Hyperbolicity & Invariant Manifolds for Finite-Time Processes

Dissertation
zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

Dipl.-Math. Daniel Karrasch
geboren am 09.06.1982 in Berlin

Gutachter:
Prof. Dr. rer. nat. habil. Stefan Siegmund
Prof. Dr. Kenneth J. Palmer

Eingereicht am: 25.06.2012
Tag der Disputation: 27.09.2012

To the many beautiful girls in my life

Acknowledgments

I would like to thank Prof. Dr. Stefan Siegmund for suggesting the research direction of finite-time dynamics and for supervising my work. I express my gratitude to Prof. Dr. Ken Palmer for writing the second review on my PhD thesis.

The work on my Ph.D. has been supported by the European Social Fund via a grant "Landesinnovationspromotion No. 080942988", which gave me the opportunity to do research, to travel to workshops and conferences and in turn to meet many inspiring people, and, last but not least, to support my family. This financial support is most gratefully acknowledged.

I am indefinitely indebted to the following colleagues: Maik Gröger for being a never-dwindling source of challenging questions; Marcus Köhler for his unbelievable proficiency in matrix theory; Sascha Trostorff and Marcus Waurick for their unstoppable strive for clarity and mathematical beauty. With all of them and with Prof. Dr. Jürgen Voigt, Prof. Dr. Rainer Picard, Jun.-Prof. Dr. Kathrin Padberg-Gehle, Luu Hoang Duc, Gabriel Fuhrmann and Martin Rasmussen I had innumerous enlightening discussions on scientific and non-scientific topics, for which I am very thankful. I am grateful to Maik Gröger and Marcus Waurick once more for proofreading the whole manuscript and for making several suggestions that have led to improvements in the manuscript.

Furthermore, I would like to thank the whole Institute of Analysis at the TU Dresden and especially the "Kaffeerunde" for creating such a pleasant and funny ambiance. Also, I am indebted to the two secretaries Mrs. Gaede-Samat and Mrs. Heinke for taking much load off of me.

I am grateful to my parents and my sister for their support and for their love. I would like to thank my beloved wife Julia and our two daughters Kristina and Wolke for just being there . . . with all its consequences.

Contents

Nomenclature

$\mathbb{N}$	set of non-negative integers
$\mathbb{R}$	set of reals
$\mathbb{R}_{>a}$, $\mathbb{R}_{\geq b}$	$\{x \in \mathbb{R};\ x > a\}$, $\{x \in \mathbb{R};\ x \geq b\}$
$\mathbb{I}$	compact subset of the real numbers, time-set
t_-	$\min \mathbb{I}$
t_+	$\max \mathbb{I}$
$\neq_{\mathbb{I}\times\mathbb{I}}$	unequal-relation on $\mathbb{I} \times \mathbb{I}$
$\mathbb{C}$	set of complex numbers
$\mathcal{S}$	unit sphere in $\mathbb{R}^n$ with respect to $\lvert\cdot\rvert$
$\mathbb{P}^{n-1}$	real projective space
2^M	power set of M
id_M	identity function on M
$\mathcal{K}(M)$	compact subsets of M
d_{H}	Hausdorff metric
$\lvert\cdot\rvert$, $\lVert\cdot\rVert$	norm
$\lVert\cdot\rVert_\infty$	supremum norm with respect to time
$B(x, \delta)$	open ball around x with radius δ
$B[x, \delta]$	closed ball around x with radius δ

$L(\mathbb{R}^n)$	set of linear operators from $\mathbb{R}^n$ to $\mathbb{R}^n$
$GL(n, \mathbb{R})$	set of invertible operators in $L(\mathbb{R}^n)$
Sym_n	symmetric operators in $L(\mathbb{R}^n)$
$\mathbb{P}(\mathbb{R}^n)$	linear projections on $\mathbb{R}^n$
$\operatorname{im} P$, $\ker P$	image and null space of $P \in L(\mathbb{R}^n)$
$U \oplus V$	direct sum of the vector spaces U and V
$\otimes$	(formal) Kronecker product
$[x]_\sim$	equivalence class of x
$C(I, \mathbb{R}^n)$	continuous functions from $\mathbb{I}$ to $\mathbb{R}^n$
$C^{0,2}(I \times \mathbb{R}^n, \mathbb{R}^n)$	continuous functions which are twice continuously differentiable in the second argument
$\lvert f\rvert_{\mathrm{Lip}}$	Lipschitz constant of f
$\partial_j g(x)$	partial derivative of g with respect to the $j+1$-th variable at x
$\Delta_{\mathbb{I}}(f)(t,s)$, $\Delta(f)(t,s)$	logarithmic difference quotient of f at t and s
grad	gradient
$\mathcal{P}(\mathbb{I}, \mathbb{R}^n)$	set of invertible processes on $\mathbb{I}$ with state space $\mathbb{R}^n$
$\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$	set of linear invertible processes on $\mathbb{I}$ with state space $\mathbb{R}^n$
$\tilde{d}_{\mathbb{I}}$, $d_{\mathbb{I}}$	(semi-)metric on $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$
$\underline{\lambda}^{\mathbb{I}}(X, \Phi)$, $\underline{\lambda}(X, \Phi)$	lower growth rate of X under Φ
$\overline{\lambda}^{\mathbb{I}}(X, \Phi)$, $\overline{\lambda}(X, \Phi)$	upper growth rate of X under Φ
$\underline{\lambda}_k^{\mathbb{I}}(\Phi)$, $\underline{\lambda}_k(\Phi)$	maximal lower k-growth rate of Φ
$\overline{\lambda}_k^{\mathbb{I}}(\Phi)$, $\overline{\lambda}_k(\Phi)$	minimal upper k-growth rate of Φ

$\Sigma^{\mathbb{I}}(\Phi), \Sigma(\Phi)$	(finite-time dichotomy) spectrum
$\varrho(\Phi)$	(finite-time) resolvent set
$V^{\mathrm{s}}, V^{\mathrm{u}}$	stable and unstable cone function
W_y^{s}	domain of (finite-time) attraction w.r.t. y
W_y^{u}	domain of (finite-time) repulsion w.r.t. y
$\mathcal{V}_\Phi^{\mathrm{s}}$	extended stable cone
$\mathcal{V}_\Phi^{\mathrm{u}}$	extended unstable cone
$\mathcal{W}_k(\Phi)$	k-th spectral cone of Φ
$\mathcal{W}_0^{\mathrm{s}}$	extended domain of attraction
$\mathcal{W}_0^{\mathrm{u}}$	extended domain of repulsion
$\underline{\mu}, \overline{\mu}$	lower/upper "nonlinear" growth rates
ρ	repulsion rate
ν	repulsion ratio
λ_i	i-th eigenvalue of the (left) Cauchy-Green strain tensor
$v_1, \dots, v_n$	orthonormal eigenbasis of the (left) Cauchy-Green strain tensor
Σ_{FTL}	finite-time Lyapunov spectrum
σ	(maximal) finite-time Lyapunov exponent
$\vec{\sigma}(C)$	a vectorization of the spectrum of C
$\mathrm{Gr}(k, \mathbb{R}^n)$	Grassmann manifold of k-dimensional subspaces in $\mathbb{R}^n$
$\mathrm{St}(k, \mathbb{R}^n)$	Stiefel manifold of k-frames in $\mathbb{R}^n$
$T_x M$	tangent space of M at x

$T_x^{\perp}M$	normal space of M at x
∇	affine connection, Riemannian connection
$\exp_p$	exponential map at p
$\mathcal{L}$	Lie derivative

List of Figures

List of Algorithms

1 Introduction

The aim of this thesis is to introduce a general framework for what is informally referred to as *finite-time dynamics*. Within this framework, we study *hyperbolicity of reference trajectories*, *existence of invariant manifolds* as well as *normal hyperbolicity of invariant manifolds* called *Lagrangian Coherent Structures*. We focus on a simple derivation of analytical results. At the same time, our approach together with the analytical results has strong impact on the numerical implementation by providing calculable expressions for known functions and continuity results that ensure robust computation.

In this work, we consider evolutionary processes on $\mathbb{R}^n$ on arbitrary compact time-sets. The processes can be generated by ordinary differential equations (ODEs) on compact time-intervals, i.e.

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

where, for instance, $f\colon I \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and $I = [t_-, t_+] \subset \mathbb{R}$. Since we consider the dynamics on an arbitrary compact time-set $\mathbb{I} \subseteq I$, the evolutionary process can equally well be generated by a difference equation or, in general, by a dynamic equation on the bounded *time scale* $\mathbb{I}$. The need to analyze such processes arises in many applications such as transport problems in fluid, ocean or atmosphere dynamics, but also increasingly in biological applications; see [88] for a recent review on the first and [5, 96] for the latter. There are at least two reasons why one is interested in dynamics on bounded time-sets. One is the interest in *transient* behavior of solutions, although the process or its generating equation might be given on an unbounded time-set. The second one is the simple fact that the right-hand side f is given only on a bounded time-set, e.g. when it is deduced from observations. In any case, classical, asymptotic concepts do not apply to the finite-time situation directly.

The main objects of investigation, reference trajectories and Lagrangian Coherent Structures, correspond to two different views on ODEs like (1.1). The first one is to consider certain reference trajectories and to study the dynamic behavior in their vicinity. Roughly speaking, this corresponds to fixing the initial value and leaving the time argument variable. In the complementary approach, one fixes the initial and the final time and considers the impact of the process $\varphi(t_+, t_-, \cdot)$ on an ensemble of initial values, where the reader should think of φ being a solution operator of Eq. (1.1) for a moment. This corresponds to fixing the time-parameters and leaving the space argument variable. Given that f is differentiable with respect to the second argument, i.e. in the space variable, this reduces to the analysis of the diffeomorphism $\varphi(t_+, t_-, \cdot)$ and differential-geometric tools come into play.

As indicated above, we study general evolutionary processes, e.g. solution operators of *nonautonomous dynamic/differential equations*; in more applied literature one also speaks of *unsteady* or *aperiodic (fluid) flows*. Therefore, throughout this work, *invariance* refers to *φ-invariance* as in [69, Definition 2.14]. As is mentioned there, φ-invariant families of subsets of the state space consist of entire trajectories. These in turn are uniquely described by their value at, for instance, the initial time-point t_-. That is why we formulate our results in terms of sets of the time-fiber $\{t_-\} \times \mathbb{R}^n$, i.e. in terms of initial values, knowing that these can be extended to the time-set under consideration via the process φ. In the literature, sometimes the term *Lagrangian* is used to distinguish invariance in state space from φ-invariance.

Hyperbolicity has been recognized as a fundamental concept in dynamical systems theory and still plays a vital role in nowadays research [68]. There exist several hyperbolicity notions in the asymptotic context, such as exponential dichotomy [78, 26, 60], nonuniform exponential dichotomy [10, 11], generalized exponential dichotomy [77], (h, k)-dichotomy [82], (μ_1, μ_2)-dichotomy [81], exponential trichotomy [26] and partial hyperbolicity [89]. During the last years efforts were made to establish finite-time analogues to asymptotic notions, such as hyperbolicity of linearizations [57, 17, 95], Lyapunov exponents [54, 101, 56], stable and unstable manifolds [57, 54, 33, 15] and Lyapunov functions and Lyapunov's second method [47].

To establish a reasonable finite-time hyperbolicity notion, most authors agree on that the dichotomic behavior should be with respect to monotonicity, a realization first made in [57]. The notion of *uniform finite-time hyperbolicity* developed there has stimulated much research; see [59, 74, 17, 95, 32]. The existing approaches consider solutions either on intervals [57, 59, 17, 32] or restrictions to an initial and a final time-point [54, 101, 74, 95, 47]. A desirable property of hyperbolicity is its robustness under certain perturbations, which is known to hold in the time-asymptotic case. Robustness for finite-time hyperbolicity was first established in [14]. The contributions of this thesis concerning hyperbolicity are:

- a simplified introduction of a unifying and generalizing finite-time hyperbolicity notion,
- introduction of a metric on the finite-time processes allowing regularity investigations,
- proof of continuity for (extremal) growth rates and the spectrum,
- a direct proof of robustness, including sharp perturbation bounds.

As is well-known and as we show in this thesis, (extremal) growth rates play an important role in the analysis of linear finite-time processes. In [17] their computation is indicated to be highly desirable. However, to the best of the author's knowledge, there exists no general numerical approach for that, except for a theoretical one in [31, Remark 21] for the special case of space dimension 2. Although the definitions of (extremal) growth rates look bulky, they can be literally considered as an optimization problem on a Riemannian manifold, namely the Grassmann manifold. Only recently, theoretical work and numerical algorithms on optimization on Riemannian manifolds have been published; see [3] for a textbook including a historical review and for further references. As a first approach, we approximate (extremal) growth rates for solution operators of linear ODEs by a Newton method. The contributions of this thesis concerning the computation of growth rates are:

- an extensive regularity investigation of growth rate functions, including sufficient conditions for differentiability,
- the performance of several example calculations.

A fundamental implication of hyperbolicity is the existence of local stable and unstable manifolds. Previous work on local finite-time stable and unstable manifolds [57, 54, 15] used (non-unique) infinite-time extensions of the ordinary differential equation under consideration. Then, the finite-time local stable and unstable manifolds are derived with methods from the classical theory. For an alternative approach which proves a weaker version, see [33]. The contributions of this thesis concerning linearization theory are:

- an intrinsic proof of finite-time local stable and unstable cones/manifolds,

- a complete local description of the finite-time process based on the linearization,
- a (conditional) Hartman-Grobman-like theorem.

As we show in this thesis, non-uniqueness of stable and unstable subspaces is an inherent feature of finite-time hyperbolicity.

Apart from the study of reference trajectories we will, as indicated above, also take an alternative viewpoint on finite-time processes. It is about the investigation of some diffeomorphism which could be the evaluation of a solution operator $\varphi(t_+, t_-, \cdot)$ to Eq. (1.1) at two time-points. This point of view originates from the application of dynamical systems theory to fluid mechanics. It was introduced in [7] and coined *chaotic advection*; see also [87, 109] for textbooks on mixing and transport (of fluids) and [8] for a historical review. The idea is to interpret the paths of particles, also referred to as *passive tracers*, as solutions of an ODE like (1.1). Although just a matter of interpretation, this has a strong impact on the type of questions that are asked, and notions like mixing regions, transport, coherent sets and transport barriers make particularly sense in this context. In the first years after its introduction, this concept was applied to ODEs given on $\mathbb{R}_{\geq 0}$, i.e. in particular systems that allow of using asymptotic analysis. It was only in the 1990's that ODEs like (1.1), given on finite time-intervals, were considered in the context of chaotic advection.

Some of the typical questions posed in that context are the following: identify coherent sets or mixing regions, i.e. sets of initial values that are dispersed as little as possible under the dynamics; find transport barriers between them, i.e. impermeable (or very weakly permeable) hypersurfaces separating dynamically similar sets of initial values; quantify transport between two different coherent regions. Among the tools that have been developed to approach these questions there are *finite-time Lyapunov exponents (FTLEs)*.

In some sense, finite-time Lyapunov exponents exist since the introduction of Lyapunov exponents in [75]: as the expression in which one lets time tend to infinity in the definition of Lyapunov exponents. However, for decades the focus was on the investigation of the Lyapunov exponents as a result of the limit process; see, for instance, [86, 97, 13, 61], to mention but a few. Apparently in [50], the first steps towards an independent investigation of finite-time Lyapunov exponents were taken. [90] introduces finite-time Lyapunov exponents to the geophysical community, and hence, to the community of chaotic advection. Apparently in [91], finite-time Lyapunov exponents appear for the first time in a closed-form and on their own right. They are considered to be "useful in identifying mixing regions and transport barriers" [91, p. 2465]. That gave rise to many numerical simulations and applications to real-world data sets.

Some of the aforementioned questions can be approached by the study of certain *invariant (or almost-invariant) manifolds*. This led to the investigation of *Lagrangian Coherent Structures (LCSs)*, which were introduced in [59]. While the notion of FTLE is well-established and widely-used, there exist many notions of distinguished invariant manifolds, only few of them rigorously defined. One approach is defining finite-time analogues to stable and unstable manifolds from the classical, asymptotic theory with their practical shortcomings like, for instance, a very limited extent.

The other and more often applied approach is the following geometrical one. The idea is to assign to each initial value a scalar derived from the action of the solution operator on a neighborhood of the initial value. This gives rise to a functional/scalar field on the initial values. The desired manifolds now consist of points maximizing (or minimizing) the scalar field, which is made precise in terms of ridges (or troughs, respectively) in the literature. During the last decade many heuristically motivated functionals have been introduced to quantify attraction and repulsion of nearby solutions of a reference solution in finite time. Among these are, for instance, the hyperbolicity time approach [59], finite-size Lyapunov exponents [9, 62], arc length approach [76], phase space warping [22], finite-time entropy [46] and the aforementioned, probably most popular, approach of finite-time Lyapunov exponents [54, 55, 101, 74, 56]. The geometrical ridge approach has been prominently advocated in [101, 74], where repelling LCS are defined as ridges of the FTLE-field. However, it was left unspecified which dynamical effect these manifolds actually have. Despite this and other critical issues, see [56, 83], LCS have since become a popular method in applied finite-time analysis; see, for instance, the aforementioned review [88]. Only recently, [56] gave a rigorous definition and proved an analytical characterization of hyperbolic LCSs, starting with some observable dynamical property and only then deriving geometric descriptions. Interestingly, LCSs are considered as boundaries of *coherent regions* and are motivated by simple analytical examples as such; for a direct approach to the identification of coherent sets, which is based on the transfer operator, see, for instance, [45, 29, 43]. However, a formal connection of the two concepts is still missing; for a numerical study see [44].

In this work, we extend the variational approach to hyperbolic LCS of [56] to introduce hyperbolic LCS of higher-codimension, thereby rigorously carrying out a note in the last paragraph of [56, p. 594].

The main contributions of this thesis concerning finite-time Lyapunov exponents and Lagrangian Coherent Structures are:

- a short and direct derivation of FTLEs from the general theory presented before,
- a rigorous investigation of their regularity,
- an extension of Haller's variational approach to hyperbolic LCS towards structures of higher codimension,
- a formal embedding principle, which allows the introduction of hyperbolic LCS on manifolds and the derivation of filtrations of hyperbolic LCS,
- the establishment of the connection of hyperbolic LCS and sets of generalized extremal points.

1.1 Outline of the Thesis

The thesis is organized as follows: after introducing notation in Section 1.2, we propose an abstract framework for dynamics on compact time-sets $\mathbb{I}$, not necessarily intervals, in Chapter 2. Our framework is based on the well-known notion of processes. Thus, we state explicitly all required regularity conditions and introduce a topology on linear finite-time processes. To the best of the author's knowledge, this is the first time that such a first-principles approach to finite-time dynamics is proposed. As already mentioned, a finite-time process can be generated by a well-posed dynamic equation on $\mathbb{I}$; see, for instance, [20] for a textbook on time scale calculus and dynamic equations. In Chapter 3 we develop a spectral theory for linear finite-time processes which is essentially known for solution operators on time-sets consisting of two points [95] and for intervals [17, 32]. The underlying hyperbolicity notion is motivated by several approaches [57, 59, 17, 95, 31]. These are generalized and unified. We call this kind of finite-time hyperbolicity suggestively *exponential monotonicity dichotomy (EMD)*. We make use of the introduced topology to establish the continuity of the spectrum and the robustness of EMD easily. In Chapter 4 we discuss regularity for the growth rate functions and present a Newton-like algorithm to approximate the extremal growth rates and hence the spectrum of a linear finite-time process.

In Chapter 5 we introduce linearizations of finite-time processes along trajectories and establish finite-time analogues of classical linearization theorems: Local Stable and Unstable Manifold Theorem, a finite-time analogue of the Theorem of Linearized Asymptotic Stability and a finite-time Hartman-Grobman-like Theorem. In Chapter 6 we show that the general results apply to ordinary differential equations on compact time-intervals. Chapter 7 is devoted to the investigation of FTLEs and hyperbolic LCSs. We conclude the thesis with an overview and outlook in Chapter 8. In the appendices, we first introduce classical notions from differential geometry, essentially to fix notation, and second give a self-consistent introduction to Grassmann manifolds.

1.2 Preliminaries and Notation

Throughout this thesis, $\mathbb{I} \subset \mathbb{R}$ denotes a compact subset of the real numbers with at least two elements. We set $t_- := \min \mathbb{I}$ and $t_+ := \max \mathbb{I}$. We will have occasions to mention the set of all pairs of numbers in $\mathbb{I}$ with unequal components. In accordance with the notion of a relation, we denote by $\neq_{\mathbb{I}\times\mathbb{I}} := \{(t,s) \in \mathbb{I} \times \mathbb{I};\ t \neq s\} \subset \mathbb{I} \times \mathbb{I}$ the unequal-relation on $\mathbb{I} \times \mathbb{I}$.

For a set M we write id_M for the identity function on M and 2^M for its power set. To distinguish between the evaluation of a function $f\colon A \to B$ at $x \in A$ from the image of a subset $A' \subseteq A$ we write $f(x)$ for the former and $f[A']$ for the latter. We denote the choice function defined on singleton sets by $\in$, i.e. for a singleton set $M = \{x\}$ we have $\in (M) = x$.

