. The width of the funnel is described by the positive function $z - y^*$, where y^* is the minimum value of ψ. Under the assumptions of Theorem 3.3, the so-called "ultimate width"; i.e., the width of the funnel in the limit $t \to \infty$, is equal to zero. The modification in Remark 3.6 leads to the positive "ultimate width" $z^* - y^*$. ◊

3.4 A numerical Example

As an example, we consider the following particular case of (3.8) and (3.9) on the state manifold $M := \mathbb{R}^2$. The control-affine system (3.8) consists of the drift vector field g_0, given by

$$g_0(t, (x_1, x_2)) \;:=\; \big(\sin(2\,t),\ \cos(t)\big)^\top,$$

and $m := 2$ time-invariant control vector fields g_1, g_2, given by

$$\begin{aligned} g_1(x_1, x_2) \;&:=\; \big(\ \ \cos(x_1 + x_2),\ \sin(x_1 + x_2)\ \big)^\top, \\ g_2(x_1, x_2) \;&:=\; \big(-\sin(x_1 + x_2),\ \cos(x_1 + x_2)\ \big)^\top. \end{aligned}$$

The output function ψ on $\mathbb{R}^2$ is given by

$$\psi(x_1, x_2) \;:=\; (x_1 - 1)^2 + (x_2 - 1)^2 + 2018.$$

It follows from Remark 3.4 that the assumptions of Theorem 3.3 are satisfied with $y^+ := \infty$. This means that control law (3.10)-(3.12) will steer the system to the optimal state $x^* := (1, 1)$ at which ψ attains its global minimum value $y^* = 2018$. The numerical simulations in Figure 3.1 confirm this statement. All results are generated with the frequency coefficients $\omega_1 := 1$, $\omega_2 := 2$ in (3.10), and for the initial conditions $x(0) = (0, 0)$, $z(0) = \psi(x(0)) + 1$, $J(0) = 0$. The three plots in the left column of Figure 3.1 show the trajectory x, the output y, and the right-hand side $j := 1/(z - y)^5$ of (3.12) on the time interval $[0, 4.5]$. One can observe that with increasing time parameter t, the trajectory $x(t)$ converges to x^*, the output $y(t)$ converges to y^*, and the adaptive frequency parameter $j(t)$ tends to infinity. We also provide numerical data under the modification of Remark 3.6 in the right column of Figure 3.1. The frequency coefficients and the initial conditions are the same as before. Following Remark 3.6, equation (3.11) is replaced by $\dot{z} = 0$ as soon as $z(t)$ reaches a desired value z^*. The results are generated for the choice $z^* := 2018.5$. The function z reaches the value z^* at $t^* \approx 2.55$. This causes the output $y(t)$ to stay below the value ≈ 2018.2 for $t \geq t^*$. The adaptive frequency parameter $j(t)$ remains below the value 400 for $t \geq t^*$.

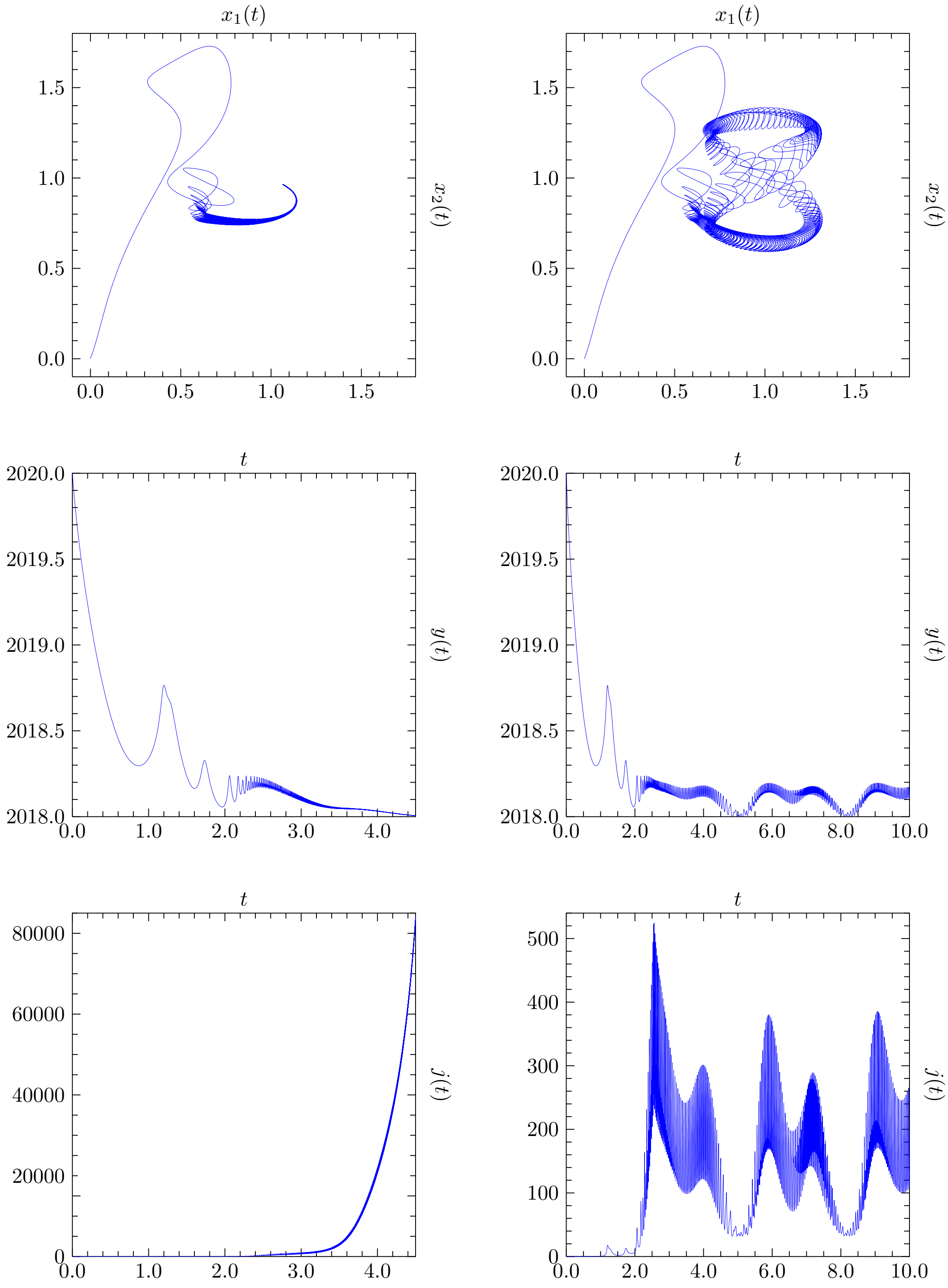


Figure 3.1: *Simulation results for the example in Section 3.4. Left column: implementation of control law* (3.10)-(3.12) *to enforce convergence to the optimal value* $y^* = 2018$*. Right column: control law* (3.10)-(3.12) *with the modification in Remark 3.6 for* $z^* = 2018.2$*.*

3.5 Convergence analysis

Throughout this section, we suppose that the assumptions of Theorem 3.3 are satisfied. We are interested in solutions of system (3.8), (3.9) under control law (3.10)-(3.12) for which the condition $\psi(x) < z$ is always satisfied. For this reason, we introduce the open submanifold

$$P \;:=\; \big\{(x,z) \in \psi^{-1}(\leq y^+, x^*) \times \mathbb{R} \;\big|\; \psi(x) < z < y^+\big\}$$

of the product manifold $M \times \mathbb{R}$. For the proof of Theorem 3.3, we study the closed-loop system

$$\dot{x} = g_0(t,x) + \sum_{k=1}^{m} \frac{\sqrt{\omega_k}}{(z-\psi(x))^2}\,\sin\Big(\omega_k\,J + \frac{1}{z-\psi(x)}\Big)\,g_k(t,x), \tag{3.14}$$

$$\dot{z} \;=\; -(z-\psi(x)), \tag{3.15}$$

$$\dot{J} \;=\; \frac{1}{(z-\psi(x))^5} \tag{3.16}$$

on the state space $P \times \mathbb{R}$. It follows from standard existence and uniqueness theorems for ordinary differential equations that system (3.14)-(3.16) on $P \times \mathbb{R}$ has a unique maximal solution for any initial value in $P \times \mathbb{R}$ at any initial time in $\mathbb{R}$. At this point, we collect several properties of the closed-loop system (3.14)-(3.16), which are easy to check.

Remark 3.8. Suppose that $((x,z),J)\colon I \to P \times \mathbb{R}$ is a maximal solution of (3.14)-(3.16). Then, the following statements hold.

(i) The function z is strictly decreasing (meaning that $z(t_1) > z(t_2)$ for all $t_1 < t_2$ in I).

(ii) The function J is strictly increasing (meaning that $J(t_1) < J(t_2)$ for all $t_1 < t_2$ in I).

(iii) For every $t \in I$, we have $y^+ > z(t) > \psi(x(t)) \geq y^*$.

(iv) If there is a finite escape time $\sup I < \infty$, then there exists $z_\infty > y^*$ such that $z(t) \geq z_\infty$ for every $t \in I$. ◊

In the next step, we rewrite equations (3.14) and (3.15) in a more suitable form. For this purpose, we define two smooth design functions h_1, h_2 on $(0,\infty)$ by

$$h_1(s) \;:=\; s\,\sin(1/s), \tag{3.17a}$$

$$h_2(s) \;:=\; s\,\cos(1/s). \tag{3.17b}$$

Note that the functions h_1, h_2 satisfy the property

$$[h_1,h_2](s) \;:=\; h_2'(s)\,h_1(s) - h_1'(s)\,h_2(s) \;=\; +1 \tag{3.18}$$

for every $s \in (0,\infty)$, which is (up to the sign) the same as for the functions h_1, h_2 in (1.32), (1.33), and (1.34). Define a positive smooth function φ on P by

$$\varphi(p) \;:=\; z - \psi(x) \tag{3.19}$$

for every $p = (x,z) \in P$. Next, we extend the time-varying vector fields $g_0, g_1, \ldots, g_m$ in the canonical way from M to the product manifold $M \times \mathbb{R}$, and then we restrict them to the submanifold P. This results in time-varying vector fields $\hat{g}_0, \hat{g}_1, \ldots, \hat{g}_m$ on P. In the same way, we extended the standard vector field on $\mathbb{R}$ to $M \times \mathbb{R}$, and denote its restriction to P by $\frac{\partial}{\partial z}$. Define a time-varying vector field f_0 on P by

$$f_0(t,p) \;:=\; \hat{g}_0(t,p) - \varphi(p)\frac{\partial}{\partial z}(p). \tag{3.20}$$

Then, we can combine (3.8) and (3.11) to the control-affine system

$$\dot{p} \;=\; f_0(t,p) + \sum_{k=1}^{m} u_k\, \hat{g}_k(t,p) \tag{3.21}$$

on P. For every $k \in \{1,\ldots,m\}$ and every $\nu \in \{1,2\}$, define a time-varying vector field $f_{(k,\nu)}$ on P by

$$f_{(k,\nu)}(t,p) \;:=\; h_\nu(\varphi(p))\, \hat{g}_k(t,p). \tag{3.22}$$

Now, using the trigonometric identity (1.49), we can combine (3.14) and (3.15) to a single equation so that the closed-loop system (3.14)-(3.16) can be written as the time-varying system

$$\dot{p} \;=\; f_0(t,p) + \sum_{k=1}^{m} \frac{\sqrt{\omega_k}}{\varphi(p)^3}\Big(\cos(\omega_k\, J)\, f_{(k,1)}(t,p) + \sin(\omega_k\, J)\;\, f_{(k,2)}(t,p)\Big), \tag{3.23}$$

$$\dot{J} \;=\; \frac{1}{\varphi(p)^5} \tag{3.24}$$

on $P \times \mathbb{R}$. By a *(maximal) solution of* (3.23), we mean a curve $p\colon I \to P$ for which there exists a function $J\colon I \to \mathbb{R}$ such that $(p,J)\colon I \to P \times \mathbb{R}$ is a (maximal) solution of (3.23), (3.24).

Let $\gamma\colon I \to P$ be a solution of (3.23). We define two functions λ and j on I by

$$\lambda(t) \;:=\; \frac{1}{\varphi(\gamma(t))} \qquad \text{and} \qquad j(t) \;:=\; \dot{J}(t) \;=\; \frac{1}{\varphi(\gamma(t))^5},$$

which will play the role of an *adaptive amplitude* and an *adaptive frequency parameter*, respectively. Similar to (3.3), for every $k \in \{1,\ldots,m\}$, we define two functions $u^j_{(k,1)}$, $u^j_{(k,2)}$ on I by

$$u^j_{(k,1)}(t) \;:=\; \sqrt{\lambda(t)\, j(t)\, \omega_k}\, \cos\big(\omega_k\, J(t)\big), \tag{3.25a}$$

$$u^j_{(k,2)}(t) \;:=\; \sqrt{\lambda(t)\, j(t)\, \omega_k}\, \sin\big(\omega_k\, J(t)\big). \tag{3.25b}$$

Note that, in first order approximation, the frequency of cosine and sine in (3.25) is $\omega_k j(t)$. Using (3.17), (3.25), and the trigonometric identity (1.49), we can write control law (3.10) also as

$$u_k \;=\; u^j_{(k,1)}(t)\, h_1(\varphi(p)) + u^j_{(k,2)}(t)\, h_2(\varphi(p)), \tag{3.26}$$

which is basically of the same form as (3.4). If we apply (3.26) to control system (3.21), then we obtain the closed-loop system

$$\dot{p} \;=\; f_0(t,p) + \sum_{k=1}^{m} \big(u^j_{(k,1)}(t) f_{(k,1)}(t,p) + u^j_{(k,2)}(t) f_{(k,2)}(t,p)\big), \tag{3.27}$$

which has the same structure as the closed-loop system (3.5). Note that, in contrast to (3.5), the amplitude λ and the frequency parameter j are not constant but depend on the curve γ. Both $\lambda(t)$ and $j(t)$ tend to $+\infty$ when $\varphi(\gamma(t))$ approaches the value 0. As explained in Section 3.1, the trajectories of (3.5) approximate the trajectories of (3.7) in the large-amplitude, high-frequency limit. It turns out that a similar statement also holds for (3.27) when $j(t)$ becomes large. A detailed averaging analysis will reveal that the Lie bracket system for (3.27) reads

$$\dot{p} \;=\; f_0(t,p) + \frac{\lambda(t)}{2}\sum_{k=1}^{m}[f_{(k,1)}, f_{(k,2)}](t,p), \tag{3.28}$$

which is the counterpart to equation (3.6). Note that the time-varying amplitude λ in (3.28) still depends on the solution γ of (3.23). A direct computation of the Lie brackets in (3.28) shows that if we write the system state of (3.28) component-wise as $p = (x, z)$, then the differential equation for the component x on M can be written as

$$\dot{x} \;=\; g_0(t,x) - \frac{\lambda(t)}{2}\sum_{k=1}^{m}(g_k\psi)(t,x)\, g_k(t,x),$$

which is the counterpart to equation (3.7). Thus, a sufficiently small value of $\varphi \circ \gamma$ causes a large value of the amplitude λ for the descent vector field $-(g_k\psi)g_k$ of ψ.

As an abbreviation, we introduce the indexing set Λ of all pairs (k, ν) with $k \in \{1, \ldots, m\}$ and $\nu \in \{1, 2\}$. Then, the closed-loop system (3.27) can be written more compactly as

$$\dot{p} \;=\; f_0(t,p) + \sum_{\ell\in\Lambda} u_\ell^j(t)\, f_\ell(t,p), \tag{3.29}$$

which may be interpreted as a control-affine system with drift f_0 and control vector fields f_ℓ under open-loop controls u_ℓ^j. In the following, we will frequently use the property that every solution γ of (3.23) is also a solution of (3.29) if the (state-dependent) functions u_ℓ^j are defined as in (3.25).

3.5.1 Estimates for the Lie derivatives

Note that the function φ on P in (3.19) is smooth. Moreover, the time-varying vector fields f_ℓ on P in (3.22) are smooth. Recall that the drift vector field f_0 on P is defined in (3.20). This allows us to compute the following (iterated) Lie derivatives along φ explicitly. For every $t \in \mathbb{R}$, every $p = (x, z) \in P$, and all $\ell_i = (k_i, \nu_i) \in \Lambda$ with $i = 1, 2, 3$, we have

$$\begin{aligned}
(f_0\varphi)(t,p) &= -(g_0\psi)(t,x) - \varphi(p),\\
(f_{\ell_1}\varphi)(t,p) &= -h_{\nu_1}(\varphi(p))\,(g_{k_1}\psi)(t,x),\\
(\partial_t(f_{\ell_1}\varphi))(t,p) &= -h_{\nu_1}(\varphi(p))\,(\partial_t(g_{k_1}\psi))(t,x),\\
(f_0(f_{\ell_1}\varphi))(t,p) &= +h'_{\nu_1}(\varphi(p))\,(g_0\psi)(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad - h_{\nu_1}(\varphi(p))\,(g_0(g_{k_1}\psi))(t,x)\\
&\quad + h'_{\nu_1}(\varphi(p))\,(g_{k_1}\psi)(t,x)\,\varphi(p),\\
(f_{\ell_2}(f_{\ell_1}\varphi))(t,p) &= +h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad - h_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(g_{k_2}(g_{k_1}\psi))(t,x),
\end{aligned}$$

$$
\begin{aligned}
(\partial_t(f_{\ell_2}(f_{\ell_1}\varphi)))(t,p) &= +h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_2}\psi)(t,x)\,(\partial_t(g_{k_1}\psi))(t,x)\\
&\quad + h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(\partial_t(g_{k_2}\psi))(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad - h_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(\partial_t(g_{k_2}(g_{k_1}\psi)))(t,x),\\
(f_0(f_{\ell_2}(f_{\ell_1}\varphi)))(t,p) &= -h_{\nu_2}(\varphi(p))\,h''_{\nu_1}(\varphi(p))\,(g_0\psi)(t,x)\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad - h'_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_0\psi)(t,x)\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad + h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_2}\psi)(t,x)\,(g_0(g_{k_1}\psi))(t,x)\\
&\quad + h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_0(g_{k_2}\psi))(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad + h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_0\psi)(t,x)\,(g_{k_2}(g_{k_1}\psi))(t,x)\\
&\quad + h'_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(g_0\psi)(t,x)\,(g_{k_2}(g_{k_1}\psi))(t,x)\\
&\quad - h_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(g_0(g_{k_2}(g_{k_1}\psi)))(t,x),\\
&\quad - h_{\nu_2}(\varphi(p))\,h''_{\nu_1}(\varphi(p))\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\,\varphi(p)\\
&\quad - h'_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\,\varphi(p)\\
&\quad + h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_2}(g_{k_1}\psi))(t,x)\,\varphi(p)\\
&\quad + h'_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(g_{k_2}(g_{k_1}\psi))(t,x)\,\varphi(p),\\
(f_{\ell_3}(f_{\ell_2}(f_{\ell_1}\varphi)))(t,p) &= -h_{\nu_3}(\varphi(p))\,h_{\nu_2}(\varphi(p))\,h''_{\nu_1}(\varphi(p))\,(g_{k_3}\psi)(t,x)\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad - h_{\nu_3}(\varphi(p))\,h'_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_3}\psi)(t,x)\,(g_{k_2}\psi)(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad + h_{\nu_3}(\varphi(p))\,h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_2}\psi)(t,x)\,(g_{k_3}(g_{k_1}\psi))(t,x)\\
&\quad + h_{\nu_3}(\varphi(p))\,h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_3}(g_{k_2}\psi))(t,x)\,(g_{k_1}\psi)(t,x)\\
&\quad + h_{\nu_3}(\varphi(p))\,h_{\nu_2}(\varphi(p))\,h'_{\nu_1}(\varphi(p))\,(g_{k_3}\psi)(t,x)\,(g_{k_2}(g_{k_1}\psi))(t,x)\\
&\quad + h_{\nu_3}(\varphi(p))\,h'_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(g_{k_3}\psi)(t,x)\,(g_{k_2}(g_{k_1}\psi))(t,x)\\
&\quad - h_{\nu_3}(\varphi(p))\,h_{\nu_2}(\varphi(p))\,h_{\nu_1}(\varphi(p))\,(g_{k_3}(g_{k_2}(g_{k_1}\psi)))(t,x).
\end{aligned}
$$

The functions h_ν in (3.17) and their derivatives can be easily computed. By doing so, one can verify that, for every $\tilde{y} \in (y^*, y^+)$, there exists $c_h > 0$ such that

$$
\begin{aligned}
|h_\nu(s)| &\leq c_h\, s,\\
|h'_\nu(s)| &\leq c_h/s,\\
|h''_\nu(s)| &\leq c_h/s^3
\end{aligned}
$$

for every $s \in (0, \tilde{y}-y^*)$. Moreover, by Assumption 3.1 and the assumptions of Theorem 3.3, for every $\tilde{y} \in (y^*, y^+)$, there exists $c_g > 0$ such that

$$
\begin{aligned}
|(g_0\psi)(t,x)| &\leq c_g,\\
|(g_{k_1}\psi)(t,x)| &\leq c_g,\\
|(\partial_t(g_{k_1}\psi))(t,x)| &\leq c_g,\\
|(g_0(g_{k_1}\psi))(t,x)| &\leq c_g,\\
|(g_{k_2}(g_{k_1}\psi))(t,x)| &\leq c_g,\\
|(\partial_t(g_{k_2}(g_{k_1}\psi)))(t,x)| &\leq c_g,\\
|(g_0(g_{k_2}(g_{k_1}\psi)))(t,x)| &\leq c_g,\\
|(g_{k_3}(g_{k_2}(g_{k_1}\psi)))(t,x)| &\leq c_g
\end{aligned}
$$

for every $t \in \mathbb{R}$, every $x \in \psi^{-1}(\leq \tilde{y}, x^*)$, and all $k_1, k_2, k_3 \in \{1, \ldots, m\}$. Now it is straight forward to derive the following estimates for the above Lie derivatives of φ.

Lemma 3.9. *For every $\tilde{y} \in (y^*, y^+)$, there exists $c > 0$ such that*

$$\begin{aligned} |(f_0\varphi)(t,p)| &\leq c, \\ |(f_\ell\varphi)(t,p)| &\leq c\,\varphi(p), \\ |(\partial_t(f_\ell\varphi))(t,p)| &\leq c\,\varphi(p), \\ |(f_0(f_\ell\varphi))(t,p)| &\leq c/\varphi(p), \\ |(f_{\ell_1}(f_{\ell_2}\varphi))(t,p)| &\leq c, \\ |(\partial_t(f_{\ell_1}(f_{\ell_2}\varphi)))(t,p)| &\leq c, \\ |(f_0(f_{\ell_1}(f_{\ell_2}\varphi)))(t,p)| &\leq c/\varphi(p)^2, \\ |(f_{\ell_1}(f_{\ell_2}(f_{\ell_3}\varphi)))(t,p)| &\leq c/\varphi(p) \end{aligned}$$

for all $\ell, \ell_1, \ell_2, \ell_3 \in \Lambda$, every $t \in \mathbb{R}$, and every $p = (x, z) \in P$ with $z \leq \tilde{y}$.

The sum of Lie brackets on the right-hand side of (3.28) motivates us to define the time-varying vector field

$$f^\infty \;:=\; \frac{1}{2}\sum_{k=1}^{m}[f_{(k,1)}, f_{(k,2)}] \tag{3.30}$$

on P. A direct computation, using (3.18), (3.19) and (3.22), shows that the Lie derivative of φ along f^∞ is given by

$$(f^\infty\varphi)(t,p) \;=\; \frac{1}{2}\sum_{k=1}^{m}(g_k\psi)(t,x)^2 \tag{3.31}$$

for every $t \in \mathbb{R}$, every $p = (x, z) \in P$.

3.5.2 Averaging of the sinusoids

Let $\gamma\colon I \to P$ be a solution of (3.23). Define two functions j and ι on I by

$$j(t) \;:=\; \frac{1}{\varphi(\gamma(t))^5} \tag{3.32}$$

and

$$\iota(t) \;:=\; -5\,j(t)^{\frac{6}{5}}\Big((f_0\varphi)(t,\gamma(t)) + \sum_{\ell\in\Lambda} u_\ell^j(t)\,(f_\ell\varphi)(t,\gamma(t))\Big), \tag{3.33}$$

respectively. Since γ is a solution of (3.23), the function j is at least locally absolutely continuous and therefore its derivative exists almost everywhere. Using that γ is also a solution of (3.29), we obtain by the chain rule that the derivative of j coincides with ι almost everywhere on I. In the following, we introduce the notation from [68, 69]. For every $\ell = (k, \nu) \in \Lambda$, define two complex-valued constants $\eta_{\pm\omega_k,\ell}$ as follows. If $\nu = 1$, let $\eta_{\pm\omega_k,\ell} := \sqrt{\omega_k}/2$, and otherwise, i.e., if $\nu = 2$, let $\eta_{\pm\omega_k,\ell} := \sqrt{\omega_k}/(2\mathrm{i})$, where i denotes the imaginary unit. Moreover, let $\Omega(\ell) := \{\pm\omega_k\}$.

Let $\ell \in \Lambda$. Using the above notation, we can write $u_\ell^j(t)$ in (3.25) as

$$u_\ell^j(t) \;=\; j(t)^{\frac{3}{5}}\sum_{\omega\in\Omega(\ell)}\eta_{\omega,\ell}\,\mathrm{e}^{\mathrm{i}\omega J(t)} \;=:\; -\widetilde{U}v_\ell^j(t) \tag{3.34}$$

for every $t \in I$, where e denotes Euler's number. Using integration by parts and $\dot{J} = j$, we get

$$\int_{t_1}^{t_2} \widetilde{Uv}_\ell^j(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_\ell^j(t)\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \widetilde{uV}_\ell^j(t)\,\mathrm{d}t,$$

where

$$\begin{aligned} \widetilde{UV}_\ell^j(t) &:= -j(t)^{-\frac{2}{5}} \sum_{\omega\in\Omega(\ell)} \frac{\eta_{\omega,\ell}}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}\omega J(t)}, \\ \widetilde{uV}_\ell^j(t) &:= \frac{2}{5}\,\iota(t)\,j(t)^{-\frac{7}{5}} \sum_{\omega\in\Omega(\ell)} \frac{\eta_{\omega,\ell}}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}\omega J(t)}. \end{aligned} \tag{3.35}$$

Finally, we let

$$\begin{aligned} r_\ell^j(t) &:= \widetilde{uV}_\ell^j(t), \\ \widetilde{uv}_\ell^j(t) &:= \widetilde{uV}_\ell^j(t) + \widetilde{Uv}_\ell^j(t). \end{aligned} \tag{3.36}$$

Then we have

$$u_\ell^j(t) \;=\; v_\ell^j(t) + r_\ell^j(t) - \widetilde{uv}_\ell^j(t), \tag{3.37}$$

$$\int_{t_1}^{t_2} \widetilde{uv}_\ell^j(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_\ell^j(t)\Big]_{t=t_1}^{t=t_2}. \tag{3.38}$$

This completes the definitions for a single index ℓ.