In the following, we consider dynamics on $\mathbb{R}^n$ with an arbitrary vector norm $|\cdot|$, which in turn induces an operator norm denoted by $\|\cdot\|$. By $\mathcal{S}$ we denote the unit sphere with respect to the given norm on $\mathbb{R}^n$. We denote by $L(\mathbb{R}^n)$ the set of linear operators on $\mathbb{R}^n$ and by $GL(n,\mathbb{R})$ the subset of invertible operators on $\mathbb{R}^n$.

For a Lipschitz continuous function f we denote by $|f|_{\mathrm{Lip}}$ the least Lipschitz constant of f. For time-dependent functions, whether vector- or operator-valued, the notation $\|\cdot\|_\infty$ denotes the supremum norm over time with respect to the respective norms. For a continuously differentiable function $g\colon B_0 \times \cdots \times B_k \to B$, where $B, B_0, \ldots, B_k$ are (open subsets of) Banach spaces, we denote the partial derivative of g with respect to the $(j+1)$-th variable evaluated at $x \in B_0 \times \cdots \times B_k$ by $\partial_j g(x)$. Note that the first argument has index 0. In case g has only one argument, the derivative is also denoted by $g'(x)$, or $\dot{g}(t)$ if the argument is the time variable.

We write $B(x, \delta)$ and $B[x, \delta]$, respectively, for the open and closed ball around x with radius $\delta \in \mathbb{R}_{>0}$.

1.3 Associated Publications

Parts of this thesis have been published. These are as follows:

[31] T. S. Doan, D. Karrasch, T. Y. Nguyen, and S. Siegmund. A unified approach to finite-time hyperbolicity which extends finite-time Lyapunov exponents. *Journal of Differential Equations*, 252(10), 5535–5554, 2012.

[64] D. Karrasch. Comment on "A variational theory of hyperbolic Lagrangian Coherent Structures, Physica D 240 (2011) 574–598". *Physica D*, 241(17):1470–1473, 2012.

[65] D. Karrasch. Linearization of hyperbolic finite-time processes. *Journal of Differential Equations*, 254(1), 256–282, 2013.

Note that in this thesis we do not refer explicitly to results that were published in the last two articles.

2 Finite-time Processes & Growth Rates

In this chapter, we specify the setting that we investigate in this work. Furthermore, we introduce notions that will be of large benefit particularly in the following three chapters.

2.1 Finite-time Processes

The notion of a *process* or *two-parameter semi-flow* was originally introduced in [27]. However, we require slightly different conditions. In particular, we include invertibility in the definition.

2.1 Definition (Process, linear process, smooth process)**.** We call a continuous function

$$\varphi\colon \mathbb{I} \times \mathbb{I} \times \mathbb{R}^n \to \mathbb{R}^n,$$

an *invertible process on* $\mathbb{I}$ *with state space* $\mathbb{R}^n$ if it is Lipschitz continuous with respect to the first argument and if for any $t, s, r \in \mathbb{I}$ and $x \in \mathbb{R}^n$ we have $\varphi(t, t, \cdot) = \mathrm{id}_{\mathbb{R}^n}$ and $\varphi(t, s, \cdot) \circ \varphi(s, r, \cdot) = \varphi(t, r, \cdot)$. We denote by $\mathcal{P}(\mathbb{I}, \mathbb{R}^n)$ the set of invertible processes on $\mathbb{I}$ with state space $\mathbb{R}^n$. We call an invertible process $\Phi \in \mathcal{P}(\mathbb{I}, \mathbb{R}^n)$ a *linear invertible process on* $\mathbb{I}$ if for any $t, s \in \mathbb{I}$ we have $\Phi(t, s, \cdot) \in GL(n, \mathbb{R})$ and the function $t \mapsto \Phi(t, s, \cdot) \in L(\mathbb{R}^n)$ is Lipschitz continuous with respect to the operator norm. To emphasize the linearity with respect to the last argument, we will write $\Phi(t, s)x$ instead of $\Phi(t, s, x)$ for $t, s \in \mathbb{I}$, $x \in \mathbb{R}^n$. We denote by $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ the set of linear invertible processes on $\mathbb{I}$ with state space $\mathbb{R}^n$. Let $k \in \mathbb{N}_{>0}$. We call an invertible process φ a C^k*-process on* $\mathbb{I}$ if for any $t, s \in \mathbb{I}$ and $x \in \mathbb{R}^n$ we have $\varphi(t, s, \cdot) \in C^k(\mathbb{R}^n, \mathbb{R}^n)$ and $\partial_2 \varphi(\cdot, s, x) \in L(\mathbb{R}^n)^{\mathbb{I}}$ is Lipschitz continuous.

In the following, we always consider invertible processes and hence skip the word invertible. The required Lipschitz continuity of $t \mapsto \Phi(t, t_-)$ in Definition 2.1 for a linear process Φ is rather a technical assumption than an integral part of the notion of a process. In principle, it could be replaced by the weaker, but again technical, assumption of absolute continuity. For convenience, we will assume Lipschitz continuity in the sequel and point out which implications the weaker assumption would have. However, the necessity for a stronger continuity assumption than just continuity will become clear in the course of this work.

We start our investigations with linear processes on $\mathbb{I}$.

2.2 Lemma. *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$. Then the function*

$$\mathbb{I} \times \mathbb{I} \to L(\mathbb{R}^n), \qquad (t,s) \mapsto \Phi(t,s),$$

is uniformly continuous with respect to the (induced) operator norm.

Proof. This is a direct consequence of the uniform continuity of $\Phi|_{\mathbb{I}\times\mathbb{I}\times B[0,1]}$. □

By the compactness of $\mathbb{I}$ we obtain in the next lemma the continuous dependence of (the norm of) trajectories on the initial value, uniformly in time.

2.3 Lemma. *Let $\varphi \in \mathcal{P}(\mathbb{I}, \mathbb{R}^n)$ and $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$. Then the function*

$$f \colon \mathbb{R}^n \to C(\mathbb{I}, \mathbb{R}^n), \qquad x \mapsto (t \mapsto \varphi(t, t_-)x)\,,$$

is continuous and the functions

$$\begin{aligned} g &\colon \mathbb{R}^n \to C(\mathbb{I}, \mathbb{R}^n), & x &\mapsto (t \mapsto \Phi(t, t_-)x)\,,\\ h &\colon \mathbb{R}^n \to C(\mathbb{I}, \mathbb{R}_{>0}), & x &\mapsto (t \mapsto |\Phi(t, t_-)x|) \end{aligned}$$

are Lipschitz continuous.

Proof. To see the continuity of f in some $x \in \mathbb{R}^n$ it suffices to restrict the continuous function φ to the compact set $\mathbb{I} \times \mathbb{I} \times B[x, 1] \subset \mathbb{R}^{2+n}$, thereby turning it into a uniformly continuous function. The continuity of f in x follows directly from that uniform continuity. As for the continuity of g, let $x, y \in \mathbb{R}^n$ and estimate

$$\|g(x) - g(y)\|_\infty = \sup\{|\Phi(t, t_-)(x - y)|\,;\ t \in \mathbb{I}\} \le \|\Phi(\cdot, t_-)\|_\infty\,|x - y|\,,$$

where $\|\Phi(\cdot, t_-)\|_\infty < \infty$ follows from Lemma 2.2 and the compactness of $\mathbb{I}$. The Lipschitz continuity of h follows from the Lipschitz continuity of the norm. □

Note that both continuity results did not use the Lipschitz continuity assumption in Definition 2.1. Even more, the continuity of $x \mapsto (t \mapsto \varphi(t, t_-)x)$ relies only on the continuity of φ as a function defined on $\mathbb{I} \times \mathbb{I} \times \mathbb{R}^n$ and mapping to $\mathbb{R}^n$. Although seemingly simple, Lemma 2.3 indicates an important difference in asymptotic and finite-time analysis: on unbounded (to the right) time-sets the continuous dependence of whole trajectories on the initial value corresponds to the definition of stability in the sense of Lyapunov, which is a nontrivial feature of certain trajectories. In the finite-time case, as Lemma 2.3 shows, it holds under very general assumptions.

2.2 Logarithmic Difference Quotient

The following concept will help us to introduce some of the forthcoming notions and to present the theory in an elegant and coherent way.

2.4 Definition (Logarithmic difference quotient)**.** We define

$$\Delta_{\mathbb{I}} \colon C(\mathbb{I}, \mathbb{R}_{>0}) \to C(\neq_{\mathbb{I}\times\mathbb{I}}, \mathbb{R}), \qquad f \mapsto \left((t,s) \mapsto \frac{\ln f(t) - \ln f(s)}{t-s} \right),$$

and we call $\Delta_{\mathbb{I}}(f)(t,s)$ the *logarithmic difference quotient of f at t and s* for $f \in C(\mathbb{I}, \mathbb{R}_{>0})$ and $(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}$. For notational convenience we will write Δ for $\Delta_{\mathbb{I}}$ when there is no risk of confusion.

2.5 Remark. In correspondence to the multiplicative calculus one can define by ${}^*\Delta(f)(t,s) = \left(\frac{f(t)}{f(s)}\right)^{\frac{1}{t-s}}$ the *difference quotient; see [12] and the references therein for more information on the multiplicative calculus. It is readily confirmed that

$$\Delta(f)(t,s) = \ln {}^*\Delta(f)(t,s).$$

Therefore, the logarithmic difference quotient is in some sense in between the (standard) additive and the multiplicative calculus.

2.6 Remark. Note that due to compactness of $\mathbb{I}$ any function $f \in C(\mathbb{I}, \mathbb{R}_{>0})$ is uniformly continuous and $f[\mathbb{I}] \subseteq [a,b]$ with $0 < a < b$, i.e. f is bounded above and bounded away from zero. Furthermore, $\ln\big|_{[a,b]}$ is Lipschitz continuous with Lipschitz constant $\frac{1}{a}$ and bounded. Suppose f is additionally Lipschitz continuous, then

$$|\sup\{\Delta(f)(t,s);\ (t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}\}|, |\inf\{\Delta(f)(t,s);\ (t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}\}| \leq \frac{|f|_{\mathrm{Lip}}}{a}.$$

Furthermore, we have that $\Delta(\cdot)(t,s)$ is continuous for any $(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}$. Summarizing, Δ considered as

$$\Delta\colon C(\mathbb{I}, \mathbb{R}_{>0}) \times \neq_{\mathbb{I}\times\mathbb{I}} \to \mathbb{R}, \qquad (f,t,s) \mapsto \frac{\ln f(t) - \ln f(s)}{t-s},$$

is continuous in the first argument and jointly continuous in the last two arguments, but, in general, not jointly continuous in all three arguments. However, for a linear process $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ one can show that the family of functions $(\Delta(|\Phi(\cdot, t_-)x|))_{x\in\mathcal{S}}$ is equicontinuous. This can be calculated directly or, alternatively, can be deduced from Lemma 2.3 and the Theorem of Arzelà-Ascoli.

When applied to a linear process we recover joint continuity for the logarithmic difference quotient.

2.7 Lemma. *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$. Then*

$$\mathbb{R}^n \setminus \{0\} \times \neq_{\mathbb{I}\times\mathbb{I}} \to \mathbb{R}, \qquad (x,t,s) \mapsto \Delta(|\Phi(\cdot, t_-)x|)(t,s),$$

is continuous.

Note that the function defined in Lemma 2.7 is well-defined, since we assume the linear process to be invertible (no trajectory starting in $\mathbb{R}^n \setminus \{0\}$ at t_- attains zero on $\mathbb{I}$).

Proof. Let $x \in \mathbb{R}^n \setminus \{0\}$, $(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}$ and $\varepsilon \in \mathbb{R}_{>0}$. By equicontinuity with respect to the initial value there exists $\delta_1 \in \mathbb{R}_{>0}$ such that for all $y \in \mathbb{R}^n \setminus \{0\}$

$$\max\left\{\left|t-t'\right|, \left|s-s'\right|\right\} < \delta_1 \;\Rightarrow\; \left|\Delta(|\Phi(\cdot,t_-)y|)(t,s) - \Delta(|\Phi(\cdot,t_-)y|)(t',s')\right| < \frac{\varepsilon}{2}.$$

By Lemma 2.3 and hence continuity of $\mathbb{R}^n \setminus \{0\} \ni y \mapsto \Delta(|\Phi(\cdot,t_-)y|)(t,s)$ there exists $\delta_2 \in \mathbb{R}_{>0}$ such that for all $x' \in \mathbb{R}^n \setminus \{0\}$

$$\left|x - x'\right| < \delta_2 \;\Rightarrow\; \left|\Delta(|\Phi(\cdot,t_-)x|)(t,s) - \Delta(\left|\Phi(\cdot,t_-)x'\right|)(t,s)\right| < \frac{\varepsilon}{2}.$$

Combining these two estimates, we obtain for $(t',s') \in \neq_{\mathbb{I}\times\mathbb{I}}$, $x' \in \mathbb{R}^n \setminus \{0\}$

$$\left|\Delta(|\Phi(\cdot,t_-)x|)(t,s) - \Delta(\left|\Phi(\cdot,t_-)x'\right|)(t',s')\right| < \varepsilon,$$

whenever $\max\{|t-t'|, |s-s'|, |x-x'|\} < \min\{\delta_1, \delta_2\}$. $\square$

2.8 Remark. The logarithmic difference quotient can be regarded as a *finite-time exponential growth rate* as introduced in [23]. With the notation used there, we find that

$$\lambda^{(t-s)}(s, \Phi(s, t_-)x) = \Delta(|\Phi(\cdot, t_-)x|)(t, s).$$

Definition 2.4 and Lemma 2.7 will have a major impact on the following theory and allow a development of the known finite-time spectral theory, starting from the notion of the logarithmic difference quotient alone. Another useful observation is the following, which holds obviously by linearity of $\Phi(t, s)$ and Definition 2.4.

2.9 Lemma. *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$. Then for any $x \in \mathbb{R}^n \setminus \{0\}$ and $\lambda \in \mathbb{R} \setminus \{0\}$ one has*

$$\Delta(|\Phi(\cdot, t_-)(\lambda x)|) = \Delta(|\Phi(\cdot, t_-)x|).$$

As a consequence, we can equally well define $\Delta(|\Phi(\cdot, t_-)\cdot|)$ on the quotient set $\mathbb{P}^{n-1} \cong \mathcal{S}$, the (real) projective space obtained by identifying vectors $x, y \in \mathbb{R}^n \setminus \{0\}$ if they are multiples of each other. We denote by $[x]_\sim$ the corresponding equivalence class associated to $x \in \mathbb{R}^n \setminus \{0\}$. The projective space $\mathbb{P}^{n-1}$ can be identified to the set of one-dimensional subspaces, or lines through the origin, in $\mathbb{R}^n$. A generalization of that concept is the Grassmann manifold $\mathrm{Gr}(k, \mathbb{R}^n)$ of k-dimensional subspaces of $\mathbb{R}^n$, which will turn out to be very useful. For an introduction to the Grassmann manifold see Appendix B and the references mentioned there.

Some of the following results are concerned with continuity of functions derived from the logarithmic difference quotient. For some results it appears to be necessary to put the following assumption on linear processes, which, for convenience, we suppose to be satisfied by the processes considered in this work.

2.10 Hypothesis. Throughout, we assume that for any $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ the function

$$\mathcal{S} \times \neq_{\mathbb{I}\times\mathbb{I}} \to \mathbb{R}, \qquad (x, t, s) \mapsto \Delta(|\Phi(\cdot, t_-)x|)(t, s)$$

can be extended continuously to $\mathcal{S} \times \overline{\neq_{\mathbb{I}\times\mathbb{I}}}$, where $\overline{\neq_{\mathbb{I}\times\mathbb{I}}}$ denotes the closure of $\neq_{\mathbb{I}\times\mathbb{I}}$ in $\mathbb{R}^2$.

As a consequence, $\overline{\neq_{\mathbb{I}\times\mathbb{I}}}$ is compact and the extended function is uniformly continuous. In the following remark we discuss the consequences of Hypothesis 2.10.

2.11 Remark. As $\mathbb{I}$ is compact, so is $\mathbb{I} \times \mathbb{I} \subset \mathbb{R}^2$. Clearly, $\neq_{\mathbb{I}\times\mathbb{I}}$ is still bounded, but not necessarily closed. We need to distinguish two cases: either $\mathbb{I}$ has no limit points, i.e. $\mathbb{I}$ is finite, or $\mathbb{I}$ has limit points, e.g. when $\mathbb{I}$ is a (nontrivial) compact interval.

(a) If $\mathbb{I}$ does not have any limit points then so does $\neq_{\mathbb{I}\times\mathbb{I}}$. Thus, $\neq_{\mathbb{I}\times\mathbb{I}}$ is closed already and Hypothesis 2.10 does not pose any restrictions on the finite-time processes under consideration.

(b) Let $t^* \in \mathbb{I}$ be a limit point. Then Hypothesis 2.10 requires that the limit

$$\lim_{t,s\to t^*} \frac{\ln|\Phi(t,t_-)x| - \ln|\Phi(s,t_-)x|}{t-s} = \frac{|\Phi(\cdot,t_-)x|'(t^*)}{|\Phi(t^*,t_-)x|}$$

exists for any $x \in \mathcal{S}$ and is continuous in x. A sufficient (but not necessary) condition for the derivative on the right-hand side to exist is the continuous differentiability of the norm (not at 0, of course) and the differentiability of the trajectories $t \mapsto \Phi(t,t_-)x$ for any $x \in \mathcal{S}$ at the (time-)limit point t^*. Clearly, solution operators of linear ordinary differential equations with continuous right-hand side as usually considered together with continuously differentiable norms do satisfy Hypothesis 2.10.

The following simple observation associates the notions introduced in the present work to notions with the same nomenclature appearing in the references [17, 32, 31].

2.12 Lemma. *Let $f \in C(\mathbb{I}, \mathbb{R}_{>0})$. Then the following statements are equivalent:*

(i) *there exists $\delta \in \mathbb{R}_{>0}$ such that for any $t, s \in \mathbb{I}$, $t \geq s$, one has $f(t) \leq \mathrm{e}^{-\delta(t-s)} f(s)$;*

(ii) *there exists $\delta \in \mathbb{R}_{>0}$ such that $t \mapsto \mathrm{e}^{\delta t} f(t)$ is decreasing;*

(iii) $\sup\{\Delta(f)(t,s);\ (t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}\} < 0$.

Moreover, if $\mathbb{I}$ is an interval and f is differentiable, then each of the above statements is equivalent to:

(iv) *there exists $\delta \in \mathbb{R}_{>0}$ such that $f' \leq -\delta f$ holds.*

Analogously, the following statements are equivalent:

(v) *there exists $\delta \in \mathbb{R}_{>0}$ such that for any $t, s \in \mathbb{I}$, $t \geq s$, one has $f(t) \geq \mathrm{e}^{\delta(t-s)} f(s)$;*

(vi) *there exists $\delta \in \mathbb{R}_{>0}$ such that $t \mapsto \mathrm{e}^{-\delta t} f(t)$ is increasing;*

(vii) $\inf\{\Delta(f)(t,s);\ (t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}\} > 0$.

Moreover, if $\mathbb{I}$ is an interval and f is differentiable, then each of the statements (v)–(vii) is equivalent to:

(viii) there exists $\delta \in \mathbb{R}_{>0}$ such that $f' \geq \delta f$ holds.

Proof. $(i) \iff (ii)$: Suppose that for some $\delta \in \mathbb{R}_{>0}$ and for any $s, t \in \mathbb{I}$, $s \leq t$, holds
$$f(t) \leq \mathrm{e}^{-\delta(t-s)} f(s),$$
which, by multiplying both sides by $\mathrm{e}^{\delta t} > 0$, is equivalent to
$$\mathrm{e}^{\delta t} f(t) \leq \mathrm{e}^{\delta s} f(s),$$
which corresponds to the desired monotonicity of the function $t \mapsto \mathrm{e}^{\delta t} f(t)$.
$(i) \iff (iii)$: Suppose that for some $\delta \in \mathbb{R}_{>0}$ and for any $s, t \in \mathbb{I}$, $s \leq t$, holds
$$f(t) \leq \mathrm{e}^{-\delta(t-s)} f(s),$$
which, by taking the logarithm of both sides and rearrangement of terms, is equivalent to
$$\frac{f(t) - f(s)}{t - s} \leq -\delta$$
for any $s, t \in \mathbb{I}$, $s \leq t$, and hence
$$\sup \{\Delta(f)(t,s);\ t, s \in \mathbb{I},\ s \leq t\} \leq -\delta < 0.$$
Now assume $\mathbb{I}$ is an interval and f is differentiable.
$(ii) \iff (iv)$: We assume that $t \mapsto \mathrm{e}^{\delta t} f(t)$ is decreasing for some $\delta \in \mathbb{R}_{>0}$. This is equivalent to the fact that for any $t \in \mathbb{I}$ holds
$$(s \mapsto \mathrm{e}^{\delta s} f(s))'(t) = \mathrm{e}^{\delta t}(\delta f(t) + f'(t)) \leq 0,$$
which is equivalent to
$$f' \leq -\delta f.$$
The proof of the equivalence of (v)-(viii) is completely analogous. □

2.3 Growth Rates

In this and the next section, we introduce (extremal) growth rates based on the logarithmic difference quotient only. This provides an easy to implement expression, in contrast to the corresponding definitions in [32, 31].