Let $\ell_1, \ell_2 \in \Lambda$. When we multiply $u_{\ell_1}^j(t)$ by $\widetilde{UV}_{\ell_2}^j(t)$, we get

$$u_{\ell_1}^j(t)\,\widetilde{UV}_{\ell_2}^j(t) \;=\; v_{\ell_1,\ell_2}^j(t) - \widetilde{Uv}_{\ell_1,\ell_2}^j(t),$$

where

$$\begin{aligned} v_{\ell_1,\ell_2}^j(t) &:= -j(t)^{\frac{1}{5}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2=0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}\omega_2}, \\ \widetilde{Uv}_{\ell_1,\ell_2}^j(t) &:= j(t)^{\frac{1}{5}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2\neq0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)J(t)}. \end{aligned} \tag{3.39}$$

Using integration by parts and $\dot{J} = j$, we get

$$\int_{t_1}^{t_2} \widetilde{Uv}_{\ell_1,\ell_2}^j(t)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_{\ell_1,\ell_2}^j(t)\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \widetilde{uV}_{\ell_1,\ell_2}^j(t)\,\mathrm{d}t,$$

where

$$\begin{aligned} \widetilde{UV}_{\ell_1,\ell_2}^j(t) &:= j(t)^{-\frac{4}{5}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2\neq0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}^2\omega_2(\omega_2+\omega_1)}\,\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)J(t)}, \\ \widetilde{uV}_{\ell_1,\ell_2}^j(t) &:= -\frac{4}{5}\,\iota(t)\,j(t)^{-\frac{9}{5}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2\neq0}} \frac{\eta_{\omega_1,\ell_1}\,\eta_{\omega_2,\ell_2}}{\mathrm{i}^2\omega_2(\omega_2+\omega_1)}\,\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)J(t)}. \end{aligned} \tag{3.40}$$

Finally, we let

$$\begin{aligned} r^j_{\ell_1,\ell_2}(t) &:= \widetilde{uV}^j_{\ell_1,\ell_2}(t), \\ \widetilde{uv}^j_{\ell_1,\ell_2}(t) &:= \widetilde{uV}^j_{\ell_1,\ell_2}(t) + \widetilde{Uv}^j_{\ell_1,\ell_2}(t). \end{aligned} \tag{3.41}$$

Then we have

$$u^j_{\ell_1}(t)\,\widetilde{UV}^j_{\ell_2}(t) = v^j_{\ell_1,\ell_2}(t) + r^j_{\ell_1,\ell_2}(t) - \widetilde{uv}^j_{\ell_1,\ell_2}(t), \tag{3.42}$$

$$\int_{t_1}^{t_2} \widetilde{uv}^j_{\ell_1,\ell_2}(t)\,\mathrm{d}t = \Big[\widetilde{UV}^j_{\ell_1,\ell_2}(t)\Big]_{t=t_1}^{t=t_2}. \tag{3.43}$$

This completes the definitions for two indices ℓ_1, ℓ_2.

The functions in (3.33)-(3.36), (3.40) and (3.41) satisfy the following estimates.

Lemma 3.10. *For every $\tilde{y} \in (y^*, y^+)$, there exists $c > 0$ such that*

$$\begin{aligned} |u^j_\ell(t)| &\leq c/\varphi(\gamma(t))^3, \\ |\iota(t)| &\leq c/\varphi(\gamma(t))^8, \\ |\widetilde{UV}^j_\ell(t)| &\leq c\,\varphi(\gamma(t))^2, \\ |r^j_\ell(t)| &\leq c/\varphi(\gamma(t)), \\ |\widetilde{UV}^j_{\ell_1,\ell_2}(t)| &\leq c\,\varphi(\gamma(t))^4, \\ |r^j_{\ell_1,\ell_2}(t)| &\leq c\,\varphi(\gamma(t)) \end{aligned}$$

for all $\ell, \ell_1, \ell_2 \in \Lambda$, every $t_0 \in \mathbb{R}$, every $p_0 = (x_0, z_0) \in P$ with $z_0 \leq \tilde{y}$, every solution $\gamma\colon I \to P$ of (3.23) *with $\gamma(t_0) = p_0$, and every $t \geq t_0$ in I.*

Proof. The estimates for u^j_ℓ, $\widetilde{UV}^j_\ell$, and $\widetilde{UV}^j_{\ell_1,\ell_2}$ follow from their definitions and the definition of j in (3.32). We know that the Lie derivatives $f_0\varphi$ and $f_\ell\varphi$ satisfy the estimates in Lemma 3.9. Using the estimate for u^j_ℓ, we easily obtain the estimate for ι. This in turn implies the estimates for $\widetilde{uV}^j_\ell$ and $\widetilde{uV}^j_{\ell_1,\ell_2}$. □

A direct computation shows that the $v^j_{\ell_1,\ell_2}$ in (3.39) are given as follows.

Lemma 3.11. *For all $\ell_1 = (k_1, \nu_1), \ell_2 = (k_2, \nu_2) \in \Lambda$, we have*

$$v^j_{\ell_1,\ell_2} = \frac{1}{2}\,j^{\frac{1}{5}} \cdot \begin{cases} +1 & \textit{if } k_1 = k_2 \textit{ and } \nu_1 = 1 \textit{ and } \nu_2 = 2, \\ -1 & \textit{if } k_1 = k_2 \textit{ and } \nu_1 = 2 \textit{ and } \nu_2 = 1, \\ 0 & \textit{otherwise.} \end{cases}$$

Because of Lemma 3.11, we have

$$j^{\frac{1}{5}}\,f^\infty\varphi = \frac{1}{2}\,j^{\frac{1}{5}}\sum_{k=1}^{m}[f_{(k,1)}, f_{(k,2)}]\varphi = \sum_{\ell_1,\ell_2\in\Lambda} v^j_{\ell_1,\ell_2}\,f_{\ell_1}(f_{\ell_2}\varphi) \tag{3.44}$$

on $I \times P$, where the time-varying vector field f^∞ is defined in (3.30).

3.5.3 Integral expansion

For the moment, fix an arbitrary $j > 0$, let $\gamma\colon \mathbb{R} \to P$ be a solution of (3.23), and let $t_1, t_2 \in \mathbb{R}$. Then, the curve γ is locally absolutely continuous and solves (3.29). The fundamental theorem of calculus applied to the composition of γ and φ implies that

$$\varphi(\gamma(t_2)) \;=\; \varphi(\gamma(t_1)) + \int_{t_1}^{t_2} (f_0\varphi)(t,\gamma(t))\,\mathrm{d}t + \sum_{\ell\in\Lambda}\int_{t_1}^{t_2} u_\ell^j(t)\,(f_\ell\varphi)(t,\gamma(t))\,\mathrm{d}t.$$

Note that each of the functions $t \mapsto (f_\ell\varphi)(t,\gamma(t))$ is locally absolutely continuous. Thus, we may apply integration by parts, which leads to

$$\begin{aligned}\varphi(\gamma(t_2)) \;=\;& \varphi(\gamma(t_1)) + \int_{t_1}^{t_2} (f_0\varphi)(t,\gamma(t))\,\mathrm{d}t + \sum_{\ell\in\Lambda}\int_{t_1}^{t_2} r_\ell^j(t)\,(f_\ell\varphi)(t,\gamma(t))\,\mathrm{d}t\\
&- \sum_{\ell\in\Lambda}\Big[\widetilde{UV}_\ell^j(t)\,(f_\ell\varphi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \sum_{\ell\in\Lambda}\int_{t_1}^{t_2}\widetilde{UV}_\ell^j(t)\,(\partial_t(f_\ell\varphi))(t,\gamma(t))\,\mathrm{d}t\\
&+ \sum_{\ell\in\Lambda}\int_{t_1}^{t_2}\widetilde{UV}_\ell^j(t)\,(f_0(f_\ell\varphi))(t,\gamma(t))\,\mathrm{d}t + \sum_{\ell_1,\ell_2\in\Lambda}\int_{t_1}^{t_2} u_{\ell_1}^j(t)\,\widetilde{UV}_{\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\varphi))(t,\gamma(t))\,\mathrm{d}t,\end{aligned}$$

where we have used first (3.37) and then (3.38) as well as that γ is a solution of (3.29). Note that each of the functions $t \mapsto (f_{\ell_1}(f_{\ell_2}\varphi))(t,\gamma(t))$ is locally absolutely continuous. Next, we replace the product of $u_{\ell_1}^j(t)$ and $\widetilde{UV}_{\ell_2}^j(t)$ in the last integral of the above equation by (3.42), and then we apply integration by parts. Using (3.43) and the property that γ is a solution of (3.29), this leads to

$$\begin{aligned}\varphi(\gamma(t_2)) \;=\;& \varphi(\gamma(t_1)) - \Big[(D_1^j\varphi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2}(D_2^j\varphi)(t,\gamma(t))\,\mathrm{d}t\\
&+ \sum_{\ell_1,\ell_2\in\Lambda}\int_{t_1}^{t_2} v_{\ell_1,\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\varphi))(t,\gamma(t))\,\mathrm{d}t,\end{aligned}$$

where the functions $D_1^j\varphi$, $D_2^j\varphi$ on $I \times P$ are defined by

$$(D_1^j\varphi)(t,p) \;:=\; \sum_{\ell\in\Lambda}\widetilde{UV}_\ell^j(t)\,(f_\ell\varphi)(t,p) + \sum_{\ell_1,\ell_2\in\Lambda}\widetilde{UV}_{\ell_1,\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\varphi))(t,p), \tag{3.45}$$

$$\begin{aligned}(D_2^j\varphi)(t,p) \;:=\;& (f_0\varphi)(t,p) + \sum_{\ell\in\Lambda} r_\ell^j(t)(t)\,(f_\ell\varphi)(t,p) + \sum_{\ell_1,\ell_2\in\Lambda} r_{\ell_1,\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\varphi))(t,p)\\
&+ \sum_{\ell\in\Lambda}\widetilde{UV}_\ell^j(t)\,(\partial_t(f_\ell\varphi))(t,p) + \sum_{\ell_1,\ell_2\in\Lambda}\widetilde{UV}_{\ell_1,\ell_2}^j(t)\,(\partial_t(f_{\ell_1}(f_{\ell_2}\varphi)))(t,p)\\
&+ \sum_{\ell\in\Lambda}\widetilde{UV}_\ell^j(t)\,(f_0(f_\ell\varphi))(t,p) + \sum_{\ell_1,\ell_2\in\Lambda}\widetilde{UV}_{\ell_1,\ell_2}^j(t)\,(f_0(f_{\ell_1}(f_{\ell_2}\varphi)))(t,p)\\
&+ \sum_{\ell_1,\ell_2,\ell_3\in\Lambda} u_{\ell_1}^j(t)\,\widetilde{UV}_{\ell_2,\ell_3}^j(t)\,(f_{\ell_1}(f_{\ell_2}(f_{\ell_3}\varphi)))(t,p).\end{aligned} \tag{3.46}$$

Because of (3.44), we have derived the following integral expansion for the propagation of φ along solutions of (3.23).

Proposition 3.12. *For every solution $\gamma\colon I \to P$ of* (3.23) *and all $t_1, t_2 \in I$, we have*

$$\begin{aligned}\varphi(\gamma(t_2)) \;&=\; \varphi(\gamma(t_1)) - \Big[(D_1^j\varphi)(t,\gamma(t))\Big]_{t=t_1}^{t=t_2} \\ &\quad + \int_{t_1}^{t_2} \frac{1}{\varphi(\gamma(t))}\,(f^\infty\varphi)(t,\gamma(t))\,\mathrm{d}t + \int_{t_1}^{t_2} (D_2^j\varphi)(t,\gamma(t))\,\mathrm{d}t,\end{aligned}$$

where $f^\infty\varphi$, $D_1^j\varphi$ and $D_2^j\varphi$ are given by (3.31), (3.45), *and* (3.46), *respectively.*

Since we assume that estimate (3.13) holds for the sum on the right-hand side of (3.31) with some continuous function b on $\psi^{-1}(\leq y^+, x^*)$, we obtain the following result.

Proposition 3.13. *For all $y_\infty, \tilde{y} \in (y^*, y^+)$ with $y_\infty < \tilde{y}$, there exists $c_0 > 0$ such that*

$$(f^\infty\varphi)(t,p) \;\geq\; c_0$$

for every $t \in \mathbb{R}$ and every $p = (x,z) \in P$ with $y_\infty \leq \psi(x) \leq \tilde{y}$.

Next, we derive estimates for the remainders $D_1^j\varphi$ and $D_2^j\varphi$ in Proposition 3.12.

Proposition 3.14. *For every $\tilde{y} \in (y^*, y^+)$, there exist $c_1, c_2 > 0$ such that*

$$\begin{aligned}|(D_1^j\varphi)(t,\gamma(t))| \;&\leq\; c_1\,\varphi(\gamma(t))^2, \\ |(D_2^j\varphi)(t,\gamma(t))| \;&\leq\; c_2\end{aligned}$$

for every $t_0 \in \mathbb{R}$, every $p_0 = (x_0, z_0) \in P$ with $z_0 \leq \tilde{y}$, every solution $\gamma\colon I \to P$ of (3.23) *with $\gamma(t_0) = p_0$, and every $t \geq t_0$ in I.*

Proof. It follows from Lemmas 3.9 and 3.10 that there exists $c > 0$ such that

$$\begin{aligned}|\widetilde{UV}_\ell^j(t)\,(f_\ell\varphi)(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t))^3, \\ |\widetilde{UV}_{\ell_1,\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\varphi))(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t))^4, \\ |r_\ell^j(t)\,(f_\ell\varphi)(t,\gamma(t))| \;&\leq\; c, \\ |r_{\ell_1,\ell_2}^j(t)\,(f_{\ell_1}(f_{\ell_2}\varphi))(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t)), \\ |\widetilde{UV}_\ell^j(t)\,(\partial_t(f_\ell\varphi))(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t))^3, \\ |\widetilde{UV}_{\ell_1,\ell_2}^j(t)\,(\partial_t(f_{\ell_1}(f_{\ell_2}\varphi)))(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t))^4, \\ |\widetilde{UV}_\ell^j(t)\,(f_0(f_\ell\varphi))(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t)), \\ |\widetilde{UV}_{\ell_1,\ell_2}^j(t)\,(f_0(f_{\ell_1}(f_{\ell_2}\varphi)))(t,\gamma(t))| \;&\leq\; c\,\varphi(\gamma(t))^2, \\ |u_{\ell_1}^j(t)\,\widetilde{UV}_{\ell_2,\ell_3}^j(t)(f_{\ell_1}(f_{\ell_2}(f_{\ell_3}\varphi)))(t,\gamma(t))| \;&\leq\; c\end{aligned}$$

for every $t_0 \in \mathbb{R}$, every $p_0 = (x_0, z_0) \in P$ with $z_0 \leq \tilde{y}$, every solution $\gamma\colon I \to P$ of (3.23) with $\gamma(t_0) = p_0$, and every $t \geq t_0$ in I. The asserted estimates follow by applying the above estimates to the terms in the definitions of $D_1^j\varphi$ and $D_2^j\varphi$ in (3.45) and (3.46), respectively. (Note that we could get a slightly stronger estimate for $D_1^j\varphi$ involving the third power of φ instead of the second, but we do not need this in the following.) $\square$

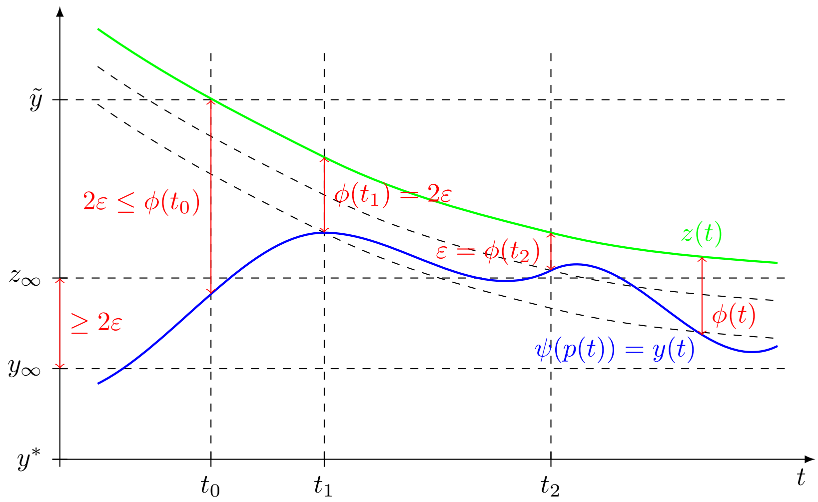


Figure 3.2: *Illustration of the parameters and functions in Subsection 3.5.4.*

3.5.4 Proof of Theorem 3.3

The proof of Theorem 3.3 is complete if we can show that, for every maximal solution $p = (x, z)\colon I \to P$ of (3.23), we have $\sup I = \infty$ and $z(t) \to y^*$ as $t \to \infty$. Assume for the sake of contradiction that this is not satisfied. Then, by Remark 3.8, there exists a maximal solution $p = (x, z)\colon I \to P$ of (3.23) and some $z_\infty > y^*$ such that $z(t) \ge z_\infty$ for every $t \in I$. Fix an arbitrary $t_0 \in I$ and define $\tilde{y} := z(t_0) > z_\infty$. Moreover, let $y_\infty := y^* + (z_\infty - y^*)/2 < z_\infty$. Then, there exists a constant $c_0 > 0$ as in Proposition 3.13 and there exist constants $c_1, c_2 > 0$ as in Proposition 3.14. Define $\phi := \varphi \circ p\colon I \to (0, \infty)$. From Proposition 3.12, we conclude that, for all $t_2 > t_1 \ge t_0$ in I, the following implication holds: if $\psi(x(t)) \ge y_\infty$ for every $t \in [t_1, t_2]$, then

$$\phi(t_2) \;\ge\; \frac{1 - c_1\,\phi(t_1)}{1 + c_1\,\phi(t_2)}\,\phi(t_1) + \int_{t_1}^{t_2} \frac{c_0/\phi(t) - c_2}{1 + c_1\,\phi(t_2)}\,\mathrm{d}t. \tag{3.47}$$

Define

$$\varepsilon \;:=\; \frac{1}{2}\,\min\left\{\frac{1}{3c_1},\; \phi(t_0),\; \frac{c_0}{c_2},\; z_\infty - y_\infty\right\} \;>\; 0.$$

Now we show that $\phi(t) \ge \varepsilon$ for every $t \ge t_0$ in I. Assume for the sake of contradiction that this is not true. Then, since ϕ is continuous with $\phi(t_0) \ge 2\varepsilon$, there exist $t_2 > t_1 \ge t_0$ in I such that $\phi(t_1) = 2\varepsilon$, $\phi(t_2) = \varepsilon$, and $\varepsilon < \phi(t) < 2\varepsilon$ for every $t \in (t_1, t_2)$. In particular, it follows that $\phi(t) \le 2\varepsilon \le z_\infty - y_\infty$, and therefore $\psi(x(t)) = z(t) - \phi(t) \ge y_\infty$ for every $t \in [t_1, t_2]$. Now (3.47) and the definition of ε, lead to the contradiction $\phi(t_2) > \varepsilon$. Thus, $\phi(t) \ge \varepsilon$ for every $t \ge t_0'$ in I. This in turn implies that the curve p stays in a compact subset of P on $I \cap [t_0, \infty)$. It follows that $\sup I = \infty$. Now $\phi(t) \ge \varepsilon$ for every $t \ge t_0$ and (3.15) lead to the contradiction $z(t) \to -\infty$ as $t \to \infty$.

4 Extremum seeking control for a class of nonholonomic systems

The content of this chapter is an extended version of [110].

4.1 Introduction and motivation

We return to the extremum seeking control problem (1.39), (1.40) in Section 1.3. This means that we consider again a multiple-input single-output system of the form

$$\dot{x} = \sum_{i=1}^{m} u_i f_i(x), \tag{4.1}$$

$$y = \psi(x), \tag{4.2}$$

where the u_i are real-valued input channels and y is a real-valued output channel. The control vector fields f_i and the output function ψ are assumed to be smooth. We have seen in Section 1.3 that a suitable output feedback law for (4.1), which involves highly oscillatory time-varying functions, leads to the effect that the trajectories of the closed-loop system, denoted by Σ^j, approximate the trajectories of an averaged system of the form

$$\dot{x} = -\sum_{i=1}^{m} (f_i\psi)(x)\, f_i(x) \tag{4.3}$$

if the frequency parameter j in Σ^j is sufficiently large. The approximation property can be explained by the well-developed Lie bracket averaging theory from [58, 69, 68]; cf. Section 1.2. A consequence of the approximation property is that if a point x^* of the state manifold is asymptotically stable for (4.3), then one can achieve that an arbitrary small neighborhood of x^* is asymptotically stable for Σ^j by choosing the frequency parameter j sufficiently large. Thus, asymptotic stability for the averaged system implies practical asymptotic stability for the closed-loop system. For the purpose of extremum seeking control it is therefore important to ensure that a minimum point x^* of ψ is asymptotically stable for the averaged system (4.3).

Suppose that the output function ψ attains a local minimum value at some point x^* of the state manifold, and assume that ψ has no other critical point than x^* in a certain neighborhood of x^*. As explained in the preceding paragraph, we may conclude that x^* is practically asymptotically stable for (4.1) under the Lie bracket-based extremum seeking control law if x^* is asymptotically stable for the averaged system (4.3). This is indeed the case if (4.1) is *fully actuated* at x^*; i.e., if the control vector fields span the entire tangent space at x^*. Then, ψ is in fact (up to an additive constant) a local Lyapunov function for (4.3) around x^*. The situation changes if (4.1) is *underactuated*; i.e., not fully actuated. Then, it might happen that there are undesired equilibria of (4.3) arbitrary close to x^*, and therefore x^* is not asymptotically stable for (4.3). In this case stabilization cannot be guaranteed by the existing results. Many underactuated systems of the form (4.1) are *nonholonomic*; i.e., the distribution generated by the control vector fields is regular but not closed under Lie bracketing. We refer to [64, 83] for examples of nonholonomic systems. To the best of our knowledge, there is no universal approach to extremum seeking control for nonholonomic control-affine systems in the literature so far. The intention of the chapter is to propose a solution to the problem, at least for a certain subclass of control systems of the form (4.1). Many nonholonomic systems have the additional property that their control vector fields satisfy the *Lie algebra rank condition*. This means that the tangent space is spanned by Lie brackets of the control vector fields (see again [64, 83] for examples of such systems). Under the assumption that the Lie algebra rank condition is satisfied at the optimal point x^*, the proposed method can ensure that an arbitrary small neighborhood of x^* becomes locally asymptotically stable for the closed-loop system. In contrast to the existing methods for extremum seeking by Lie bracket approximations, the novel control law does not only steer the system into descent directions of the output function along the control vector fields but also along Lie brackets of the control vector fields. The approach uses a suitable combination of the approximation algorithm from [68] (to generate Lie brackets of the control vector fields) and the Lie bracket-based extremum seeking strategies from [31, 95] (to get access to descent directions of the output function). This leads to the first extremum seeking control law that gives access to descent directions of the output function along Lie brackets of any desired degree.

The chapter is organized as follows. The subsequent Section 4.2 contains some algebraic concepts which are required to state the approximation algorithm from [68]. In Section 4.3 we describe the control objective and outline the proposed strategy. The extremum seeking control law is presented in Section 4.4. The main results on approximation and stability are stated in Section 4.5. We also illustrate the stability result by a numerical simulation. Since the proof of the approximation result requires some lengthy computations, it is carried out at the end in Sections 4.6 and 4.7.

4.2 Local definitions and notation for the chapter

We begin by recalling several algebraic concepts from [13, 14, 104]. Let $\mathbf{X} = \{X_1, \ldots, X_m\}$ be a nonempty finite set of m pairwise distinct objects $X_1, \ldots, X_m$, called *indeterminates*. A sequence of sets $\mathrm{M}_\ell(\mathbf{X})$ is defined by induction on $\ell = 1, 2, \ldots$ as follows. For $\ell = 1$, we let $\mathrm{M}_1(\mathbf{X}) := \mathbf{X}$. For $\ell = 2, 3, \ldots$, the set $\mathrm{M}_\ell(\mathbf{X})$ is defined as the disjoint union of $\mathrm{M}_k(\mathbf{X}) \times \mathrm{M}_{\ell-k}(\mathbf{X})$ with $k = 1, \ldots, \ell-1$. The disjoint union of all the sets $\mathrm{M}_\ell(\mathbf{X})$ is denoted by $\mathrm{M}(\mathbf{X})$. Each of the sets $\mathrm{M}_\ell(\mathbf{X})$ is identified with its canonical image in $\mathrm{M}(\mathbf{X})$. To give an example: if $m \geq 2$, then $((X_1, X_1), X_2)$ is considered to be an element of both $\mathrm{M}_3(\mathbf{X})$ and $\mathrm{M}(\mathbf{X})$. The set $\mathrm{M}(\mathbf{X})$ is called the *free magma generated by* $\mathbf{X}$. However, for our purposes, it is more suitable to refer to $\mathrm{M}(\mathbf{X})$ as the *set of formal brackets generated by* $\mathbf{X}$. For every $B \in \mathrm{M}(\mathbf{X})$, there exists a unique positive integer, denoted by $\delta(B)$, such that $B \in \mathrm{M}_{\delta(B)}(\mathbf{X})$, called the *degree of* B. For every $i \in \{1, \ldots, m\}$ and every $B \in \mathrm{M}(\mathbf{X})$, the *degree of* B *in* X_i, denoted by $\delta_i(B)$, is the nonnegative integer that counts the number of appearances of X_i in B. To give an example: if $B = ((X_1, X_1), X_2)$, then we have $\delta(B) = 3$, $\delta_1(B) = 2$, and $\delta_2(B) = 1$. For all $B, B' \in \mathrm{M}(\mathbf{X})$, the image of (B, B') under the canonical injection of $\mathrm{M}_{\delta(B)}(\mathbf{X}) \times \mathrm{M}_{\delta(B')}(\mathbf{X})$ into $\mathrm{M}(\mathbf{X})$ is denoted by the same symbol and is called the *formal bracket of* B *and* B'. Conversely, for every $B \in \mathrm{M}(\mathbf{X})$ with $\delta(B) > 1$, there exist unique $B_1, B_2 \in \mathrm{M}(\mathbf{X})$ such that $B = (B_1, B_2)$, where B_1 and B_2 are called the *left and* *right factors of* B, respectively. To give an example: if $B = ((X_1, X_1), X_2)$, then the left factor of B is (X_1, X_1) and the right factor of B is X_2. A *P. Hall set of* $\mathrm{M}(\mathbf{X})$ is a subset $\mathcal{B}$ of $\mathrm{M}(\mathbf{X})$, endowed with a total order[1] $\preceq$, that satisfies the following properties:

PH1. if $B_1 \in \mathcal{B}$, $B_2 \in \mathcal{B}$, and $\delta(B_1) < \delta(B_2)$, then $B_1 \prec B_2$;

PH2. every indeterminate $X_i \in \mathrm{M}_1(\mathbf{X})$ belongs to $\mathcal{B}$, and a pair $(X_{i_1}, X_{i_2}) \in \mathrm{M}_2(\mathbf{X})$ belongs to $\mathcal{B}$ if and only if $X_{i_1} \prec X_{i_2}$;

PH3. an element $B \in \mathrm{M}_k(\mathbf{X})$ of degree $k \geq 3$ belongs to $\mathcal{B}$ if and only if there exist $B_1, B_2, B_3 \in \mathcal{B}$ such that $B = (B_1, (B_2, B_3))$, $(B_2, B_3) \in \mathcal{B}$, $B_2 \preceq B_1 \prec (B_2, B_3)$, and $B_2 \prec B_3$.