2.13 Definition (Growth rate, cf. [17, 32, 31]). We define

$$\underline{\lambda}^{\mathbb{I}}\colon \bigcup_{k=1}^{n} \mathrm{Gr}(k,\mathbb{R}^n) \times \mathcal{LP}(\mathbb{I},\mathbb{R}^n) \to \mathbb{R}, \quad (X,\Phi) \mapsto \min_{\substack{x\in X\cap\mathcal{S}\\ (t,s)\in\neq_{\mathbb{I}\times\mathbb{I}}}} \left\{\Delta(|\Phi(\cdot,t_-)x|)(t,s)\right\},$$

$$\overline{\lambda}^{\mathbb{I}}\colon \bigcup_{k=1}^{n} \mathrm{Gr}(k,\mathbb{R}^n) \times \mathcal{LP}(\mathbb{I},\mathbb{R}^n) \to \mathbb{R}, \quad (X,\Phi) \mapsto \max_{\substack{x\in X\cap\mathcal{S}\\ (t,s)\in\neq_{\mathbb{I}\times\mathbb{I}}}} \left\{\Delta(|\Phi(\cdot,t_-)x|)(t,s)\right\},$$

and call $\underline{\lambda}^{\mathbb{I}}(X,\Phi)$ and $\overline{\lambda}^{\mathbb{I}}(X,\Phi)$, respectively, the *lower* and *upper growth rate of X under Φ*. For convenience, we drop the index $\mathbb{I}$ in case the time-set is clear from the context. We extend the definition naturally by $\underline{\lambda}(\{0\},\Phi) = \infty$ and $\overline{\lambda}(\{0\},\Phi) = -\infty$ for any $\Phi \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$.

2.14 Remark. In the last definition, we made explicit use of Hypothesis 2.10 in defining the growth rates to be minimum and maximum, respectively. However, we could have defined the growth rates equally well as infimum and supremum over $(t,s) \in\neq_{\mathbb{I}\times\mathbb{I}}$. Thus, Hypothesis 2.10 is not necessary for the definition.

One readily verifies that for any $\Phi \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ and $X \in \mathrm{Gr}(1,\mathbb{R}^n)$ one has $|\underline{\lambda}(X,\Phi)|, |\overline{\lambda}(X,\Phi)| \leq L_1L_2 < \infty$, where $L_1 \in \mathbb{R}_{>0}$ is the Lipschitz constant of the logarithm restricted to some interval as in Remark 2.6 and $L_2 \in \mathbb{R}_{>0}$ is the (global) Lipschitz constant of $t \mapsto \Phi(t,s,\cdot) \in L(\mathbb{R}^n)$, cf. Definition 2.1.

2.15 Remark. One can consider the one-dimensional growth rates as a finite-time analogue to *Bohl exponents*. Bohl exponents have been first introduced in [19]; see also [28, p. 118, pp. 146–148] for the definition and a historical review. Recall that for a linear solution operator Φ of a linear ordinary differential equation defined on $\mathbb{R}_{\geq 0}$ the (upper) Bohl exponent $\overline{\kappa_{\mathrm{B}}}(x)$ for a solution starting in $x \in \mathbb{R}^n$ at time 0 is defined as

$$\overline{\kappa_{\mathrm{B}}}(x) := \inf\left\{\beta \in \mathbb{R};\ \bigvee_{N\in\mathbb{R}_{>0}}\ \bigwedge_{t\geq s\geq 0}\ |\Phi(t,0)x| \leq Ne^{\beta(t-s)}\,|\Phi(s,0)x|\right\}.$$

Analogously, the lower Bohl exponent is defined as

$$\underline{\kappa_{\mathrm{B}}}(x) := \sup\left\{\beta \in \mathbb{R};\ \bigvee_{N\in\mathbb{R}_{>0}}\ \bigwedge_{t\geq s\geq 0}\ |\Phi(t,0)x| \geq Ne^{\beta(t-s)}\,|\Phi(s,0)x|\right\}.$$

In the finite-time setting, the general coefficient N in the definition of the Bohl exponents would render the concept meaningless. By setting $N = 1$ in the definition and modifying the time-set appropriately, we recover with Lemma 2.12, (i) and (iii), the one-dimensional growth rates; see also the formulas given after Eq. (4.4′) in [28, p. 118].

The following simple observations follow directly from the definition.

2.16 Lemma (cf. [32, Remark 6]). *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ and $X, Y \subseteq \mathbb{R}^n$ be subspaces. Then*

(i) $X \subseteq Y$ *implies* $\underline{\lambda}(Y, \Phi) \leq \underline{\lambda}(X, \Phi)$ *and* $\overline{\lambda}(Y, \Phi) \geq \overline{\lambda}(X, \Phi)$;

(ii) $X \cap Y \neq \{0\}$ *implies* $\underline{\lambda}(X, \Phi) \leq \overline{\lambda}(Y, \Phi)$ *and* $\underline{\lambda}(Y, \Phi) \leq \overline{\lambda}(X, \Phi)$.

Proof. (i) This is clear since, with respect to Y, minimum and maximum, respectively, are taken over a larger set.

(ii) Suppose $X \cap Y \neq \{0\}$ and let $x \in (X \cap Y) \setminus \{0\}$. Then by definition we have for any $(t, s) \in \neq_{\mathbb{I} \times \mathbb{I}}$

$$\underline{\lambda}(X, \Phi) \leq \Delta(|\Phi(\cdot, t_-)x|)(t, s) \leq \overline{\lambda}(Y, \Phi)$$

and

$$\underline{\lambda}(Y, \Phi) \leq \Delta(|\Phi(\cdot, t_-)x|)(t, s) \leq \overline{\lambda}(X, \Phi). \qquad \square$$

Next, we establish continuity for the growth rates when restricted to some $\mathrm{Gr}(k, \mathbb{R}^n)$. This extends [32, Lemma 7] and its proof relies on Hypothesis 2.10.

2.17 Proposition. *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$. Then for each $k \in \{1, \ldots, n\}$, the restrictions*

$$\underline{\lambda}(\cdot, \Phi)\big|_{\mathrm{Gr}(k, \mathbb{R}^n)} \quad \textit{and} \quad \overline{\lambda}(\cdot, \Phi)\big|_{\mathrm{Gr}(k, \mathbb{R}^n)}$$

are continuous and bounded.

Proof. We prove only the lower growth rate case. Let $\varepsilon \in \mathbb{R}_{>0}$. By Hypothesis 2.10, we have that

$$\mathcal{S} \times \overline{\neq_{\mathbb{I} \times \mathbb{I}}} \ni (x, t, s) \mapsto \Delta(|\Phi(\cdot, t_-)x|)(t, s) \in \mathbb{R}$$

is uniformly continuous. Hence, there exists $\delta \in \mathbb{R}_{>0}$ such that for any $x_1, x_2 \in \mathcal{S}$, $t_1, t_2, s_1, s_2 \in \mathbb{I}$ the inequality

$$\max\left\{|x_1 - x_2|, |t_1 - t_2|, |s_1 - s_2|\right\} < \delta$$

implies

$$|\Delta(|\Phi(\cdot, t_-)x_1|)(t_1, s_1) - \Delta(|\Phi(\cdot, t_-)x_2|)(t_2, s_2)| < \varepsilon.$$

Now, let $X, Y \in \mathrm{Gr}(k, \mathbb{R}^n)$ with $\Theta(X, Y) < \delta/2$ and

$$\underline{\lambda}(X, \Phi) = \Delta(|\Phi(\cdot, t_-)x_1|)(t_1, s_1), \quad \text{and} \quad \underline{\lambda}(Y, \Phi) = \Delta(|\Phi(\cdot, t_-)x_2|)(t_2, s_2),$$

for $x_1 \in X \cap \mathcal{S}$, $x_2 \in Y \cap \mathcal{S}$, $t_1, t_2, s_1, s_2 \in \mathbb{I}$. By Lemma B.8 there exist $\hat{x}_1 \in Y \cap \mathcal{S}$ and $\hat{x}_2 \in X \cap \mathcal{S}$ such that $|x_1 - \hat{x}_1|, |x_2 - \hat{x}_2| < \delta$. Finally, to prove that $|\underline{\lambda}(Y, \Phi) - \underline{\lambda}(X, \Phi)| < \varepsilon$, we consider two cases: if $\underline{\lambda}(Y, \Phi) \geq \underline{\lambda}(X, \Phi)$ we have

$$\begin{aligned} |\underline{\lambda}(Y, \Phi) - \underline{\lambda}(X, \Phi)| &= \Delta(|\Phi(\cdot, t_-)x_2|)(t_2, s_2) - \Delta(|\Phi(\cdot, t_-)x_1|)(t_1, s_1) \\ &\leq \Delta(|\Phi(\cdot, t_-)\hat{x}_1|)(t_1, s_1) - \Delta(|\Phi(\cdot, t_-)x_1|)(t_1, s_1) < \varepsilon, \end{aligned}$$

if $\underline{\lambda}(Y, \Phi) < \underline{\lambda}(X, \Phi)$ we have

$$\begin{aligned} |\underline{\lambda}(Y, \Phi) - \underline{\lambda}(X, \Phi)| &= \Delta(|\Phi(\cdot, t_-)x_1|)(t_1, s_1) - \Delta(|\Phi(\cdot, t_-)x_2|)(t_2, s_2) \\ &\leq \Delta(|\Phi(\cdot, t_-)\hat{x}_2|)(t_2, s_2) - \Delta(|\Phi(\cdot, t_-)x_2|)(t_2, s_2) < \varepsilon, \end{aligned}$$

which finishes the proof. □

For later robustness investigations the following (semi-)metric and the induced topology on $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ will turn out to be very helpful.

2.18 Definition (Metric on $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$)**.** We define $\tilde{d}_{\mathbb{I}}, d_{\mathbb{I}} \colon \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)^2 \to \mathbb{R}_{\geq 0}$ by

$$\tilde{d}_{\mathbb{I}} \colon (\Phi, \Psi) \mapsto \sup_{X \in \mathrm{Gr}(1, \mathbb{R}^n)} \left\{ \max \left\{ |\underline{\lambda}(X, \Phi) - \underline{\lambda}(X, \Psi)|, |\overline{\lambda}(X, \Phi) - \overline{\lambda}(X, \Psi)| \right\} \right\},$$

$$d_{\mathbb{I}} \colon (\Phi, \Psi) \mapsto \max \left\{ \sup_{x \in \mathcal{S}} \|(\Phi(\cdot, t_-) - \Psi(\cdot, t_-))x\|_\infty, \tilde{d}_{\mathbb{I}}(\Phi, \Psi) \right\}.$$

Obviously, $\tilde{d}_{\mathbb{I}}$ is a semimetric, i.e. it satisfies non-negativity, symmetry and the triangle inequality (but not definiteness), and $d_{\mathbb{I}}$ is a metric on $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$.

Note that, in general, $d_{\mathbb{I}}$ cannot be extended to a proper metric on the set of absolutely continuous linear processes due to possibly unbounded growth rates. Nevertheless, the (semi-)metric can be used to define open balls around absolutely continuous linear processes.

In the following, we endow $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ with the topology induced by $d_{\mathbb{I}}$.

By construction of the topology we obtain the continuous dependence of growth rates on the linear process.

2.19 Corollary. *Let $X \subseteq \mathbb{R}^n$ be a subspace. Then $\underline{\lambda}(X, \cdot), \overline{\lambda}(X, \cdot) \colon \mathcal{LP}(\mathbb{I}, \mathbb{R}^n) \to \mathbb{R} \cup \{-\infty, \infty\}$ are Lipschitz continuous with Lipschitz constant 1.*

2.4 Extremal Growth Rates

The following notion is connected to subspaces with optimal growth rates among subspaces of the same dimension. It will play an important role in the linear spectral theory as well as in robustness issues.

2.20 Definition (Extremal k-growth rates, cf. [17, 32, 31])**.** For $k \in \{1, \dots, n\}$ we define

$$\begin{aligned} \underline{\lambda}_k^{\mathbb{I}} &: \mathcal{LP}(\mathbb{I}, \mathbb{R}^n) \to \mathbb{R}, & \Phi &\mapsto \max \underline{\lambda}^{\mathbb{I}}([\mathrm{Gr}(k, \mathbb{R}^n)], \Phi), \\ \overline{\lambda}_k^{\mathbb{I}} &: \mathcal{LP}(\mathbb{I}, \mathbb{R}^n) \to \mathbb{R}, & \Phi &\mapsto \min \overline{\lambda}^{\mathbb{I}}([\mathrm{Gr}(k, \mathbb{R}^n)], \Phi), \end{aligned}$$

and call $\underline{\lambda}_k^{\mathbb{I}}(\Phi)$ and $\overline{\lambda}_k^{\mathbb{I}}(\Phi)$, respectively, the *maximal lower* and *minimal upper k-growth rate of* Φ. For convenience, we drop the index $\mathbb{I}$ in case the time-set is clear from the context. We extend the above definition naturally by $\underline{\lambda}_0^{\mathbb{I}}(\Phi) = \infty$ and $\overline{\lambda}_0(\Phi) = -\infty$.

2.21 Remark. Note that the extremal k-growth rate functions are well-defined due to the continuity from Proposition 2.17 and the compactness of $(\mathrm{Gr}(k, \mathbb{R}^n), \Theta)$. That means in particular that for any $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ and $k \in \{0, \dots, n\}$ there exist $X, Y \in \mathrm{Gr}(k, \mathbb{R}^n)$ such that $\underline{\lambda}(X, \Phi) = \underline{\lambda}_k(\Phi)$ and $\overline{\lambda}(Y, \Phi) = \overline{\lambda}_k(\Phi)$. We refer to such subspaces as *extremal subspaces*. Of course, in general, extremal subspaces need not be unique.

2.22 Lemma (cf. [17, Remark 11], [32, Remark 9])**.** *For any $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$, the extremal growth rates are ordered and nested as follows:*

$$\begin{aligned} -\infty = \overline{\lambda}_0(\Phi) < \underline{\lambda}_n(\Phi) \leq \overline{\lambda}_1(\Phi) \leq \underline{\lambda}_{n-1}(\Phi) \leq \dots \\ \dots \leq \overline{\lambda}_{n-1}(\Phi) \leq \underline{\lambda}_1(\Phi) \leq \overline{\lambda}_n(\Phi) < \underline{\lambda}_0(\Phi) = \infty. \end{aligned}$$

Proof. The ordering properties follow directly from Lemma 2.16(i) and the nesting property from Lemma 2.16(ii). □

Another observation is the continuity of the extremal growth rate functions, which easily follows from Corollary 2.19.

2.23 Lemma. *For any $k \in \{1, \dots, n\}$ the extremal k-growth rate functions $\underline{\lambda}_k, \overline{\lambda}_k$ are Lipschitz continuous with Lipschitz constant* 1.

Note that [31, Theorem 20] is essentially the ε-δ-notation of the continuity established in Lemma 2.23 for the special case that $\mathbb{I}$ is finite.

More interesting is the following convergence result with respect to the time-set. We denote by d_{H} the Hausdorff metric on the space of compact subsets of $M \subseteq \mathbb{R}$, denoted by $\mathcal{K}(M)$.

2.24 Lemma ([31, Theorem 17]). *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$. Then for each $k \in \{1, \dots, n\}$ and any sequence $\mathbb{J} = (\mathbb{J}_i)_{i\in\mathbb{N}} \in \mathcal{K}(\mathbb{I})^{\mathbb{N}}$ of compact subsets of $\mathbb{I}$ with $d_{\mathrm{H}}(\mathbb{I}, \mathbb{J}_i) \xrightarrow{i\to\infty} 0$ one has*

$$\lim_{i\to\infty} \left| \underline{\lambda}_k^{\mathbb{J}_i}(\Phi\big|_{\mathbb{J}_i^2}) - \underline{\lambda}_k^{\mathbb{I}}(\Phi) \right| = 0, \qquad \textit{and} \qquad \lim_{i\to\infty} \left| \overline{\lambda}_k^{\mathbb{J}_i}(\Phi\big|_{\mathbb{J}_i^2}) - \overline{\lambda}_k^{\mathbb{I}}(\Phi) \right| = 0.$$

Proof. First note that the statement is trivial if $\mathbb{I}$ is finite, that is why we assume that $\mathbb{I}$ is infinite in the following. Furthermore, without loss of generality, we may assume that $\operatorname{card}(\mathbb{J}_i) > 2$ and $t_- \in \mathbb{J}_i$ for all $i \in \mathbb{N}$. For convenience, we abbreviate $\Phi_i := \Phi\big|_{\mathbb{J}_i^2}$. Clearly, from the assumption $\mathbb{J}_i \subseteq \mathbb{I}$, Definition 2.13 and Definition 2.20 it follows that

$$\liminf_{i\to\infty} \underline{\lambda}_k^{\mathbb{J}_i}(\Phi_i) \geq \underline{\lambda}_k^{\mathbb{I}}(\Phi), \qquad \limsup_{i\to\infty} \overline{\lambda}_k^{\mathbb{J}_i}(\Phi_i) \leq \overline{\lambda}_k^{\mathbb{I}}(\Phi).$$

Hence, it remains to prove that for all $k \in \{1, \dots, n\}$ we have

$$\limsup_{i\to\infty} \underline{\lambda}_k^{\mathbb{J}_i}(\Phi_i) \leq \underline{\lambda}_k^{\mathbb{I}}(\Phi), \qquad \liminf_{i\to\infty} \overline{\lambda}_k^{\mathbb{J}_i}(\Phi_i) \geq \overline{\lambda}_k^{\mathbb{I}}(\Phi). \tag{2.1}$$

We prove only the first inequality in Eq. (2.1) since the second is proved analogously. To this end, let $k \in \{1, \dots, n\}$ and $(X_i)_{i\in\mathbb{N}} \in \operatorname{Gr}(k, \mathbb{R}^n)^{\mathbb{N}}$ satisfy $\underline{\lambda}^{\mathbb{J}_i}(X_i, \Phi_i) = \underline{\lambda}_k^{\mathbb{J}_i}(\Phi_i) =: \alpha_i$, i.e. $(X_i)_{i\in\mathbb{N}}$ is a sequence of subspaces realizing the maximal lower k-growth rate with respect to $\mathbb{J}_i$. Suppose to the contrary that

$$\alpha' := \limsup_{i\to\infty} \alpha_i > \underline{\lambda}_k^{\mathbb{I}}(\Phi) =: \alpha.$$

Define $\delta := \frac{\alpha' - \alpha}{2}$. Then there exists a subsequence $\left(\alpha_{i_j}\right)_{j\in\mathbb{N}}$ such that

(i) $\alpha_{i_j} > \alpha + \delta$ for all $j \in \mathbb{N}$, and

(ii) the associated sequence of subspaces $\left(X_{i_j}\right)_{j\in\mathbb{N}} \in \operatorname{Gr}(k, \mathbb{R}^n)^{\mathbb{N}}$ is convergent with $\lim_{j\to\infty} X_{i_j} =: X \in \operatorname{Gr}(k, \mathbb{R}^n)$, due to compactness of $\operatorname{Gr}(k, \mathbb{R}^n)$.

Let $(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}$ and $x \in X \cap \mathcal{S}$. By assumption, there exist sequences $(t_j, s_j)_{j\in\mathbb{N}} \in (\neq_{\mathbb{I}\times\mathbb{I}})^{\mathbb{N}}$ and $(x_j)_{j\in\mathbb{N}} \in \mathcal{S}^{\mathbb{N}}$ satisfying

(i) for each $j \in \mathbb{N}$ that $t_j, s_j \in \mathbb{J}_{i_j}$, $x_j \in X_{i_j}$,

(ii) $\lim_{j\to\infty}(t_j, s_j) = (t,s)$ and $\lim_{j\to\infty} x_j = x$, and

(iii) for all $j \in \mathbb{N}$ we have $\Delta(\left|\Phi(\cdot, t_-)x_j\right|)(t_j, s_j) \geq \alpha_{i_j}$.

Due to Lemma 2.7 we obtain in the limit

$$\Delta(|\Phi(\cdot, t_-)x|)(t,s) \geq \liminf_{j\to\infty} \alpha_{i_j} \geq \alpha + \delta,$$

which implies that $\underline{\lambda}^{\mathbb{I}}(X, \Phi) \geq \alpha + \delta$. In other words, there exists a subspace of dimension k with its lower growth rate being larger than α. This is a contradiction to the maximality of $\alpha = \underline{\lambda}_k^{\mathbb{I}}(\Phi)$. Similarly, we have $\liminf_{i\to\infty} \overline{\lambda}_k^{\mathbb{J}_i}(\Phi) \geq \overline{\lambda}_k^{\mathbb{I}}(\Phi)$ for each $k \in \{1, \dots, n\}$ and Eq. (2.1) is proved. □

3 Finite-time Hyperbolicity

This chapter is devoted to the introduction and investigation of a notion of finite-time hyperbolicity for linear finite-time processes, which is in the spirit of [98, 103, 17]. We start with a notion based on a certain dynamical dichotomy. We denote by

$$\mathbb{P}(\mathbb{R}^n) := \{P \in L(\mathbb{R}^n);\ P^2 = P\}$$

the set of (linear) projections on $\mathbb{R}^n$. Throughout this chapter, let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ denote a linear process on $\mathbb{I}$.

3.1 Definition (Invariant projector)**.** Suppose we have subspaces $U, V \subseteq \mathbb{R}^n$ with $\mathbb{R}^n = U \oplus V$, which define a unique projection $Q \in \mathbb{P}(\mathbb{R}^n)$ by $U = \operatorname{im} Q$ and $V = \ker Q$. In turn, Q defines a so-called *invariant projector* $P \colon \mathbb{I} \to \mathbb{P}(\mathbb{R}^n)$ and vice versa in the following way:

$$P := \Phi(\cdot, t_-) Q \Phi(t_-, \cdot).$$

3.2 Remark. By definition an invariant projector satisfies

$$\bigwedge_{s,t \in \mathbb{I}} P(t)\Phi(t,s) = \Phi(t,s)P(s),$$

which is often used as the definition of invariant projectors. Due to the continuity of Φ an invariant projector is continuous. Furthermore, by the fact that Φ is isomorphism-valued we have $\dim \operatorname{im} P(t) = \dim \operatorname{im} Q$ and $\dim \ker P(t) = \dim \ker Q$ for all $t \in \mathbb{I}$. For that reason, we define the rank and the deficiency of the invariant projector P as the rank and the deficiency of the projection Q, respectively.

3.1 Exponential Monotonicity Dichotomy

3.3 Definition (Exponential monotonicity dichotomy)**.** Φ admits an *exponential monotonicity dichotomy (EMD) on* $\mathbb{I}$ if there exists $k \in \{0, \dots, n\}$ such that

$$\overline{\lambda}_k(\Phi) < 0 < \underline{\lambda}_{n-k}(\Phi). \tag{3.1}$$

For brevity, we sometimes call a linear process Φ on $\mathbb{I}$ admitting an EMD also *(finite-time) hyperbolic*. We call a hyperbolic linear process Φ on $\mathbb{I}$ *(finite-time) attractive* if $k = n$ and *(finite-time) repulsive* if $k = 0$.

3.4 Remark. Since the logarithmic difference quotient depends on the norm, so do the growth rates and hence the property of admitting an EMD. In general, it might be possible that a linear process on $\mathbb{I}$ is hyperbolic with respect to one norm, but is not with respect to another norm; see [17, Example 14].

In the next lemma, we characterize an EMD on $\mathbb{I}$.

3.5 Lemma (cf. [17])**.** *The following statements are equivalent:*

(i) Φ *admits an EMD on* $\mathbb{I}$ *with* $k \in \{0, \dots, n\}$.

(ii) There exist subspaces $X \in \mathrm{Gr}(k, \mathbb{R}^n)$ *and* $Y \in \mathrm{Gr}(n-k, \mathbb{R}^n)$ *such that* $\overline{\lambda}(X, \Phi) < 0 < \underline{\lambda}(Y, \Phi)$.

(iii) There exists a projection $Q \in \mathbb{P}(\mathbb{R}^n)$ *with* $\mathrm{rk}\, Q = k$ *such that* $\overline{\lambda}(\mathrm{im}\, Q, \Phi) < 0 < \underline{\lambda}(\ker Q, \Phi)$.

(iv) There exist a projection $Q \in \mathbb{P}(\mathbb{R}^n)$ *and* $\alpha, \beta \in \mathbb{R}_{>0}$ *such that for any* $t, s \in \mathbb{I}$, $t \geq s$, *and* $x \in \mathrm{im}\, Q$, $y \in \ker Q$ *one has*

$$\begin{aligned} |\Phi(t, t_-)x| &\leq \mathrm{e}^{-\alpha(t-s)} \, |\Phi(s, t_-)x| \, , \\ |\Phi(t, t_-)y| &\geq \mathrm{e}^{\beta(t-s)} \, |\Phi(s, t_-)y| \, . \end{aligned} \tag{3.2}$$

(v) There exist an invariant projector $P \in C(\mathbb{I}, \mathbb{P}(\mathbb{R}^n))$ *and* $\alpha, \beta \in \mathbb{R}_{>0}$ *such that for any* $t, s \in \mathbb{I}$, $t \geq s$, *and* $x, y \in \mathbb{R}^n$ *one has*

$$\begin{aligned} |\Phi(t, s)P(s)x| &\leq \mathrm{e}^{-\alpha(t-s)} \, |P(s)x| \, , \\ |\Phi(s, t)(\mathrm{id}_{\mathbb{R}^n} - P(t))x| &\leq \mathrm{e}^{\beta(s-t)} \, |(\mathrm{id}_{\mathbb{R}^n} - P(t))x| \, , \end{aligned}$$

see [15, Definition 1].

Proof. (i) $\Rightarrow$ (ii): By Remark 2.21 there exist subspaces $X \in \mathrm{Gr}(k, \mathbb{R}^n)$ and $Y \in \mathrm{Gr}(n-k, \mathbb{R}^n)$ satisfying $\overline{\lambda}(X, \Phi) = \overline{\lambda}_k(\Phi) < 0 < \underline{\lambda}_{n-k}(\Phi) = \underline{\lambda}(Y, \Phi)$.

(ii) $\Rightarrow$ (i): By Definition 2.20, we obtain

$$\overline{\lambda}_k(\Phi) \leq \overline{\lambda}(X, \Phi) < 0 < \underline{\lambda}(Y, \Phi) \leq \underline{\lambda}_{n-k}(\Phi).$$

(ii) $\Leftrightarrow$ (iii): The subspaces X and Y define a projection by $\operatorname{im} Q := X$ and $\ker Q := Y$ and vice versa.

(iii) $\Leftrightarrow$ (iv): This follows directly from Lemma 2.12.

(iv) $\Leftrightarrow$ (v): Holds by Definition 3.1. □

For a hyperbolic linear process Φ we refer to subspaces and projections satisfying (ii) and (iii) of Lemma 3.5 as EMD-subspaces and EMD-projections, respectively. Although we prefer the definition of EMD to the equivalent characterizations in terms of EMD-subspaces or -projections, we will sometimes speak synonymously of Φ admitting an EMD on $\mathbb{I}$ with respect to a projection $Q \in \mathbb{P}(\mathbb{R}^n)$, for instance.

The basic concept of a dynamical dichotomy has a predecessor in the *exponential dichotomy (ED)* in the asymptotic analysis of dynamical systems. Next, we recall that notion. Let $J \subseteq \mathbb{R}$ be an interval and $A \in L(R^n)^J$ an admissible operator-valued function defined on J, where admissible means that

$$\dot{x} = A(t)x$$

is well-posed and hence admits a solution operator $\Phi\colon J \times J \to L(\mathbb{R}^n)$. Now Φ is said to admit an *exponential dichotomy on J* if there exist an invariant projector $P \in C(J, \mathbb{P}(\mathbb{R}^n))$ and $K \in \mathbb{R}_{\geq 1}$, $\alpha, \beta \in \mathbb{R}_{>0}$ such that for any $t, s \in J$, $t \geq s$ one has

$$\begin{aligned} \|\Phi(t,s)P(s)\| &\leq K\mathrm{e}^{-\alpha(t-s)}, \\ \|\Phi(s,t)(\mathrm{id}_{\mathbb{R}^n} - P(t))\| &\leq K\mathrm{e}^{\beta(s-t)}. \end{aligned}$$

For more details on exponential dichotomy we refer the reader to [78, 28, 26].

3.6 Remark. We want to emphasize the following points concerning Definition 3.3.

1. The two essential features that both ED and EMD share are:

   a) the existence of a decomposition of the state space into the direct sum of two subspaces such that trajectories under Φ are bounded in the norm by exponentially decaying/growing functions,

   b) and the exponential decay/growth rates α and β can be chosen uniformly with respect to the initial state and time.

   In both cases the decomposition of the state space is also referred to as *hyperbolic splitting*.

2. We want to emphasize that ED is a notion aiming at the asymptotic analysis of linear processes, whereas EMD aims at the transient behavior of linear processes. Therefore, the essential difference between the two notions lies in the constant K in the estimates in (3.2), just as discussed for the one-dimensional growth rates in Remark 2.15. In the general case of ED the constant is allowed to be larger than 1, in order to "conceal" transient behavior and to capture asymptotic exponential growth rates only. In the two-sided infinite-time situation with state space $\mathbb{R}^n$ or $\mathbb{C}^n$ the constant K can be additionally linked to the infimum of the angles between image and kernel of the invariant projector over time. An estimate of the form (3.2) for the solution operator of a linear differential equation with $K = 1$ corresponds to orthogonality of $\operatorname{im} P(t)$ and $\ker P(t)$ for all $t \in \mathbb{R}$ and, in this special case of exponential dichotomy, one speaks of *strong dichotomy*; see [78, 107]. As we see, and as was already discussed in [17], notions of finite-time hyperbolicity must be more than just time-restrictions of infinite-time hyperbolicity notions.
3. In contrast to the two-sided infinite-time situation, in the finite-time context the projection Q of a linear process on I admitting an EMD needs not be unique; see, for instance, [17, Example 4]. Consequently, stable/unstable subspace and linear integral manifold are neither. In fact, in Lemma 3.18 we are going to show that EMD-projections are always non-unique.

As can be seen directly from Lemma 3.5, Definition 3.3 includes some other finite-time hyperbolicity notions as special cases. In the finite-time context so far only solution operators $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ of linear differential equations on a compact time-interval $\mathbb{I} = [\tau, \tau + T]$, $\tau \in \mathbb{R}$, $T \in \mathbb{R}_{>0}$, i.e.

$$\dot{x} = A(t)x,$$

where usually $A \in C(\mathbb{I}, L(\mathbb{R}^n))$, have been considered. In this case, Φ regarded as a linear process satisfies even continuous differentiability where we required just Lipschitz continuity in Definition 2.1. From Lemma 2.12 and Lemma 3.5 it follows directly that, setting $\mathbb{I} = [\tau, \tau + T]$, our definition of EMD coincides with the definition of *M-hyperbolicity* as defined in [16, 32], *(finite-time) hyperbolicity* as defined in [17, 15] and *uniform hyperbolicity* as in [53]. By setting $\mathbb{I} = J$ it is obvious that our definition of an EMD corresponds essentially to the finite-time hyperbolicity proposed in [31], which generalizes in particular the *nonhyperbolic (τ, T)-dichotomy* as suggested in [95]. From Lemma 2.12 and taking into account the monotonicity preserving property of the logarithm one easily concludes that nonhyperbolic (τ, T)-dichotomy is equivalent to Definition 3.3 with

$\mathbb{I} = \{\tau, \tau + T\}$. Another finite-time hyperbolicity notion of the EMD-type which is based only on the start and end time-points can be found in the finite-time Lyapunov exponent approach; see, for instance, [54, 101] and Section 7.1. As we are going to show in Proposition 7.1, the FTLE approach is closely related to our finite-time hyperbolicity notion.

Note that usually EMD-like finite-time hyperbolicity definitions are expressed in terms of invariant projectors; see for instance [17, 14, 15, 32, 31].

By the nesting property of the extremal growth rates, see Lemma 2.16, we have that $k \in \{0, \dots, n\}$ in Definition 3.3 is uniquely defined. Of course, this does not imply that the subspaces/projections mentioned in Lemma 3.5 are unique. Consequently, the notion of EMD is well-defined, at least up to rank of the projection, in contrast to the definition of a finite-time exponential dichotomy.

In the following, we investigate for a given linear process Φ the associated family of linear processes $(\Phi_\gamma)_{\gamma\in\mathbb{R}} \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)^{\mathbb{R}}$ which is defined for any $\gamma \in \mathbb{R}$ and $s, t \in \mathbb{I}$ by:

$$\Phi_\gamma(t, s) := \mathrm{e}^{-\gamma(t-s)}\Phi(t, s). \tag{3.3}$$

The motivation to study Φ_γ comes from the fact that they arise naturally in the study of linear differential equations

$$\dot{x} = A(t)x,$$

where, for instance, $A \in C(\mathbb{I}, L(\mathbb{R}^n))$, $\mathbb{I}$ an interval, and the corresponding *shifted* linear differential equations for $\gamma \in \mathbb{R}$

$$\dot{x} = (A(t) - \gamma\,\mathrm{id}_{\mathbb{R}^n})x. \tag{3.4}$$

Clearly, we have by definition that for any $x \in \mathbb{R}^n$ and $(t, s) \in \neq_{\mathbb{I}\times\mathbb{I}}$ holds

$$\Delta(|\Phi_\gamma(\cdot, t_-)x|)(t, s) = \Delta(|\Phi(\cdot, t_-)x|)(t, s) - \gamma.$$

As trivial consequences of this observation we obtain the following results. Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ and $\gamma \in \mathbb{R}$. Φ_γ admits an EMD on $\mathbb{I}$ with projection $Q \in \mathbb{P}(\mathbb{R}^n)$ if and only if $\overline{\lambda}(\operatorname{im} Q, \Phi) < \gamma < \underline{\lambda}(\ker Q, \Phi)$. Furthermore, the extremal growth rates behave as the logarithmic difference quotient under the exponential weight, i.e.

$$\underline{\lambda}_k(\Phi_\gamma) + \gamma = \underline{\lambda}_k(\Phi) \quad \text{and} \quad \overline{\lambda}_k(\Phi_\gamma) + \gamma = \overline{\lambda}_k(\Phi). \tag{3.5}$$

Consequently, the following characterization of EMD for weighted linear processes holds.

3.7 Corollary. *Φ_γ admits an EMD on $\mathbb{I}$ with $k \in \{0, \ldots, n\}$ if and only if $\overline{\lambda}_k(\Phi) < \gamma < \underline{\lambda}_{n-k}(\Phi)$.*

Since the semimetric $\tilde{d}_{\mathbb{I}}$ depends only on the growth rates, we get the following normalization property: $\tilde{d}_{\mathbb{I}}(\Phi, \Phi_\gamma) = \gamma$.

3.2 Spectrum of Linear Finite-time Processes

Next, we introduce a finite-time spectral notion which is based on the EMD.

3.8 Definition (Finite-time dichotomy spectrum, cf. [17, 95, 32, 31])**.** We define

$$\Sigma^{\mathbb{I}}\colon \mathcal{LP}(\mathbb{I}, \mathbb{R}^n) \to 2^{\mathbb{R}}, \qquad \Phi \mapsto \{\gamma \in \mathbb{R};\ \Phi_\gamma \text{ does not admit an EMD on } \mathbb{I}\},$$

and call $\Sigma^{\mathbb{I}}(\Phi)$ the *(finite-time dichotomy) spectrum* and $\varrho(\Phi) := \mathbb{R} \setminus \Sigma^{\mathbb{I}}(\Phi)$ the *(finite-time) resolvent set of* Φ, respectively.

3.9 Remark. By Definition 3.8 it is clear that for two compact sets $\mathbb{J}, \mathbb{I} \subset \mathbb{R}$ with $\mathbb{J} \subseteq \mathbb{I}$ and $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ we have that $\Sigma^{\mathbb{J}}(\Phi|_{\mathbb{J}^2}) \subseteq \Sigma^{\mathbb{I}}(\Phi)$.

3.10 Theorem (Spectral Theorem, [31, Theorem 10], cf. also [17, 32])**.** *Denote*

$$\{i_0, \ldots, i_d\} := \left\{ j \in \{0, \ldots, n\}\,;\ \overline{\lambda}_j(\Phi) < \underline{\lambda}_{n-j}(\Phi) \right\},\ i_k < i_{k+1},\ k \in \{0, \ldots, d-1\}. \tag{3.6}$$

Then the spectrum of Φ is the union of d disjoint compact intervals, i.e.

$$\Sigma^{\mathbb{I}}(\Phi) = \bigcup_{k=1}^{d} [\underline{\lambda}_{n-i_{k-1}}(\Phi), \overline{\lambda}_{i_k}(\Phi)],$$

and we call $[\underline{\lambda}_{n-i_{k-1}}(\Phi), \overline{\lambda}_{i_k}(\Phi)]$ the k-th spectral interval.

Proof. This follows directly from Corollary 3.7, i.e. from the fact that

$$\varrho^{\mathbb{I}}(\Phi) = \bigcup_{k=0}^{n} (\overline{\lambda}_k(\Phi), \underline{\lambda}_{n-k}(\Phi)) = \bigcup_{k=0}^{d} (\overline{\lambda}_{i_k}(\Phi), \underline{\lambda}_{n-i_k}(\Phi)),$$

since $(\overline{\lambda}_k(\Phi), \underline{\lambda}_{n-k}(\Phi)) = \varnothing$ for $k \in \{0, \ldots, n\} \setminus \{i_0, \ldots, i_d\}$, and the nesting property of the extremal growth rates stated in Lemma 2.22. □

The assertion of Theorem 3.10 remains basically true if we require only absolute continuity of Φ. However, in this case, the left-most and right-most spectral intervals may be unbounded; see also the remark after Definition 2.13 and [17, Theorem 17]. Theorem 3.10 together with Lemma 2.22 implies that $\Sigma^{\mathbb{I}}(\Phi)$ is non-empty for any linear process Φ.

Due to the simple interval-structure of the spectrum where the endpoints of the intervals are the extremal growth rates, we obtain the continuous dependence of $\Sigma^{\mathbb{I}}$ on the linear process from Lemma 2.23.

3.11 Proposition. *The spectrum function* $\Sigma^{\mathbb{I}}\colon (\mathcal{LP}(\mathbb{I},\mathbb{R}^n), d_{\mathbb{I}}) \to (\mathcal{K}(\mathbb{R}), d_{\mathrm{H}})$ *is Lipschitz continuous with Lipschitz constant* 1.

Additionally, we obtain a kind of continuity result of the spectrum with respect to the time-set, corresponding to Lemma 2.24.

3.12 Corollary ([31, Theorem 17])**.** *For any* $(\mathbb{J}_i)_{i\in\mathbb{N}} \in \mathcal{K}(\mathbb{I})^{\mathbb{N}}$ *one has*

$$\lim_{i\to\infty} d_{\mathrm{H}}(\mathbb{I},\mathbb{J}_i) = 0 \;\Rightarrow\; \lim_{i\to\infty} d_{\mathrm{H}}(\Sigma^{\mathbb{I}}(\Phi), \Sigma^{\mathbb{J}_i}(\Phi|_{\mathbb{J}_i^2})) = 0.$$

Proof. This follows from Theorem 3.10 and Lemma 2.24. □

3.3 Robustness of Hyperbolicity

A desired property of hyperbolicity is its robustness under perturbations. Due to our previous work on continuous dependence of growth rates this is established easily.

3.13 Theorem (Robustness of EMD)**.** *Let* Φ *admit an EMD on* $\mathbb{I}$ *with* $k \in \{0,\dots,n\}$. *Then any* $\Psi \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ *with* $\tilde{d}_{\mathbb{I}}(\Phi,\Psi) < \min\{-\overline{\lambda}_k(\Phi), \underline{\lambda}_{n-k}(\Phi)\}$ *admits an EMD on* $\mathbb{I}$ *with* k *(and the same extremal subspaces).*

Proof. This is a simple consequence of Lemma 2.23 and Definition 2.18, i.e. the fact that the extremal growth rates of Φ and Ψ differ at most as much as Φ and Ψ do. By Corollary 2.19 the same holds for the subspaces realizing the extremal growth rates, so these can be chosen to define a hyperbolic splitting. □

3.14 Remark. For proofs of robustness in the infinite-time situation see [26, Lecture 4] for solution operators of linear ODEs on $\mathbb{R}$, also [107, 108] for (semi-)strong exponential dichotomy, [93] for linear skew-product semiflows on general Banach spaces and [63, 94] for further improvements of the aforementioned results.

3.15 Corollary. *For given* $\mathbb{I}$ *the set*

$$\{\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n);\ \Phi \text{ admits an EMD on } \mathbb{I}\}$$

is open in $\mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ *with respect to the topology induced by* $d_{\mathbb{I}}$*.*

Next, we are going to show that the estimate in Theorem 3.13 is sharp and hence gives the maximal perturbation bound for EMD persistence. This can be interpreted as a hyperbolicity radius or, in case Φ is an attractive linear process, as a stability radius.

3.16 Theorem (Hyperbolicity radius)**.** *Let* Φ *admit an EMD on* $\mathbb{I}$*. Then*

$$\theta := \operatorname{dist}(\{0\}, \Sigma^{\mathbb{I}}(\Phi)) \in \mathbb{R}_{>0}$$

is the largest number such that any $\Psi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ *with* $\tilde{d}_{\mathbb{I}}(\Phi, \Psi) < \theta$ *admits an EMD on* $\mathbb{I}$*.*

Proof. The fact that θ is a number with the asserted property is clear by Theorem 3.13, thus it remains to show that it is the largest one. This is indeed clear by the normalization property of $\tilde{d}_{\mathbb{I}}$, i.e. $\tilde{d}_{\mathbb{I}}(\Phi, \Phi_\theta) = \theta$ and that Φ_θ does not admit an EMD on $\mathbb{I}$. The latter is due to the fact that $\theta \in \Sigma^{\mathbb{I}}(\Phi)$. □

3.4 Spectral Cones

In this section, we change the focus from subspaces to cones. This was originally proposed in [32].