As in [68], we will additionally require the convenient property that

PH4. for all $i_1, i_2 \in \{1, \ldots, m\}$ with $i_1 < i_2$, we have $X_{i_1} \prec X_{i_2}$.

Note that PH4. can always be established by simply relabeling the indeterminates.

Proposition 4.1 (Proposition II.2.11 in [14])**.** *For every finite set* $\mathbf{X}$ *of indeterminates, there exists a P. Hall set of* $\mathrm{M}(\mathbf{X})$.

The *free non-unital associative algebra generated by* $\mathbf{X}$ *over* $\mathbb{R}$, denoted by $\mathrm{A}_0(\mathbf{X})$, is the non-unital associative algebra[2] of all linear combinations of *monomials*

$$X_I \;:=\; X_{i_1} \cdots X_{i_k}, \tag{4.4}$$

where $I = (i_1, \ldots, i_k)$ is any multi-index of length $k > 0$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$. As usual, for all $p, q \in \mathrm{A}_0(\mathbf{X})$, the *Lie bracket of* p *and* q is defined by $[p, q] := pq - qp$. It is well-known that the Lie bracket turns $\mathrm{A}_0(\mathbf{X})$ into a Lie algebra[3]. Let $\mathrm{L}(\mathbf{X})$ be the Lie subalgebra of $\mathrm{A}_0(\mathbf{X})$ generated by $\mathbf{X}$; i.e., the smallest Lie subalgebra of $\mathrm{A}_0(\mathbf{X})$ that contains $\mathbf{X}$. This Lie algebra is called the *free Lie algebra generated by* $\mathbf{X}$ *over* $\mathbb{R}$. Let $\mu\colon \mathrm{M}(\mathbf{X}) \to \mathrm{L}(\mathbf{X})$ denote the canonical map that replaces round brackets "(", ")" by square brackets "[", "]". For instance, $\mu((X_1, X_1)) = [X_1, X_1] = 0 = [X_2, X_2] = \mu((X_2, X_2))$, but, of course, $(X_1, X_1) \neq (X_2, X_2)$.

Theorem 4.2 (Theorem II.2.1 in [14])**.** *Let* $(\mathcal{B}, \preceq)$ *be a P. Hall set of* $\mathrm{M}(\mathbf{X})$*. Then, the above map* μ *is injective on* $\mathcal{B}$ *and the image of* $\mathcal{B}$ *under* μ *is a basis of the vector space* $\mathrm{L}(\mathbf{X})$.

Let M be a smooth manifold. The commutative algebra of smooth functions on M is denoted by $C^\infty(M)$. The set of smooth vector fields on M is denoted by $\mathfrak{X}(M)$. If we apply a smooth vector field f to a smooth function φ, then the Lie derivative $f\varphi$ of φ along f is again a smooth function. In this sense, every smooth vector field can be considered as a vector space endomorphism on $C^\infty(M)$, which gives $\mathfrak{X}(M)$ the structure of a non-unital associative algebra. It is well-known that the Lie bracket (cf. Section 1.5) turns $\mathfrak{X}(M)$

[1] A *partial order on* $\mathcal{B}$ is a relation $\preceq$ on $\mathcal{B}$ with the following properties for all $B_1, B_2, B_3 \in \mathcal{B}$: (i) $B_1 \preceq B_1$; (ii) if $B_1 \preceq B_2$ and $B_2 \preceq B_1$, then $B_1 = B_2$; (iii) if $B_1 \preceq B_2$ and $B_2 \preceq B_3$, then $B_1 \preceq B_3$. A *total order on* $\mathcal{B}$ is a partial order $\preceq$ on $\mathcal{B}$ such that $B_1 \preceq B_2$ or $B_2 \preceq B_1$ for all $B_1, B_2 \in \mathcal{B}$. If we have $B_1 \preceq B_2$ and $B_1 \neq B_2$ for $B_1, B_2 \in \mathcal{B}$, then we write $B_1 \prec B_2$.

[2] A *non-unital associative algebra over* $\mathbb{R}$ is a vector space A_0 over $\mathbb{R}$ endowed with an associative bilinear map $A_0 \times A_0 \to A_0$, $(p, q) \mapsto pq$, which is called the *multiplication* on A_0.

[3] A *Lie algebra over* $\mathbb{R}$ is a vector space L over $\mathbb{R}$ endowed with a bilinear map $L \times L \to L$, $(x, y) \mapsto [x, y]$ such that $[x, x] = 0$ and $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in L$

into a Lie algebra. Let $\mathbf{f}$ be a nonempty finite subset of $\mathfrak{X}(M)$. We denote by $\mathrm{L}(\mathbf{f})$ the *Lie subalgebra of* $\mathfrak{X}(M)$ *generated by* $\mathbf{f}$; i.e. the smallest Lie subalgebra of $\mathfrak{X}(M)$ that contains $\mathbf{f}$. For every positive integer r, we define a subspace $\mathrm{L}^r(\mathbf{f})$ of $\mathfrak{X}(M)$ as follows. For $r = 1$, let $\mathrm{L}^1(\mathbf{f})$ be the $\mathbb{R}$-span of $\mathbf{f}$. By induction on r, define $\mathrm{L}^{r+1}(\mathbf{f})$ to be the span of all elements in $\mathrm{L}^r(\mathbf{f})$ and all Lie brackets $[g_1, g_2]$ with $g_1 \in \mathrm{L}^{r_1}(\mathbf{f})$, $g_2 \in \mathrm{L}^{r_2}(\mathbf{f})$, and $r_1 + r_2 = r + 1$. Clearly, the sets $\mathrm{L}^1(\mathbf{f}) \subseteq \mathrm{L}^2(\mathbf{f}) \subseteq \cdots$ form an increasing sequence of subspaces of $\mathrm{L}(\mathbf{f})$.

4.3 Problem statement and motivation of the proposed control strategy

Throughout the chapter, we consider a control-affine system of the form

$$\dot{x} \;=\; \sum_{i=1}^{m} u_i\, f_i(x) \tag{4.5}$$

on a smooth *state manifold* M with smooth *control vector fields* $f_1, \ldots, f_m \in \mathfrak{X}(M)$, real-valued *input channels* $u_1, \ldots, u_m$, and a real-valued *output channel*

$$y \;=\; \psi(x), \tag{4.6}$$

where $\psi \in C^\infty(M)$ is called the *output function*. We assume that the current value of (4.6) can be measured constantly while the system state of (4.5) is an unknown quantity. Moreover, information about descent directions of ψ are not assumed to be known. The goal is to derive an output-feedback control law for (4.5) that asymptotically stabilizes the closed-loop system around states at which the output function attains a local minimum value. In the following paragraphs, we explain how the problem can be solved under suitable assumptions on the control vector fields and the output function. We use the notation and definitions from Sections 1.5 and 4.2.

Suppose that the output function ψ attains a local minimum value $y^* \in \mathbb{R}$ at some point $x^* \in M$. For each $x \in M$, a *descent direction of* ψ *at* x is any tangent vector v_x to M at x such that $\mathrm{d}\varphi(x)v_x < 0$. For our goal to asymptotically stabilize (4.5) around x^*, it is certainly desirable to steer the control system (at least approximately) into descent directions of ψ. In order to do so, we need two additional assumptions. First, to ensure the existence of descent directions around x^*, we assume that ψ has no other critical point than x^* in a certain neighborhood of x^*. Second, to get access to those descent directions, we assume that the control vector fields of (4.5) satisfy the *Lie algebra rank condition* at x^*; i.e., the elements of the Lie algebra generated by the vector fields in $\mathbf{f} := \{f_1, \ldots, f_m\}$ span the entire tangent space to M at x^*. The latter assumption ensures that there exist a sufficiently large positive integer r and vector fields $g_\ell \in \mathrm{L}^r(\mathbf{f})$, $\ell \in \Lambda^r$, indexed by some finite set Λ^r such that, for every x in some neighborhood of x^*, the vectors $g_\ell(x)$ with $\ell \in \Lambda^r$ span the tangent space to M at x. Because of the first assumption on ψ, it follows that there exists some $y^+ > y^*$ such that $\psi^{-1}(\leq y^+, x^*)$ is compact and such that, for every $x \in \psi^{-1}(\leq y^+, x^*)$ with $x \neq x^*$, there exists $\ell \in \Lambda^r$ such that $(g_\ell\psi)(x) \neq 0$. Note that if $(g_\ell\psi)(x) \neq 0$, then $-(g_\ell\psi)(x)^2 < 0$, and therefore the tangent vector $-(g_\ell\psi)(x)\, g_\ell(x)$ is a descent direction of ψ at x. Using a standard Lyapunov argument, it is now easy to verify that x^* is locally asymptotically stable for the system

$$\dot{x} \;=\; -\sum_{\ell\in\Lambda^r} (g_\ell\psi)(x)\, g_\ell(x). \tag{4.7}$$

Thus, if we can find an output-feedback control law for (4.5) such that the trajectories of the closed-loop system approximate the trajectories of (4.7) sufficiently well, then we can expect that this closed-loop system has similar desirable stability properties as (4.7). This is exactly the intention of the control law that we present Section 4.4. An outline of the approximation approach is given in the subsequent paragraphs.

As motivated above, we are interested in an output-feedback control law for (4.5) such that the trajectories of the closed-loop system approximate the trajectories of (4.7). Note that information about the Lie derivatives $g_\ell\psi$ on the right-hand side of (4.7) is not directly accessible from measurements of current value of ψ. To circumvent this problem, we use the extremum seeking control approach from Section 1.3. For the sake of completeness and for later references, we recall the steps in the following. As in (1.42), we write each of the vector fields $-(g_\ell\psi)\, g_\ell$ as the Lie bracket of two vector fields $\bar{f}_{(\ell,1)}, \bar{f}_{(\ell,2)}$ that only depend on g_ℓ and the current value of ψ. As in (1.41), for every $\ell \in \Lambda^r$, we define $\bar{f}_{(\ell,1)}, \bar{f}_{(\ell,2)} \in \mathfrak{X}(M)$ by

$$\bar{f}_{(\ell,1)}(x) \;:=\; h_1\big(\psi(x)\big)\, g_\ell(x), \tag{4.8a}$$

$$\bar{f}_{(\ell,2)}(x) \;:=\; h_2\big(\psi(x)\big)\, g_\ell(x), \tag{4.8b}$$

where the smooth design functions h_1, h_2 on $\mathbb{R}$ are given by (1.33); i.e.,

$$h_1(y) \;:=\; \sin(y), \tag{4.9a}$$

$$h_2(y) \;:=\; \cos(y). \tag{4.9b}$$

A direct computation shows that

$$[\bar{f}_{(\ell,1)}, \bar{f}_{(\ell,2)}](x) \;=\; -(g_\ell\psi)(x)\, g_\ell(x) \tag{4.10}$$

for every $x \in M$. Next, choose pairwise distinct positive real frequency coefficients $\bar{\omega}_\ell$, $\ell \in \Lambda^r$, and, as in (1.44), for every $j > 0$, define sinusoids $\bar{u}^j_{(\ell,1)}, \bar{u}^j_{(\ell,2)} \colon \mathbb{R} \to \mathbb{R}$ by

$$\bar{u}^j_{(\ell,1)}(t) \;:=\; (2\,\bar{j}\,\bar{\omega}_\ell)^{\frac{1}{2}}\, \cos(\bar{j}\,\bar{\omega}_\ell t), \tag{4.11a}$$

$$\bar{u}^j_{(\ell,2)}(t) \;:=\; (2\,\bar{j}\,\bar{\omega}_\ell)^{\frac{1}{2}}\, \sin(\bar{j}\,\bar{\omega}_\ell t), \tag{4.11b}$$

where the precise dependence of $\bar{j}$ on j will be defined later in (4.16). The amplitudes and frequencies of the sinusoids $\bar{u}^j_{(\ell,1)}, \bar{u}^j_{(\ell,2)}$ grow with increasing parameter j. Then, we know from Section 1.3 that the trajectories of

$$\dot{x} \;=\; \sum_{\ell\in\Lambda^r} \Big(\bar{u}^j_{(\ell,1)}(t)\, \bar{f}_{(\ell,1)}(x) + \bar{u}^j_{(\ell,2)}(t)\, \bar{f}_{(\ell,2)}(x)\Big) \tag{4.12}$$

approximate the trajectories of (4.7) if j is sufficiently large. One can interpret (4.12) as the closed-loop system that originates from the fictitious control-affine system

$$\dot{x} \;=\; \sum_{\ell\in\Lambda^r} u_\ell\, g_\ell(x) \tag{4.13}$$

under the control law

$$u_\ell \;=\; \bar{u}^j_{(\ell,1)}(t)\, h_1(y) + \bar{u}^j_{(\ell,2)}(t)\, h_2(y) \;=\; \big(2\bar{j}\bar{\omega}_\ell\big)^{\frac{1}{2}}\, \sin(\bar{j}\bar{\omega}_\ell t + y) \tag{4.14}$$

with $y = \psi(x)$, where we have used the trigonometric identity (1.49) in the last equality.

As indicated in the preceding paragraph, a suitable choice of vector fields $\bar{f}_{(\ell,\nu)}$ and highly oscillatory signals $\bar{u}^{j}_{(\ell,\nu)}$ ensures that the trajectories of (4.12) approximate the trajectories of (4.7) if the parameter j is sufficiently large. Note that system (4.12) has the beneficial feature that its right-hand side does not depend on Lie derivatives of ψ but only on the current value of ψ, which is a first step toward an output-feedback law for (4.5). However, the vector fields in (4.8) involve the elements g_ℓ from the Lie algebra generated by the control vector fields f_i, which are not necessarily in the linear span of the f_i. To circumvent this problem, we can use again highly oscillatory signals to approximate the vector fields g_ℓ. Based on the approximation algorithm in [68], we will construct a time-varying output-feedback control law $u_i = u_i^j(t;y)$ for (4.5) such that the trajectories of the closed-loop system

$$\dot{x} \;=\; \sum_{i=1}^{m} u_i^j(t;y)\, f_i(x) \tag{4.15}$$

approximate the trajectories of (4.12) if the parameter j is sufficiently large. In summary, our approach to extremum seeking control involves the following two *approximation properties* for sufficiently large values of the parameter j:

AP1. system (4.15) approximates the behavior of system (4.12), and

AP2. system (4.12) approximates the behavior of system (4.7).

At this point we merely note that the parameter $\bar{j}$ will be defined in such a way that it automatically increases with increasing j. This is made precise later in (4.16). A consequence of AP1 and AP2 is that (4.15) approximates the behavior of system (4.7) for sufficiently large j. Thus, if a point $x^* \in M$ is asymptotically stable for (4.7), we can expect that x^* is at least practically asymptotically stable for (4.15), where the word *practically* indicate the dependence on the parameter j.

The above approximation properties AP1 and AP2 require oscillating signals with sufficiently large amplitudes and frequencies. However, an approximation of the vector fields in (4.8) also requires that the output value varies sufficiently slow compared to oscillations that lead to AP1. Otherwise, the approach would lead to an approximation of vector fields which involve undesired Lie derivatives of ψ. On the other hand, an approximation of the Lie brackets in (4.10) requires that the output value varies sufficiently fast compared to the oscillations that lead to AP2. To avoid undesired resonances, the control law in Section 4.4 is designed in such a way that it induces a certain *separation of time scales*. With increasing parameter j, the amplitudes and frequencies that lead to AP1 grow faster than the amplitudes and frequencies that lead to AP2. For this reason, the signals associated with AP1 will be called "*fast oscillations*" and the signals associated with AP2 will be called "*slow oscillations*". It is clear that this is a slightly misleading terminology because the amplitudes and frequencies of both types of signals grow with increasing parameter j. In order to "slow down" the variations of the output value sufficiently, we introduce a first-order hold of the output signal with a suitable sampling rate. With increasing parameter j, the sampling rate becomes slow compared to the fast oscillations and fast compared to the slow oscillations. The precise definitions are given in the next section.

4.4 Control law

Let m be the number of input channels in (4.5) and let r be the maximum degree of Lie brackets of the control vector fields in (4.5) that shall be approximated. As explained in

Section 4.3, the approximation approach involves slow and fast oscillations as well as a first-order hold of the output signal.

4.4.1 Definition of the slow oscillations and the first-order hold

As in Section 4.2, let $\mathbf{X} = \{X_1, \ldots, X_m\}$ be a finite set of m indeterminates, and let $(\mathcal{B}, \preceq)$ be a P. Hall set of $\mathrm{M}(\mathbf{X})$. For every positive integer n, let $\mathcal{B}_n$ denote the elements of $\mathcal{B}$ of degree n. Divide each $\mathcal{B}_n$ into equivalence classes by declaring two of its members B_1, B_2 equivalent if, for every $i \in \{1, \ldots, m\}$, we have $\delta_i(B_1) = \delta_i(B_2)$. Let $\mathcal{E}_n$ be the set of all equivalence classes of $\mathcal{B}_n$. For every $E \in \mathcal{E}_n$ and every $i \in \{1, \ldots, m\}$, we can define $\delta(E)$, $\delta_i(E)$ to be $\delta(B)$, $\delta_i(B)$ for an arbitrary representative B of E. Note that $\mathcal{E}_1$ consists precisely of the singletons $\{X_i\}$ with $i \in \{1, \ldots, m\}$. Because of PH4., the set $\mathcal{E}_2$ consists precisely of the singletons $\{(X_{i_1}, X_{i_2})\}$ with $i_1, i_2 \in \{1, \ldots, m\}$ and $i_1 < i_2$. The set

$$\Lambda^r \;:=\; \big\{\ell = (E, \rho) \;\big|\; E \in \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_r,\ \rho \in \{1, \ldots, |E|\}\big\}$$

will play the role of the indexing set for the vector fields g_ℓ in the averaged system (4.7), where the g_ℓ are specified later in Section 4.5.

The time scale of the fast oscillations will be defined by a parameter $j > 0$, and

$$\bar{j} \;:=\; j^{\frac{1}{r+1}} \tag{4.16}$$

will define the time scale of the slow oscillations. Choose a positive real constant Δ. Let $\mathcal{Y}$ denote the set of all functions on half-open intervals of the form $[0, T)$, where T is either a positive real number and otherwise $+\infty$. For every such function $y\colon [0, T) \to \mathbb{R}$ and every $j > 0$, we define a function $\bar{y}^j\colon [0, T) \to \mathbb{R}$, called the *first-order hold of y with sampling time $\Delta/\bar{j}^2$*, by

$$\bar{y}^j(t) \;:=\; y(\tau_{k-1}^j) + \frac{t - \tau_k^j}{\Delta/\bar{j}^2}\big(y(\tau_k^j) - y(\tau_{k-1}^j)\big) \tag{4.17}$$

for every integer $k \geq 0$ and every $t \in [\tau_k^j, \tau_{k+1}^j)$ with $t < T$, where $\tau_{-1}^j := 0$ and

$$\tau_k^j \;:=\; k\,\Delta/\bar{j}^2.$$

This means that $\bar{y}^j$ is just a linear interpolation of the values of y at $\tau_0^j, \tau_1^j, \ldots$. For large j, we have the separation of time scales $1 \ll \bar{j} \ll \bar{j}^2/\Delta \ll \bar{j}^{r+1} = j$ of the control-vector fields, the slow oscillations, the first-order hold, and the fast oscillations.[4]

Choose pairwise distinct positive real frequency coefficients $\bar{\omega}_\ell$ with $\ell \in \Lambda^r$. For every $\ell = (E, \rho) \in \Lambda^r$, every $k \in \{1, \ldots, \delta(E)\}$, every $j > 0$, and every $y \in \mathcal{Y}$, define a function $\zeta_{E,\rho,k}^j(\cdot; y)$ on the domain of y by

$$\zeta_{E,\rho,k}^j(t; y) \;:=\; 2^{\frac{\delta(E)-1}{\delta(E)}}\,(2\,\bar{j}\,\bar{\omega}_\ell)^{\frac{1}{2\,\delta(E)}}\,\sin\Big(\frac{\bar{j}\,\bar{\omega}_\ell\,t + \bar{y}^j(t) + (k-1)\pi}{\delta(E)}\Big). \tag{4.18}$$

Using standard trigonometric identities[5], one easily obtains that

$$\zeta_\ell^j(t; y) \;:=\; \prod_{k=1}^{\delta(E)} \zeta_{E,\rho,k}^j(t; y) \;=\; \bar{u}_{(\ell,1)}^j(t)\,h_1(\bar{y}^j(t)) + \bar{u}_{(\ell,2)}^j(t)\,h_2(\bar{y}^j(t)), \tag{4.19}$$

[4] In the "trivial" case $r = 1$, we do not need fast oscillations to generate the vector fields g_ℓ, and therefore we only have the three time scales $1 \ll \bar{j} \ll \bar{j}^2/\Delta$.

[5] To be more precise, we use the trigonometric identities (1.49) and $\sin(\alpha) = 2^{n-1}\prod_{k=1}^{n}\sin\big(\frac{\alpha+(k-1)\pi}{n}\big)$ with $n = \delta(E)$ and $\alpha = \bar{j}\,\bar{\omega}_\ell\,t + \bar{y}^j(t)$.

where h_1, h_2 are defined by (4.9) and the $\bar{u}^j_{(\ell,1)}, \bar{u}^j_{(\ell,2)}$ are defined by (4.11). The sinusoids $\bar{u}^j_{(\ell,1)}, \bar{u}^j_{(\ell,2)}$ serve as the slow oscillations to induce the approximation property AP2 in Section 4.3. The first-order hold $\bar{y}^j$ in (4.19) has the purpose to "slow down" the variations of the output signal sufficiently. It is left to specify the fast oscillations to induce the approximation property AP1. For this purpose, we use the approximation algorithm from [68], which is summarized in the following subsection.

4.4.2 Definition of the fast oscillations

The cardinality of a finite set F is denoted by $|F|$. Let F be a finite subset of $\mathbb{R} \setminus \{0\}$. The set F is said to be *canceling* if the sum of all the members of F is equal to 0. The set F is said to be *properly noncanceling* (PNC) if every proper nonempty subset of F is not canceling. The set F is said to be *minimally canceling* (MC) if, for all integers a_ω with $\sum_{\omega \in F} |a_\omega| \leq |F|$, the following equivalence holds: $\sum_{\omega \in F} a_\omega \omega = 0$ if and only if the a_ω are all equal (in which case, of course, they all have to be equal to 0, 1, or -1). It is clear that MC implies PNC. Finally, F is said to be *symmetrically minimally canceling* (SMC) if it is *symmetric*, i.e. $F = -F$, and if it contains an MC subset of cardinality $|F|/2$.

For every $i \in \{1, \ldots, m\}$ and every $\omega \in \mathbb{R} \setminus \{0\}$, define $\hat{g}_{X_i}(\omega) := 1$. For a formal bracket $B \in \mathcal{B}$ of degree ≥ 2, we define $\hat{g}_B$ inductively as follows. Since $(\mathcal{B}, \preceq)$ is a P. Hall set, there exists a unique positive integer κ and unique $B_1, B_2 \in \mathcal{B}$ such that $B = \mathrm{ad}^{\kappa}_{B_1}(B_2)$, where $B_1 \prec B_2$ and either $\delta(B_2) = 1$ or the left factor $B_3 \in \mathcal{B}$ of B_2 satisfies $B_3 \preceq B_1$ (here $\mathrm{ad}^{\kappa}_{B_1}$ denotes the κ-fold application of the map $\mathrm{ad}_{B_1} \colon \mathrm{M}(\mathbf{X}) \to \mathrm{M}(\mathbf{X})$ defined by $\mathrm{ad}_{B_1}(Z) := (B, Z)$)[6]. Then, for all pairwise distinct $\omega_1, \ldots, \omega_{\delta(B)} \in \mathbb{R} \setminus \{0\}$ that form a PNC set, the assignment

$$\hat{g}_B(\omega_1, \ldots, \omega_{\delta(B)}) \;:=\; \hat{g}_{B_2}(\omega_{\kappa\delta(B_1)+1}, \ldots, \omega_{\delta(B)}) \, \frac{1}{\kappa!} \prod_{q=1}^{\kappa} \frac{\hat{g}_{B_1}(\omega_{(q-1)\delta(B_1)+1}, \ldots, \omega_{q\delta(B_1)})}{\omega_{(q-1)\delta(B_1)+1} + \cdots + \omega_{q\delta(B_1)}}$$

gives a well-defined real number. For every $B \in \mathcal{B}$, we also define a map g_B as follows. Let Σ_B denote the list of $\delta(B)$ indeterminates that originates from B by deleting all round brackets. For each $i \in \{1, \ldots, m\}$, let $\theta_{B,i}$ be the (possibly empty) set of those $k \in \{1, \ldots, \delta(B)\}$ for which the indeterminate X_i is at the kth position of Σ_B. For each $i \in \{1, \ldots, m\}$, let $\mathcal{I}_{B,i}$ denote the (possibly empty) set of integers from $\delta_1(B) + \cdots + \delta_{i-1}(B) + 1$ to $\delta_1(B) + \cdots + \delta_i(B)$. Let P_B be the set of all permutations of $\{1, \ldots, \delta(B)\}$ that map $\theta_{B,i}$ to $\mathcal{I}_{B,i}$ for every $i \in \{1, \ldots, m\}$. Let $S_1, \ldots, S_m$ be (possibly empty) pairwise disjoint subsets of $\mathbb{R} \setminus \{0\}$ such that $|S_i| = \delta_i(B)$ for every $i \in \{1, \ldots, m\}$ and such that $S_1 \cup \cdots \cup S_m$ is PNC. Then, the assignment

$$g_B(S_1, \ldots, S_m) \;:=\; \sum_{\pi \in P_B} \hat{g}_B(\omega_{\pi(1)}, \ldots, \omega_{\pi(\delta(B))}) \tag{4.20}$$

gives a well-defined real number, where $\omega_1, \ldots, \omega_{\delta(B)}$ is any listing of the elements of $S_1 \cup \cdots \cup S_m$ such that $S_i = \{\omega_k, k \in \mathcal{I}_{B,i}\}$ for every $i \in \{1, \ldots, m\}$.

[6] To obtain the unique decomposition $B = \mathrm{ad}^{\kappa}_{B_1}(B_2)$, one can use the following procedure from [104]. Let L_1 and R_1 denote the left and the right factor of B, respectively. If $\delta(R_1) = 1$, then the procedure stops with $\kappa := 1$, $B_1 := L_1$, and $B_2 := R_1$. Otherwise, i.e. if $\delta(R_1) > 1$, then we may define L_{i+1} and R_{i+1} inductively to be the left and the right factor of R_i, respectively, for $i = 1, 2, \ldots$ as long as the conditions $\delta(R_i) > 1$ and $L_i = L_{i+1}$ are satisfied. The number κ is the smallest positive integer for which $\delta(R_\kappa) = 1$ or $L_\kappa \neq L_{\kappa+1}$. Then, we define $B_1 := L_1 = \cdots = L_\kappa$ and $B_2 := R_\kappa$.