3.17 Definition (Stable/unstable cone, cf. [32, p. 11])**.** We define

$$\begin{aligned} V^{\mathrm{s}} \colon \mathcal{LP}(\mathbb{I}, \mathbb{R}^n) &\to 2^{\mathbb{R}^n}, \qquad & \Phi &\mapsto \bigcup \{X \in \operatorname{Gr}(1, \mathbb{R}^n);\ \overline{\lambda}(X, \Phi) < 0\}, \\ V^{\mathrm{u}} \colon \mathcal{LP}(\mathbb{I}, \mathbb{R}^n) &\to 2^{\mathbb{R}^n}, \qquad & \Phi &\mapsto \bigcup \{X \in \operatorname{Gr}(1, \mathbb{R}^n);\ \underline{\lambda}(X, \Phi) > 0\}, \end{aligned}$$

where we call $V^{\mathrm{s}}(\Phi)$ and $V^{\mathrm{u}}(\Phi)$ the *stable* and *unstable cone of* Φ, respectively.

Clearly, we can equivalently consider $V^{\mathrm{s}}(\Phi)$ and $V^{\mathrm{u}}(\Phi)$ as subsets of $\mathrm{Gr}(1,\mathbb{R}^n)$ (by not taking the union in Definition 3.17). In the following, we identify the set of points lying in a one-dimensional subspace $X \subseteq \mathbb{R}^n$ with the point $X \in \mathrm{Gr}(1,\mathbb{R}^n)$. Consequently, we obtain

$$V = \bigcup \{X \in \mathrm{Gr}(1,\mathbb{R}^n);\ X \subseteq V\} \cong \{X \in \mathrm{Gr}(1,\mathbb{R}^n);\ X \subseteq V\} = \mathrm{Gr}(1,V), \quad (3.7)$$

and in that sense we can speak of a subspace being compact. By the continuity of the growth rate functions for fixed Φ proved in Proposition 2.17, we obtain the following result.

3.18 Lemma. *The stable and unstable cones of Φ are open subsets of $\mathrm{Gr}(1,\mathbb{R}^n)$.*

The next proposition generalizes [32, Theorem 14] to compact time-sets and partially removes the differentiability assumption on the norm, see Remark 2.11.

3.19 Proposition (cf. [32, Theorem 14])**.** *The following statements hold:*

(i) If Φ admits an EMD on $\mathbb{I}$ with $Q \in \mathbb{P}(\mathbb{R}^n)$, then

$$\operatorname{im} Q \subseteq V^{\mathrm{s}}(\Phi), \qquad \text{and} \qquad \ker Q \subseteq V^{\mathrm{u}}(\Phi).$$

(ii) Suppose there exist subspaces $U_1 \subseteq V^{\mathrm{s}}(\Phi)$ and $U_2 \subseteq V^{\mathrm{u}}(\Phi)$ with $U_1 \oplus U_2 = \mathbb{R}^n$. Then Φ admits an EMD with projection Q defined by $\operatorname{im} Q = U_1$ and $\ker Q = U_2$.

Proof. (i): This follows directly from Definition 3.3 and Definition 3.17.
(ii): We need to show that $\overline{\lambda}(U_1,\Phi) < 0 < \underline{\lambda}(U_2,\Phi)$. This is clear by the continuity of the one-dimensional growth rate functions from Proposition 2.17 and the compactness of U_1 and U_2 considered as subsets of $\mathrm{Gr}(1,\mathbb{R}^n)$. □

3.20 Remark. From the proof of part (ii) we can read off directly that the supremum (infimum) over upper (lower) growth rates with respect to an arbitrary compact subset of $V^{\mathrm{s}}(\Phi)$ ($V^{\mathrm{u}}(\Phi)$), considered as subsets of $\mathrm{Gr}(1,\mathbb{R}^n)$, gives a negative (positive) number.

In the following, we are going to study the behavior of stable and unstable cones under variation of the weighting parameter $\gamma \in \mathbb{R}$ for fixed linear process Φ, i.e. the functions

$$\mathbb{R} \to 2^{\mathbb{R}^n}, \qquad \begin{cases} \gamma \mapsto V^{\mathrm{s}}(\Phi_\gamma), \\ \gamma \mapsto V^{\mathrm{u}}(\Phi_\gamma). \end{cases} \quad (3.8)$$

The sets $V^{\mathrm{s}}(\Phi_\gamma)$ and $V^{\mathrm{u}}(\Phi_\gamma)$ can be considered as finite-time analogues of the sets of so-called γ^+- and γ^--quasibounded solutions in the two-sided infinite-time ODE case; see [103, p. 247]. Clearly, for any $\gamma \in \mathbb{R}$ we have

$$\begin{aligned} V^{\mathrm{s}}(\Phi_\gamma) &= \left\{x \in \mathbb{R}^n;\ \overline{\lambda}(\operatorname{span}\{x\}, \Phi) < \gamma\right\}, \\ V^{\mathrm{u}}(\Phi_\gamma) &= \{x \in \mathbb{R}^n;\ \underline{\lambda}(\operatorname{span}\{x\}, \Phi) > \gamma\}. \end{aligned} \tag{3.9}$$

Less formally spoken, $V^{\mathrm{s}}(\gamma)$ and $V^{\mathrm{u}}(\gamma)$ contain those initial values, whose evolutions "proceed" at an exponential rate of at most and at least γ, respectively. The following result is clear by definition in Eq. (3.8).

3.21 Lemma. *The functions $\gamma \mapsto V^{\mathrm{s}}(\Phi_\gamma)$ and $\gamma \mapsto V^{\mathrm{u}}(\Phi_\gamma)$ are increasing and decreasing with respect to set-inclusion, respectively.*

Thus, the families of cones $(V^{\mathrm{s}}(\Phi_\gamma))_{\gamma\in\mathbb{R}}$ and $(V^{\mathrm{u}}(\Phi_\gamma))_{\gamma\in\mathbb{R}}$ can be regarded as (reversed) filtrations of cones.

Theorem 3.10 clarifies the structure of the spectrum and consequently of the resolvent set. The following relationship between the parameter $\gamma \in \varrho(\Phi)$ and the rank of the associated EMD-projections is easily seen. We provide a direct and short proof for [17, Lemma 16].

3.22 Lemma (cf. [17, Lemma 16])**.** *The following function is well-defined:*

$$r\colon \varrho(\Phi) \to \mathbb{N}, \qquad\qquad \gamma \mapsto\in \big(\operatorname{rk}[P(\gamma)]\big),$$

where

$$P\colon \varrho(\Phi) \to 2^{\mathbb{P}(\mathbb{R}^n)}, \quad \gamma \mapsto \{Q \in \mathbb{P}(\mathbb{R}^n));\ \Phi_\gamma \textit{ admits an EMD on } \mathbb{I} \textit{ with } Q\},$$

and one has

$$\gamma \in (\overline{\lambda}_{i_k}, \underline{\lambda}_{n-i_k}) \subseteq \varrho(\Phi) \iff r(\gamma) = i_k. \tag{3.10}$$

Proof. By Lemma 2.22 the set $\operatorname{rk}[p(\gamma)]$ is a singleton set and hence the function r is well-defined. The equivalence in Eq. (3.10) now follows from Eq. (3.5), Corollary 3.7 and Lemma 3.5, which states that the rank of an EMD-projection can be read off as the index of the largest maximal lower growth rate smaller than γ. □

Clearly, $r(\gamma)$ also corresponds to the maximal dimension of subspaces contained in $V^{\mathrm{s}}(\Phi_\gamma)$ for $\gamma \in \varrho(\Phi)$. Opposing to this piecewise constancy we have the following strict monotonicity result.

3.23 Lemma. *The functions* $V^{\mathrm{s}}\big|_{[\underline{\lambda}_n(\Phi),\overline{\lambda}_n(\Phi)]}$ *and* $V^{\mathrm{u}}\big|_{[\underline{\lambda}_n(\Phi),\overline{\lambda}_n(\Phi)]}$ *are, respectively, strictly increasing and strictly decreasing with respect to set-inclusion.*

Proof. We prove the assertion for $V^{\mathrm{s}}\big|_{[\underline{\lambda}_n(\Phi),\overline{\lambda}_n(\Phi)]}$ since the other case is completely analogous. First we observe, that $V^{\mathrm{s}}(\Phi_\gamma) = \{0\}$ if and only if $\gamma \in (-\infty, \underline{\lambda}_n(\Phi)]$. Consequently, $V^{\mathrm{s}}(\Phi_\gamma)$ is open if and only if $\gamma \in (\underline{\lambda}_n(\Phi), \infty) \supset (\underline{\lambda}_n(\Phi), \overline{\lambda}_n(\Phi)]$. On the other hand, $V^{\mathrm{s}}(\Phi_\gamma)$ has non-empty boundary if and only if $\gamma \in (-\infty, \overline{\lambda}_n(\Phi)]$, i.e. if and only if $V^{\mathrm{s}}(\Phi_\gamma) \neq \mathbb{R}^n$. In summary, $V^{\mathrm{s}}(\Phi_\gamma)$ is open and has non-empty boundary if and only if $\gamma \in (\underline{\lambda}_n(\Phi), \overline{\lambda}_n(\Phi)]$. From Eq. (3.9) it follows immediately that $V^{\mathrm{s}}(\Phi_{\gamma'})$ contains the boundary of $V^{\mathrm{s}}(\Phi_\gamma)$ whenever $\gamma' > \gamma$. Moreover, due to continuous dependence of Φ on the initial value, for any $x \in V^{\mathrm{s}}(\Phi_\gamma)$ there exists a neighborhood U of x (in $\mathbb{R}^n$) such that $U \subsetneq V^{\mathrm{s}}(\Phi_{\gamma'})$ for $\gamma' > \gamma$. □

Observe that this sensitive dependence on the parameter γ within each interval of the resolvent set $\varrho(\Phi)$ does not hold in the two-sided infinite-time ODE case; see [103, Lemma 3.2]. This indicates a difference in the parameter dependence, which is not surprising in finite time. As a consequence of the proof of Lemma 3.23 we obtain the following result.

3.24 Corollary. *For any* $\gamma \in (\underline{\lambda}_n(\Phi), \overline{\lambda}_n(\Phi))$ *we have*

$$\operatorname{span}(V^{\mathrm{s}}(\Phi_\gamma)) = \operatorname{span}(V^{\mathrm{u}}(\Phi_\gamma)) = \mathbb{R}^n.$$

A natural question is what happens with $V^{\mathrm{s}}(\Phi_\gamma)$ and $V^{\mathrm{u}}(\Phi_\gamma)$ when changing γ within one of the middle intervals of $\varrho(\Phi)$? On the one hand, Lemma 3.23 says that $V^{\mathrm{s}}(\Phi_\gamma)$ increases and $V^{\mathrm{u}}(\Phi_\gamma)$ decreases strictly when increasing γ. On the other hand, the maximal dimension of contained subspaces does not change. Roughly speaking, by increasing γ the stable cones $V^{\mathrm{s}}(\gamma)$ fatten up without gaining higher-dimensional subspaces exactly as do the unstable cones $V^{\mathrm{u}}(\Phi_\gamma)$ by decreasing γ. In this sense, we obtain the "purest" or "thinnest" stable (unstable) cone of a connected component of $\varrho(\Phi)$ by taking the limit towards the lower (higher, respectively) boundary with respect to γ. This is how we introduce the finite-time analogue of spectral manifolds in the two-sided infinite-time case; see [98, 103].

3.25 Definition (Spectral cones)**.** With the notation in Eq. (3.6) let

$$\Sigma^{\mathbb{I}}(\Phi) = \biguplus_{k=1}^{d} [\underline{\lambda}_{n-i_{k-1}}(\Phi), \overline{\lambda}_{i_k}(\Phi)].$$

Then we define for $k \in \{1, \dots, d\}$

$$\mathcal{W}_k(\Phi) := \bigcap_{\gamma < \underline{\lambda}_{n-i_{k-1}}(\Phi)} V^{\mathrm{u}}(\gamma) \cap \bigcap_{\gamma > \overline{\lambda}_{i_k}(\Phi)} V^{\mathrm{s}}(\gamma),$$

and we call $\mathcal{W}_k(\Phi)$ the k-th *spectral cone of* Φ, associated to the spectral interval $[\underline{\lambda}_{n-i_{k-1}}(\Phi), \overline{\lambda}_{i_k}(\Phi)]$.

Definition 3.25 generalizes the notion of spectral manifolds, which are introduced in [32, p. 4183] and which are in general non-unique. The following theorem strengthens [32, Theorem 11(ii)] towards a characterization of spectral cones.

3.26 Theorem (Characterization of spectral cones)**.** *Let d denote the number of the spectral intervals and $k, l \in \{1, \dots, d\}$. Then the k-th spectral cone can be characterized by*

$$\mathcal{W}_k(\Phi) = \left\{ x \in \mathbb{R}^n;\ \bigwedge_{(t,s) \in \neq_{\mathbb{I} \times \mathbb{I}}} \underline{\lambda}_{n-i_{k-1}}(\Phi) \le \Delta(|\Phi(\cdot, t_-)x|)(t,s) \le \overline{\lambda}_{i_k}(\Phi) \right\}.$$

As a consequence, for $k \neq l$ we have that $\mathcal{W}_k(\Phi) \cap \mathcal{W}_l(\Phi) = \{0\}$. Furthermore, there exists a subspace $U \subseteq \mathcal{W}_k(\Phi)$ with $\dim U \ge i_k - i_{k-1}$.

Proof. The characterization of the spectral cones follows directly from their Definition 3.25 and (3.9). It remains to show the existence of the subspace. To this end, let $k \in \{1, \dots, d\}$, $X \in Gr(n - i_{k-1}, n)$, $Y \in Gr(i_k, n)$ such that $\underline{\lambda}(X, \Phi) = \underline{\lambda}_{n-i_{k-1}}(\Phi)$ and $\overline{\lambda}(Y, \Phi) = \overline{\lambda}_{i_k}(\Phi)$. It is easily seen by (3.9) that

$$X \subseteq V^{\mathrm{u}}(\underline{\lambda}_{n-i_{k-1}}(\Phi) - 0) \qquad \text{and} \qquad Y \subseteq V^{\mathrm{s}}(\overline{\lambda}_{i_k}(\Phi) + 0).$$

For all $\gamma < \underline{\lambda}_{n-i_{k-1}}(\Phi)$ with $(\gamma, \underline{\lambda}_{n-i_{k-1}}(\Phi)) \cap \varrho(\Phi) = \varnothing$ ($\gamma > \overline{\lambda}_{i_k}(\Phi)$ with $(\overline{\lambda}_{i_k}(\Phi), \gamma) \cap \varrho(\Phi) = \varnothing$) the (un-)stable cone $V^{\mathrm{s}}(\Phi_\gamma)$ ($V^{\mathrm{u}}(\gamma)$, respectively) contains subspaces of dimension at most i_k ($n - i_{k-1}$, respectively). Thus, by the dimension formula for intersections of subspaces, we obtain for $U := X \cap Y \subseteq \mathcal{W}_k$

$$\begin{aligned} \dim(X \cap Y) &= \dim X + \dim Y - \dim(X + Y) \\ &= n - i_{k-1} + i_k - \dim(X + Y) \ge i_k - i_{k-1} > 0. \end{aligned}$$ □

A desirable result is that the spectral cones span the whole space $\mathbb{R}^n$. A further question is if one can choose a subspace from each spectral cone such that these subspaces span the whole state space $\mathbb{R}^n$. Both questions turn out to be very difficult and, to the best of the author's knowledge, a solution has not been found yet. In the following, we want to sketch two approaches, which could answer the two questions.

The first approach is to show that the span of all spectral cones is $\mathbb{R}^n$. To this end, assume to the contrary that $\sum_{j=1}^{d} \mathcal{W}_j(\Phi) \neq \mathbb{R}^n$ and thus $\dim(\sum_{j=1}^{d} \mathcal{W}_j(\Phi)) = k < n$. Then there exists a direct summand $V \subseteq \mathbb{R}^n$, i.e. $\sum_{j=1}^{d} \mathcal{W}_j(\Phi) \oplus V = \mathbb{R}^n$ and $\dim V = n - k$. Necessarily, we have $V \cap \bigcup_{j=1}^{d} \mathcal{W}_j(\Phi) = \{0\}$ and consequently $V \cap \mathcal{W}_j(\Phi) = \{0\}$ for each $j \in \{1, \ldots, d\}$. By the characterization in Theorem 3.26(i) it follows that $[\underline{\lambda}(V, \Phi), \overline{\lambda}(V, \Phi)] \subset \varrho(\Phi)$. One would hope that this leads to some contradiction. However, such subspaces exist, obviously one-dimensional subspaces, i.e. lines that "exit" the unstable cone when increasing the parameter γ within one interval of the resolvent set $\varrho(\Phi)$ and that "enter" the stable cone when further increasing γ (cf. the last paragraph of Section 6.3), but also higher-dimensional subspaces are possible. A solution might be obtained by a careful investigation of the dimension of V and the rank jumps as considered in Lemma 3.22.

A second approach is to answer the second question directly by finding subspaces in the spectral cones which span $\mathbb{R}^n$. The difference to the first approach is that in the case of subspaces, a classical yet elementary argument, which is often referred to as *algebraic lemma*, becomes available; cf. [98], see [102, p. 58] for a proof.

3.27 Lemma. *Let X be a vector space, $A, B, C \subseteq X$ subspaces and $C \subseteq A$, then*

$$A \cap ([B] + [C]) = [A \cap B] + [C].$$

Evidently, for $X_k \in \mathrm{Gr}(n - i_k, \mathbb{R}^n)$, $Y_k \in \mathrm{Gr}(i_k, \mathbb{R}^n)$, $k \in \{0, \ldots, d\}$, with $\underline{\lambda}(X_k, \Phi) = \underline{\lambda}_{n-i_k}(\Phi)$ and $\overline{\lambda}(Y_k, \Phi) = \overline{\lambda}_{i_k}(\Phi)$ we have $X_k \oplus Y_k = \mathbb{R}^n$. In order to apply the algebraic lemma successively, we need to prove that for *any* choice of $X_k \in \mathrm{Gr}(n - i_k, \mathbb{R}^n)$ such that $\underline{\lambda}(X_k, \Phi) = \underline{\lambda}_{n-i_{k-1}}(\Phi)$ there exists $X_{k+1} \in Gr(n - i_{k+1}, \mathbb{R}^n)$ with $\underline{\lambda}(X_{k+1}, \Phi) = \underline{\lambda}_{n-i_{k+1}}(\Phi)$ and $X_{k+1} \subset X_k$. However, it seems to be unclear if such a reversed filtration of subspaces with respect to the spectral cones exists. The difficulty that appears here is the required independence of the choice.

For the above mentioned reasons and as is discussed in [17, Section 4], the role of spectral cones (or manifolds) remains unclear.

4 Computation of Extremal Growth Rates

> Under a practical perspective, it is essential to have efficient tools available for the computation of the spectrum. [...] The finite-time nature of $\Sigma(A)$, however, should make a computational approach feasible. Moreover, any such approach should work for rather large classes of functions $t \mapsto A(t)$ alike [...]. The examples above indicate that $\Sigma(A)$ can be quite hard to compute directly from its definition, even for very simple A.
>
> (Berger, Doan & Siegmund [17])

This chapter is devoted to an approach to the computation of the spectrum via the computation of extremal growth rates, as desired in the above quotation. Indeed, we suggest to compute (extremal) growth rates directly from their definition, recall Definition 2.13, Definition 2.13 and Definition 2.20, as a (consecutive) optimization problem on $\mathbb{I}$ and the Riemannian manifolds $\mathrm{Gr}(1, X)$ and $\mathrm{Gr}(k, \mathbb{R}^n)$, respectively. Optimization on Riemannian manifolds is a relatively new field in continuous optimization.

Two classical optimization approaches for unconstrained optimization problems are linesearch and trust-region methods; see [52]. In the classical formulation, the basic problem is to minimize a (twice) continuously differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}$. Following [52], the general scheme for linesearch algorithms is as follows: for an initial point x, firstly, a search direction $p \in \mathbb{R}^n$ is calculated. This direction is required to be a descent direction, i.e.,

$$\langle p, \operatorname{grad} f(x)\rangle < 0, \qquad \text{if } \operatorname{grad} f(x) \neq 0;$$

so that, for small steps along p, Taylor's theorem guarantees that the objective function may be reduced. Secondly, a suitable step size $\alpha > 0$ is calculated so that

$$f(x + \alpha p) < f(x).$$

The computation of α is the linesearch, and may itself be an iteration. Finally, given both search direction and step size, the iteration concludes by setting

$$\hat{x} = x + \alpha p,$$

and repeating the iteration with initial point $\hat{x}$. Two specific versions of linesearch algorithms are the *method of steepest descent* with search direction $p = -\operatorname{grad} f(x)$ and the *Newton method* with search direction p solving

$$H(x)p := \partial \otimes \operatorname{grad} f(x)p = -\operatorname{grad} f(x),$$

i.e. $H(x)$ is the Hessian of f at x.

Trust-region methods, by contrast, pick the overall step $s := \alpha p$ to deduce a "model" of $f(x + \cdot)$, often a linear or quadratic approximation (but not necessarily obtained from a Taylor expansion), and accept $\hat{x} = x + s$ if the decrease predicted by the model is realized by $f(x + s)$. Since there is no guarantee that this will always be so, the fall-back mechanism is to set $\hat{x} = x$, and to refine the model when the existing model produces a poor step; for a textbook on trust-region methods see [25].

Both approaches, the linesearch and the trust-region approach, have been extended to optimization problems on Riemannian manifolds; see [3, Chapter 4] and [3, Chapter 7], respectively, and the references therein.