Let Λ^r_+ denote the set of $(E,\rho) \in \Lambda^r$ with $\delta(E) \geq 2$. For every $(E,\rho) \in \Lambda^r_+$ and every $i \in \{1,\ldots,m\}$, let $\Omega_{E,\rho,i}$ be a symmetric subset of $\mathbb{R} \setminus \{0\}$ of cardinality $2\delta_i(E)$ such that

FC1. for each fixed $(E,\rho) \in \Lambda^r_+$, the following holds:

(a) if $\delta(E) = 2$; i.e., $E = \{(X_{i_1}, X_{i_2})\}$ for some $i_1, i_2 \in \{1,\ldots,m\}$ with $i_1 < i_2$, then $\Omega_{E,\rho,i_1} = \Omega_{E,\rho,i_2}$;

(b) if $\delta(E) > 2$, then the sets $\Omega_{E,\rho,1},\ldots,\Omega_{E,\rho,m}$ are pairwise disjoint.

For every $(n,i) \in \{2,\ldots,r\} \times \{1,\ldots,m\}$, define $\Omega(n,i) := \bigcup_{E\in\mathcal{E}_n} \bigcup_{\rho=1}^{|E|} \Omega_{E,\rho,i}$, which is a finite symmetric subset of $\mathbb{R} \setminus \{0\}$. We require that

FC2. the sets $\Omega(n,i)$ with $(n,i) \in \{3,\ldots,r\} \times \{1,\ldots,m\}$ are pairwise disjoint.

For every $(E,\rho) \in \Lambda^r_+$, define $\Omega_{E,\rho} := \bigcup_{i=1}^{m} \Omega_{E,\rho,i}$, which is a finite symmetric subset of $\mathbb{R} \setminus \{0\}$. Note that $|\Omega_{E,\rho}| = 2$ for $\delta(E) = 2$ and that $|\Omega_{E,\rho}| = 2\delta(E)$ for $\delta(E) > 2$. We require that

FC3. for every $(E,\rho) \in \Lambda^r_+$ with $\delta(E) > 2$, the set $\Omega_{E,\rho}$ is SMC.

It is also required that the sets $\Omega_{E,\rho}$ with $(E,\rho) \in \Lambda^r_+$ are *independent with respect to* r in the following sense.

FC4. The sets $\Omega_{E,\rho}$ with $(E,\rho) \in \Lambda^r_+$ are pairwise disjoint and the following implication holds: if $\sum_{(E,\rho)\in\Lambda^r_+} \sum_{\omega\in\Omega_{E,\rho}} |a_\omega| \leq r$ and $\sum_{(E,\rho)\in\Lambda^r_+} \sum_{\omega\in\Omega_{E,\rho}} a_\omega \omega = 0$ for any integers a_ω, then $\sum_{\omega\in\Omega_{E,\rho}} a_\omega \omega = 0$ for every $(E,\rho) \in \Lambda^r_+$.

For every $E \in \bigcup_{n=1}^{r} \mathcal{E}_n$, every $B \in E$, and every $\rho \in \{1,\ldots,|E|\}$, we define real constants $\hat{\xi}_{B,\rho}$ as follows. If $\delta(E) = 1$; i.e., $E = \{B\}$ and $B = X_i$ for some $i \in \{1,\ldots,m\}$, then we let $\hat{\xi}_{B,\rho} := 1$. If $\delta(E) = 2$; i.e., $E = \{B\}$ and $B = (X_{i_1}, X_{i_2})$ for some $i_1, i_2 \in \{1,\ldots,m\}$ with $i_1 < i_2$, then there exists $\omega_E > 0$ such that $\Omega_{E,\rho} = \{\pm\omega_E\}$, and we let $\hat{\xi}_{B,\rho} := \frac{1}{\omega_E}$. Finally suppose that $\delta(E) > 2$. Because of FC3, there exists an MC subset $F_{E,\rho}$ of $\Omega_{E,\rho}$ of cardinality $|\Omega_{E,\rho}|/2 = \delta(E)$. Because of FC1, the intersections $F_{E,\rho} \cap \Omega_{E,\rho,1},\ldots,F_{E,\rho} \cap \Omega_{E,\rho,m}$ are pairwise disjoint subsets of $\mathbb{R} \setminus \{0\}$ with $|F_{E,\rho} \cap \Omega_{E,\rho,i}| = \delta_i(E)$ for every $i \in \{1,\ldots,m\}$ and their union coincides with $F_{E,\rho}$. Thus, for every $B \in E$, we can define

$$\hat{\xi}_{B,\rho} \; := \; g_B(F_{E,\rho} \cap \Omega_{E,\rho,1},\ldots,F_{E,\rho} \cap \Omega_{E,\rho,m}),$$

where the right-hand side is given by (4.20). We require that

FC5. for every $E \in \bigcup_{n=3}^{r} \mathcal{E}_n$, the square matrix $(\hat{\xi}_{B,\rho})_{B\in E, 1\leq\rho\leq|E|}$ is invertible[7].

Theorem 4.3 ([68]). *It is always possible to satisfy FC1-FC5.*

Let $\mathrm{i} = \sqrt{-1}$ denote the imaginary unit. For every $E \in \bigcup_{n=1}^{r} \mathcal{E}_n$, every $j > 0$, and every $y \in \mathcal{Y}$, we use the real-valued functions $\zeta^j_{E,\rho,k}(\cdot;y)$ from (4.18) to define complex-valued functions on the domain of y according to the subsequent choices CH1-CH3.

CH1. If $\delta(E) = 1$; i.e., $E = \{X_i\}$ for some $i \in \{1,\ldots,m\}$, then we define

$$\eta^j_{i,0}(t;y) \; := \; \zeta^j_{E,1,1}(t;y).$$

CH2. If $\delta(E) = 2$; i.e., $E = \{(X_{i_1}, X_{i_2})\}$ for some $i_1, i_2 \in \{1,\ldots,m\}$ with $i_1 < i_2$, then we define

$$\eta^j_{\omega_E,i_1}(t;y) \; := \; (-1)^{\delta(E)-1}\eta^j_{-\omega_E,i_1}(t;y) \; := \; \mathrm{i}^{\delta(E)-1}\, 2^{-\frac{1}{\delta(E)}}\, \zeta^j_{E,1,1}(t;y),$$

[7] A natural order of the rows of the matrix $(\hat{\xi}_{B,\rho})_{B\in E, 1\leq\rho\leq|E|}$ is given by the total order $\preceq$ on $\mathcal{B}$.

$$\eta^j_{\omega_E,i_2}(t;y) \;:=\; \eta^j_{-\omega_E,i_2}(t;y) \;:=\; 2^{-\frac{1}{\delta(E)}}\,\zeta^j_{E,1,2}(t;y).$$

CH3. If $\delta(E) > 2$, then, for every $\rho \in \{1,\ldots,|E|\}$, we let $\omega_{E,\rho,1},\ldots,\omega_{E,\rho,\delta(E)}$ denote the $\delta(E)$ elements of $F_{E,\rho}$ (in an arbitrary order), and define

$$\eta^j_{\omega_{E,\rho,1}}(t;y) \;:=\; (-1)^{\delta(E)-1}\eta^j_{-\omega_{E,\rho,1}}(t;y) \;:=\; \mathrm{i}^{\delta(E)-1}\,2^{-\frac{1}{\delta(E)}}\,\zeta^j_{E,\rho,1}(t;y),$$
$$\eta^j_{\omega_{E,\rho,k}}(t;y) \;:=\; \eta^j_{-\omega_{E,\rho,k}}(t;y) \;:=\; 2^{-\frac{1}{\delta(E)}}\,\zeta^j_{E,\rho,k}(t;y)$$

for every $k \in \{2,\ldots,\delta(E)\}$.

4.4.3 Control law and closed-loop system

For the moment, fix arbitrary $i \in \{1,\ldots,m\}$ and $j > 0$. Note that by CH2, for every $\omega \in \Omega(2,i)$, the complex conjugate of $\eta^j_{\omega,i}$ is given by $\eta^j_{-\omega,i}$. Moreover, by CH3, for every $n \in \{3,\ldots,r\}$ and every $\omega \in \Omega(n,i)$, the complex conjugate of η^j_{ω} is given by $\eta^j_{-\omega}$. Thus, for every $y \in \mathcal{Y}$, a purely real-valued function $u^j_i(\cdot;y)$ on the domain of y is given by

$$u^j_i(t;y) \;:=\; \eta^j_{i,0}(t;y) + j^{\frac{1}{2}} \sum_{\omega\in\Omega(2,i)} \eta^j_{\omega,i}(t;y)\,\mathrm{e}^{\mathrm{i}j\omega t} + \sum_{n=3}^{r} j^{\frac{n-1}{n}} \sum_{\omega\in\Omega(n,i)} \eta^j_{\omega}(t;y)\,\mathrm{e}^{\mathrm{i}j\omega t}, \tag{4.21}$$

where e denotes Euler's number. For the ith input channel u_i of (4.5), we propose the j-dependent time-varying output-feedback control law

$$u_i \;=\; u^j_i(t;y) \quad \text{with} \qquad y \;=\; \psi(x), \tag{4.22}$$

where $y \in \mathcal{Y}$ is the measured output signal (4.6).

It follows from standard existence and uniqueness theorems for ordinary differential equations that, for every *frequency parameter* $j > 0$ and every *initial state* $x_0 \in M$, there exists a unique *maximal solution* of the initial value problem

$$\dot{x} \;=\; \sum_{i=1}^{m} u^j_i(t;y)\,f_i(x), \qquad x(0) \;=\; x_0 \tag{4.23}$$

with $y = \psi(x)$. We denote this maximal solution by

$$\gamma^j_{x_0}\colon [0,T^j_{x_0}) \to M, \qquad t \mapsto \gamma^j_{x_0}(t), \tag{4.24}$$

where the maximum time of existence $T^j_{x_0}$ is either a positive real number and otherwise $+\infty$. The output signal associated with (4.24) is denoted by

$$y^j_{x_0}\colon [0,T^j_{x_0}) \to \mathbb{R}, \qquad t \mapsto \psi(\gamma^j_{x_0}(t)), \tag{4.25}$$

which is an element of $\mathcal{Y}$.

4.5 Main results

For the rest of the chapter, we fix constants and functions as described in Section 4.4. The only control parameter that remains variable is the frequency parameter j. For every formal bracket $B \in \mathrm{M}(\mathbf{X})$, let $[f_B]$ denote the smooth vector field on M that originates from B by replacing the round brackets "(", ")" by square brackets "[", "]" and by "plugging in" the f_i for the X_i. For instance, if $B = (X_1,(X_1,X_2))$, then $[f_B] = [f_1,[f_1,f_2]]$. An immediate consequence of Theorem 4.2 is the following statement.

Corollary 4.4. *The vector fields* $[f_B]$ *with* $B \in \bigcup_{n=1}^{r} \mathcal{B}_n$ *span* $\mathrm{L}^r(\mathbf{f})$.

In the notation of Section 4.4, for every $\ell = (E, \rho) \in \Lambda^r$, define $g_\ell \in \mathfrak{X}(M)$ by

$$g_\ell(x) \;:=\; \sum_{B \in E} \hat{\xi}_{B,\rho}\,[f_B](x). \tag{4.26}$$

Since, for each $E \in \bigcup_{n=1}^{r} \mathcal{E}_n$, the matrix $(\hat{\xi}_{B,\rho})_{B\in E, 1\leq\rho\leq|E|}$ is assumed to be invertible, we can also state Corollary 4.4 as follows.

Corollary 4.5. *The vector fields* g_ℓ *with* $\ell \in \Lambda^r$ *span* $\mathrm{L}^r(\mathbf{f})$.

Our first main result states that the trajectories of (4.5) under control law (4.22) approximate the trajectories of (4.7) with the g_ℓ given by (4.26). As an abbreviation, let f^∞ denote the vector field on the right-hand side of (4.7); i.e.,

$$f^\infty(x) \;:=\; -\sum_{\ell\in\Lambda^r} (g_\ell\psi)(x)\, g_\ell(x) \tag{4.27}$$

for every $x \in M$. Using the notation (4.24) for the maximal solutions of (4.23), the following holds.

Theorem 4.6. *Let* $\varphi \in C^\infty(M)$ *and* $\varepsilon_1, \varepsilon_2 > 0$. *Suppose that* $f_1, \ldots, f_m$ *are compactly supported.*[8] *Then, there exists* $j_0 > 0$ *such that, for every* $j \geq j_0$, *every* $x_0 \in M$, *and all* $t_2 \geq t_1 \geq 0$, *we have*

$$\left|\varphi(\gamma^j_{x_0}(t_2)) - \varphi(\gamma^j_{x_0}(t_1)) - \int_{t_1}^{t_2} (f^\infty\varphi)(\gamma^j_{x_0}(t))\right| \;\leq\; \varepsilon_1 + \varepsilon_2(t_2 - t_1).$$

The proof of Theorem 4.6 is given in Sections 4.6 and 4.7. A consequence of Theorem 4.6 is the subsequent Theorem 4.7 that addresses stability properties of the closed-loop system. Since we are interested in minimizing the output value, we do not state the result in terms of a distance function on the state manifold but in terms of the sublevel sets of the output function (using the notation of Sections 1.5 and 4.2).

Theorem 4.7. *Suppose that* ψ *attains a local minimum value* $y^* \in \mathbb{R}$ *at some point* $x^* \in M$. *Assume that there exists* y^+ *with* $y^* < y^+ \leq +\infty$ *such that, for every* $x \in \psi^{-1}(\leq y^+, x^*)$ *with* $x \neq x^*$, *we have* $(g\psi)(x) \neq 0$ *for some* $g \in \mathrm{L}^r(\mathbf{f})$. *Moreover, assume that* $\psi^{-1}(\leq \tilde{y}, x^*)$ *is compact for every* $\tilde{y} < y^+$. *Then, the point* x^* *is* practically asymptotically stable for (4.5) under (4.22), *meaning that, for all* $\varepsilon, \delta > 0$ *and every* $\tilde{y} \in (y^*, y^+)$, *there exist* $j_0, \sigma > 0$ *such that, for every* $j \geq j_0$ *and every* $x_0 \in \psi^{-1}(\leq \tilde{y}, x^*)$, *we have*

- $\gamma^j_{x_0}(t) \in \psi^{-1}(\leq \psi(x_0) + \varepsilon, x^*)$ *for every* $t \geq 0$ (stability and boundedness),
- $\gamma^j_{x_0}(t) \in \psi^{-1}(\leq y^* + \delta, x^*)$ *for every* $t \geq \sigma$ (attraction).

The proof is given in Subsection 4.7.4.

Remark 4.8. Suppose that ψ attains a local minimum value at $x^* \in M$. From Theorem 4.7 we can conclude that x^* is (locally) practically asymptotically stable for (4.5) under (4.22) if the following condition is satisfied: for each $x \neq x^*$ in a neighborhood of x^*, we have

[8] A vector field on M is said to be *compactly supported* if it vanishes outside a compact subset of M.

$(g\psi)(x) \neq 0$ for some $g \in \mathrm{L}^r(\mathbf{f})$. A similar statement can be found in [95] for a closely related control law under the above condition for $r = 1$. In general, however, the above condition is not satisfied for $r = 1$ if (4.5) is nonholonomic; for instance, in the situation of Example 4.10. In this case, Lie brackets of higher order (i.e. $r > 1$) are needed to get access to descent directions of ψ. ◊

A local version[9] of Theorem 4.7 can be given as follows.

Corollary 4.9. *Assume that*

1. *the function ψ attains a local minimum value y^* at some point x^* of M,*
2. *the second derivative of ψ at x^* is positive definite,*
3. *the vectors $g(x^*)$ with $g \in \mathrm{L}^r(\mathbf{f})$ span the tangent space to M at x^*.*

Then, there exists $y^+ > y^*$ such that the conclusions of Theorem 4.7 hold.

Example 4.10. As a toy example, we consider Brockett's integrator from [15]; i.e.,

$$\dot{x}_1 \;=\; u_1, \qquad \dot{x}_2 \;=\; u_2, \qquad \dot{x}_3 \;=\; u_1\,x_2 - u_2\,x_1 \tag{4.28}$$

on $M := \mathbb{R}^3$. Note that (4.28) is of the form (4.5) if we define $m := 2$ control vector fields f_1 and f_2 by $f_1(x) := (1, 0, x_2)^\top$ and $f_2(x) := (0, 1, -x_1)^\top$, respectively. Since we have $[f_1, f_2](x) = (0, 0, -2)^\top$ for every $x \in \mathbb{R}^3$, the Lie brackets of degree $\leq r := 2$ span $\mathbb{R}^3$. Let $x^* \in \mathbb{R}^3$ and $y^* \in \mathbb{R}$. Suppose that the output function ψ is given by

$$\psi(x) \;:=\; y^* + \frac{1}{2}\|x - x^*\|^2, \tag{4.29}$$

where $\|\cdot\|$ denotes the Euclidean norm. Clearly, ψ has no other critical point than x^* and all sublevel sets of ψ are connected and compact. Therefore, the assumptions of Theorem 4.7 are satisfied with $y^+ := +\infty$. In the terminology of Theorem 4.7, we conclude that the optimal point x^* is practically asymptotically stable for (4.28) under (4.22).

To generate numerical data, we choose the following control parameters. For the slow oscillations we choose the frequency coefficients $\bar{\omega}_{(\{X_1\},1)} := 1$, $\bar{\omega}_{(\{X_2\},1)} := 2$, $\bar{\omega}_{(\{(X_1,X_2)\},1)} := 3$. The only frequency coefficient for the fast oscillations is chosen as $\omega_{\{(X_1,X_2)\}} := 1$. For the simulation, we choose the optimal value $y^* := 0$, the optimal point $x^* := (1, 1, 1)$, the frequency parameter $j := 10^3$, the sampling time $2\pi/j$, and the initial state $x_0 := -x^*$. The result is shown in Figure 4.1. ◊

4.6 Averaging of the fast oscillations

4.6.1 Iterated Lie derivatives and Lie brackets

Recall from Section 4.2 that $\mathrm{M}(\mathbf{X})$ denotes the free magma generated by the set of indeterminates $\mathbf{X} = \{X_1, \ldots, X_m\}$ over $\mathbb{R}$. As in Theorem 4.2, we denote by μ the canonical map from $\mathrm{M}(\mathbf{X})$ into free Lie algebra $\mathrm{L}(\mathbf{X})$ generated by $\mathbf{X}$ over $\mathbb{R}$. The surrounding free non-unital associative algebra generated by $\mathbf{X}$ is denoted by $\mathrm{A}_0(\mathbf{X})$. On the other hand,

[9] By *local version*, we mean sufficient conditions to ensure the existence of some (possibly small) sublevel $y^+ > y^*$ as in Theorem 4.7.

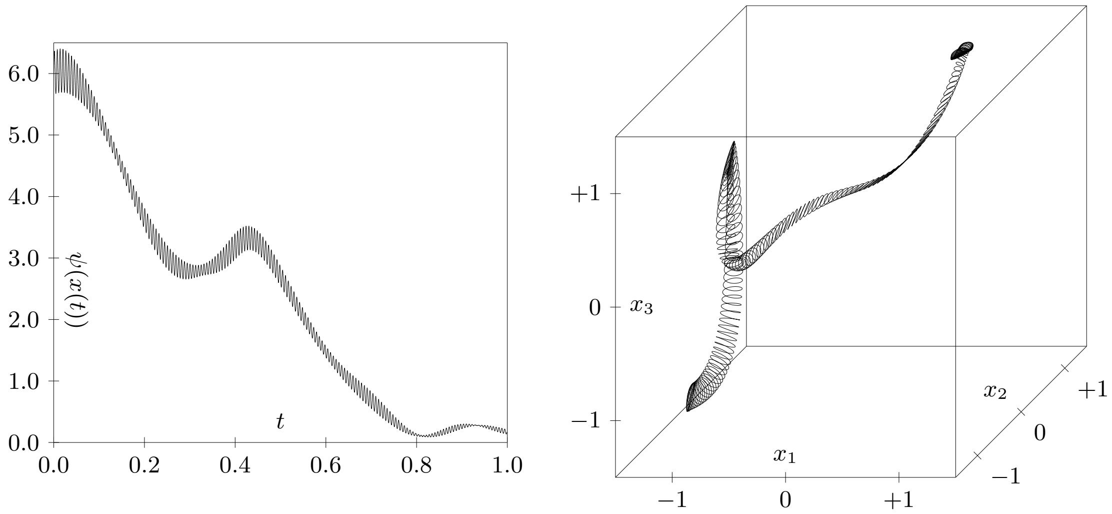


Figure 4.1: *Numerical simulations for the particular situation in Example 4.10. Left: decay of the output* $y(t) = \psi(x(t))$ *as a function of the time parameter* t*. Right: image of the trajectory* $t \mapsto x(t)$ *on the time interval* $[0, 1]$*.*

also the set $\mathfrak{X}(M)$ of smooth vector fields on M is a non-unital associative algebra if we consider each $f \in \mathfrak{X}(M)$ as the differential operator on the algebra $C^\infty(M)$ of smooth functions on M that assigns to each $\varphi \in C^\infty(M)$ to the Lie derivative $f\varphi \in C^\infty(M)$ of φ along f. For every $\varphi \in C^\infty(M)$ and every multi-index $I = (i_1, \ldots, i_k)$ of length $|I| := k > 0$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$, we use the notation

$$f_I\varphi \;:=\; f_{i_1} \cdots f_{i_k}\varphi \;:=\; f_{i_1}(\cdots(f_{i_k}\varphi)\cdots) \tag{4.30}$$

for the iterated Lie derivative of φ along the control vector fields $f_{i_1}, \ldots, f_{i_k}$. Since $\mathrm{A}_0(\mathbf{X})$ has the universal property, there exists a unique algebra homomorphism $\mathrm{Ev}\colon \mathrm{A}_0(\mathbf{X}) \to \mathfrak{X}(M)$, called *evaluation map*, such that $\mathrm{Ev}(X_i) = f_i$ for every $i \in \{1, \ldots, m\}$. In other words, for every $v \in \mathrm{A}_0(\mathbf{X})$, the differential operator $\mathrm{Ev}(v)$ is obtained from v by "plugging in the f_i for the X_i". For example, if $\varphi \in C^\infty(M)$ and if $v_I \in \mathbb{R}$ for every multi-index $I = (i_1, \ldots, i_k)$ of length $|I| = k \in \{1, \ldots, r\}$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$, then we have

$$\left(\mathrm{Ev}\left(\sum_{0<|I|\leq r} v_I X_I\right)\right)\varphi \;=\; \sum_{0<|I|\leq r} v_I\,(f_I\varphi). \tag{4.31}$$

Recall from Section 4.5 that, for every formal bracket $B \in \mathrm{M}(\mathbf{X})$, we denote by $[f_B]$ the (iterated) Lie bracket that originates from B by replacing the round brackets "(", ")" by square brackets "[", "]" and by "plugging in" the f_i for the X_i. By composing the canonical map $\mu\colon \mathrm{M}(\mathbf{X}) \to \mathrm{L}(\mathbf{X}) \subset \mathrm{A}_0(\mathbf{X})$ and the evaluation map $\mathrm{Ev}\colon \mathrm{A}_0(\mathbf{X}) \to \mathfrak{X}(M)$, we can write

$$[f_B] \;=\; (\mathrm{Ev} \circ \mu)(B) \tag{4.32}$$

for every $B \in \mathrm{M}(\mathbf{X})$.

Since the control vector fields $f_1, \ldots, f_m$ are assumed to be smooth, we can give estimates for iterated Lie derivatives as follows.

Lemma 4.11. *Suppose that $f_1, \ldots, f_m$ are compactly supported. Then, for every $\varphi \in C^\infty(M)$ and every multi-index $I = (i_1, \ldots, i_k)$ of length $k > 0$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$, there exists $c > 0$ such that*

$$|(f_I\varphi)(x)| \;\leq\; c$$

for every $x \in M$.

Since the vector fields g_ℓ, $\ell \in \Lambda^r$, in (4.26) are $\mathbb{R}$-linear combinations of Lie brackets of the vector fields $f_1, \ldots, f_m$, we immediately conclude the following estimates.

Lemma 4.12. *Suppose that $f_1, \ldots, f_m$ are compactly supported. Then, for every $\varphi \in C^\infty(M)$, there exists $c > 0$ such that*

$$|(g_\ell\varphi)(x)| \;\leq\; c$$

for every $\ell \in \Lambda^r$ and every $x \in M$.

4.6.2 Averaging of the sinusoids

Let $y\colon [0,T) \to \mathbb{R}$ be an element of $\mathcal{Y}$. For every $j > 0$, the first-order hold $\bar{y}^j$ of y is piecewise continuously differentiable. It is clear that we can extend its derivative to a function on $[0,T)$ by taking the right-handed limit of the difference quotient at every $\tau_k^j \in [0,T)$ with $k \in \{0,1,\ldots\}$. The resultant piecewise constant derivative of $\bar{y}^j$ on $[0,T)$ is denoted by $\dot{\bar{y}}^j$. We also use the notation $\mathrm{dom}(y)$ for the domain $[0,T)$ of y.

In this subsection, let $j > 0$ and $y \in \mathcal{Y}$. For every multi-index $I = (i_1, \ldots, i_k)$ of length $k \in \{1, \ldots, r\}$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$, we define certain functions $v_I^j(\cdot;y)$, $\widetilde{UV}_I^j(\cdot;y)$, $r_I^j(\cdot;y)$ on $\mathrm{dom}(y)$ that will appear again in Subsection 4.6.3 in the integral expansion for the solutions of (4.5) under (4.22). As in [68], the functions will be chosen in such a way that they satisfy

$$\begin{aligned}
\Big[\widetilde{UV}_{i_1}^j(t;y)\Big]_{t=t_1}^{t=t_2} \;&=\; \int_{t_1}^{t_2} \Big(v_{i_1}^j(t;y)\,\mathrm{d}t + r_{i_1}^j(t;y)\,\mathrm{d}t - u_{i_1}^j(t;y)\Big)\mathrm{d}t,\\
\Big[\widetilde{UV}_{i_1,\ldots,i_k}^j(t;y)\Big]_{t=t_1}^{t=t_2} \;&=\; \int_{t_1}^{t_2} \Big(v_{i_1,\ldots,i_k}^j(t;y) + r_{i_1,\ldots,i_k}^j(t;y) - u_{i_1}^j(t;y)\,\widetilde{UV}_{i_2,\ldots,i_k}^j(t;y)\Big)\mathrm{d}t
\end{aligned}$$

for all $t_1, t_2 \in \mathrm{dom}(y)$. The definitions coincide in large parts with the ones in [68] up to some slight modifications of the $\widetilde{UV}_I^j(\cdot;y)$ and $r_I^j(\cdot;y)$ in order to derive suitable estimates at the end in Lemma 4.13. We begin with the definitions for indices I of length $k = 1$.