Note that necessarily, when running computer programs, we face discretization in time and space, as well as numerical errors in the integration scheme. As with respect to time discretization, recall Lemma 2.24 saying that the computation of extremal growth rates is robust under time-set approximations. Secondly, Lemma 2.23 says that the computation of extremal growth rates is robust under approximations of the solution operator. In summary, a finer discretization will result in improved approximations of the true (extremal) growth rates. This justifies any computational approach to the computation of the extremal growth rates and hence of the spectrum of linear processes generated by linear ODEs.

As a consequence of the approximation results, we consider in this chapter linear finite-time processes that are generated by a linear ODE on a compact interval $\mathbb{I} = [t_-, t_+]$. In this chapter, we assume that $\mathbb{R}^n$ is endowed with a C^∞-norm (naturally, except for the origin) such as the Euclidean norm $|\cdot| = \left(x \mapsto \langle x, x\rangle^{\frac{1}{2}}\right)$ or the Γ-norm $|\cdot|_\Gamma = \left(x \mapsto \langle x, \Gamma x\rangle^{\frac{1}{2}}\right)$ with a symmetric positive definite matrix

$\Gamma \in \mathbb{R}^{n\times n}$, see also Section 6.3. In particular, that turns the unit sphere $\mathcal{S} = |\cdot|^{-1}[\{1\}]$ into a C^∞ submanifold of $\mathbb{R}^n$ by Proposition A.10, since $|\cdot||_{\mathbb{R}^n\setminus\{0\}}\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}_{>0}$ is a surjective submersion.

The scope of this chapter is to apply one of the discussed approaches to the computation of (extremal) growth rates of linear processes on $\mathbb{I}$, namely the Newton algorithm. A comparison of different methods and an extensive study of different finite-time processes is beyond the scope of this work and is reserved for future research. To apply the Newton algorithm we first have to examine conditions to establish the necessary regularity of the growth rate functions. Later, we present the Grassmann-Newton algorithm as derived in [2, 3], before we describe conceptually how we evaluate the respective objective function. We close with a simple numerical example and some tests in the discussion section.

4.1 Regularity of Growth Rates

Throughout this chapter, we require the right-hand side $A\colon \mathbb{I} \to L(\mathbb{R}^n, \mathbb{R}^n)$ of the linear ODE under consideration to be four times continuously differentiable, such that the associated linear solution operator $\Phi\colon \mathbb{I}\times\mathbb{I} \to L(\mathbb{R}^n,\mathbb{R}^n)$ is five times continuously differentiable; see, for instance, [6, Section 9] for more details on the regularity theory for solution operators.

Regularity of one-dimensional growth rates

In a first step, we investigate the regularity of one-dimensional growth rates. We make use of the equivalent definition of growth rates; see Lemma 2.12 and Eq. (6.5). For a subspace $X \subseteq \mathbb{R}^n$ we introduce

$$f\colon \mathbb{I}\times X\setminus\{0\} \to \mathbb{R}, \qquad (t,x) \mapsto \frac{(|\Phi(\cdot,t_-)x|)'(t)}{|\Phi(t,t_-)x|}.$$

Clearly, f is four times continuously differentiable in the first argument and indefinitely continuously differentiable in the second argument. We have by definition

$$\min\{f(t,x);\ t\in\mathbb{I}\} = \underline{\lambda}(\operatorname{span}\{x\},\Phi)$$

and

$$\max\{f(t,x);\ t\in\mathbb{I}\} = \overline{\lambda}(\operatorname{span}\{x\},\Phi).$$

In the following, we make use of the identification stated in the introduction to Section 2.3 and use vectors and one-dimensional subspaces as well as neighborhoods in $\mathbb{R}^n$ and corresponding neighborhoods in $\operatorname{Gr}(1,X)$ synonymously. The neighborhood correspondence is justified by Lemma B.8.

Next, we discuss sufficient conditions for threefold continuous differentiability of one-dimensional growth rates. To this end, let $X\subseteq\mathbb{R}^n$ be a subspace, $U\in\operatorname{Gr}(1,X)$ and $x\in U$ such that $U=\operatorname{span}\{x\}$. Then $\min\{f(t,x);\ t\in\mathbb{I}\}$ is attained at some $t^*\in\mathbb{I}$. The most important requirement for regularity is the uniqueness of the (global) minimum point. Moreover, we need to distinguish between two cases.

(i) $t^*\in(t_-,t_+)$: In this case, if t^* is the unique time point at which the minimum is attained and t^* is a non-degenerate minimum, i.e. $\partial_0 f(t^*,x)=0$ and $\partial_0^2 f(t^*,x)>0$, then by the Implicit Function Theorem there exists a locally defined three times continuously differentiable function $T\colon V\subseteq\mathbb{R}^n\to\mathbb{I}$, V a neighborhood of x, such that $(v\mapsto\underline{\lambda}(\operatorname{span}\{v\},\Phi))=(v\mapsto f(T(v),v))$ on V. Hence, $\underline{\lambda}(\operatorname{span}\{\cdot\},\Phi)$ is three times continuously differentiable on V.

(ii) $t^*\in\{t_-,t_+\}$: In this case, if the minimal point t^* on the boundary of $\mathbb{I}$ is unique and has a non-vanishing one-sided derivative, then the minimum point $t^*(y)$ will remain there constantly for $y\in U$, U a neighborhood of x, due to the continuity of f.

In the case of non-uniqueness of the extremal point t^*, a general analysis is impossible due to “generalized” (in the sense of multi-parameter dependence) crossings of different local minima, which interchange the role of global minima, compare also the discussion in Theorem 7.7 in the simpler two time-point setting. The discussion for the upper growth rate is completely analogous.

Regularity of general growth rates

By similar arguments we can discuss differentiability for the general growth rates. As before, we assume that $\Phi\in\mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ is five times continuously differentiable. Let $k\in\{2,\dots,n\}$.

As a first preparatory step we need a parametrization for the elements in $\mathrm{Gr}(1, Y)$, for $Y \in U_A$ around some fixed $X \in \mathrm{Gr}(k, \mathbb{R}^n)$ and $A \in \mathrm{St}(k, \mathbb{R}^n)$ with $\pi(A) = X$, i.e. for the lines in the k-dimensional subspaces Y that are close to X. Locally, this can be done most naturally by

$$p\colon \mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{k-1} \to \mathrm{Gr}(1, \mathbb{R}^n), \qquad (K, u) \mapsto (A + A_\perp K)[p_S(u)],$$

where

- p_S is a local parametrization of $\mathrm{Gr}(1, \mathbb{R}^k)$ (in the sense of p_A in Section B.2),
- p is smooth in both arguments, and
- for any $K \in \mathbb{R}^{(n-k)\times k}$ with $p_A(K) = Y \in U_A \subset \mathrm{Gr}(k, \mathbb{R}^n)$ and $u \in \mathbb{R}^k \setminus \{0\}$ we have $p(K, u) \in \mathrm{Gr}(1, Y)$.

For each $X \in \mathrm{Gr}(k, \mathbb{R}^n)$ the lower growth rate $\underline{\lambda}(X, \Phi)$ is attained at some $U^* \in \mathrm{Gr}(1, X)$. Now fix X, choose $A \in \mathrm{St}(k, \mathbb{R}^n)$ such that $\pi(A) = X$ and assume that U^* is the unique minimal point. Let p_S be a local parametrization around $u^* \in \mathbb{R}^{k-1}$ such that $A[p_S(u^*)] = U^*$. In the parametrization above this means that there is a unique $u^* \in \mathbb{R}^{k-1}$ such that $p(0, u^*) = U^*$. Locally, i.e. on U_A and around U^*, we can rewrite our optimization problem introducing

$$g_X\colon \mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{k-1} \to \mathbb{R}, \qquad (K, u) \mapsto \underline{\lambda}(p(K, u), \Phi),$$

as

$$\underline{\lambda}(X, \Phi) = \min\left\{g_X(0, u);\ u \in \mathbb{R}^{k-1}\right\} = g_X(0, u^*).$$

By assumption u^* is unique, furthermore assume that u^* is "well-behaved" in the sense of (i) or (ii) of the preceding section. Then $g_X(0, \cdot)$ is three times continuously differentiable in u^*. Additionally, assume that u^* is a non-degenerate minimum point, i.e.

$$\partial_1 g_X(0, u^*) = 0, \qquad \partial_1^2 g_X(0, u^*) > 0.$$

Then by the Implicit Function Theorem there exists a neighborhood V of the origin in $\mathbb{R}^{(n-k)\times k}$ and a twice continuously differentiable function $v\colon V \to \mathbb{R}^{k-1}$ such that the value $\underline{\lambda}(\pi(A + A_\perp K), \Phi)$ is uniquely attained by $g_X(K, v(K))$ for $K \in V$. This means that the lower growth rate can be represented as the composition of twice continuously differentiable functions.

For a discussion of the uniqueness of extremal subspaces see also [17, Section 4]. It seems to be an open problem to find sufficient conditions for the presence of unique extremal subspaces. Nevertheless, the above regularity considerations show that the application of a Newton-like algorithm on the Grassmann manifold is feasible. We have pointed out the obstacles for differentiability, which can be the reason for a malfunctioning computation.

4.2 The Grassmann-Newton Algorithm

As mentioned in the introduction to this chapter, two fundamental methods in continuous optimization are linesearch methods and trust-region methods. As a first computational approach we describe in the following a special linesearch method, the Newton method, specified to the Riemannian manifold $\mathrm{Gr}(k, \mathbb{R}^n)$.

The classical Newton method

The Newton algorithm, also referred to as Newton-Raphson method, is one of the classical iterative algorithms for solving nonlinear equations

$$F(x) = 0,$$

where $F \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. It is treated in probably every introductory textbook on numerical analysis and we refer the reader to [30]; a historical review can be found in [111]. In this standard setting, the Newton iteration is given by

$$F'(x_k)y_k = \partial_{y_k} F(x_k) = -F(x_k), \qquad x_{k+1} = x_k + y_k, \qquad k \in \mathbb{N}, \tag{4.1}$$

with initial value $x_0 \in \mathbb{R}^n$ and supposed invertibility of $F'(x_k)$ for $k \in \mathbb{N}$. Under appropriate conditions, cf. [30, Chapters 2 & 3], the Newton algorithm is well-known to converge (locally) quadratically.

The Newton algorithm has been generalized and modified in several directions. One of these is the establishment of a Newton algorithm on manifolds; see [3, Section 6.6] for a historical review and further references, where [105] and [106] where the first to formulate the method on Riemannian manifolds; [105] provides a first proof of quadratic convergence.

Newton method on Riemannian manifolds

Let (M, g) be a Riemannian manifold with Riemannian connection ∇. The idea of transferring the classical Newton method to Riemannian manifolds is to give meaning to the objects appearing in Eq. (4.1) in the manifold context by the following substitutions: we replace $F \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ by a C^1 vector field $F \in \mathcal{X}(M)$, the directional derivative ∂_y by the Riemannian connection ∇, and the linear update rule $x_{k+1} = x_k + y_k$ by the exponential map $x_{k+1} = \exp_{x_k}(y_k)$. Thus, the Riemann-Newton method reads as follows:

$$\nabla(y_k, F) = -F(x_k), \qquad x_{k+1} = \exp_{x_k}(y_k), \qquad k \in \mathbb{N},$$

with $y_k \in T_{x_k}M$ and initial value $x_0 \in M$. Again and for the rest of this chapter, we assume solvability of the first equation. For modifications in the absence of invertibility, also referred to as *Quasi-Newton-methods*, see, for instance, [30] and [3, Chapter 8].

Newton method for real-valued functions

In the case that the C^1 vector field F is the gradient field of a C^2 cost function $f\colon M \to \mathbb{R}$, the aim of applying the Newton algorithm is to find critical points of f, and we obtain the Newton method for real-valued functions as

$$\nabla(y_k, \operatorname{grad} f) = -\operatorname{grad} f(x_k), \qquad x_{k+1} = \exp_{x_k}(y_k), \qquad k \in \mathbb{N}.$$

The lifted Grassmann-Newton method for real-valued functions

So far, we stated the Newton method on manifolds in terms of the affine (or the Riemannian) connection. However, though mathematically sufficient, this formulation leaves open how to actually *perform* the Newton algorithm numerically. For our concrete problem of finding extreme points of (one-dimensional) growth rates, which are in particular critical points, we make use of the matrix representations of the elements in $\mathrm{Gr}(k, \mathbb{R}^n)$ and their construction by horizontal lifts to calculate in $\mathrm{St}(k, \mathbb{R}^n)$; see Appendix B.

For the Grassmann case, the Newton algorithm reads as follows:

$$\nabla_{\mathrm{Gr}}(Y_k, \operatorname{grad}_{\mathrm{Gr}} f) = -\operatorname{grad}_{\mathrm{Gr}} f(X_k), \qquad X_{k+1} = \exp_{X_k}(Y_k), \qquad k \in \mathbb{N}.$$

Since the first equation holds in the Grassmann tangent space, we lift it to the horizontal space (omitting the subindex k for convenience) at $A \in \pi^{-1}[\{X\}]$:

$$\begin{aligned}
\overline{\nabla_{\mathrm{Gr}}(Y, \operatorname{grad}_{\mathrm{Gr}} f)}_A &\overset{\text{(B.24)}}{=} \Pi_{\pi(A_\perp)} \nabla_{\mathrm{St}} \left(\overline{Y}_A, \overline{\operatorname{grad}_{\mathrm{Gr}} f \circ \pi} \right) \\
&\overset{\text{(B.19)}}{=} \Pi_{\pi(A_\perp)} \nabla_{\mathrm{St}} \left(\overline{Y}_A, \left(B \mapsto \operatorname{grad}_{\mathrm{St}}(f \circ \pi)(B) B^\top B \right) \right) \\
&\overset{\text{(B.18)}}{=} \Pi_{\pi(A_\perp)} \nabla_{\mathrm{St}} \left(\overline{Y}_A, \left(B \mapsto \partial \otimes (f \circ \pi)(B) B^\top B \right) \right) \\
&\overset{\text{(B.25)}}{=} \Pi_{\pi(A_\perp)} \partial_{\overline{Y}_A} \left(B \mapsto \partial \otimes (f \circ \pi)(B) B^\top B \right) (A) \\
&= -\partial \otimes (f \circ \pi)(A) A^\top A \\
&= -\operatorname{grad}_{\mathrm{St}}(f \circ \pi)(A) A^\top A \\
&\overset{\text{(B.19)}}{=} -\overline{\operatorname{grad}_{\mathrm{Gr}} f(X)}_A,
\end{aligned}$$

i.e. the corresponding Newton equation on the Stiefel manifold is

$$\Pi_{\pi(A_\perp)} \partial_Y \left(B \mapsto \partial \otimes (f \circ \pi)(B) B^\top B \right) (A) = -\partial \otimes (f \circ \pi)(A) A^\top A.$$

This needs to be solved for $Y \in H_A$ and the solution maps to the Newton direction on the Grassmann manifold via $D\pi\big|_A$. The last equation can be further simplified: we make use of the fact that we solve in the horizontal space, i.e. $Y^\top A = 0$, so after applying the product rule

$$\begin{aligned}
&\partial_Y \left(B \mapsto \partial \otimes (f \circ \pi)(B) B^\top B \right) (A) = \\
&\quad = \partial_Y (\partial \otimes (f \circ \pi))(A) A^\top A + \partial \otimes (f \circ \pi)(A) Y^\top A + \partial \otimes (f \circ \pi)(A) A^\top Y \\
&\qquad = \partial_Y (\partial \otimes (f \circ \pi))(A) A^\top A
\end{aligned}$$

we obtain

$$\Pi_{\pi(A_\perp)} \partial_Y (\partial \otimes (f \circ \pi))(A) A^\top A = -\partial \otimes (f \circ \pi)(A) A^\top A$$

and, since $A^\top A$ is invertible, we finally get the Newton equation

$$\Pi_{\pi(A_\perp)} \partial_Y (\partial \otimes (f \circ \pi))(A) = -\partial \otimes (f \circ \pi)(A)$$

with the update step

$$\hat{A} = AV \cos \Sigma + U \sin \Sigma,$$

where $Y = U \Sigma V^\top$ is a thin SVD, see Proposition B.27.

Algorithm 1 Pseudocode for the Grassmann-Newton algorithm

Require: Cost function f and initial value $A \in \mathrm{St}(k, \mathbb{R}^n)$
repeat
compute the gradient by finite differences
compute the Hessian by finite differences
matricize the Hessian and vectorize the gradient to form the Grassmann-Newton equations
solve for a horizontal Y
take a geodesic step along the direction given by Y to obtain a new iterate A
until some tolerance condition is satisfied
return $[f(Y^*), Y^*]$, i.e. critical value and critical point of the cost function

The pseudocode for the lifted Grassmann-Newton algorithm is displayed in Algorithm 1.

An implementation of the Grassmann-Newton method can be found in the `sg_min`-package by R. A. Lippert[1], which follows the slightly different approach in [36] by restricting to $\mathrm{St}^*(k, \mathbb{R}^n)$, thereby incorporating orthogonality constraints; for a documentation of the involved procedures of the Matlab-package, for pseudocodes and examples see [37].

4.3 The Computation of Growth Rates

In this section, we want to describe how we evaluate the different growth rate functions. We start with the one-dimensional growth rates, which are computed straightforwardly using the code displayed in Algorithm 2: solve the ODE on an interval adaptively, compute the logarithmic difference quotient for subsequent time points and return the minimum of these values.

The general idea for computing k-dimensional growth rates is the following: construct a discretization of the unit sphere in $\mathbb{R}^k$, map it into the k-dimensional subspace under consideration of $\mathbb{R}^n$ by a subspace spanning matrix, compute the lower/upper growth rate for each initial value thus obtained and return the minimum/maximum, respectively, of the growth rates. The code for the 2-dimensional lower growth rate is displayed in Algorithm 3, the one for the 3-dimensional lower growth rate in Algorithm 4.

[1] `http://web.mit.edu/~ripper/www/sgmin.html`

Algorithm 2 MATLAB-code for the computation of lower growth rates of one-dimensional subspaces

```
function lmin = lgr1d(ode,tstart,tfinal,iv,options)

[T,Y] = ode45(ode,[tstart tfinal],iv,options);
n = size(Y,1);
Ynorm = zeros(1,n);
for j = 1:n
    Ynorm(j) = norm(Y(j,:));
end
logY = log(Ynorm);
rate = diff(logY)./diff(T');
lmin = min(rate);
end
```

Algorithm 3 MATLAB-code for the computation of lower growth rates of two-dimensional subspaces

```
function lmin = lgr2d(ode,tstart,tfinal,IS,options)

t = 0:0.01:1;
m = size(t,2);
X0 = [cos(2*pi*t); sin(2*pi*t)];
I0 = IS*X0;
l = zeros(1,m);
for j=1:m
    l(j) = lgr1d(ode,tstart,tfinal,I0(:,j),options);
end
lmin = min(l);
end
```

Algorithm 4 MATLAB-code for the computation of lower growth rates of three-dimensional subspaces

```
function lmin = lgr3d(ode,tstart,tfinal,options)

[A,B,C] = sphere(100);
[m,n] = size(A);
l = zeros(m,n);
for i = 1:m
    for j = 1:n
        l(i,j) = lgr1d(ode,tstart,tfinal,[A(i,j); B(i,j); C(i,j)],options);
    end
end
lmin = min(min(l));
end
```

Finally, the extremal growth rates are computed by applying the Grassmann-Newton algorithm to the respective growth rate functions.

4.4 Numerical Example

We consider the following time-translation invariant ODE

$$\dot{x} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} x.$$

This system serves as a three-dimensional toy model to perform the simplest but non-trivial task: compute extremal 2-growth rates. The computation of the minimal upper 1-growth rate starting with a random normalized initial vector gives the following output:

```
>> Y = orth(random('Normal',0,1,3,1))

Y =

   -0.1817
   -0.6198
    0.7634
```

```
>> [f,B]=sg_min(Y,1e-4,1e-4)
iter grad F(Y) step type
0 9.285488e+00 3.386564e+00       110 one
  invdgrad: Hessian not positive definite, CG terminating early
1 9.270664e-01 1.648862e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
2 1.129089e+01 1.232478e+00        78 ewton step
3 1.062883e+00 1.004275e+00        78 ewton step
4 1.739459e-03 1.000000e+00        78 ewton step
5 1.300902e-04 1.000000e+00        78 ewton step

f =

    1.0000

B =

   -0.0000
   -0.0000
    1.0000
```

The maximal lower 1-growth rate is computed as

```
>> Y = orth(random('Normal',0,1,3,1))

Y =

   -0.7110
    0.0618
   -0.7005

>> [f,B]=sg_min(Y,1e-4,1e-4)
iter grad F(Y) step type
0 3.978992e+00 -3.039832e+00       110 one
  invdgrad: max iterations reached inverting the hessian by CG
1 1.282777e-01 -4.998619e+00       115 teepest step
2 5.070674e-07 -4.999988e+00        78 ewton step
3 3.863432e-08 -4.999988e+00        78 ewton step
```

```
f =

   -5.0000

B =

   -1.0000
   -0.0000
    0.0000
```

Note that the optimizers are computed correctly up to 4 decimals and that the sign of the function values is reversed due to the fact that we solve a maximization problem with a minimizing Newton algorithm. The computation of the minimal upper 2-growth rate yields

```
>> Y = [.2 -.1; 1 0; 0 1];
>> [f,B]=sg_min(Y,1e-4,1e-4)
iter grad F(Y) step type
0 7.607323e+00 3.573170e+00       110 one
  invdgrad: Hessian not positive definite, CG terminating early
1 1.737347e-03 1.999996e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
2 5.683726e-03 1.999996e+00       115 teepest step
  invdgrad: Hessian not positive definite, CG terminating early
3 1.720800e-03 1.999996e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
4 3.615272e-04 2.000000e+00        78 ewton step

f =

    2.0000

B =

   -0.0000   -0.0666
   -1.0000   -0.0001
    0.0001   -0.9978
```

where the computed minimizer is close to the known minimizer

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The computation of the maximal lower 2-growth rate yields

```
>> Y = [1 0; 0 1; .15 -.1];
>> [f,B]=sg_min(Y,1e-4,1e-4)
iter grad F(Y) step type
0 1.955473e-01 -1.991239e+00       110 one
  invdgrad: Hessian not positive definite, CG terminating early
1 6.272228e-02 -2.002850e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
  invdgrad: max iterations reached inverting the hessian by CG
2 1.335748e-01 -2.002865e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
3 4.048595e-03 -1.999996e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
4 1.856702e-04 -2.000000e+00        78 ewton step
  invdgrad: Hessian not positive definite, CG terminating early
5 5.775199e-06 -2.000000e+00        78 ewton step

f =

   -2.0000

B =

   -0.9886   -0.0000
   -0.0000   -1.0000
   -0.1508    0.0000
```

where again the maximizer is close to the known one. The call of the upper growth rate function for $\mathbb{R}^3$ gives

```
>> ugr3d(@linear3d,0,.5,options)
```

```
ans =

    5.0000
```

whereas the lower growth rate of $\mathbb{R}^3$ is computed as

```
>> lgr3d(@linear3d,0,.5,options)

ans =

    1.0000
```

In summary, the computed approximation of the spectrum of $e^{\cdot A}$ on $[0, 0.5]$ is

$$\Sigma(e^{\cdot A}) = \{1\} \cup \{2\} \cup \{5\},$$

which correctly matches the analytically known result, cf. Section 6.3. In summary, this example shows the applicability of the Newton algorithm for the computation of extremal growth rates.

4.5 Discussion

In this section, we want to apply the algorithm to more challenging examples. The following time-translation invariant examples are all considered on the time-interval $[0, 0.5]$.

4.1 Example (Non-unique extremal subspaces)**.** Consider the following example with some obviously non-unique extremal subspaces

$$\dot{x} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} x.$$

The unique one-dimensional subspace realizing the minimal upper 1-growth rate together with its growth rate is correctly computed. For the non-unique one-dimensional subspace realizing the maximal lower growth rate we obtain for a randomly chosen initial value

```
f =

   -2.0000

B =

   -0.8588
    0.5123
    0.0000
```

with a correctly computed vanishing third component. For the non-unique two-dimensional subspace realizing the minimal upper 2-growth rate we obtain the correct growth rate with a minimizer, which is not aligned with the canonical axes,

```
f =

    2.0000

B =

    0.7001   -0.5143
    0.6982    0.3478
    0.1495    0.7839
```

The maximal growth rate within the subspace spanned by B is attained for an initial vector lying in the $\mathbb{R}^2 \times \{0\}$-plane. For the unique extremal subspace with respect to the maximal lower 2-growth rate the calculation, initialized with a random initial plane, finishes after 3 steps with the desired result:

```
f =

   -2.0000

B =

   -0.8656   -0.5007
    0.5007   -0.8656
   -0.0000   -0.0000
```

In summary, the extremal growth rates are computed correctly.

4.2 Example (EMD detection)**.** Consider the following test example for a correct detection of EMD,

$$\dot{x} = \begin{pmatrix} 0.001 & 0 \\ 0 & -0.001 \end{pmatrix} x.$$

For randomly chosen initial values both extremal 1-growth rates are computed correctly and hence EMD is detected.

4.3 Example (Non-diagonal right-hand side)**.** Consider the following time-translation invariant ODE

$$\dot{x} = \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} x, \tag{4.2}$$

where the linear operator has eigenvalues -1 and 1 with corresponding eigenvectors $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, respectively. In this case, both computations for the minimal upper and the maximal lower 1-growth rate do not terminate when started with the parameters used in Section 4.4. A glimpse on the lower growth rates in Figure 4.1(a) and on the upper growth rates in Figure 4.1(b) of the lines parametrized over the (normalized) angle indicates that at the respective optimizers the growth rate functions may be not differentiable.

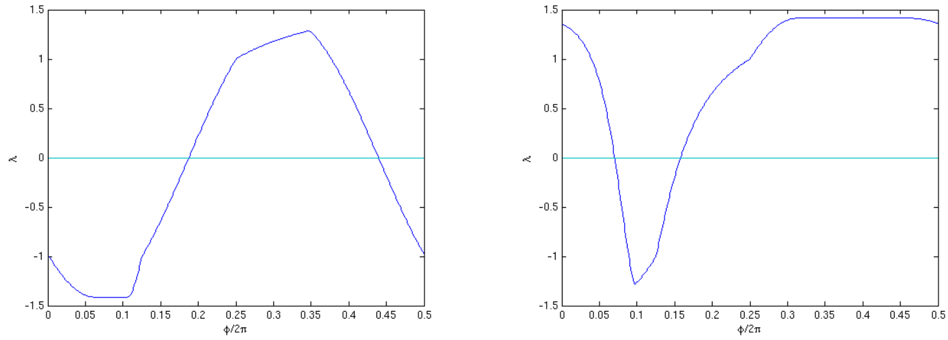
The same problem seems to occur when we consider the following parametrized version of Eq. (4.2): let

$$S_\alpha := \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}, \qquad A_\alpha := S_\alpha \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} S_\alpha^{-1}$$

and consider

$$\dot{x} = A_\alpha x.$$

Observe that A_0 is diagonal and A_1 yields Eq. (4.2). For $\alpha > 0$ the output of the function call indicates that the algorithm jumps far away from the minimum when approaching it, possibly due to non-differentiability or a bad approximation of the derivative. Thus, the computations do not terminate. However, for small positive α one can make the calculations terminate at a reasonable good approximation of the optimizer by relaxing the exit condition. In general, there are a lot of parameters that one could tune and whose influence on the performance on the algorithm need to be further studied: the step-size for the approximation of the derivatives, the relative and absolute error tolerances for the ODE solver, the number of points to discretize the unit circle and the sphere, and finally the exit conditions for the Newton algorithm on the norm of the gradient and on the number of digits of the function value which changed in the last iteration.



(a) lower growth rates **(b)** upper growth rates

Figure 4.1: The growth rates for Eq. (4.2) and integration time 0.5

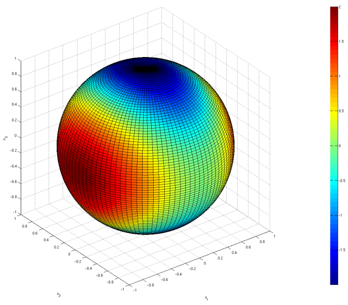
4.4 Example (Sharp peaks of the growth rates)**.** Consider the following linear time-translation invariant ODE

$$\dot{x} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} x. \tag{4.3}$$

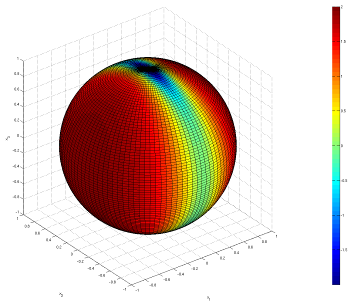
For increasing integration times one observes an increasing dominance of the largest eigenvalue on the one-dimensional upper growth rates; see Figure 4.2(a) for a graph of the upper growth rates for integration time 0.05, Figure 4.2(b) for integration time 0.5 and Figure 4.2(c) for integration time 5. Analogously, one would observe an increasing dominance of the smallest eigenvalue on the one-dimensional lower growth rates with increasing backwards integration time.

In the case of the minimal upper growth rate, the strong dominance of the largest eigenvalue turns the minimal upper growth rate-field very flat on large domains on the level of 2 and a sharp "peak" to the level of -2. This means we have an almost vanishing gradient on the large domain and a large gradient around the optimizer. The same flattening effect can be observed in Example 4.3 in Figure 4.1(b).

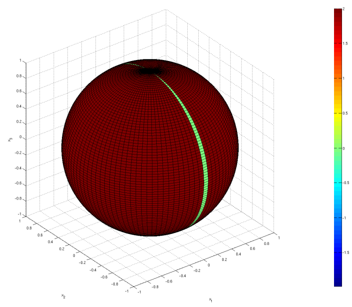
Another obvious critical issue is the under- and overestimation of the growth rates, respectively, which comes from the discretization and hence from taking the sup and inf on a proper subset only.



(a) integration time 0.05



(b) integration time 0.5



(c) integration time 5

Figure 4.2: The upper growth rates for Eq. (4.3) and different integration times

5 Linearization of Finite-time Processes

This chapter is devoted to the local investigation of C^1-processes on $\mathbb{I}$ in the vicinity of reference trajectories by means of linearizations. This technique is classical in the asymptotic stability analysis of differentiable dynamical systems. Throughout this chapter, let φ denote a C^1-process on $\mathbb{I}$ on $\mathbb{R}^n$.

5.1 Linearization and Hyperbolicity of Finite-time Processes

Motivated by the classical theory, we introduce the following notion.

5.1 Definition (Linearization)**.** Let $x \in \mathbb{R}^n$. We define

$$\Phi_{(t_-,x)} \colon \mathbb{I} \times \mathbb{I} \to L(\mathbb{R}^n), \qquad (t,s) \mapsto \Phi_{(t_-,x)}(t,s) := \partial_2 \varphi(t,s,\varphi(s,t_-,x)).$$

and call $\Phi_{(t_-,x)}$ the *linearization of φ along $\varphi(\cdot,t_-,x)$*.

5.2 Lemma. *For any $x \in \mathbb{R}^n$ the function $\Phi_{(t_-,x)}$ as defined in Definition 5.1 is a linear process on $\mathbb{I}$.*

Proof. The cocycle properties including the invertibility are easily checked with the cocycle properties of φ and the chain rule of differentiation. Lipschitz continuity of $\Phi_{(t_-,x)}$ holds by definition. □

5.3 Definition (Finite-time hyperbolicity, attraction and repulsion)**.** Let $x_0 \in \mathbb{R}^n$ and $\Phi_{(t_-,x_0)} \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ be the linearization of φ along $\varphi(\cdot,t_-,x_0)$. We call $\varphi(\cdot,t_-,x_0)$ *(finite-time) hyperbolic* if $\Phi_{(t_-,x_0)}$ admits an EMD on $\mathbb{I}$. We call $\varphi(\cdot,t_-,x_0)$ *(finite-time) attractive/repulsive* if $\Phi_{(t_-,x_0)}$ is attractive/repulsive.

For an extensive study of finite-time attractivity with respect to the two-point time-set $\mathbb{I} = \{t_-,t_+\}$ we refer the reader to [47].

Finite-time local stable and unstable manifolds are studied for different problem classes in [57, 53, 54, 17]. In these works, the manifolds are introduced as manifolds in the extended phase space, depending on some non-unique extension of a given differential equation on some compact time-interval to the whole real line. The obtained manifolds have, for any given extension, indeed a C^1-manifold structure. In [33, Definition 35] for two-dimensional ODEs so-called stable and unstable manifolds are introduced, which do not have a manifold structure. However, these objects are defined in an "intrinsic" way, by requiring some decay and growth property of trajectories on the compact time-interval with respect to a reference trajectory, respectively. In [47, Definition 3.1] domains of attraction are introduced intrinsically with a decay requirement with respect to the two-point time-set $\mathbb{I} = \{t_-, t_+\}$. We take this as a motivation for the next definition.

5.4 Definition (Domains of finite-time attraction/repulsion)**.** Let $y \in \mathbb{R}^n$. Then we define

$$W^{\mathrm{s}}_y := \left\{ x \in \mathbb{R}^n \setminus \{y\}\, ; \sup_{(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}} \{\Delta(|\varphi(\cdot, t_-, x) - \varphi(\cdot, t_-, y)|)(t,s)\} < 0 \right\} \cup \{y\}\,,$$
$$W^{\mathrm{u}}_y := \left\{ x \in \mathbb{R}^n \setminus \{y\}\, ; \inf_{(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}} \{\Delta(|\varphi(\cdot, t_-, x) - \varphi(\cdot, t_-, y)|)(t,s)\} > 0 \right\} \cup \{y\}\,,$$

and call W^{s}_y and W^{u}_y the *domains of (finite-time) attraction* and *repulsion* with respect to $\varphi(\cdot, t_-, y)$, respectively.

5.5 Remark. We call W^{s}_y and W^{u}_y domains of attraction and repulsion, respectively, to emphasize their set structure and to avoid terms like manifold or cone.

It is easy to see that under a time-dependent coordinate shift $(t, x) \mapsto (t, x - \varphi(t, t_-, y))$ the linearization $\Phi_{(t_-,y)}$ (along $\varphi(\cdot, t_-, y)$ in the original coordinates) coincides with the linearization $\tilde{\Phi}_{(t_-,0)}$ (along $\mathbb{I} \times \{0\}$ in the transformed coordinates). Without loss of generality we assume the reference trajectory to be the zero trajectory for the rest of this chapter.

As a next step, we prove that hyperbolic trajectories have, under some condition on the approximation quality of the process by the linearization, non-empty domains of attraction and repulsion. We are going to show that locally cones and domains look very similar, which is a result of local persistence and we adapt the reasoning that led to the robustness of EMD to the nonlinear case. To this end, we first introduce analogues to the one-dimensional growth rates as follows:

$$\underline{\mu}\colon \mathbb{R}^n \setminus \{0\} \times \mathcal{P}(\mathbb{I},\mathbb{R}^n) \to \mathbb{R}, \quad (x,\varphi) \mapsto \inf_{(t,s)\in\neq_{\mathbb{I}\times\mathbb{I}}} \Delta(|\varphi(\cdot,t_-,x)|)(t,s), \tag{5.1}$$

$$\overline{\mu}\colon \mathbb{R}^n \setminus \{0\} \times \mathcal{P}(\mathbb{I},\mathbb{R}^n) \to \mathbb{R}, \quad (x,\varphi) \mapsto \sup_{(t,s)\in\neq_{\mathbb{I}\times\mathbb{I}}} \Delta(|\varphi(\cdot,t_-,x)|)(t,s). \tag{5.2}$$

Note that for $\Phi \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ and $x \in \mathbb{R}^n \setminus \{0\}$ we have $\underline{\mu}(x,\Phi) = \underline{\lambda}(\operatorname{span}\{x\},\Phi)$ and $\overline{\mu}(x,\Phi) = \overline{\lambda}(\operatorname{span}\{x\},\Phi)$. Based on $\underline{\mu}$ and $\overline{\mu}$ we introduce a measure of approximation of the C^1-process φ by the linearization Φ along the zero reference trajectory

$$m\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \tag{5.3a}$$

$$\eta \mapsto \begin{cases} 0, & \eta = 0, \\ \sup\limits_{x\in B[0,\eta]\setminus\{0\}} \max\left\{\left|\underline{\mu}(x,\varphi) - \underline{\mu}(x,\Phi)\right|, |\overline{\mu}(x,\varphi) - \overline{\mu}(x,\Phi)|\right\}, & \text{otherwise.} \end{cases} \tag{5.3b}$$

With this notation at hand, the domains of attraction and repulsion of the zero reference trajectory take the simple form

$$W_0^{\mathrm{s}} := \{x \in \mathbb{R}^n \setminus \{0\}\,;\, \overline{\mu}(x,\varphi) < 0\} \cup \{0\},$$
$$W_0^{\mathrm{u}} := \left\{x \in \mathbb{R}^n \setminus \{0\}\,;\, \underline{\mu}(x,\varphi) > 0\right\} \cup \{0\},$$

from which the similarity to the stable and unstable cones of Definition 3.17 becomes already visible.

Next we give a sufficient condition for m to be continuous at 0.

5.6 Lemma. *Let $\Phi \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ be the linearization of φ along $\varphi(\cdot,t_-,0) = 0$. If $\mathbb{I}$ is finite then the function m as in Eq.* (5.3) *is continuous at* 0*.*

Proof. By the inverse triangle inequality it suffices to prove

$$D(t,s,x) := |\Delta(|\varphi(\cdot,t_-,x)|)(t,s) - \Delta(|\Phi(\cdot,t_-)x|)(t,s)| \xrightarrow{|x|\to 0} 0,$$

uniformly in $(t,s) \in\neq_{\mathbb{I}\times\mathbb{I}}$. Since $\mathbb{I}$ is finite we have $M := \min_{(t,s)\in\neq_{\mathbb{I}\times\mathbb{I}}} |t-s| > 0$. Observe that

$$R(t,x) := \varphi(t,t_-,x) - \partial_2\varphi(t,t_-,0)x = \varphi(t,t_-,x) - \Phi(t,t_-)x$$

is continuous and

$$|R(t,x)| \xrightarrow{|x|\to 0} 0, \quad \text{uniformly in } t \in \mathbb{I}. \tag{5.4}$$

We estimate

$$\begin{aligned}
D(t,s,x) &= \left| \frac{\ln|\varphi(t,t_-,x)| - \ln|\varphi(s,t_-,x)|}{t-s} - \frac{\ln|\Phi(t,t_-)x| - \ln|\Phi(s,t_-)x|}{t-s} \right| \\
&\leq \frac{1}{M}\left(\left|\ln\frac{|\varphi(t,t_-,x)|}{|\Phi(t,t_-)x|}\right| + \left|\ln\frac{|\varphi(s,t_-)x|}{|\Phi(s,t_-,x)|}\right| \right) \\
&= \frac{1}{M}\left(\left|\ln\frac{|\Phi(t,t_-)x + R(t,x)|}{|\Phi(t,t_-)x|}\right| + \left|\ln\frac{|\Phi(s,t_-)x + R(s,x)|}{|\Phi(s,t_-)x|}\right| \right) \\
&\xrightarrow{|x|\to 0} \frac{1}{M}(\ln 1 + \ln 1) = 0,
\end{aligned}$$

uniformly in $(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}$ by Lemma 2.2 and Eq. (5.4). □

Note that in the previous lemma we did not impose any extra regularity conditions neither on the norm nor on the process. The next lemma gives sufficient conditions for the ODE case, which requires both additional regularity of the norm and of the process. We state and prove it for the Euclidean norm, making use of the following facts: the Euclidean norm is continuously differentiable on $\mathbb{R}^n \setminus \{0\}$, and the modulus of continuity of its derivative when restricted to a compact domain D not containing the origin is $\omega(t) = t/\alpha$, where $\alpha := \min\{|x|\,;\, x \in D\}$ (intercept theorem). The following result can be clearly transferred to other norms by requiring continuous differentiability and a certain behavior of the modulus of continuity of the derivative close to the origin, which will become clear in the course of the proof.

5.7 Lemma. *Let $\mathbb{I}$ be a compact interval, $|\cdot|$ denote the Euclidean norm, $\varphi \in \mathcal{P}(\mathbb{I},\mathbb{R}^n)$ be a C^2-process on $\mathbb{I}$ and $\Phi \in \mathcal{LP}(\mathbb{I},\mathbb{R}^n)$ be the linearization of φ along $\varphi(\cdot,t_-,0) = 0$. Suppose that φ and Φ are continuously differentiable in the first argument. Then the function m as in Eq.* (5.3) *is continuous at* 0.

Proof. As in the previous lemma, we show that $D(t,s,x) \to 0$ as $|x| \to 0$ uniformly in $(t,s) \in \neq_{\mathbb{I}\times\mathbb{I}}$, this time by applying the mean value theorem to the continuously differentiable function $t \mapsto \ln\frac{|\varphi(t,t_-,x)|}{|\Phi(t,t_-)x|}$. To this end, note that for $R(t,x) := \varphi(t,t_-,x) - \Phi(t,t_-)x$, $t \in \mathbb{I}$, $x \in \mathbb{R}^n$, we have that $R(t,\cdot)$ together with

$\partial_0 R(t,\cdot)$ is of class $o(|x|)$ for $|x| \to 0$ uniformly in $t \in \mathbb{I}$ due to the twice continuous differentiability of φ. Besides elementary calculations and estimates, the crucial ingredient of the proof is to show that

$$\left\| |\cdot|' \left(\Phi(t,t_-)x + R(t,x)\right) - |\cdot|' \left(\Phi(t,t_-)x\right) \right\| \xrightarrow{|x|\to 0} 0, \tag{5.5}$$

uniformly in $t \in \mathbb{I}$. To show this, we first observe that there exist constants $\varepsilon, \delta \in \mathbb{R}_{>0}$ such that $|R(t,x)| \leq \varepsilon |x|^2 \leq \varepsilon\delta |x|$ whenever $|x| \leq \delta$, due to the aforementioned observation on the convergence of R. Since Φ is invertible and $\mathbb{I}$ is compact, we have that $\alpha := \min \|\Phi([\mathbb{I}], t_-)[\mathcal{S}]\|, \beta := \max \|\Phi([\mathbb{I}], t_-)\| > 0$, where α is the absolute value closest to zero that a trajectory starting on the unit circle attains on $\mathbb{I}$ and, analogously, β is the largest such value. We may assume w.l.o.g. that $\delta < \alpha/\varepsilon$ and hence $\alpha - \varepsilon\delta > 0$. Now choose $\eta < \delta$, then for any $|x| = \eta$ we have that $|\Phi(t,t_-)x| \in [\alpha\eta, \beta\eta]$ and $|\varphi(t,t_-,x)| \in [(\alpha - \varepsilon\delta)\eta, \beta\eta + \varepsilon\eta^2]$. When restricted to the compact annulus $B[0, \beta\eta + \varepsilon\eta^2] \setminus B(0, (\alpha - \varepsilon\delta)\eta)$, the derivative of the Euclidean norm is uniformly continuous with a modulus of continuity of $\omega(t) = t/((\alpha - \varepsilon\delta)\eta)$. On the other hand, on this annulus and all annuli constructed for smaller η we have the quadratic estimate on R, yielding $\omega(\varepsilon\eta^2) = \varepsilon\eta^2/((\alpha - \varepsilon\delta)\eta) \xrightarrow{\eta \to 0} 0$ and in turn proving Eq. (5.5). $\square$

5.2 Local Stable/Unstable Cones and Manifolds

In the next two sections we introduce a new, "intrinsic" approach to local stable and unstable cones and manifolds, which uses information about φ on $\mathbb{I}$ only and hence does not rely on classical asymptotic methods. The essential assumption is the continuity of m at 0.