Let $i \in \{1, \ldots, m\}$. We write (4.21) as

$$u_i^j(t;y) \;=\; v_i^j(t;y) - \widetilde{Uv}_i^j(t;y),$$

where

$$v_i^j(t;y) \;:=\; \eta_{i,0}^j(t;y), \tag{4.33}$$

$$\widetilde{Uv}_i^j(t;y) \;:=\; -j^{\frac{1}{2}} \sum_{\omega\in\Omega(2,i)} \eta_{\omega,i}^j(t;y)\,\mathrm{e}^{\mathrm{i}j\omega t} - \sum_{n=3}^{r} j^{\frac{n-1}{n}} \sum_{\omega\in\Omega(n,i)} \eta_\omega^j(t;y)\,\mathrm{e}^{\mathrm{i}j\omega t}. \tag{4.34}$$

Since $\omega \neq 0$ for all of the terms in (4.34), integration by parts leads to

$$\int_{t_1}^{t_2} \widetilde{Uv}_i^j(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_i^j(t;y)\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \widetilde{uV}_i^j(t;y)\,\mathrm{d}t,$$

where

$$\widetilde{UV}_i^j(t;y) \;:=\; -j^{-\frac{1}{2}} \sum_{\omega\in\Omega(2,i)} \frac{\eta_{\omega,i}^j(t;y)}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}j\omega t} - \sum_{n=3}^{r} j^{-\frac{1}{n}} \sum_{\omega\in\Omega(n,i)} \frac{\eta_{\omega}^j(t;y)}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}j\omega t}, \tag{4.35}$$

$$\widetilde{uV}_i^j(t;y) \;:=\; -j^{-\frac{1}{2}} \sum_{\omega\in\Omega(2,i)} \frac{\dot\eta_{\omega,i}^j(t;y)}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}j\omega t} - \sum_{n=3}^{r} j^{-\frac{1}{n}} \sum_{\omega\in\Omega(n,i)} \frac{\dot\eta_{\omega}^j(t;y)}{\mathrm{i}\omega}\,\mathrm{e}^{\mathrm{i}j\omega t}. \tag{4.36}$$

In (4.36), the symbols $\dot\eta_{\omega,i}^j(t;y)$ and $\dot\eta_{\omega}^j(t;y)$ denote the derivatives of the functions $\eta_{\omega,i}^j(\cdot;y)$ and $\eta_{\omega}^j(\cdot;y)$ at t, respectively. Finally, we let

$$r_i^j(t;y) \;:=\; \widetilde{uV}_i^j(t;y), \tag{4.37}$$

$$\widetilde{uv}_i^j(t;y) \;:=\; \widetilde{uV}_i^j(t;y) + \widetilde{Uv}_i^j(t;y). \tag{4.38}$$

Then we have

$$u_i^j(t;y) \;=\; v_i^j(t;y) + r_i^j(t;y) - \widetilde{uv}_i^j(t;y), \tag{4.39}$$

$$\int_{t_1}^{t_2} \widetilde{uv}_i^j(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}_i^j(t;y)\Big]_{t=t_1}^{t=t_2}. \tag{4.40}$$

This completes the definitions for a single index i.

Now in order to show the idea how the general v_I^j, r_I^j, and $\widetilde{UV}_I^j$ are defined, let us proceed one more step and construct v_{i_1,i_2}^j and $\widetilde{UV}_{i_1,i_2}^j$ explicitly for two indices $i_1, i_2 \in \{1,\dots,m\}$. When we multiply $u_{i_1}^j(t;y)$ by $\widetilde{UV}_{i_2}^j(t;y)$, we get

$$u_{i_1}^j(t;y)\,\widetilde{UV}_{i_2}^j(t;y) \;=\; \eta_{i_1,0}^j(t;y)\,\widetilde{UV}_{i_2}^j(t;y) + b_{i_1,i_2}^j + c_{i_1,i_2}^j,$$

where

$$b_{i_1,i_2}^j \;:=\; -\sum_{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(2,i_2)} \frac{(\eta_{\omega_1,i_1}^j\eta_{\omega_2,i_2}^j)(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t},$$

$$\begin{aligned} c_{i_1,i_2}^j \;:=\; & -\sum_{n=3}^{r} j^{1-\frac{1}{2}-\frac{1}{n}} \sum_{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(n,i_2)} \frac{(\eta_{\omega_1,i_1}^j\eta_{\omega_2}^j)(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\ & -\sum_{n=3}^{r} j^{1-\frac{1}{n}-\frac{1}{2}} \sum_{(\omega_1,\omega_2)\in\Omega(n,i_1)\times\Omega(2,i_2)} \frac{(\eta_{\omega_1}^j\eta_{\omega_2,i_2}^j)(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\ & -\sum_{n_1,n_2=3}^{r} j^{1-\frac{1}{n_1}-\frac{1}{n_2}} \sum_{(\omega_1,\omega_2)\in\Omega(n_1,i_1)\times\Omega(n_2,i_2)} \frac{(\eta_{\omega_1}^j\eta_{\omega_2}^j)(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t}. \end{aligned}$$

The terms in b_{i_1,i_2}^j that correspond to $\omega_1+\omega_2=0$ are denoted by

$$v_{i_1,i_2}^j(t;y) \;:=\; -\sum_{\substack{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(2,i_2)\\ \omega_1+\omega_2=0}} \frac{(\eta_{\omega_1,i_1}^j\eta_{\omega_2,i_2}^j)(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t}. \tag{4.41}$$

Note that, for the first two summations in the right-hand side of $c^j_{i_1,i_2}$, we always have $\omega_1 + \omega_2 \neq 0$. The third summation in the right-hand side of $c^j_{i_1,i_2}$ can contain terms with $\omega_1 + \omega_2 = 0$. However, we will show that they add up to 0, that is

$$\sum_{n_1,n_2=3}^{r} j^{1-\frac{1}{n_1}-\frac{1}{n_2}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(n_1,i_1)\times\Omega(n_2,i_2)\\ \omega_1+\omega_2=0}} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2})(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \;=\; 0.$$

To see this, we use the following symmetry argument: If $\frac{\eta^j_{\omega_1}\eta^j_{-\omega_1}}{\mathrm{i}\omega_1} = \frac{|\eta^j_{\omega_1}|^2}{\mathrm{i}\omega_1}$ is in the summation, then, by the symmetry of the sets $\Omega(n,i)$, also $\frac{\eta^j_{-\omega_1}\eta^j_{\omega_1}}{-\mathrm{i}\omega_1} = -\frac{|\eta^j_{\omega_1}|^2}{\mathrm{i}\omega_1}$ is in the summation. So they add up to 0. Thus, we have

$$b^j_{i_1,i_2} + c^j_{i_1,i_2} \;=\; v^j_{i_1,i_2}(t;y) - \widetilde{Uv}^j_{i_1,i_2}(t;y),$$

where

$$\begin{aligned}
\widetilde{Uv}^j_{i_1,i_2}(t;y) \;:=\; & \sum_{n=3}^{r} j^{1-\frac{1}{2}-\frac{1}{n}} \sum_{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(n,i_2)} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2})(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \sum_{n=3}^{r} j^{1-\frac{1}{n}-\frac{1}{2}} \sum_{(\omega_1,\omega_2)\in\Omega(n,i_1)\times\Omega(2,i_2)} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2,i_2})(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \sum_{n_1,n_2=3}^{r} j^{1-\frac{1}{n_1}-\frac{1}{n_2}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(n_1,i_1)\times\Omega(n_2,i_2)\\ \omega_1+\omega_2\neq 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2})(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \widetilde{Uw}^j_{i_1,i_2}(t;y)
\end{aligned} \tag{4.42}$$

with the additional contribution

$$\widetilde{Uw}^j_{i_1,i_2}(t;y) \;:=\; \sum_{\substack{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(2,i_2)\\ \omega_1+\omega_2\neq 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2,i_2})(t;y)}{\mathrm{i}\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \tag{4.43}$$

compared to [68]. Since $\omega_1 + \omega_2 \neq 0$ for all of the terms in (4.42) and (4.43), integration by parts leads to

$$\int_{t_1}^{t_2} \widetilde{Uv}^j_{i_1,i_2}(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}^j_{i_1,i_2}(t;y)\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \widetilde{uV}^j_{i_1,i_2}(t;y)\,\mathrm{d}t,$$

where

$$\begin{aligned}
\widetilde{UV}^j_{i_1,i_2}(t;y) \;:=\; & \sum_{n=3}^{r} j^{-\frac{1}{2}-\frac{1}{n}} \sum_{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(n,i_2)} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2})(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \sum_{n=3}^{r} j^{-\frac{1}{n}-\frac{1}{2}} \sum_{(\omega_1,\omega_2)\in\Omega(n,i_1)\times\Omega(2,i_2)} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2,i_2})(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \sum_{n_1,n_2=3}^{r} j^{-\frac{1}{n_1}-\frac{1}{n_2}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(n_1,i_1)\times\Omega(n_2,i_2)\\ \omega_1+\omega_2\neq 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2})(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t}
\end{aligned}$$

$$+ \widetilde{UW}^j_{i_1,i_2}(t;y), \tag{4.44}$$

$$\begin{aligned}
\widetilde{uV}^j_{i_1,i_2}(t;y) \;:=\; & \sum_{n=3}^{r} j^{-\frac{1}{2}-\frac{1}{n}} \sum_{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(n,i_2)} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2})^{\cdot}(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \sum_{n=3}^{r} j^{-\frac{1}{n}-\frac{1}{2}} \sum_{(\omega_1,\omega_2)\in\Omega(n,i_1)\times\Omega(2,i_2)} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2,i_2})^{\cdot}(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \sum_{n_1,n_2=3}^{r} j^{-\frac{1}{n_1}-\frac{1}{n_2}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(n_1,i_1)\times\Omega(n_2,i_2)\\ \omega_1+\omega_2\neq 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\omega_2})^{\cdot}(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \\
& + \widetilde{uW}^j_{i_1,i_2}(t;y)
\end{aligned} \tag{4.45}$$

with the additional contributions

$$\widetilde{UW}^j_{i_1,i_2}(t;y) \;:=\; j^{-1} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(2,i_2)\\ \omega_1+\omega_2\neq 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2,i_2})(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t}, \tag{4.46}$$

$$\widetilde{uW}^j_{i_1,i_2}(t;y) \;:=\; j^{-1} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(2,i_1)\times\Omega(2,i_2)\\ \omega_1+\omega_2\neq 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2,i_2})^{\cdot}(t;y)}{\mathrm{i}^2(\omega_1+\omega_2)\omega_2}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\omega_2)t} \tag{4.47}$$

compared to [68]. In the above definitions, the expression $(\eta^j_{\omega_1,i_1}\eta^j_{\omega_2})^{\cdot}(t;y)$ denotes the derivative of the product of $\eta^j_{\omega_1,i_1}(\cdot;y)$ and $\eta^j_{\omega_2}(\cdot;y)$ at t. The expressions $(\eta^j_{\omega_1}\eta^j_{\omega_2,i_2})^{\cdot}(t;y)$ and $(\eta^j_{\omega_1}\eta^j_{\omega_2})^{\cdot}(t;y)$ are defined correspondingly. Finally, we let

$$r^j_{i_1,i_2}(t;y) \;:=\; \eta^j_{i_1,0}(t;y)\,\widetilde{UV}^j_{i_2}(t;y) + \widetilde{uV}^j_{i_1,i_2}(t;y), \tag{4.48}$$

$$\widetilde{uv}^j_{i_1,i_2}(t;y) \;:=\; \widetilde{uV}^j_{i_1,i_2}(t;y) + \widetilde{Uv}^j_{i_1,i_2}(t;y). \tag{4.49}$$

Then, we have

$$u^j_{i_1}(t;y)\,\widetilde{UV}^j_{i_2}(t;y) \;=\; v^j_{i_1,i_2}(t;y) + r^j_{i_1,i_2}(t;y) - \widetilde{uv}^j_{i_1,i_2}(t;y), \tag{4.50}$$

$$\int_{t_1}^{t_2} \widetilde{uv}^j_{i_1,i_2}(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}^j_{i_1,i_2}(t;y)\Big]_{t=t_1}^{t=t_2}. \tag{4.51}$$

This completes the definitions for two indices i_1, i_2.

To the state the definitions for multi-indices of length > 2, we introduce the following notation from [68] for each $k \in \{2,\dots,r\}$.

(1) For every $\hat{n} = (n_1,\dots,n_k) \in \{2,\dots,r\}^k$ and every $\hat{\omega} = (\omega_1,\dots,\omega_k) \in \mathbb{R}^k$, we write

$$\begin{aligned}
\alpha_{\hat{n}} \;&:=\; \frac{1}{n_1} + \cdots + \frac{1}{n_k}, \\
\Sigma(\hat{\omega}) \;&:=\; \omega_1 + \cdots + \omega_k, \\
\Pi(\hat{\omega}) \;&:=\; (\omega_1 + \cdots + \omega_k)\cdots(\omega_{k-1}+\omega_k)\,\omega_k.
\end{aligned}$$

(2) We define the sets

$$\begin{aligned}
\Omega_1(k) \;&:=\; \big\{\hat{n} \in \{2,\dots,r\}^k \mid \alpha_{\hat{n}} < 1 \text{ and precisely one entry of } \hat{n} \text{ is equal to } 2\big\}, \\
\Omega_2(k) \;&:=\; \big\{\hat{n} \in \{3,\dots,r\}^k \mid \alpha_{\hat{n}} < 1\big\}.
\end{aligned}$$

(3) For every $\hat{n} = (n_1, \ldots, n_k) \in \{2, \ldots, r\}^k$, every $I = (i_1, \ldots, i_k) \in \{1, \ldots, m\}^k$, let

$$\Omega(\hat{n}, I) \;:=\; \big\{(\omega_1, \ldots, \omega_k) \in \Omega(n_1, i_1) \times \cdots \times \Omega(n_k, i_k) \;\big|\; \Pi(\hat{\omega}) \neq 0\big\}.$$

(4) For every $\hat{n} = (n_1, \ldots, n_k) \in \Omega_1(k)$, every $I = (i_1, \ldots, i_k) \in \{1, \ldots, m\}^k$, and every $\hat{\omega} = (\omega_1, \ldots, \omega_k) \in \Omega(\hat{n}, I)$, we write

$$\eta^j_{\hat{\omega}}(t; y) \;:=\; \eta^j_{\omega_1}(t; y) \cdots \eta^j_{\omega_{\tau-1}}(t; y)\, \eta^j_{\omega_\tau, i_\tau}(t; y)\, \eta^j_{\omega_{\tau+1}}(t; y) \cdots \eta^j_{\omega_k}(t; y),$$

where τ is the unique element of $\{1, \ldots, k\}$ for which $n_\tau = 2$. For every $\hat{n} \in \Omega_2(k)$, every $I \in \{1, \ldots, m\}^k$, and every $\hat{\omega} = (\omega_1, \ldots, \omega_k) \in \Omega(\hat{n}, I)$, we write

$$\eta^j_{\hat{\omega}}(t; y) \;:=\; \eta^j_{\omega_1}(t; y) \cdots \eta^j_{\omega_k}(t; y).$$

(5) If $k > 2$, then, for every multi-index $I = (i_1, \ldots, i_k)$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$, define

$$\widetilde{UV}^j_I(t; y) \;:=\; (-1)^k \sum_{\hat{n} \in \Omega_1(k)} j^{-\alpha_{\hat{n}}} \sum_{\hat{\omega} \in \Omega(\hat{n}, I)} \frac{\eta^j_{\hat{\omega}}(t; y)}{\mathrm{i}^k\, \Pi(\hat{\omega})}\, \mathrm{e}^{\mathrm{i} j \Sigma(\hat{\omega}) t} \tag{4.52a}$$

$$+ (-1)^k \sum_{\hat{n} \in \Omega_2(k)} j^{-\alpha_{\hat{n}}} \sum_{\hat{\omega} \in \Omega(\hat{n}, I)} \frac{\eta^j_{\hat{\omega}}(t; y)}{\mathrm{i}^k\, \Pi(\hat{\omega})}\, \mathrm{e}^{\mathrm{i} j \Sigma(\hat{\omega}) t} + \widetilde{UW}^j_I(t; y) \tag{4.52b}$$

with an additional contribution $\widetilde{UW}^j_I(t; y)$ (compared to [68]) that will be specified later in (4.60).

Note that $\widetilde{UV}^j_I(t; y)$ in (4.52) is well-defined since the denominators $\Pi(\hat{\omega})$ are nonzero by definition of the sets $\Omega(\hat{n}, I)$. Moreover, for $k = 2$, we have already defined $\widetilde{UV}^j_{i_1, i_2}(t; y)$ of the form (4.52) in (4.44) with the additional contribution $\widetilde{UW}^j_{i_1, i_2}(t; y)$ given by (4.46).

We proceed with the definitions by induction on the length k of multi-indices. The base case $k = 2$ was already carried out. For the induction step, let $k \in \{3, \ldots, r\}$ and $I = (i_1, \ldots, i_k)$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$. Define the multi-index $\bar{I} := (i_2, \ldots, i_k)$ of length $k - 1$. When we multiply $u^j_{i_1}(t; y)$ by $\widetilde{UV}^j_{\bar{I}}(t; y)$ we get

$$\begin{aligned} u^j_{i_1}(t; y)\, \widetilde{UV}^j_{\bar{I}}(t; y) \;&=\; \eta^j_{i_1, 0}(t; y)\, \widetilde{UV}^j_{\bar{I}}(t; y) + a^j_I + b^j_I + c^j_I + d^j_I \\ &\quad + \big(u^j_{i_1}(t; y) - \eta^j_{i_1, 0}(t; y)\big)\, \widetilde{UW}^j_{\bar{I}}(t; y), \end{aligned}$$

where

$$a^j_I \;:=\; (-1)^{k-1} \sum_{\hat{n} \in \Omega_1(k-1)} j^{1 - \frac{1}{2} - \alpha_{\hat{n}}} \sum_{(\omega_1, \hat{\omega}) \in \Omega(2, i_1) \times \Omega(\hat{n}, \bar{I})} \frac{(\eta^j_{\omega_1, i_1} \eta^j_{\hat{\omega}})(t; y)}{\mathrm{i}^{k-1}\, \Pi(\hat{\omega})}\, \mathrm{e}^{\mathrm{i} j (\omega_1 + \Sigma(\hat{\omega})) t}, \tag{4.53}$$

$$b^j_I \;:=\; (-1)^{k-1} \sum_{\hat{n} \in \Omega_2(k-1)} j^{1 - \frac{1}{2} - \alpha_{\hat{n}}} \sum_{(\omega_1, \hat{\omega}) \in \Omega(2, i_1) \times \Omega(\hat{n}, \bar{I})} \frac{(\eta^j_{\omega_1, i_1} \eta^j_{\hat{\omega}})(t; y)}{\mathrm{i}^{k-1}\, \Pi(\hat{\omega})}\, \mathrm{e}^{\mathrm{i} j (\omega_1 + \Sigma(\hat{\omega})) t},$$

$$c^j_I \;:=\; (-1)^{k-1} \sum_{n_1 = 3}^{r} \sum_{\hat{n} \in \Omega_1(k-1)} j^{1 - \frac{1}{n_1} - \alpha_{\hat{n}}} \sum_{(\omega_1, \hat{\omega}) \in \Omega(n_1, i_1) \times \Omega(\hat{n}, \bar{I})} \frac{(\eta^j_{\omega_1} \eta^j_{\hat{\omega}})(t; y)}{\mathrm{i}^{k-1}\, \Pi(\hat{\omega})}\, \mathrm{e}^{\mathrm{i} j (\omega_1 + \Sigma(\hat{\omega})) t},$$

$$d^j_I \;:=\; (-1)^{k-1} \sum_{n_1 = 3}^{r} \sum_{\hat{n} \in \Omega_2(k-1)} j^{1 - \frac{1}{n_1} - \alpha_{\hat{n}}} \sum_{(\omega_1, \hat{\omega}) \in \Omega(n_1, i_1) \times \Omega(\hat{n}, \bar{I})} \frac{(\eta^j_{\omega_1} \eta^j_{\hat{\omega}})(t; y)}{\mathrm{i}^{k-1}\, \Pi(\hat{\omega})}\, \mathrm{e}^{\mathrm{i} j (\omega_1 + \Sigma(\hat{\omega})) t}.$$

The definition of the set $\Omega_1(k-1)$ implies that

$$1 - \frac{1}{2} - \alpha_{\hat{n}} \;\le\; -\frac{1}{r} \tag{4.54}$$

for every $\hat{n} \in \Omega_1(k-1)$. In particular, all j-powers in a_I^j are negative. Because of the independence of the sets $\Omega_{E,\rho}$ in FC4, for all the terms in b_I^j and c_I^j, we know that $\omega_1 + \Sigma(\hat{\omega}) \neq 0$. The terms in d_I^j that correspond to $\omega_1 + \Sigma(\hat{\omega}) = 0$ can be written as the sum of the following three contributions. The first contribution originates from terms with j-powers $1 - \frac{1}{n_1} - \alpha_{\hat{n}} = 0$, which is denoted by

$$v_I^j(t;y) \;:=\; (-1)^{k-1} \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_2(k-1)} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I}) \\ \omega_1+\Sigma(\hat{\omega})=0 \\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}=0}} \frac{(\eta_{\omega_1}^j \eta_{\hat{\omega}}^j)(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}. \tag{4.55}$$

The remaining two contributions

$$\bar{d}_I^j \;:=\; (-1)^{k-1} \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_2(k-1)} j^{\,1-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I}) \\ \omega_1+\Sigma(\hat{\omega})=0 \\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}>0}} \frac{(\eta_{\omega_1}^j \eta_{\hat{\omega}}^j)(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})},$$

$$\widehat{d}_I^j \;:=\; (-1)^{k-1} \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_2(k-1)} j^{\,1-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I}) \\ \omega_1+\Sigma(\hat{\omega})=0 \\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}<0}} \frac{(\eta_{\omega_1}^j \eta_{\hat{\omega}}^j)(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})} \tag{4.56}$$

contain the positive and the negative j-powers, respectively.

In order to take care of the terms $\bar{d}_I^j$ and $\widehat{d}_I^j$, we notice the following fact. Let $(\omega_1, \ldots, \omega_k)$ be a k-tuple of real numbers such that $\{\omega_1, \ldots, \omega_k\}$ is a subset of $\cup_{i=1}^m \cup_{n=2}^r \Omega(n,i)$. Because of the minimal cancellation requirement in FC3 of each $F \in Q_{E,\rho} := \{\pm F_{E,\rho}\}$ and the linear independence of the sets $\Omega_{E,\rho}$ in FC4, a cancellation $\omega_1 + \cdots + \omega_k = 0$ is possible only in the following three cases (cf. [68]):

(a) k is even and each ω_i is canceled out by its negative $-\omega_i$. This case will be referred to as *pure cancellation by pairs*;

(b) there exist $E \in \cup_{n=3}^r \mathcal{E}_n$, $\rho \in \{1, \ldots, |E|\}$, and $F \in Q_{E,\rho}$ such that $F = \{\omega_1, \ldots, \omega_k\}$; this case will be referred to as *pure cancellation by F*;

(c) *mixed cancellation*; i.e., some of the ω_i are canceled out by $-\omega_i$, and some others are canceled out because the set of them is equal to some $F \in Q_{E,\rho}$.

Suppose that $n_1 \in \{3, \ldots, r\}$, $\hat{n} \in \Omega_2(k-1)$, $\omega_1 \in \Omega(n_1, i_1)$, and $\hat{\omega} \in \Omega(\hat{n}, \bar{I})$ such that $\omega_1 + \Sigma(\hat{\omega}) = 0$. If we have pure cancellation by some $F \in Q_{E,\rho}$, then this implies $\frac{1}{n_1} + \alpha_{\hat{n}} = 1$. Thus, pure cancellation cannot happen in $\bar{d}_I^j$ or in $\widehat{d}_I^j$. Next, suppose that we have pure cancellation by pairs. Then, in particular the length k of I has to be even. But now, by the symmetry of the $\Omega(n_1, i_1)$ and the $\Omega(\hat{n}, \bar{I})$, we know that if $\frac{\eta_{\omega_1}^j \eta_{\hat{\omega}}^j}{\mathrm{i}^{k-1}\Pi(\hat{\omega})}$ is in $\bar{d}_I^j$, then since $\eta_{-\omega_1}\eta_{-\hat{\omega}} = \eta_{\omega_1}\eta_{\hat{\omega}}$ and $\Pi(-\hat{\omega}) = -\Pi(\hat{\omega})$, also

$$\frac{\eta_{-\omega_1}^j \eta_{-\hat{\omega}}^j}{\mathrm{i}^{k-1}\,\Pi(-\hat{\omega})} \;=\; -\frac{\eta_{\omega_1}^j \eta_{\hat{\omega}}^j}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}$$

is in $\bar{d}_I^j$. So they add up 0; i.e., the contribution of each term and that of its negative cancel. The same argument also can be applied to $\widehat{d}_I^j$. Thus, only mixed cancellation can give a possibly non-vanishing contribution in $\bar{d}_I^j$ or in $\widehat{d}_I^j$. So finally, suppose that we have a mixed cancellation. Then, there exists an ω_i that is canceled out by $-\omega_i$, and there exists a set of some others that is canceled by some $F \in Q_{E,\rho}$. Thus, $|F| \leq k-2$, which implies that $\frac{1}{n_1} + \alpha_{\hat{n}} \geq \frac{2}{r} + 1$. Thus, mixed cancellation cannot occur in $\bar{d}_I^j$, and therefore, we have $\bar{d}_I^j = 0$. Moreover, we have

$$1 - \frac{1}{n_1} - \alpha_{\hat{n}} \;\leq\; -\frac{2}{r} \tag{4.57}$$

for the j-powers of the noncanceling contributions in $\widehat{d}_I^j$.

In summary, we have

$$b_I^j + c_I^j + d_I^j \;=\; \widetilde{Uv}_I^j(t;y) + v_I^j(t;y) + \widehat{d}_I^j,$$

where

$$\begin{aligned}
\widetilde{Uv}_I^j(t;y) := (-1)^k & \sum_{\hat{n}\in\Omega_2(k-1)} j^{1-\frac{1}{2}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(2,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{2}-\alpha_{\hat{n}}>0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t} \\
&+ (-1)^k \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_1(k-1)} j^{1-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}>0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t} \\
&+ (-1)^k \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_2(k-1)} j^{1-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ \omega_1+\Sigma(\hat{\omega})\neq 0\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}>0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t} \\
&+ \widetilde{Uw}_I^j(t;y)
\end{aligned} \tag{4.58}$$

with the additional contribution

$$\begin{aligned}
\widetilde{Uw}_I^j(t;y) := (-1)^k & \sum_{\hat{n}\in\Omega_2(k-1)} j^{1-\frac{1}{2}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(2,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{2}-\alpha_{\hat{n}}\leq 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t} \\
&+ (-1)^k \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_1(k-1)} j^{1-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}\leq 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t} \\
&+ (-1)^k \sum_{n_1=3}^{r} \sum_{\hat{n}\in\Omega_2(k-1)} j^{1-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ \omega_1+\Sigma(\hat{\omega})\neq 0\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}\leq 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^{k-1}\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}
\end{aligned} \tag{4.59}$$

compared to [68]. Since $\omega_1 + \Sigma(\hat{\omega}) \neq 0$ for all terms in (4.59), integration by parts leads to

$$\int_{t_1}^{t_2} \widetilde{Uw}^j_{i_1,i_2}(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UW}^j_{i_1,i_2}(t;y)\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \widetilde{uW}^j_{i_1,i_2}(t;y)\,\mathrm{d}t,$$

where

$$
\begin{aligned}
\widetilde{UW}_I^j(t;y) \;:=\; &(-1)^k \sum_{\hat{n}\in\Omega_2(k-1)} j^{-\frac{1}{2}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(2,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{2}-\alpha_{\hat{n}}\le 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+(-1)^k \sum_{n_1=3}^{r}\sum_{\hat{n}\in\Omega_1(k-1)} j^{-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}\le 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+(-1)^k \sum_{n_1=3}^{r}\sum_{\hat{n}\in\Omega_2(k-1)} j^{-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ \omega_1+\Sigma(\hat{\omega})\neq 0\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}\le 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}
\end{aligned}
\tag{4.60}
$$

is the additional contribution to (4.52) compared to [68] and

$$
\begin{aligned}
\widetilde{uW}_I^j(t;y) := &(-1)^k \sum_{\hat{n}\in\Omega_2(k-1)} j^{-\frac{1}{2}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(2,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{2}-\alpha_{\hat{n}}\le 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\hat{\omega}})\dot{}(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+(-1)^k \sum_{n_1=3}^{r}\sum_{\hat{n}\in\Omega_1(k-1)} j^{-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}\le 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})\dot{}(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+(-1)^k \sum_{n_1=3}^{r}\sum_{\hat{n}\in\Omega_2(k-1)} j^{-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ \omega_1+\Sigma(\hat{\omega})\neq 0\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}\le 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})\dot{}(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}.
\end{aligned}
\tag{4.61}
$$

In the above definition, the expression $(\eta^j_{\omega_1,i_1}\eta^j_{\hat{\omega}})\dot{}(t;y)$ denotes the derivative of the product of $\eta^j_{\omega_1,i_1}(\cdot;y)$ and $\eta^j_{\hat{\omega}}(\cdot;y)$ at t. The expression $(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})\dot{}(t;y)$ is defined correspondingly. By definition of the sets $\Omega_1(k)$ and $\Omega_2(k)$, we can write (4.52) also as

$$
\begin{aligned}
\widetilde{UV}_I^j(t;y) \;=\; &(-1)^k \sum_{\hat{n}\in\Omega_2(k-1)} j^{-\frac{1}{2}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(2,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{2}-\alpha_{\hat{n}}> 0}} \frac{(\eta^j_{\omega_1,i_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+(-1)^k \sum_{n_1=3}^{r}\sum_{\hat{n}\in\Omega_1(k-1)} j^{-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}> 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+(-1)^k \sum_{n_1=3}^{r}\sum_{\hat{n}\in\Omega_2(k-1)} j^{-\frac{1}{n_1}-\alpha_{\hat{n}}} \sum_{\substack{(\omega_1,\hat{\omega})\in\Omega(n_1,i_1)\times\Omega(\hat{n},\bar{I})\\ \omega_1+\Sigma(\hat{\omega})\neq 0\\ 1-\frac{1}{n_1}-\alpha_{\hat{n}}> 0}} \frac{(\eta^j_{\omega_1}\eta^j_{\hat{\omega}})(t;y)}{\mathrm{i}^k\,(\omega_1+\Sigma(\hat{\omega}))\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j(\omega_1+\Sigma(\hat{\omega}))t}\\
&+\widetilde{UW}_I^j(t;y).
\end{aligned}
$$

Thus, integration by parts leads to

$$\int_{t_1}^{t_2} \widetilde{Uv}^j_{i_1,i_2}(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}^j_{i_1,i_2}(t;y)\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \widetilde{uV}^j_{i_1,i_2}(t;y)\,\mathrm{d}t,$$

where

$$\widetilde{uV}^j_I(t;y) \;:=\; (-1)^k \sum_{\hat{n}\in\Omega_1(k)} j^{-\alpha_{\hat{n}}} \sum_{\hat{\omega}\in\Omega(\hat{n},I)} \frac{\dot{\eta}^j_{\hat{\omega}}(t;y)}{\mathrm{i}^k\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j\Sigma(\hat{\omega})t} \tag{4.62a}$$

$$+\,(-1)^k \sum_{\hat{n}\in\Omega_2(k)} j^{-\alpha_{\hat{n}}} \sum_{\hat{\omega}\in\Omega(\hat{n},I)} \frac{\dot{\eta}^j_{\hat{\omega}}(t;y)}{\mathrm{i}^k\,\Pi(\hat{\omega})}\,\mathrm{e}^{\mathrm{i}j\Sigma(\hat{\omega})t} + \widetilde{uW}^j_I(t;y) \tag{4.62b}$$

with the additional contribution $\widetilde{uW}^j_I(t;y)$ in (4.61) compared to [68]. Finally, we let

$$r^j_I(t;y) \;:=\; \eta^j_{i_1,0}(t;y)\,\widetilde{UV}^j_{\bar{I}}(t;y) + a^j_I + \widetilde{uV}^j_I(t;y) + \widehat{d}^j_I + \big(u^j_{i_1}(t;y) - \eta^j_{i_1,0}(t;y)\big)\,\widetilde{UW}^j_{\bar{I}}(t;y), \tag{4.63}$$

$$\widetilde{uv}^j_I(t;y) \;:=\; \widetilde{uV}^j_I(t;y) + \widetilde{Uv}^j_I(t;y). \tag{4.64}$$

Then, we have

$$u^j_{i_1}(t;y)\,\widetilde{UV}^j_{\bar{I}}(t;y) \;=\; v^j_I(t;y) + r^j_I(t;y) - \widetilde{uv}^j_I(t;y), \tag{4.65}$$

$$\int_{t_1}^{t_2} \widetilde{uv}^j_I(t;y)\,\mathrm{d}t \;=\; \Big[\widetilde{UV}^j_I(t;y)\Big]_{t=t_1}^{t=t_2}. \tag{4.66}$$

This completes the definitions for multi-indices of length $3,\ldots,r$.

We will need the following estimates for the above functions in the next subsection.

Lemma 4.13. *There exists $c>0$ such that*

$$\begin{aligned}
\big|u^j_i(t;y)\big| &\le c\,j^{\frac{r-1}{r}}\,\bar{j}^{\frac{1}{2r}},\\
\big|\widetilde{UV}^j_I(t;y)\big| &\le c\,j^{-\frac{k}{r}}\,\bar{j}^{\frac{k}{2r}},\\
\big|r^j_I(t;y)\big| &\le c\,\bar{j}^{-1-\frac{1}{2r}}\max\big\{\bar{j},\dot{\bar{y}}^j(t)\big\}
\end{aligned}$$

for every $i\in\{1,\ldots,m\}$, $k\in\{1,\ldots,r\}$, every $I\in\{1,\ldots,m\}^k$, every $j\ge 1$, every $y\in\mathcal{Y}$, and every $t\in\mathrm{dom}(y)$.

Proof. It follows from the definitions of the functions $\eta^j_{i,0}(\cdot;y)$, $\eta^j_{\omega,i}(\cdot;y)$, and $\eta^j_{\omega}(\cdot;y)$ in Subsection 4.4.2 that there exists some sufficiently large $c_\eta>0$ such that

$$\big|\eta^j_{i,0}(t;y)\big| \le c_\eta\,\bar{j}^{\frac{1}{2}} \qquad\text{and}\qquad \big|\dot{\eta}^j_{i,0}(t;y)\big| \le c_\eta\,\bar{j}^{\frac{1}{2}}\max\big\{\bar{j},\dot{\bar{y}}^j(t)\big\}$$

for every $i\in\{1,\ldots,m\}$, every $j\ge 1$, every $y\in\mathcal{Y}$, and every $t\in\mathrm{dom}(y)$;

$$\big|\eta^j_{\omega,i}(t;y)\big| \le c_\eta\,\bar{j}^{\frac{1}{4}} \qquad\text{and}\qquad \big|\dot{\eta}^j_{\omega,i}(t;y)\big| \le c_\eta\,\bar{j}^{\frac{1}{4}}\max\big\{\bar{j},\dot{\bar{y}}^j(t)\big\}$$

for every $i\in\{1,\ldots,m\}$, every $\omega\in\Omega(2,i)$, every $j\ge 1$, every $y\in\mathcal{Y}$, and every $t\in\mathrm{dom}(y)$;

$$\big|\eta^j_{\omega}(t;y)\big| \le c_\eta\,\bar{j}^{\frac{1}{2n}} \qquad\text{and}\qquad \big|\dot{\eta}^j_{\omega}(t;y)\big| \le c_\eta\,\bar{j}^{\frac{1}{2n}}\max\big\{\bar{j},\dot{\bar{y}}^j(t)\big\}$$

for every $n \in \{3,\ldots,r\}$, every $\omega \in \cup_{i=1}^m \Omega(n,i)$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$. Using the above estimates, we conclude from the definition of the functions $\eta_{\hat\omega}^j(\cdot;y)$ that there exists $c_{\hat\eta} > 0$ such that

$$\left|\eta_{\hat\omega}^j(t;y)\right| \;\leq\; c_{\hat\eta}\,\bar{j}^{\alpha_{\hat n}/2} \qquad \text{and} \qquad \left|\dot\eta_{\hat\omega}^j(t;y)\right| \;\leq\; c_{\hat\eta}\,\bar{j}^{\alpha_{\hat n}/2} \max\left\{\bar{j}, \dot{\bar{y}}^j(t)\right\}$$

for every $k \in \{2,\ldots,r\}$, every $I \in \{1,\ldots,m\}^k$, every $\hat n \in \Omega_1(k) \cup \Omega_2(k)$, every $\hat\omega \in \Omega(\hat n, I)$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$.

Because of (4.16), we obtain for the functions in (4.21) that there exists $c_u > 0$ such that

$$\left|u_i^j(t;y)\right| \;\leq\; c_u\, j^{\frac{r-1}{r}}\, \bar{j}^{\frac{1}{2r}} \;=\; c_u\, \bar{j}^{r-\frac{1}{2r}}$$

for every $i \in \{1,\ldots,m\}$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$. Similarly, for the functions in (4.46), (4.47), (4.60) and (4.61), we obtain that there exist $c_{UW}, c_{uW} > 0$ such that

$$\left|\widetilde{UW}_I^j(t;y)\right| \;\leq\; c_{UW}\,\bar{j}^{-r} \qquad \text{and} \qquad \left|\widetilde{uW}_I^j(t;y)\right| \;\leq\; c_{uW}\,\bar{j}^{-r} \max\left\{\bar{j}, \dot{\bar{y}}^j(t)\right\}$$

for every $k \in \{2,\ldots,r\}$, every $I \in \{1,\ldots,m\}^k$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$. Since $\alpha_{\hat n} \geq k/r$ for every $\hat n \in \{2,\ldots,r\}^k$, we obtain for the functions in (4.35), (4.36), (4.44), (4.45), (4.52) and (4.62) that there exist $c_{UV}, c_{uV} > 0$ such that

$$\left|\widetilde{UV}_I^j(t;y)\right| \;\leq\; c_{UV}\,\bar{j}^{-k\frac{r+1}{r}+\frac{k}{2r}} \qquad \text{and} \qquad \left|\widetilde{uV}_I^j(t;y)\right| \;\leq\; c_{uV}\,\bar{j}^{-k\frac{r+1}{r}+\frac{k}{2r}} \max\left\{\bar{j}, \dot{\bar{y}}^j(t)\right\}$$

for every $k \in \{1,\ldots,r\}$, every $I \in \{1,\ldots,m\}^k$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$. This in turn implies that

$$\begin{aligned} \left|\eta_{i_1,0}^j(t;y)\,\widetilde{UV}_{\bar I}^j(t;y)\right| \;&\leq\; c_\eta\, c_{UV}\,\bar{j}^{\frac12-(k-1)\frac{r+1}{r}+\frac{k-1}{2r}} \;\leq\; c_\eta\, c_{UV}\,\bar{j}^{-\frac12-\frac{1}{2r}}, \\ \left(u_{i_1}^j(t;y) - \eta_{i_1,0}^j(t;y)\right)\widetilde{UW}_{\bar I}^j(t;y) \;&\leq\; (c_u + c_\eta)\, c_{UW}\,\bar{j}^{-\frac{1}{2r}} \end{aligned}$$

for every $k \in \{2,\ldots,r\}$, every $I = (i_1,\ldots,i_k)$, every $i_1,\ldots,i_k \in \{1,\ldots,m\}$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$. Now the asserted estimates for (4.21), (4.35), (4.37), (4.44), (4.48) and (4.52) are clear. To complete the proof, we have to show the asserted estimate for (4.63). For this purpose, it is left to derive suitable estimates for the terms a_I^j and $\widehat{d}_I^j$ in the right-hand side of (4.63), which are given by (4.53) and (4.56), respectively. To this end, we use the above estimates and also the estimates (4.54) and (4.57) for the j-powers in (4.53) and (4.56), respectively. This leads to the existence of $c_a, c_d > 0$ such that

$$|a_I^j| \;\leq\; c_a\,\bar{j}^{-\frac12-\frac{1}{2r}} \qquad \text{and} \qquad \left|\widehat{d}_I^j\right| \;\leq\; c_d\,\bar{j}^{-\frac32-\frac{1}{2r}},$$

for every $k \in \{3,\ldots,r\}$, every $I = (i_1,\ldots,i_k)$, every $i_1,\ldots,i_k \in \{1,\ldots,m\}$, every $j \geq 1$, every $y \in \mathcal{Y}$, and every $t \in \mathrm{dom}(y)$, where the dependence on t and y is suppressed in the notation. Now the asserted estimate for (4.63) follows and the proof is complete. □

As in [68], for every $j > 0$, every $E \in \bigcup_{n=1}^r \mathcal{E}_n$, every $B \in E$, and every $y \in \mathcal{Y}$, we define a function $v_B^j(\cdot;y)$ on the domain of y as follows:
If $\delta(E) = 1$; i.e., $E = \{B\}$, $B = X_i$ for some $i \in \{1,\ldots,m\}$, then

$$v_B^j(t;y) \;:=\; \eta_{i,0}^j(t;y).$$

If $\delta(E) = 2$; i.e., $E = \{B\}$, $B = (X_{i_1}, X_{i_2})$ for $i_1, i_2 \in \{1, \ldots, m\}$ with $i_1 < i_2$, then

$$v_B^j(t;y) \;:=\; \mathrm{i}^{1-\delta(E)}\,\frac{1}{\omega_E}\,\big(\eta^j_{\omega_E,i_1}(t;y)\,\eta^j_{-\omega_E,i_2}(t;y) - \eta^j_{-\omega_E,i_1}(t;y)\,\eta^j_{\omega_E,i_2}(t;y)\big).$$

If $\delta(E) > 2$, then

$$v_B^j(t;y) \;:=\; \sum_{\rho=1}^{|E|}\sum_{\sigma=\pm}(\sigma\mathrm{i})^{1-\delta(E)}\,\hat{\xi}_{B,\rho}\prod_{\omega\in\sigma F_{E,\rho}}\eta^j_\omega(t;y).$$

In any case, the above definitions lead to

$$v_B^j(t;y) \;=\; \sum_{\rho=1}^{|E|}\hat{\xi}_{B,\rho}\,\zeta^j_{(E,\rho)}(t;y), \tag{4.67}$$

where $\zeta^j_{(E,\rho)}(t;y)$ is given by (4.19).

Since the above functions $v_B^j(\cdot;y)$ and the functions $v_I^j(\cdot;y)$ in (4.33), (4.41) and (4.55) are defined in the same way as in [68], one can apply the same argument as therein to conclude that following algebraic identity holds on $\mathrm{A}_0(\mathbf{X})$.

Lemma 4.14. *For every $j > 0$, every $y \in \mathcal{Y}$, every $t \in \mathrm{dom}(\mathcal{Y})$, we have*

$$\sum_{0<|I|\leq r} v_I^j(t;y)\,X_I \;=\; \sum_{n=1}^{r}\sum_{B\in\mathcal{B}_n} v_B^j(t;y)\,\mu(B),$$

where the sum on the left-hand side ranges over all multi-indices $I = (i_1, \ldots, i_k)$ of length $k \in \{1, \ldots, r\}$ with $i_1, \ldots, i_k \in \{1, \ldots, m\}$, and the X_I are given by (4.4).

4.6.3 Integral expansion

For the moment, fix a frequency parameter $j > 0$ and an initial state $x_0 \in M$. Recall that we use the notation in (4.24) and (4.25) for the maximal solution of (4.23) and the associated output signal, respectively. Let $\varphi \in C^\infty(M)$ and let $t_1, t_2 \in [0, T^j_{x_0})$. The fundamental theorem of calculus applied to the composition of $\gamma^j_{x_0}$ and φ implies that

$$\varphi(\gamma^j_{x_0}(t_2)) \;=\; \varphi(\gamma^j_{x_0}(t_1)) + \sum_{i=1}^{m}\int_{t_1}^{t_2} u_i^j(t;y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t.$$

Note that each of the functions $t \mapsto (f_i\varphi)(\gamma^j_{x_0}(t))$ is differentiable. Thus, we may apply integration by parts, which leads to

$$\begin{aligned}\varphi(\gamma^j_{x_0}(t_2)) \;=\; &\varphi(\gamma^j_{x_0}(t_1)) - \sum_{i=1}^{m}\Big[\widetilde{UV}^j_i(t;y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\Big]_{t=t_1}^{t=t_2}\\ &+\sum_{i=1}^{m}\int_{t_1}^{t_2} v_i^j(t;y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t + \sum_{i=1}^{m}\int_{t_1}^{t_2} r_i^j(t;y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t\\ &+\sum_{i_1,i_2=1}^{m}\int_{t_1}^{t_2} u^j_{i_1}(t;y^j_{x_0})\,\widetilde{UV}^j_{i_2}(t;y^j_{x_0})\,(f_{i_1}f_{i_2}\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t,\end{aligned}$$

where we have used first (4.39) and then (4.40) as well as that $\gamma^j_{x_0}$ is a solution of (4.23). Note that each of the functions $t \mapsto (f_{i_1} f_{i_2}\varphi)(\gamma^j_{x_0}(t))$ in the last integral of the above equation is differentiable. Thus, we may apply again integration by parts, which leads to

$$\begin{aligned}\varphi(\gamma^j_{x_0}(t_2)) \; = \; & \varphi(\gamma^j_{x_0}(t_1)) - \sum_{i=1}^{m} \Big[\widetilde{UV}^j_i(t; y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\Big]_{t=t_1}^{t=t_2} \\ & - \sum_{i_1,i_2=1}^{m} \Big[\widetilde{UV}^j_{i_1,i_2}(t; y^j_{x_0})\,(f_{i_1} f_{i_2}\varphi)(\gamma^j_{x_0}(t))\Big]_{t=t_1}^{t=t_2} \\ & + \sum_{i=1}^{m} \int_{t_1}^{t_2} v^j_i(t; y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t + \sum_{i_1,i_2=1}^{m} \int_{t_1}^{t_2} v^j_{i_1,i_2}(t; y^j_{x_0})\,(f_{i_1} f_{i_2}\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t \\ & + \sum_{i=1}^{m} \int_{t_1}^{t_2} r^j_i(t; y^j_{x_0})\,(f_i\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t + \sum_{i_1,i_2=1}^{m} \int_{t_1}^{t_2} r^j_{i_1,i_2}(t; y^j_{x_0})\,(f_{i_1} f_{i_2}\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t \\ & + \sum_{i_1,i_2,i_3=1}^{m} \int_{t_1}^{t_2} u^j_{i_1}(t; y^j_{x_0})\,\widetilde{UV}^j_{i_2,i_3}(t; y^j_{x_0})\,(f_{i_1} f_{i_2} f_{i_3}\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t,\end{aligned}$$

where we have used first (4.50) and then (4.51) as well as that $\gamma^j_{x_0}$ is a solution of (4.23). We repeat this procedure $(r-2)$ times. First we insert (4.65) and then we apply integration by parts using (4.66) and that $\gamma^j_{x_0}$ is a solution of (4.23). This leads to

$$\varphi(\gamma^j_{x_0}(t_2)) \; = \; \varphi(\gamma^j_{x_0}(t_1)) + \sum_{0<|I|\le r} \int_{t_1}^{t_2} v^j_I(t; y^j_{x_0})\,(f_I\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t \tag{4.68a}$$

$$- \Big[(D^j_1\varphi)(t, \gamma^j_{x_0}(t); y^j_{x_0})\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} (D^j_2\varphi)(t, \gamma^j_{x_0}(t); y^j_{x_0})\,\mathrm{d}t, \tag{4.68b}$$

where the time-varying differential operators D^j_1, D^j_2 on $C^\infty(M)$ are defined by

$$(D^j_1\varphi)(t,x;y) := \sum_{0<|I|\le r} \widetilde{UV}^j_I(t;y)\,(f_I\varphi)(x), \tag{4.69}$$

$$(D^j_2\varphi)(t,x;y) := \sum_{0<|I|\le r} r^j_I(t;y)\,(f_I\varphi)(x) + \sum_{|iI|=r+1} u^j_i(t;y)\,\widetilde{UV}^j_I(t;y)\,(f_i f_I\varphi)(x) \tag{4.70}$$

for every $y \in \mathcal{Y}$, every $t \in \mathrm{dom}(y)$, and every $x \in M$. As in [69], by "plugging in the f_i for the X_i", we obtain from Lemma 4.14 and (4.31), (4.32) that the sum of the integrals in (4.68a) can be written as

$$\sum_{0<|I|\le r} \int_{t_1}^{t_2} v^j_I(t; y^j_{x_0})\,(f_I\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t \; = \; \sum_{n=1}^{r} \sum_{B\in\mathcal{B}_n} \int_{t_1}^{t_2} v^j_B(t; y^j_{x_0})\,([f_B]\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t.$$

This motivates us to define a time-varying vector field $G^j(\cdot,\cdot;y)$ on M by

$$G^j(t,x;y) \; := \; \sum_{n=1}^{r} \sum_{B\in\mathcal{B}_n} v^j_B(t;y)\,[f_B](x) \tag{4.71}$$

for every $y \in \mathcal{Y}$, every $t \in \mathrm{dom}(y)$, and every $x \in M$. Naturally, $G^j(\cdot,\cdot;y)$ acts as a time-varying differential operator on $C^\infty(M)$ by taking (fixed-time) Lie derivatives. Using (4.26) and (4.67), we obtain that

$$G^j(t,x;y) \;=\; \sum_{\ell\in\Lambda^r} \zeta_\ell^j(t;y)\, g_\ell(x). \tag{4.72}$$

In summary, we have derived the following integral expansion for the propagation of smooth functions along solutions of (4.23).

Proposition 4.15. *For every $j>0$, every $x_0\in M$, and all $t_1,t_2\in[0,T^j_{x_0})$, we have*

$$\begin{aligned}\varphi(\gamma^j_{x_0}(t_2)) \;=\;& \varphi(\gamma^j_{x_0}(t_1)) - \Big[(D_1^j\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} (D_2^j\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\,\mathrm{d}t \\ &+ \int_{t_1}^{t_2} (G^j\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\,\mathrm{d}t,\end{aligned}$$

where D_1^j, D_2^j, and G^j are given by (4.69), (4.70), *and* (4.72), *respectively.*

The subsequent estimates for the remainder terms (4.69) and (4.70) follow immediately from Lemmas 4.11 and 4.13.

Proposition 4.16. *Let $\varphi\in C^\infty(M)$. Suppose that $f_1,\ldots,f_m$ are compactly supported. Then, there exist $c_1,c_2>0$ such that*

$$\begin{aligned}\big|(D_1^j\varphi)(t,x;y)\big| \;&\leq\; c_1\,\bar{j}^{\,-1-\frac{1}{2r}}, \\ \big|(D_2^j\varphi)(t,x;y)\big| \;&\leq\; c_2\,\bar{j}^{\,-1-\frac{1}{2r}}\,\max\{\bar{j},\dot{\bar{y}}^j(t)\}\end{aligned}$$

for every $j\geq 1$, every $y\in\mathcal{Y}$, every $t\in\operatorname{dom}(y)$, and every $x\in M$.

The averaged term (4.72) is analyzed in the next section.

4.7 Averaging of the slow oscillations

In this section, we apply a similar procedure as in Section 4.6 to the time-varying vector field defined in (4.72), which arises in Proposition 4.15 as the averaged contribution of the fast oscillations. Because of (4.19), we can write (4.72) as

$$G^j(t,x;y) \;=\; \sum_{(\ell,\nu)\in J^r} \bar{u}^j_{(\ell,\nu)}(t)\, h_\nu(\bar{y}^j(t))\, g_\ell(x) \tag{4.73}$$

for every $j>0$, every $y\in\mathcal{Y}$, every $t\in\operatorname{dom}(y)$, and every $x\in M$, where

- the indexing set J^r consists of all pairs (ℓ,ν) with $\ell\in\Lambda^r$ and $\nu\in\{1,2\}$,
- the sinusoids $\bar{u}^j_{(\ell,\nu)}\colon\mathbb{R}\to\mathbb{R}$ are defined in (4.11),
- the functions $h_\nu\colon\mathbb{R}\to\mathbb{R}$ are defined in (4.9),
- the first-order hold $\bar{y}^j\colon\operatorname{dom}(y)\to\mathbb{R}$ of y is defined in (4.17),
- the vector fields g_ℓ on M are defined in (4.26).

Note that (4.73) does not depend on any of the fast oscillations on the time scale j but only on the slow oscillations $\bar{u}^j_{(\ell,\nu)}$ on the time scale $\bar{j}$, where $\bar{j}$ is given by (4.16).

4.7.1 Averaging of the sinusoids

We repeat the procedure from Subsection 2.5.2 in a slightly different notation. Recall that the sinusoids $\bar{u}^j_{(\ell,\nu)}$ in (4.11) are determined by the pairwise distinct frequency coefficients $\bar{\omega}_\ell > 0$. For every $\iota = (\ell,\nu) \in J^r$, define two complex-valued constants $\bar{\eta}_{\pm\omega_\ell,\iota}$ as follows. If $\nu = 1$, let $\bar{\eta}_{\pm\omega_\ell,\iota} := \sqrt{2\bar{\omega}_\ell}/2$, and otherwise, i.e., if $\nu = 2$, let $\bar{\eta}_{\pm\omega_\ell,\iota} := \pm\sqrt{2\bar{\omega}_\ell}/(2\mathrm{i})$, where i denotes the imaginary unit. Moreover, let $\bar{\Omega}(\iota) := \{\pm\bar{\omega}_\ell\}$.

Let $\iota \in J^r$. Using the above notation, we can write $\bar{u}^j_\iota(t)$ in (4.11) as

$$\bar{u}^j_\iota(t) \;=\; \bar{j}^{\frac{1}{2}} \sum_{\bar{\omega}\in\bar{\Omega}(\iota)} \bar{\eta}_{\bar{\omega},\iota}\, \mathrm{e}^{\mathrm{i}\bar{j}\bar{\omega}t} \tag{4.74}$$

for every $t \in \mathbb{R}$. When we integrate $-\bar{u}^j_\iota$, we get

$$-\int_{t_1}^{t_2} \bar{u}^j_\iota(t)\,\mathrm{d}t \;=\; \Big[\widetilde{\overline{UV}}{}^j_\iota(t)\Big]_{t=t_1}^{t=t_2}, \tag{4.75}$$

where

$$\widetilde{\overline{UV}}{}^j_\iota(t) \;:=\; -\bar{j}^{-\frac{1}{2}} \sum_{\bar{\omega}\in\bar{\Omega}(\iota)} \frac{\bar{\eta}_{\bar{\omega},\iota}}{\mathrm{i}\bar{\omega}}\, \mathrm{e}^{\mathrm{i}\bar{j}\bar{\omega}t}. \tag{4.76}$$

Let $\iota_1,\iota_2 \in J^r$. When we multiply $\bar{u}^j_{\iota_1}(t)$ by $\widetilde{\overline{UV}}{}^j_{\iota_2}(t)$, we get

$$\bar{u}^j_{\iota_1}(t)\,\widetilde{\overline{UV}}{}^j_{\iota_2}(t) \;=\; \bar{v}_{\iota_1,\iota_2} - \widetilde{\overline{uv}}{}^j_{\iota_1,\iota_2}(t), \tag{4.77}$$

where

$$\bar{v}_{\iota_1,\iota_2} \;:=\; -\sum_{\substack{(\bar{\omega}_1,\bar{\omega}_2)\in\bar{\Omega}(\iota_1)\times\bar{\Omega}(\iota_2)\\ \bar{\omega}_1+\bar{\omega}_2=0}} \frac{\bar{\eta}_{\bar{\omega}_1,\iota_1}\,\bar{\eta}_{\bar{\omega}_2,\iota_2}}{\mathrm{i}\,\bar{\omega}_2}, \tag{4.78}$$

$$\widetilde{\overline{uv}}{}^j_{\iota_1,\iota_2}(t) \;:=\; \sum_{\substack{(\bar{\omega}_1,\bar{\omega}_2)\in\bar{\Omega}(\iota_1)\times\bar{\Omega}(\iota_2)\\ \bar{\omega}_1+\bar{\omega}_2\neq 0}} \frac{\bar{\eta}_{\bar{\omega}_1,\iota_1}\,\bar{\eta}_{\bar{\omega}_2,\iota_2}}{\mathrm{i}\,\bar{\omega}_2}\, \mathrm{e}^{\mathrm{i}\bar{j}(\bar{\omega}_1+\bar{\omega}_2)t} \tag{4.79}$$

for every $t \in \mathbb{R}$. When we integrate $\widetilde{\overline{uv}}{}^j_{\iota_1,\iota_2}$, we get

$$\int_{t_1}^{t_2} \widetilde{\overline{uv}}{}^j_{\iota_1,\iota_2}(t)\,\mathrm{d}t \;=\; \Big[\widetilde{\overline{UV}}{}^j_{\iota_1,\iota_2}(t)\Big]_{t=t_1}^{t=t_2}, \tag{4.80}$$

where

$$\widetilde{\overline{UV}}{}^j_{\iota_1,\iota_2}(t) \;:=\; \bar{j}^{-1} \sum_{\substack{(\bar{\omega}_1,\bar{\omega}_2)\in\bar{\Omega}(\iota_1)\times\bar{\Omega}(\iota_2)\\ \bar{\omega}_1+\bar{\omega}_2\neq 0}} \frac{\bar{\eta}_{\bar{\omega}_1,\iota_1}\,\bar{\eta}_{\bar{\omega}_2,\iota_2}}{\mathrm{i}^2\,\bar{\omega}_2(\bar{\omega}_2+\bar{\omega}_1)}\, \mathrm{e}^{\mathrm{i}\bar{j}(\bar{\omega}_1+\bar{\omega}_2)t}. \tag{4.81}$$

The functions in (4.74), (4.76) and (4.81) satisfy the following estimates.

Lemma 4.17. *There exists $c > 0$ such that*

$$\begin{aligned} \big|\bar{u}^j_\iota(t)\big| \;&\le\; c\,\bar{j}^{\frac{1}{2}},\\ \big|\widetilde{\overline{UV}}{}^j_\iota(t)\big| \;&\le\; c\,\bar{j}^{-\frac{1}{2}},\\ \big|\widetilde{\overline{UV}}{}^j_{\iota_1,\iota_2}(t)\big| \;&\le\; c\,\bar{j}^{-1} \end{aligned}$$

for every $j \geq 1$, all $\iota,\iota_1,\iota_2 \in J^r$, and every $t \in \mathbb{R}$.

This means that the functions $\widetilde{\widetilde{UV}}{}^j_\iota$ and $\widetilde{\widetilde{UV}}{}^j_{\iota_1,\iota_2}$ converge uniformly to 0 as the global frequency parameter j tends to ∞. Moreover, a direct computation shows that the $\bar{v}_{\iota_1,\iota_2}$ in (4.78) are given as follows.

Lemma 4.18. *For all $\iota_1 = (\ell_1, \nu_1), \iota_2 = (\ell_2, \nu_2) \in J^r$, we have*

$$\bar{v}_{\iota_1,\iota_2} \;=\; \begin{cases} +1 & \text{if } \ell_1 = \ell_2 \text{ and } \nu_1 = 1 \text{ and } \nu_2 = 2, \\ -1 & \text{if } \ell_1 = \ell_2 \text{ and } \nu_1 = 2 \text{ and } \nu_2 = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Because of Lemma 4.18 and equation (4.10), we have

$$\sum_{\iota_1,\iota_2 \in J^r} \bar{v}_{\iota_1,\iota_2}\, \bar{f}_{\iota_1}(\bar{f}_{\iota_2}\varphi) \;=\; \sum_{\ell \in \Lambda^r} ([\bar{f}_{(\ell,1)}, \bar{f}_{(\ell,2)}]\varphi) \;=\; f^\infty \varphi \tag{4.82}$$

for every $\varphi \in C^\infty(M)$, where the vector field f^∞ on M is given by (4.27).

4.7.2 Estimate for the output variations

Recall from Section 4.4 that measurements of the output (4.6) are conducted at the time instances $\tau^j_k = k\,\Delta/\bar{j}^2$ for $k = 0, 1, \ldots$, where $\bar{j}$ is given by (4.16) and Δ is a positive constant. Our next goal is to derive an estimate for the variation of the output signal $y^j_{x_0}$ in (4.25) on time intervals between two subsequent output measurements. For this purpose, we first need the following estimate for the Lie derivative of a smooth function along the time-varying vector field given in (4.73).

Lemma 4.19. *Suppose that $f_1, \ldots, f_m$ are compactly supported. Then, for every $\varphi \in C^\infty(M)$, there exists $c_3 > 0$ such that*

$$|(G^j\varphi)(t, x; y)| \;\le\; c_3\, \bar{j}^{\frac{1}{2}}$$

for every $y \in \mathcal{Y}$, every $t \in \operatorname{dom}(y)$, and every $x \in M$.

Proof. By the definitions in (4.9), the functions h_ν are bounded by a constant. Therefore, the asserted estimate follows immediately from equation (4.73) and Lemmas 4.12 and 4.17. □

Lemma 4.20. *Suppose that $f_1, \ldots, f_m$ are compactly supported. Then, there exists $j_0 \ge 1$ such that*

$$\frac{1}{\Delta/\bar{j}^2}\big|y^j_{x_0}(t) - y^j_{x_0}(\tau^j_k)\big| \;\le\; \bar{j}$$

for every $j \ge j_0$, every $x_0 \in M$, every nonnegative integer k, and every $t \in [\tau^j_k, \tau^j_{k+1}]$.

Proof. Let $c_1, c_2 > 0$ and $c_3 > 0$ be the constants from Proposition 4.16 and Lemma 4.19 for $\varphi = \psi$, respectively. Choose a sufficiently large $j_0 \ge 1$ such that

$$\frac{2\,c_1}{\Delta}\,\bar{j}^{1-\frac{1}{2r}} + c_2\,\bar{j}^{-\frac{1}{2r}} + c_3\,\bar{j}^{\frac{1}{2}} \;\le\; \bar{j}$$

for every $j \ge j_0$. To verify the asserted estimate for the output, fix arbitrary $j \ge j_0$, $x_0 \in M$, and write $y = y^j_{x_0}$. From Proposition 4.15 with $\varphi = \psi$, we conclude that

$$\frac{1}{\Delta/\bar{j}^2}|y(t_2) - y(t_1)| \;\le\; \frac{2\,c_1}{\Delta}\,\bar{j}^{1-\frac{1}{2r}} + \big(c_2\,\bar{j}^{-1-\frac{1}{2r}} \max\{\bar{j}, \dot{\bar{y}}^j(t)\} + c_3\,\bar{j}^{\frac{1}{2}}\big)\,\frac{t_2 - t_1}{\Delta/\bar{j}^2} \tag{4.83}$$

for all $t_2 \geq t_1 \geq 0$. By definition of the first-order hold, we have $\dot{\bar{y}}^j(t) = 0$ for $t \in [\tau_0^j, \tau_1^j]$. Using (4.83) and the choice of j_0, this implies that the asserted estimate holds for every $t \in [\tau_0^j, \tau_1^j]$. Now we proceed by induction on k. Suppose that the asserted estimate is true for every $t \in [\tau_k^j, \tau_{k+1}^j]$ with some nonnegative integer k. Then, in particular, the estimate is true for $t = \tau_{k+1}^j$. By definition of the first-order hold, this implies $\dot{\bar{y}}^j(t) \leq \bar{j}$ for every $t \in [\tau_{k+1}^j, \tau_{k+2}^j]$. Now it follows again from (4.83) and the choice of j_0 that the asserted estimate also holds for every $t \in [\tau_{k+1}^j, \tau_{k+2}^j]$. □

For every $j > 0$, the right-hand side of (4.12) defines a time-varying vector field $\bar{F}^j$ on M, which is given by

$$\bar{F}^j(t,x) \;:=\; \sum_{\iota \in J^r} \bar{u}_\iota^j(t)\, \bar{f}_\iota(x) \;=\; \sum_{(\ell,\nu)\in J^r} \bar{u}^j_{(\ell,\nu)}(t)\, h_\nu(\psi(x))\, g_\ell(x), \tag{4.84}$$

where we have used the definition of the vector fields $\bar{f}_\iota$ in (4.8). Our next goal is to replace the averaged term $(G^j\varphi)(t,\gamma_{x_0}^j(t); y_{x_0}^j)$ in Proposition 4.15 by $(\bar{F}^j\varphi)(t,\gamma_{x_0}^j(t))$. This requires a suitable estimate for their difference. For this reason, we define a time-varying differential operator $\widehat{D}_2^j$ on $C^\infty(M)$ by

$$(\widehat{D}_2^j\varphi)(t,x;y) \;:=\; (D_2^j\varphi)(t,x;y) + (G^j\varphi)(t,x;y) - (\bar{F}^j\varphi)(t,x) \tag{4.85}$$

for every $y \in \mathcal{Y}$, every $t \in \mathrm{dom}(y)$, and every $x \in M$, where D_2^j is given by (4.70).

Lemma 4.21. *Suppose that $f_1, \ldots, f_m$ are compactly supported. Then, there exists $j_0 \geq 1$ such that, for every $\varphi \in C^\infty(M)$, there exists $\widehat{c}_2 > 0$ such that*

$$\big|(\widehat{D}_2^j\varphi)(t,\gamma_{x_0}^j(t); y_{x_0}^j)\big| \;\leq\; \widehat{c}_2\, \bar{j}^{-\frac{1}{2r}}$$

for every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$.

Proof. Choose $j_0 \geq 1$ as in Lemma 4.20. Then, in particular, by definition of the first-order hold, we have $|\dot{\bar{y}}^j(t)| \leq \bar{j}$ for every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$, where y abbreviates $y_{x_0}^j$. Fix an arbitrary $\varphi \in C^\infty(M)$. Then, by the choice of j_0 and Proposition 4.16, there exists $c_2 > 0$ such that

$$\big|(D_2^j\varphi)(t,\gamma_{x_0}^j(t); y_{x_0}^j)\big| \;\leq\; c_2\, \bar{j}^{-\frac{1}{2r}}$$

for every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$. To complete the proof, it suffices to show that there exists $c_2' > 0$ such that

$$\big|(G^j\varphi)(t,\gamma_{x_0}^j(t); y_{x_0}^j) - (\bar{F}^j\varphi)(t,\gamma_{x_0}^j(t))\big| \;\leq\; c_2'\, \bar{j}^{-\frac{1}{2r}}$$

for every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$. By (4.73) and (4.84), we have

$$\big|(G^j\varphi)(t,x;y) - (\bar{F}^j\varphi)(t,x)\big| \;\leq\; \sum_{(\ell,\nu)\in J^r} \big|\bar{u}^j_{(\ell,\nu)}(t)\big|\, \big|h_\nu(\bar{y}^j(t)) - h_\nu(\psi(x))\big|\, \big|(g_\ell\varphi)(x)\big|$$

for every $j > 0$, every $y\colon [0,\infty) \to \mathbb{R}$, every $t \geq 0$, and every $x \in M$. The functions $h_\nu\colon \mathbb{R} \to \mathbb{R}$ in (4.9) are globally Lipschitz continuous with Lipschitz constant 1. From the definition of the first-order hold and Lemma 4.20, it is easy to derive that $|\bar{y}^j(t) - y(t)| \leq 2\Delta/\bar{j}$ for every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$, where y abbreviates $y_{x_0}^j$. Since $\bar{j}^{-\frac{1}{2}} \leq \bar{j}^{-\frac{1}{2r}}$ for $j \geq 1$, the asserted estimate for the difference of $G^j\varphi$ and $\bar{F}^j\varphi$ now follows from Lemma 4.12 and Lemma 4.17. □

Using (4.85), we obtain from Proposition 4.15 the equation

$$\varphi(\gamma^j_{x_0}(t_2)) \;=\; \varphi(\gamma^j_{x_0}(t_1)) + \sum_{\iota\in J^r}\int_{t_1}^{t_2} \bar{u}^j_\iota(t)\,(\bar{f}_\iota\varphi)(\gamma^j_{x_0}(t))\,\mathrm{d}t \tag{4.86a}$$

$$- \Big[(D^j_1\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} (\widehat{D}^j_2\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\,\mathrm{d}t \tag{4.86b}$$

for every $j>0$, every $x_0\in M$, and all $t_1,t_2\in[0,T^j_{x_0})$.

Remark 4.22. We already know from Proposition 4.16 and Lemma 4.21 that the contribution in (4.86b) becomes small with increasing j. Thus, only the contribution in (4.86a) remains for large values of j. Note that an integral equation of the form (4.86a) (without the contribution in (4.86b)) describes the propagation of φ along a solution of (4.12). This implies the approximation properties AP1 in Section 4.3. It is left to extract the averaged contribution of (4.86a). This is done in Subsection 4.7.3. ♢

4.7.3 Integral expansion

Define two time-varying differential operators D^j, D^j_0 on $C^\infty(M)$ by

$$(D^j\varphi)(t,x;y) \;:=\; \varphi(x) + (D^j_1\varphi)(t,x;y), \tag{4.87}$$

$$(D^j_0\varphi)(t,x;y) \;:=\; (\widehat{D}^j_2\varphi)(t,x;y) - \sum_{\iota\in J^r}\bar{u}^j_\iota(t)\,(D^j_1(\bar{f}_\iota\varphi))(t,x;y) \tag{4.88}$$

for every $y\in\mathcal{Y}$, every $t\in\operatorname{dom}(y)$, and every $x\in M$. Then, equation (4.86) can be written as

$$(D^j\varphi)(t_2,\gamma^j_{x_0}(t_2);y^j_{x_0}) \;=\; (D^j\varphi)(t_1,\gamma^j_{x_0}(t_1);y^j_{x_0}) + \int_{t_1}^{t_2}(D^j_0\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\,\mathrm{d}t \tag{4.89a}$$

$$+ \sum_{\iota\in J^r}\int_{t_1}^{t_2}\bar{u}^j_\iota(t)\,(D^j(\bar{f}_\iota\varphi))(t,\gamma^j_{x_0}(t);y^j_{x_0})\,\mathrm{d}t \tag{4.89b}$$

for every $j>0$, every $\varphi\in C^\infty(M)$, every $x_0\in M$, and all $t_1,t_2\in[0,T^j_{x_0})$. By the fundamental theorem of calculus, (4.89) means that the locally absolutely function

$$t\mapsto (D^j\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0}) \tag{4.90}$$

is an antiderivative of the locally integrable function

$$t\mapsto (D^j_0\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0}) + \sum_{\iota\in J^r}\bar{u}^j_\iota(t)\,(D^j(\bar{f}_\iota\varphi))(t,\gamma^j_{x_0}(t);y^j_{x_0}). \tag{4.91}$$

Thus, if we apply integration by parts in (4.89b), then we obtain

$$(D^j\varphi)(t_2,\gamma^j_{x_0}(t_2);y^j_{x_0}) \;=\; (D^j\varphi)(t_1,\gamma^j_{x_0}(t_1);y^j_{x_0}) + \int_{t_1}^{t_2}(D^j_0\varphi)(t,\gamma^j_{x_0}(t);y^j_{x_0})\,\mathrm{d}t$$

$$- \sum_{\iota\in J^r}\Big[\widetilde{\overline{UV}}{}^j_\iota(t)\,(D^j(\bar{f}_\iota\varphi))(t,\gamma^j_{x_0}(t);y^j_{x_0})\Big]_{t=t_1}^{t=t_2}$$

$$+ \sum_{\iota \in J^r} \int_{t_1}^{t_2} \widetilde{\overline{UV}}_\iota^j(t)(D_0^j(\bar{f}_\iota \varphi))(t, \gamma_{x_0}^j(t); y_{x_0}^j)\,\mathrm{d}t$$
$$+ \sum_{\iota_1,\iota_2 \in J^r} \int_{t_1}^{t_2} \bar{u}_{\iota_1}^j(t)\,\widetilde{\overline{UV}}_{\iota_2}^j(t)\,(D^j(\bar{f}_{\iota_1}\bar{f}_{\iota_2}\varphi))(t, \gamma_{x_0}^j(t); y_{x_0}^j)\,\mathrm{d}t,$$

where we have used (4.75) and that (4.90) is an antiderivative of (4.91). Next, we insert (4.77) in the last integral of the above equation. Then, we apply integration by parts using (4.80) and that (4.90) is an antiderivative of (4.91). This leads to

$$\begin{aligned}\varphi(\gamma_{x_0}^j(t_2)) &= \varphi(\gamma_{x_0}^j(t_1)) + \sum_{\iota_1,\iota_2 \in J^r} \int_{t_1}^{t_2} \bar{v}_{\iota_1,\iota_2}\,(\bar{f}_{\iota_1}\bar{f}_{\iota_2}\varphi)(\gamma_{x_0}^j(t))\,\mathrm{d}t \\ &\quad - \Big[(\bar{D}_1^j\varphi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\Big]_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} (\bar{D}_2^j\varphi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\,\mathrm{d}t,\end{aligned}$$

where the time-varying differential operators $\bar{D}_1^j$ and $\bar{D}_2^j$ on $C^\infty(M)$ are defined by

$$(\bar{D}_1^j\varphi)(t,x;y) := (D_1^j\varphi)(t,x;y) + \sum_{\iota \in J^r} \widetilde{\overline{UV}}_\iota^j(t)\,(D^j(\bar{f}_\iota\varphi))(t,x;y) \tag{4.92a}$$
$$+ \sum_{\iota_1,\iota_2 \in J^r} \widetilde{\overline{UV}}_{\iota_1,\iota_2}^j(t)\,(D^j(\bar{f}_{\iota_1}\bar{f}_{\iota_2}\varphi))(t,x;y) \tag{4.92b}$$

and

$$(\bar{D}_2^j\varphi)(t,x;y) := (D_0^j\varphi)(t,x;y) + \sum_{\iota_1,\iota_2 \in J^r} \bar{v}_{\iota_1,\iota_2}\,(D_1^j(\bar{f}_{\iota_1}\bar{f}_{\iota_2}\varphi))(t,x;y) \tag{4.93a}$$
$$+ \sum_{\iota \in J^r} \widetilde{\overline{UV}}_\iota^j(t)(D_0^j(\bar{f}_\iota\varphi))(t,x;y) + \sum_{\iota_1,\iota_2 \in J^r} \widetilde{\overline{UV}}_{\iota_1,\iota_2}^j(t)(D_0^j(\bar{f}_{\iota_1}\bar{f}_{\iota_2}\varphi))(t,x;y) \tag{4.93b}$$
$$+ \sum_{\iota_1,\iota_2,\iota_3 \in J^r} \bar{u}_{\iota_1}^j(t)\,\widetilde{\overline{UV}}_{\iota_2,\iota_3}^j(t)\,(D^j(\bar{f}_{\iota_1}\bar{f}_{\iota_2}\bar{f}_{\iota_3}\varphi))(t,x;y) \tag{4.93c}$$

for every $y \in \mathcal{Y}$, every $t \in \mathrm{dom}(y)$, and every $x \in M$. Because of (4.82), we have derived the following integral expansion for the propagation of smooth functions along trajectories of (4.23).

Proposition 4.23. *For every $j > 0$, every $x_0 \in M$, and all $t_1, t_2 \in [0, T_{x_0}^j)$, we have*

$$\begin{aligned}\varphi(\gamma_{x_0}^j(t_2)) &= \varphi(\gamma_{x_0}^j(t_1)) - \Big[(\bar{D}_1^j\varphi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\Big]_{t=t_1}^{t=t_2} \\ &\quad + \int_{t_1}^{t_2} (f^\infty\varphi)(\gamma_{x_0}^j(t))\,\mathrm{d}t + \int_{t_1}^{t_2} (\bar{D}_2^j\varphi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\,\mathrm{d}t,\end{aligned}$$

where f^∞, D_1^j, and D_2^j are given by (4.27), (4.92), and (4.93), respectively.

Proposition 4.24. *Suppose that the assumptions of Theorem 4.7 are satisfied. Then, for all $\tilde{y}^-, \tilde{y}^+ \in (y^*, y^+)$ with $\tilde{y}^- < \tilde{y}^+$, there exists $c_0 > 0$ such that*

$$(f^\infty\psi)(x) \;\leq\; -c_0$$

for every $x \in \psi^{-1}(\leq \tilde{y}^+, x^*)$ with $\psi(x) \geq \tilde{y}^-$.

Proof. Suppose that the assumptions of Theorem 4.7 are satisfied, and fix arbitrary $\tilde{y}^-, \tilde{y}^+ \in (y^*, y^+)$ with $\tilde{y}^- < \tilde{y}^+$. Then, the set K of all $x \in \psi^{-1}(\leq \tilde{y}^+, x^*)$ with $\psi(x) \geq \tilde{y}^-$ is compact. Moreover, Corollary 4.5 implies that $f^\infty\psi$ only takes negative values on K. Since $f^\infty\psi$ is continuous, we can choose $c_0 > 0$ as the maximum of $-f^\infty\psi$ on K. □

The remainders $\bar{D}_1^j\varphi$ and $\bar{D}_2^j\varphi$ in Proposition 4.23 satisfy the subsequent estimates.

Proposition 4.25. *Let $\varphi \in C^\infty(M)$. Suppose that $f_1, \ldots, f_m$ are compactly supported. Then, there exist $j_0, \bar{c}_1, \bar{c}_2 > 0$ such that*

$$\begin{aligned} \left|(\bar{D}_1^j\varphi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\right| &\leq \bar{c}_1\, \bar{j}^{-\frac{1}{2}}, \\ \left|(\bar{D}_2^j\varphi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\right| &\leq \bar{c}_2\, \bar{j}^{-\frac{1}{2r}} \end{aligned}$$

for every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$.

Proof. Let $\mathcal{F}$ be the finite set of all the functions φ, $\bar{f}_{\iota_1}\varphi$, $\bar{f}_{\iota_1}\bar{f}_{\iota_2}\varphi$, and $\bar{f}_{\iota_1}\bar{f}_{\iota_2}\bar{f}_{\iota_3}\varphi$ with $\iota_1, \iota_2, \iota_3 \in J^r$. We already know that the differential operators D_1^j and $\widehat{D}_2^j$ satisfy the estimates in Proposition 4.16 and Lemma 4.21 for every smooth function on M. Thus, we can find sufficiently large $j_0, c_1, \widehat{c}_2 > 0$ such that

$$\left|(D_1^j\phi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\right| \leq c_1\, \bar{j}^{-1-\frac{1}{2r}} \qquad \text{and} \qquad \left|(\widehat{D}_2^j\phi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\right| \leq \widehat{c}_2\, \bar{j}^{-\frac{1}{2r}}$$

for every $\phi \in \mathcal{F}$, every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$. Since the vector fields $f_1, \ldots, f_m$ are assumed to be compactly supported, also the vector fields $\bar{f}_\iota$, $\iota \in J^r$, are compactly supported. Using the definition of D^j in (4.87) it follows that there exists $c > 0$ such that

$$\left|(D^j\phi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\right| \leq c$$

for every $\phi \in \mathcal{F} \setminus \{\varphi\}$, every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$. Using Lemma 4.17 and the definition of D_0^j in (4.88) it follows that there exists $c_0 > 0$ such that

$$\left|(D_0^j\phi)(t, \gamma_{x_0}^j(t); y_{x_0}^j)\right| \leq c_0\, \bar{j}^{-\frac{1}{2r}}$$

for every $\phi \in \mathcal{F}$, every $j \geq j_0$, every $x_0 \in M$, and every $t \geq 0$. From Lemma 4.17, we know estimates for the sinusoids $\bar{u}_\iota^j$, $\widetilde{\overline{UV}}{}_\iota^j$, and $\widetilde{\overline{UV}}{}_{\iota_1,\iota_2}^j$ in (4.92) and (4.93). Now we have suitable estimates for all constituents of $\bar{D}_1^j\varphi$ and $\bar{D}_2^j\varphi$ to conclude that the statement is true. □

The two-step averaging procedure is complete. Theorem 4.6 follows immediately from Propositions 4.23 and 4.25.

4.7.4 Proof of Theorem 4.7

Suppose that the assumptions of Theorem 4.7 are satisfied with certain $x^* \in M$ and $y^* < y^+ \leq +\infty$ as therein. Fix arbitrary small $\varepsilon, \delta > 0$ and an arbitrary large $\tilde{y} \in (y^*, y^+)$. After possibly shrinking ε, we may assume that $\tilde{y}^+ := \tilde{y} + \varepsilon < y^+$. Then, the set $K := \psi^{-1}(\leq \tilde{y}^+, x^*)$ is compact. After multiplication by a suitable smooth bump function[10], we may suppose that the vector fields $f_1, \ldots, f_m$ are compactly supported and that they

[10] It is known from differential geometry [62] that, for every compact subset K of a smooth manifold M, there exists a compactly supported smooth function on M that is identically equal to 1 on K. Such a function is usually referred to as a *bump function*.

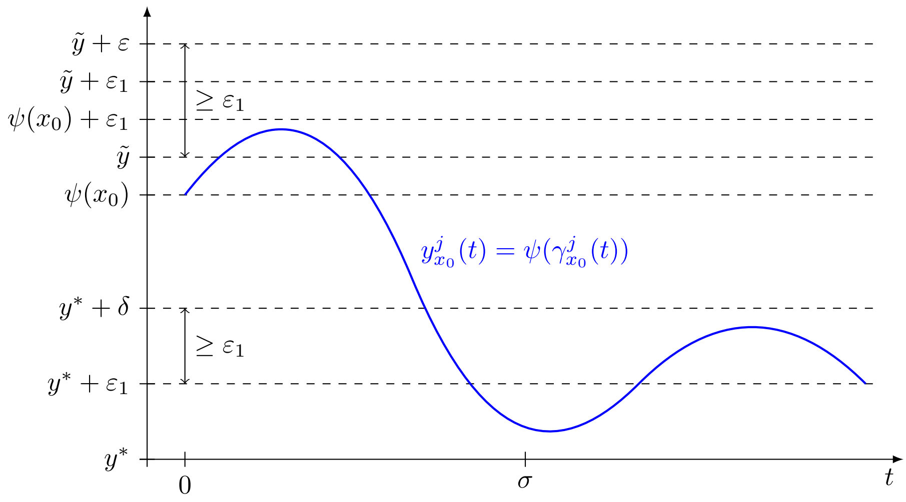


Figure 4.2: *Illustration of the parameters and functions in Subsection 4.7.4.*

coincide with the initially given control vector fields on K. After possibly shrinking δ, we may assume that $y^* + \delta < \tilde{y}$. Let $\varepsilon_1 := \min\{\varepsilon, \delta/2\}$ and $\tilde{y}^- := y^* + \varepsilon_1$. Let K' denote the compact set of all $x \in K$ with $\psi(x) \geq \tilde{y}^-$. Then, there exists $c_0 > 0$ as in Proposition 4.24. Let $\varepsilon_2 := c_0/2$. We conclude from Theorem 4.6 with $\varphi = \psi$ that there exists $j_0 > 0$ such that, for every $j \geq j_0$, every $x_0 \in K$, and all $t_2 > t_1 \geq 0$, the following implication holds: if $\gamma^j_{x_0}(t) \in K'$ for every $t \in [t_1, t_2]$, then

$$\psi(\gamma^j_{x_0}(t_2)) \;\leq\; \psi(\gamma^j_{x_0}(t_1)) + \varepsilon_1 - \varepsilon_2\,(t_2 - t_1).$$

It is now easy to see that the above inequality and the choice of $\varepsilon_1, \varepsilon_2$ imply the asserted statements on stability, boundedness, and attraction in Theorem 4.7 with $\sigma := (\tilde{y} - y^*)/\varepsilon_2$. In particular, every solution of (4.23) that starts in $\psi^{-1}(\leq \tilde{y}, x^*)$ stays in the compact set K on which the vector fields $f_1, \ldots, f_m$ coincide with the initially given control vector fields.

Possible directions for future research

Over the past years, approximations of Lie brackets have been successfully used to solve a variety of optimization problems. However, there are still many promising directions that should be investigated in the future. The following paragraphs collect some of these directions.

Nowadays, modern control systems are highly interconnected and involve analog and digital components. This goes along with a range of novel challenges such as sampled measurements, quantized inputs, discrete-time systems or local information. However, a major limitation of many extremum seeking control schemes, including the Lie bracket approach, is that they are not capable of dealing with these situations in an integrative manner. By further developing the Lie bracket approach, a long-term vision would be a general framework for analyzing and designing extremum seeking algorithms for such modern interconnected control systems consisting of digital and analog components. This would allow completely novel areas of application, which are not feasible with the existing state of the art extremum seeking schemes.

Closely related to the directions in the previous paragraph is the study of robustness against disturbances. Most of the existing theoretical studies on extremum seeking control assume that the output signal can be sensed accurately at any given time and that the control law is fed into the system without any disturbances. In particular, this is required for the existing extremum seeking results obtained by Lie bracket approximations. However, in many real-world applications, such ideal conditions are not satisfied. The question of robustness arises naturally, for example, when the output measurements are corrupted by noise. Moreover, any deviation from a prescribed continuous-time control law due to a digital implementation can be interpreted as a disturbance. An investigation of robustness properties of the Lie bracket approach could lead to helpful guidelines in practical implementations of such a control strategy.

The majority of the existing extremum seeking strategies by means of Lie bracket approximations require a first-order kinematic control system, meaning that the first-order time derivative of the system state can be directly controlled through the input channels. However, many extremum seeking control problems involve higher-order control systems such as mechanical systems or integrator chains. For example, in applications to source seeking with an acceleration-controlled robot. It is therefore desirable to extend the Lie bracket approach to a larger class of non-kinematic models. The known Lie bracket-based methods lead to rapidly increasing velocities with increasing frequency parameter. A promising goal would be a less invasive control strategy, which ensures bounded velocities even in the high-frequency limit. This in turn can lead to a reduced vulnerability to disturbances and to an improved performance in case of digital implementations.

Another promising (but very challenging) direction would be an extension to non-smooth and infinite-dimensional optimization problems. This includes extremum seeking control for systems modeled by partial differential equations. The existing Lie bracket-based methods only give access to finite-dimensional subspaces of descent directions and their stability analysis heavily relies on a certain degree of smoothness. As a first step, one could investigate the behavior of closed-loop systems with a certain class of non-smooth output functions on finite dimensional state spaces. This probably requires suitable mathematical tools, like the subdifferential, from non-smooth optimization.

Bibliography

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer, New York, second edition, 1988.

[2] V. Adetola and M. Guay. Parameter convergence in adaptive extremum-seeking control. *Automatica*, 43(1):105–110, 2007.

[3] A. A. Agrachev and Y. L. Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer, Berlin, 2004.

[4] B. D. O. Anderson and C. Yu. Range-only sensing for formation shape control and easy sensor network localization. In *Proceedings of the 2011 Chinese Control and Decision Conference*, pages 3310–3315, 2011.

[5] B. D. O. Anderson, C. Yu, B. Fidan, and J. M. Hendrickx. Rigid graph control architectures for autonomous formations. *IEEE Control Systems*, 28(6):48–63, 2008.

[6] K. B. Ariyur and M. Krstić. *Real-Time Optimization by Extremum-Seeking Control*. John Wiley & Sons, Inc., Hoboken, NJ, 2003.

[7] L. Asimow and B. Roth. The rigidity of graphs. *Transactions of the American Mathematical Society*, 245:279–289, 1978.

[8] L. Asimow and B. Roth. The rigidity of graphs, II. *Journal of Mathematical Analysis and Applications*, 68(1):171–190, 1979.

[9] K. J. Åström and B. Wittenmark. *Adaptive Control*. Dover Publications, Inc., New York, second edition, 2008. revised and corrected printing.

[10] M. Benosman. *Learning-Based Adaptive Control*. Butterworth-Heinemann, Oxford, 2016.

[11] A. S. Besicovitch. On the linear independence of fractional powers of integers. *Journal of the London Mathematical Society*, 15(1):3–6, 1940.

[12] P. F. Blackman. Extremum-seeking regulators. In J. H. Westcott, editor, *An Exposition of Adaptive Control*, pages 36–50. Pergamon Press, Oxford, 1962.

[13] N. Bourbaki. *Algebra I*. Elements of Mathematics. Springer, Berlin, 1989.

[14] N. Bourbaki. *Lie Groups and Lie Algebras I*. Elements of Mathematics. Springer, Berlin, 1989.

[15] R. W. Brockett. Asymptotic stability and feedback stabilization. In R. W. Brockett, R. S. Millman, and H. J. Sussmann, editors, *Differential Geometric Control Theory*, volume 27 of *Progress in Mathematics*, pages 181–191. Birkhäuser, Boston, 1983.

[16] S. L. Brunton, C. W. Rowley, S. R. Kulkarni, and C. Clarkson. Maximum power point tracking for photovoltaic optimization using ripple-based extremum seeking control. *IEEE Transactions on Power Electronics*, 25(10):2531–2540, 2010.

[17] F. Bullo. Averaging and vibrational control of mechanical systems. *SIAM Journal on Control and Optimization*, 41(2):542–562, 2002.

[18] F. Bullo and A. D. Lewis. *Geometric Control of Mechanical Systems*, volume 49 of *Texts in Applied Mathematics*. Springer, New York, 2005.

[19] M. Cao, C. Yu, and B. D. O. Anderson. Formation control using range-only measurements. *Automatica*, 47(4):776–781, 2011.

[20] Y. Cao, W. Yu, W. Ren, and G. Chen. An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Transactions on Industrial informatics*, 9(1):427–438, 2013.

[21] Z. Cao, H.-B. Dürr, C. Ebenbauer, F. Allgöwer, and F. Gao. Iterative learning and extremum seeking for repetitive time-varying mappings. *IEEE Transactions on Automatic Control*, 62(7):3339–3353, 2017.

[22] K.-T. Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.

[23] J.-Y. Choi, M. Krstić, K. B. Ariyur, and J. S. Lee. Extremum seeking control for discrete-time systems. *IEEE Transactions on Automatic Control*, 47(2):318–323, 2002.

[24] F. Dörfler and B. A. Francis. Formation control of autonomous robots based on cooperative behavior. In *Proceedings of the 2009 European Control Conference*, pages 2432–2437, 2009.

[25] E. Durand-Cartagena and J. A. Jaramillo. Pointwise Lipschitz functions on metric spaces. *Journal of Mathematical Analysis and Applications*, 363(2):525–548, 2010.

[26] H.-B. Dürr. Distributed Positioning of Autonomous Mobile Sensors with Application to the Coverage Problem. Master's thesis, Royal Institute of Technology, Stockholm, Sweden and University of Stuttgart, Stuttgart, Germany, 2010.

[27] H.-B. Dürr, M. Krstić, A. Scheinker, and C. Ebenbauer. Singularly perturbed Lie bracket approximation. *IEEE Transactions on Automatic Control*, 60(12):3287–3292, 2015.

[28] H.-B. Dürr, M. Krstić, A. Scheinker, and C. Ebenbauer. Extremum seeking for dynamic maps using Lie brackets and singular perturbations. *Automatica*, 83:91–99, 2017.

[29] H.-B. Dürr, M. S. Stanković, D. V. Dimarogonas, C. Ebenbauer, and K. H. Johansson. Obstacle avoidance for an extremum seeking system using a navigation function. In *Proceedings of the 2013 American Control Conference*, pages 4062–4067, 2013.

[30] H.-B. Dürr, M. S. Stanković, C. Ebenbauer, and K. H. Johansson. Examples of distance-based synchronization: An extremum seeking approach. In *Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing*, pages 366–373, 2013.

[31] H.-B. Dürr, M. S. Stanković, C. Ebenbauer, and K. H. Johansson. Lie bracket approximation of extremum seeking systems. *Automatica*, 49(6):1538–1552, 2013.

[32] H.-B. Dürr, M. S. Stanković, and K. H. Johansson. A Lie bracket approximation for extremum seeking vehicles. In *Proceedings of the 18th IFAC World Congress*, pages 11393–11398, 2011.

[33] H.-B. Dürr, M. S. Stanković, and K. H. Johansson. Distributed positioning of autonomous mobile sensors with application to coverage control. In *Proceedings of the 2011 American Control Conference*, pages 4822–4827, 2011.

[34] H.-B. Dürr, M. S. Stanković, K. H. Johansson, and C. Ebenbauer. Extremum seeking on submanifolds in the Euclidean space. *Automatica*, 50(10):2591–2596, 2014.

[35] J. Feiling, A. Zeller, and C. Ebenbauer. Derivative-free optimization algorithms based on non-commutative maps. *IEEE Control Systems Letters*, 2(4):743–748, 2018.

[36] M. Fliess. Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bulletin de la Société Mathématique de France*, 109:3–40, 1981.

[37] A. Friedman. *Stochastic Differential Equations and Applications*, volume 1. Academic Press, New York, 1975.

[38] H. Gluck. Almost all simply connected closed surfaces are rigid. In L. C. Glaser and T. B. Rushing, editors, *Geometric Topology*, pages 225–239. Springer, Berlin, 1975.

[39] V. Grushkovskaya, H.-B. Dürr, C. Ebenbauer, and A. Zuyev". Extremum seeking for time-varying functions using Lie bracket approximations. In *Proceedings of the 20th IFAC World Congress*, pages 5522–5528, 2017.

[40] V. Grushkovskaya, S. Michalowsky, A. Zuyev, M. May, and C. Ebenbauer. A family of extremum seeking laws for a unicycle model with a moving target: Theoretical and experimental studies. In *Proceedings of the 2018 European Control Conference*, pages 1–6, 2018.

[41] V. Grushkovskaya, A. Zuyev, and C. Ebenbauer. On a class of generating vector fields for the extremum seeking problem: Lie bracket approximation and stability properties. *Automatica*, 94:151–160, 2018.

[42] M. Guay. A time-varying extremum-seeking control approach for discrete-time systems. *Journal of Process Control*, 24(3):98–112, 2014.

[43] M. Guay and D. Dochain. A time-varying extremum-seeking control approach. *Automatica*, 51:356–363, 2015.

[44] M. Guay and T. Zhang. Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. *Automatica*, 39(7):1283–1293, 2003.

[45] C. M. Hackl, N. Hopfe, A. Ilchmann, M. Mueller, , and S. Trenn. Funnel control for systems with relative degree two. *SIAM Journal on Control and Optimization*, 51(2):965–995, 2013.

[46] N. W. J. Hazelton and R. B. Buckner. Distance measurements. In R. C. Dorf, editor, *The Engineering Handbook*, chapter 163. CRC Press, Boca Raton, second edition, 2004.

[47] J. Heinonen. Nonsmooth calculus. *Bulletin of the American Mathematical Society*, 44(2):163–232, 2007.

[48] U. Helmke, S. Mou, Z. Sun, and B. D. O. Anderson. Geometrical methods for mismatched formation control. In *Proceedings of the 53rd Annual Conference on Decision and Control*, pages 1341–1346, 2014.

[49] A. Ilchmann and H. Schuster. PI-funnel control for two mass systems. *IEEE Transactions on Automatic Control*, 54(4):918–923, 2009.

[50] B. Jakubczyk. Introduction to geometric nonlinear control; controllability and Lie bracket. In A. A. Agrachev, editor, *Mathematical Control Theory*, volume 8 of *ICTP Lecture Notes Series*, pages 107–168. Abdus Salam International Centre for Theoretical Physics, Trieste, 2002.

[51] B. Jiang, M. Deghat, and B. D. O. Anderson. Simultaneous velocity and position estimation via distance-only measurements with application to multi-agent system control. *IEEE Transactions on Automatic Control*, 62(2):869–875, 2017.

[52] S. Z. Khong, D. Nešić, C. Manzie, and Y. Tan. Multidimensional global extremum seeking via the DIRECT optimisation algorithm. *Automatica*, 49(7):1970–1978, 2013.

[53] N. J. Killingsworth and M. Krstić. PID tuning using extremum seeking: Online, model-free performance optimization. *IEEE Control Systems*, 26(1):70–79, 2006.

[54] L. Krick, M. E. Broucke, and B. A. Francis. Stabilization of infinitesimally rigid formations of multi-robot networks. *International Journal of Control*, 82(3):423–439, 2009.

[55] M. Krstić and H.-H. Wang. Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, 36(4):595–601, 2000.

[56] J. Kurzweil and J. Jarník. Limit processes in ordinary differential equations. *Journal of Applied Mathematics and Physics*, 38(2):241–256, 1987.

[57] J. Kurzweil and J. Jarník. A convergence effect in ordinary differential equations. In V. S. Korolyuk, editor, *Asymptotic Methods of Mathematical Physics*, pages 134–144. Naukova Dumka, Kiev, 1988.

[58] J. Kurzweil and J. Jarník. Iterated Lie brackets in limit processes in ordinary differential equations. *Results in Mathematics*, 14(1-2):125–137, 1988.

[59] C. Labar, E. Garone, M. Kinnaert, and C. Ebenbauer. Newton-based extremum seeking: A second-order Lie bracket approximation approach. *Automatica*, 105:356–367, 2019.

[60] V. Lakshmikantham, S. Leela, and A. A. Martynyuk. *Practical Stability of Nonlinear Systems*. World Scientific, Signapore, 1990.

[61] M. Leblanc. Sur l'électrification des chemins de fer au moyen de courants alternatifs de fréquence élevée. *Revue Générale de l'Électricité*, 7(8):275–277, 1922.

[62] J. M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2012.

[63] R. Leyva, C. Alonso, I. Queinnec, A. Cid-Pastor, D. Lagrange, and L. Martinez-Salamero. MPPT of photovoltaic systems using extremum-seeking control. *IEEE Transactions on Aerospace and Electronic Systems*, 42(1):249–258, 2006.

[64] Z. Li and J. F. Canny. *Nonholonomic Motion Planning*, volume 192 of *The Springer International Series in Engineering and Computer Science*. Springer, New York, 1992.

[65] S.-J. Liu and M. Krstić. *Stochastic Averaging and Stochastic Extremum Seeking*. Communications and Control Engineering. Springer, London, 2012.

[66] S.-J. Liu and M. Krstić. Stochastic averaging in discrete time and its applications to extremum seeking. *IEEE Transactions on Automatic Control*, 61(1):90–102, 2016.

[67] W. Liu. *Averaging Theorems for Highly Oscillatory Differential Equations and the Approximation of General Paths by Admissible Trajectories for Nonholonomic Systems*. Ph.D. thesis, Rutgers University, 1992.

[68] W. Liu. An approximation algorithm for nonholonomic systems. *SIAM Journal on Control and Optimization*, 35(4):1328–1365, 1997.

[69] W. Liu. Averaging theorems for highly oscillatory differential equations and iterated Lie brackets. *SIAM Journal on Control and Optimization*, 35(6):1989–2020, 1997.

[70] S. Michalowsky and C. Ebenbauer. The multidimensional n-th order heavy ball method and its application to extremum seeking. In *Proceedings of the 53rd IEEE Conference on Decision and Control*, pages 2660–2666, 2014.

[71] S. Michalowsky and C. Ebenbauer. Model-based extremum seeking for a class of nonlinear systems. In *Proceedings of the 2015 American Control Conference*, pages 2026–2031, 2015.

[72] S. Michalowsky, B. Gharesifard, and C. Ebenbauer. Distributed extremum seeking over directed graphs. In *Proceedings of the 56th IEEE Conference on Decision and Control*, pages 2095–2101, 2017.

[73] S. Michalowsky, B. Gharesifard, and C. Ebenbauer. On the Lie bracket approximation approach to distributed optimization: Extensions and limitations. In *Proceedings of the 2018 European Control Conference*, pages 119–124, 2018.

[74] J. W. Milnor. *Morse Theory*. Princeton University Press, Princeton, New Jersey, revised edition, 1973.

[75] J. M. Montenbruck, H.-B. Dürr, C. Ebenbauer, and F. Allgöwer. Extremum seeking and obstacle avoidance on the special orthogonal group. In *Proceedings of the 19th IFAC World Congress*, volume 47, pages 8229–8234, 2014.

[76] J. M. Montenbruck, H.-B. Dürr, C. Ebenbauer, and F. Allgöwer. Extremum seeking with drift. In *Proceedings of the 1st IFAC Conference on Modelling, Identification and Control of Nonlinear Systems*, volume 48-11, pages 126–130, 2015.

[77] L. Moreau and D. Aeyels. Practical stability for systems depending on a small parameter. In *Proceedings of the 37th IEEE Conference on Decision and Control*, volume 2, pages 1428–1433, 1998.

[78] L. Moreau and D. Aeyels. Asymptotic methods in stability analysis and control. In D. Aeyels, F. Lamnabhi-Lagarrigue, and A. Schaft, editors, *Stability and Stabilization of Nonlinear Systems*, volume 246 of *Lecture Notes in Control and Information Sciences*, pages 201–213. Springer, London, 1999.

[79] L. Moreau and D. Aeyels. Trajectory-based local approximations of ordinary differential equations. *SIAM Journal on Control and Optimization*, 41(6):1922–1945, 2003.

[80] P. Morin, J.-B. Pomet, and C. Samson. Design of homogeneous time-varying stabilizing control laws for driftless controllable systems via oscillatory approximation of Lie brackets in closed loop. *SIAM Journal on Control and Optimization*, 38(1):22–49, 1999.

[81] P. Morin, J.-B. Pomet, and C. Samson. Exponential stabilization of nonlinear driftless systems with robustness to unmodeled dynamics. *ESAIM: Control, Optimisation and Calculus of Variations*, 4:1–35, 1999.

[82] S. Mou, M.-A. Belabbas, A. S. Morse, Z. Sun, and B. D. O. Anderson. Undirected rigid formations are problematic. *IEEE Transactions on Automatic Control*, 61(10):2821–2836, 2016.

[83] R. M. Murray and S. S. Sastry. Nonholonomic motion planning: Steering using sinusoids. *IEEE Transactions on Automatic Control*, 38(5):700–716, 1993.

[84] W. Noll. *Finite-Dimensional Spaces*, volume 10 of *Mechanics: Analysis*. Springer, Dordrecht, 1987.

[85] K. K. Oh and H. S. Ahn. Distance-based undirected formations of single-integrator and double-integrator modeled agents in n-dimensional space. *International Journal of Robust and Nonlinear Control*, 24(12):1809–1820, 2014.

[86] K.-K. Oh, M.-C. Park, and H.-S. Ahn. A survey of multi-agent formation control. *Automatica*, 53:424–440, 2015.

[87] Y. Pan, U. Özgüner, and T. Acarman. Stability and performance improvement of extremum seeking control with sliding mode. *International Journal of Control*, 76(9-10):968–985, 2003.

[88] K. S. Peterson and A. G. Stefanopoulou. Extremum seeking control for soft landing of an electromechanical valve actuator. *Automatica*, 40(6):1063–1069, 2004.

[89] D. Popovic, M. Jankovic, S. Magner, and A. R. Teel. Extremum seeking methods for optimization of variable cam timing engine operation. *IEEE Transactions on Control Systems Technology*, 14(3):398–407, 2006.

[90] M. A. Rotea. Analysis of multivariable extremum seeking algorithms. In *Proceedings of the 2000 American Control Conference*, volume 1, pages 433–437, 2000.

[91] A. Scheinker. Bounded extremum seeking for angular velocity actuated control of nonholonomic unicycle. *Optimal Control Applications and Methods*, 38(4):575–585, 2017.

[92] A. Scheinker, M. Bland, M. Krstić, and J. Audia. Extremum seeking-based optimization of high voltage converter modulator rise-time. *IEEE Transactions on Control Systems Technology*, 22(1):34–43, 2014.

[93] A. Scheinker and M. Krstić. A universal extremum seeking-based stabilizer for unknown LTV systems with unknown control directions. In *Proceedings of the 2006 American Control Conference*, pages 1129–1136, 2012.

[94] A. Scheinker and M. Krstić. Extremum seeking-based tracking for unknown systems with unknown control directions. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 6065–6070, 2012.

[95] A. Scheinker and M. Krstić. Minimum-seeking for CLFs: Universal semiglobally stabilizing feedback under unknown control directions. *IEEE Transactions on Automatic Control*, 58(5):1107–1122, 2013.

[96] A. Scheinker and M. Krstić. Extremum seeking with bounded update rates. *Systems & Control Letters*, 63:25–31, 2014.

[97] A. Scheinker and M. Krstić. Non-C^2 Lie bracket averaging for nonsmooth extremum seekers. *Journal of Dynamic Systems, Measurements, and Control*, 136:011010, 2014.

[98] A. Scheinker and M. Krstić. *Model-Free Stabilization by Extremum Seeking.* Springer Briefs in Control, Automation and Robotics. Springer, Cham, 2017.

[99] A. Scheinker and D. Scheinker. Bounded extremum seeking with discontinuous dithers. *Automatica*, 69:250–257, 2016.

[100] A. Scheinker and D. Scheinker. Constrained extremum seeking stabilization of systems not affine in control. *International Journal of Robust and Nonlinear Control*, 28(2):568–581, 2018.

[101] E. D. Sontag. *Mathematical Control Theory*, volume 6 of *Texts in Applied Mathematics.* Springer, New York, second edition, 1998.

[102] J. Sternby. Extremum control systems: An area for adaptive control? In *Proceedings of the Joint Automatic Control Conference*, pages : Paper WA2–A, 1980.

[103] Z. Sun, S. Mou, B. D. O. Anderson, and M. Cao. Exponential stability for formation control systems with generalized controllers: A unified approach. *Systems & Control Letters*, 93:50–57, 2016.

[104] H. J. Sussmann. A product expansion for the Chen series. In C. I. Byrnes and A. Lindquist, editors, *Theory and Applications of Nonlinear Control Systems*, pages 323–335. North-Holland, New York, 1986.

[105] H. J. Sussmann. A general theorem on local controllability. *SIAM Journal on Control and Optimization*, 25(1):158–194, 1987.

[106] H. J. Sussmann. An introduction to the coordinate-free maximum principle. In B. Jakubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, number 207 in Lecture notes in pure and applied mathematics, chapter 12, pages 463–557. Marcel Dekker, New York, 1998.

[107] H. J. Sussmann and W. Liu. Limits of highly oscillatory controls and the approximation of general paths by admissible trajectories. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 437–442, 1991.

[108] R. Suttner. Extremum seeking control for an acceleration controlled unicycle. In *Proceedings of the 11th IFAC Symposium on Nonlinear Control Systems*, pages 676–681, 2019.

[109] R. Suttner. Extremum seeking control with an adaptive dither signal. *Automatica*, 101:214–222, 2019.

[110] R. Suttner. Extremum seeking control for a class of nonholonomic systems. *SIAM Journal on Control and Optimization*, 58(4):2588–2615, 2020.

[111] R. Suttner and S. Dashkovskiy. Exponential stability for extremum seeking control systems. In *Proceedings of the 20th IFAC World Congress*, pages 15464–15470, 2017.

[112] R. Suttner and M. Krstić. Acceleration-actuated source seeking without position and velocity sensing. In *Proceedings of the 21th IFAC World Congress*, page to appear, 2020.

[113] R. Suttner and Z. Sun. Formation shape control based on distance measurements using Lie bracket approximations. *SIAM Journal on Control and Optimization*, 56(6):4405–4433, 2018.

[114] R. Suttner and Z. Sun. Exponential and practical exponential stability of second-order formation control systems. In *Proceedings of the 58th IEEE Conference on Decision and Control*, pages 3521–3526, 2019.

[115] Y. Tan, W. H. Moase, C. Manzie, D. Nešić, and I. M. Y. Mareels. Extremum seeking from 1922 to 2010. In *Proceedings of the 29th Chinese Control Conference*, pages 14–26, 2010.

[116] Y. Tan, D. Nešić, and I. Mareels. On non-local stability properties of extremum seeking control. *Automatica*, 42(6):889–903, 2006.

[117] A. R. T. Teel and D. Popović. Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming. In *Proceedings 2001 American Control Conference*, pages 2394–2399, 2001.

[118] C. Yin, Y. Chen, and S.-M. Zhong. Fractional-order sliding mode based extremum seeking control of a class of nonlinear systems. *Automatica*, 50(12):3173–3181, 2014.

[119] C. Zhang and R. Ordóñez. Numerical optimization-based extremum seeking control with application to ABS design. *IEEE Transactions on Automatic Control*, 52(3):454–467, 2007.

[120] C. Zhang and R. Ordóñez. *Extremum-Seeking Control and Applications.* Advances in Industrial Control. Springer, London, 2012.

[121] A. Zuyev. Exponential stabilization of nonholonomic systems by means of oscillating controls. *SIAM Journal on Control and Optimization*, 54(3):1678–1696, 2016.