We define the two functions

$$\eta\colon \operatorname{Gr}(1,\mathbb{R}^n) \to \mathbb{R}_{\geq 0}, \qquad X \mapsto \inf\left\{r \in \mathbb{R}_{>0};\ x \in X \cap \mathcal{S},\ rx \notin W_0^{\mathrm{s}}\right\}, \tag{5.6a}$$
$$\hat{\eta}\colon \operatorname{Gr}(1,\mathbb{R}^n) \to \mathbb{R}_{\geq 0}, \qquad X \mapsto \inf\left\{r \in \mathbb{R}_{>0};\ x \in X \cap \mathcal{S},\ rx \notin W_0^{\mathrm{u}}\right\}. \tag{5.6b}$$

5.8 Theorem (Local Stable/Unstable Cone Theorem). *Let $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ be the linearization of φ along $\varphi(\cdot, t_-, 0) = 0$. If the function m as defined in Eq. (5.3) is continuous at 0 then for any $X, Y \in \operatorname{Gr}(1,\mathbb{R}^n)$ with $X \subseteq V^{\mathrm{s}}(\Phi)$ and $Y \subseteq V^{\mathrm{u}}(\Phi)$ one has $\eta(X), \hat{\eta}(Y) > 0$. Moreover, η and $\hat{\eta}$ are bounded away from zero on compact subsets of $V^{\mathrm{s}}(\Phi)$ and $V^{\mathrm{u}}(\Phi)$, respectively.*

Proof. By Definition 3.17 we have for any $X \in \mathrm{Gr}(1,\mathbb{R}^n)$ with $X \subseteq V^{\mathrm{s}}(\Phi)$ that $\overline{\lambda}(X,\Phi) < 0$. By the continuity assumption on m there exists $\delta \in \mathbb{R}_{>0}$ such that $m(\eta) < -\overline{\lambda}(X,\Phi)$ for any $\eta \in (0,\delta]$. Then for any $x \in B[0,\delta] \cap X$ we have $\overline{\mu}(x,\varphi) < 0$ and analogously the assertion for $Y \in \mathrm{Gr}(1,\mathbb{R}^n)$ with $Y \subseteq V^{\mathrm{u}}(\Phi)$. The second part follows from Remark 3.20 and the same argument as applied before to single directions $X \in \mathrm{Gr}(1,\mathbb{R}^n)$. □

By the same continuity argument as in Theorem 5.8 we can find positive radii $\delta \in \mathbb{R}_{>0}$ for the directions in the interior of $V := \mathbb{R}^n \setminus (V^{\mathrm{s}}(\Phi) \cup V^{\mathrm{u}}(\Phi))$ such that $V \cap B[0,\delta] \cap W_0^{\mathrm{s}}, V \cap B[0,\delta] \cap W_0^{\mathrm{u}} = \varnothing$. Roughly speaking, we find that stable and unstable cones of the linearization Φ and domains of attraction and repulsion of the process φ, respectively, are locally indistinguishable. Note that this is a pure continuity result and not an implication of hyperbolicity. For the ODE case in $\mathbb{R}^2$ and stronger regularity assumptions a similar approximation result has been proved in [33, Theorem 44].

As a special case we obtain the following result.

5.9 Theorem (Local Stable/Unstable Manifold Theorem)**.** *Suppose the assumptions of Theorem 5.8 are satisfied and let $Q \in \mathbb{P}(\mathbb{R}^n)$ be a projection such that* $\operatorname{im} Q \subseteq V^{\mathrm{s}}(\Phi)$ *and* $\ker Q \subseteq V^{\mathrm{u}}(\Phi)$. *Then there exist neighborhoods U and V of the origin such that* $\operatorname{im} Q \cap U \subseteq W_0^{\mathrm{s}}$ *and* $\ker Q \cap V \subseteq W_0^{\mathrm{u}}$. *Furthermore, for any $t \in \mathbb{I}$ the following equations hold:*

$$\begin{aligned} T_0\varphi(t,t_-,[\operatorname{im} Q]) &= \Phi(t,t_-)[\operatorname{im} Q], \\ T_0\varphi(t,t_-,[\ker Q]) &= \Phi(t,t_-)[\ker Q]. \end{aligned} \tag{5.7}$$

Consequently, $\varphi([\mathbb{I}],t_-,[\operatorname{im} Q \cap U])$ and $\varphi([\mathbb{I}],t_-,[\ker Q \cap V])$ can be considered as finite-time local stable *and* unstable manifolds, *respectively.*

Proof. Since $\operatorname{im} Q$ and $\ker Q$ are compact subsets of $V^{\mathrm{s}}(\Phi)$ and $V^{\mathrm{u}}(\Phi)$, respectively. Hence, Theorem 5.8 applies and the first part is proved. Furthermore, the tangencies in Eq. (5.7) are easily verified with the definition of the linearization. □

5.10 Remark. We want to comment on some issues concerning Theorem 5.9.

1. We would like to point out that Theorem 5.9 holds for general compact $\mathbb{I}$. So far, finite-time Local Stable Manifold Theorems have been proved only in the ODE case with $\mathbb{I}$ being a compact interval; see [57, 54, 15] and also Remark 6.4(2).

2. In the finite-time context, one can consider Theorem 5.8 and Theorem 5.9 as robustness results as well. As we proved, locally, the stable and unstable cones (together with the subspaces that they contain) persist under nonlinear perturbations with vanishing first order terms.

3. Despite the lack of structure for the domains of attraction and repulsion themselves, we see that the maximal dimension of manifolds contained in these domains going through the origin corresponds to the indices of the EMD growth rates, i.e. to rank and deficiency of the EMD-projection, respectively.

4. Note that by the assumption that φ is a C^k-process on $\mathbb{I}$, $k \in \{1,2\}$, we obtain directly that the extension of $\operatorname{im} Q \cap U$ and $\ker Q \cap U$ by φ to the extended state space $\mathbb{I} \times \mathbb{R}^n$ gives a C^k-manifold in each time-fiber $\{t\} \times \mathbb{R}^n$, $t \in \mathbb{I}$. Evidently, the chart is given by $\varphi(t, t_-, \cdot)$.

5. The function m can be considered as a local measure of nonlinearity of φ, in the sense that $m = 0$ if φ itself is linear and that m takes small values in case φ is only a small perturbation of a linear process. The more linear φ becomes, the larger we can choose the radius $\delta \in \mathbb{R}_{>0}$ such that $B(0,\delta) \cap \operatorname{im} Q \subseteq W_0^{\mathrm{s}}$ and $B(0,\delta) \cap \ker Q \subseteq W_0^{\mathrm{u}}$. In the "linear limit" we recover that $\operatorname{im} Q \subseteq W_0^{\mathrm{s}}$ and $\ker Q \subseteq W_0^{\mathrm{u}}$. In this sense, we believe that our version of a local finite-time stable manifold theorem is a very natural one. On the other hand, for fixed nonlinearity, clearly m is an increasing function, i.e. the further away we go from the hyperbolic reference trajectory, the weaker we expect the exponential decay/growth to be, until some point where the EMD-subspaces/cones leave the domain of attraction and repulsion, respectively.

As an easy consequence of Theorem 5.9 we obtain the following finite-time analogue of the classical Theorem of Linearized Asymptotic Stability. It is a generalization of [95, Theorem 5.1] to arbitrary compact time-sets.

5.11 Theorem (Linearized Finite-time Attraction/Repulsion)**.** *Let $\varphi \in \mathcal{P}(\mathbb{I}, \mathbb{R}^n)$ be a C^1-process on $\mathbb{I}$, $x \in \mathbb{R}^n$, $\varphi(\cdot, t_-, x)$ an attractive (repulsive) trajectory and m be continuous at 0. Then there exists a neighborhood U of x such that $U \subseteq W_x^{\mathrm{s}}$ ($U \subseteq W_x^{\mathrm{u}}$, respectively).*

5.3 Time-Extensions of Cones and Domains

Next we investigate the relationship between the cones $V^{\mathrm{s}}(\Phi)$ and $V^{\mathrm{u}}(\Phi)$ extended by the linearization Φ to the extended state space on the one hand, i.e.

$$\begin{aligned}\mathcal{V}^{\mathrm{s}}_{\Phi} &:= \Phi(\cdot, t_-)[V^{\mathrm{s}}(\Phi)] = \{(t, \Phi(t, t_-)x) \in \mathbb{I} \times \mathbb{R}^n;\ (t, x) \in \mathbb{I} \times V^{\mathrm{s}}(\Phi)\}\,,\\ \mathcal{V}^{\mathrm{u}}_{\Phi} &:= \Phi(\cdot, t_-)[V^{\mathrm{u}}(\Phi)] = \{(t, \Phi(t, t_-)x) \in \mathbb{I} \times \mathbb{R}^n;\ (t, x) \in \mathbb{I} \times V^{\mathrm{u}}(\Phi)\}\,,\end{aligned}$$

and the domains W_0^{s} and W_0^{u} extended by φ on the other hand, i.e.

$$\begin{aligned}\mathcal{W}^{\mathrm{s}}_{0} &:= \varphi(\cdot, t_-, [W_0^{\mathrm{s}}]) = \{(t, \varphi(t, t_-, x)) \in \mathbb{I} \times \mathbb{R}^n;\ (t, x) \in \mathbb{I} \times W_0^{\mathrm{s}}\}\,,\\ \mathcal{W}^{\mathrm{u}}_{0} &:= \varphi(\cdot, t_-, [W_0^{\mathrm{u}}]) = \{(t, \varphi(t, t_-, x)) \in \mathbb{I} \times \mathbb{R}^n;\ (t, x) \in \mathbb{I} \times W_0^{\mathrm{u}}\}\,.\end{aligned}$$

We denote by $\mathcal{V}^{\mathrm{s}}_{\Phi}(t)$, $\mathcal{V}^{\mathrm{u}}_{\Phi}(t)$, $\mathcal{W}^{\mathrm{s}}_{0}(t)$ and $\mathcal{W}^{\mathrm{u}}_{0}(t)$, $t \in \mathbb{I}$, the t-fiber of the respective subsets of the extended state space. The next proposition states that in each t-fiber the extended stable and unstable cones are locally contained in the domains of attraction and repulsion, respectively.

5.12 Theorem (Relationship between Extensions)**.** *Suppose the assumptions of Theorem 5.8 are satisfied. Define the time-extensions of Eq.* (5.6)

$$\begin{aligned}\eta &: \mathbb{I} \times \mathrm{Gr}(1, \mathbb{R}^n) \to \mathbb{R}_{\geq 0}, \quad (t, Y) \mapsto \inf\{r \in \mathbb{R}_{>0};\ y \in Y \cap \mathcal{S},\ ry \notin \mathcal{W}^{\mathrm{s}}_0(t)\}\,,\\ \hat{\eta} &: \mathbb{I} \times \mathrm{Gr}(1, \mathbb{R}^n) \to \mathbb{R}_{\geq 0}, \quad (t, Y) \mapsto \inf\{r \in \mathbb{R}_{>0};\ y \in Y \cap \mathcal{S},\ ry \notin \mathcal{W}^{\mathrm{u}}_0(t)\}\,.\end{aligned}$$

Then for each $t \in \mathbb{I}$ one has $\eta(t, \cdot)|_{\mathcal{V}^{\mathrm{s}}_{\Phi}(t)}, \hat{\eta}(t, \cdot)|_{\mathcal{V}^{\mathrm{u}}_{\Phi}(t)} > 0$.

Proof. We prove only $\eta(t, \cdot)|_{\mathcal{V}^{\mathrm{s}}_{\Phi}(t)} > 0$, since the second assertion can be proved completely analogously. Let $t \in \mathbb{I}$, $y \in \mathcal{V}^{\mathrm{s}}_{\Phi}(t) \cap \mathcal{S}$. We sketch the idea of the proof first: By the invariance of $\mathcal{V}^{\mathrm{s}}_{\Phi}$ under Φ it is clear that $\Phi(t_-, t)y \in \mathcal{V}^{\mathrm{s}}_{\Phi}(t_-) = V^{\mathrm{s}}(\Phi)$. Now, consider $x_r := \varphi(t_-, t, ry)$, $r \in (0, 1]$. To prove that $ry \in \mathcal{W}^{\mathrm{s}}_0(t)$ for sufficiently small r, it is sufficient to show that $x_r \in W_0^{\mathrm{s}} = \mathcal{W}^{\mathrm{s}}_0(t_-)$. Since $V^{\mathrm{s}}(\Phi)$ is open by Lemma 3.18, our aim is to show that for sufficiently small r we have that x_r is contained in a neighborhood of one-dimensional subspaces around span $\{\Phi(t_-, t)y\}$ which is a subset of $V^{\mathrm{s}}(\Phi)$. By Theorem 5.8 we then conclude that for sufficiently small r the vector x_r is contained in the domain of attraction W_0^{s}. This proves the strict positivity of $\eta(t, \cdot)|_{\mathcal{V}^{\mathrm{s}}_{\Phi}(t)}$ as claimed.

Thus, it remains to show that $x_r \in \mathcal{W}^{\mathrm{s}}_0(t_-)$ for sufficiently small r. By Lemma 3.18 there exists $\theta \in \mathbb{R}_{>0}$ such that

$$B(\Phi(t_-, t)y, \theta) \subset \mathcal{V}^{\mathrm{s}}_{\Phi}(t_-) = V^{\mathrm{s}}(\Phi).$$

Clearly, due to the positive homogeneity of the norm $|\cdot|$ on $\mathbb{R}^n$, the invariance of $\mathcal{V}^{\mathrm{s}}_{\Phi}(t_-)$ under scalar multiplication and linearity of $\Phi(t_-, t)$, we find that for all $r \in (0, 1]$ holds $B(\Phi(t_-, t)(ry), r\theta) \subset \mathcal{V}^{\mathrm{s}}_{\Phi}(t_-)$. By expanding $\varphi(t_-, t, \cdot)$ at 0 we obtain for all $z \in \mathbb{R}^n \setminus \{0\}$

$$|\varphi(t_-, t, z) - \Phi(t_-, t)z| \in o(|z|) \quad \text{for } |z| \to 0.$$

This is equivalent to the fact that for any $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that for any $z \in \mathbb{R}^n \setminus \{0\}$ with $|z| \leq \delta$ we have

$$|\varphi(t_-, t, z) - \Phi(t_-, t)z| < \varepsilon \, |z| \, . \tag{5.8}$$

In other words, for any $\varepsilon \in \mathbb{R}_{>0}$, sufficiently small $\delta \in \mathbb{R}_{>0}$ and $|z| \leq \delta$ we have

$$\varphi(t_-, t, z) \in B(\Phi(t_-, t)z, \varepsilon \, |z|) \subseteq B(z, \varepsilon\delta).$$

In particular, choosing $\varepsilon = \theta/2$ we find that for $\delta \in \mathbb{R}_{>0}$ from Eq. (5.8) and consequently for any ry with $r \in [0, \min\{\delta, 1\}]$ holds

$$x_r = \varphi(t_-, t, ry) \in B[\Phi(t_-, t)(ry), r\theta/2] \subset B(\Phi(t_-, t)(ry), r\theta) \subset \mathcal{V}^{\mathrm{s}}_{\Phi}(t_-).$$

Since $B[\operatorname{span}\{\Phi(t_-, t)(ry)\}, r\theta/2] \subset \mathrm{Gr}(k, \mathbb{R}^n)$ is closed and hence compact, we know that for sufficiently small $r \in \mathbb{R}_{>0}$ we have $x_r \in \mathcal{W}^{\mathrm{s}}_0(t_-)$ by Theorem 5.8. □

Figure 5.1 visualizes the statements of Theorem 5.9 and Theorem 5.12, namely that in directions of the stable and unstable cones the domains of attraction and repulsion have a positive extent.

5.13 Remark. Since our hyperbolicity notion is based fundamentally on the monotonicity of the norm of trajectories, a finite-time conjugacy between the linear process Φ and the general process φ should preserve the type of monotonicity of trajectories. For more information on nonautonomous conjugacy see [102, Section 3.1] and the references therein. Roughly speaking and supposing that the assumptions are satisfied for $\mathbb{I}$, Theorem 5.12 can therefore be interpreted as a finite-time Hartman-Grobman-like theorem in the following informal sense: as demonstrated in [104, p. 546] the function

$$H\colon \mathbb{I} \times \mathbb{R}^n \to \mathbb{R}^n, \qquad (t, x) \mapsto \varphi(t, t_-, \Phi(t_-, t)x),$$

with fiberwise inverse

$$H(t, \cdot)^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n, \qquad x \mapsto \Phi(t, t_-, \varphi(t_-, t)x),$$

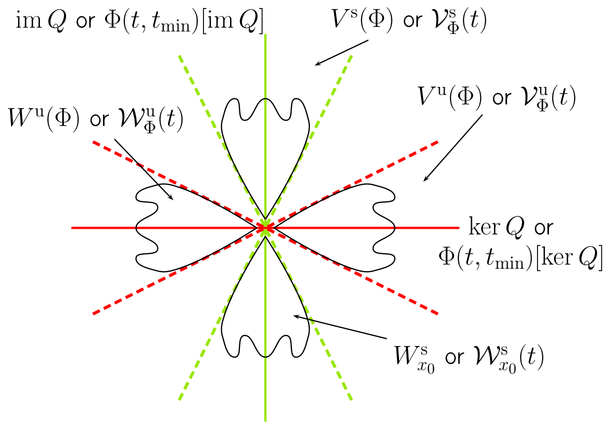


Figure 5.1: Schematic visualization of Theorem 5.9 and Theorem 5.12

maps trajectories of Φ to trajectories of φ homeomorphically and is therefore a candidate for a nonautonomous topological conjugacy between Φ and φ with respect to the two zero reference trajectories. Now consider restrictions of H to $V^{\mathrm{s}}(\Phi)$ and $V^{\mathrm{u}}(\Phi)$ (or compact subsets $S \subset V^{\mathrm{s}}(\Phi)$ and $U \subset V^{\mathrm{u}}(\Phi)$ considered as subsets of $\mathrm{Gr}(1, \mathbb{R}^n)$). By Theorem 5.12 we obtain for y from the respective set with $|y|$ sufficiently small, that H preserves the monotonicity type of $\Phi(\cdot, t_-)y$, but, in general, not the exponential rate. The task of finding a monotonicity preserving (nonautonomous) topological conjugacy for a whole neighborhood seems to be futile since too restrictive. In particular it is unclear how "monotonicity preservation" should be applied to trajectories with "indefinite" monotonicity behavior.

6 Ordinary Differential Equations on Compact Time-Intervals

In the following, we want to apply the developed notions and results to ordinary differential equations, which we define without loss of generality globally for the sake of simplicity, i.e.

$$\dot{x} = f(t, x), \tag{6.1}$$

where $f \in C^1(I \times \mathbb{R}^n, \mathbb{R}^n)$, $I \subseteq \mathbb{R}$ an interval, $\mathbb{I} \subseteq I$ compact and $(f(t, \cdot))_{t \in I}$ is uniformly Lipschitz continuous with Lipschitz constant L_f. Thus, Eq. (6.1) is well-posed and the solution operator φ is well-defined. It is well-known that φ satisfies the conditions of a C^1-process. We consider the case that $|\cdot|$ is continuously differentiable on $\mathbb{R}^n \setminus \{0\}$. In particular, this covers all norms induced by the Euclidean inner product and a symmetric positive definite matrix $\Gamma \in \mathbb{R}^{n \times n}$ as considered in [14, 15]. We fix a solution $\varphi(\cdot, t_-, x)\colon \mathbb{I} \to \mathbb{R}^n$, $x \in \mathbb{R}^n$, and perform a time-dependent coordinate shift of the form $(t, x) \mapsto (t, x - \varphi(t, t_-, x)) =: (t, y)$. Then in the new coordinates Eq. (6.1) takes the form

$$\dot{y} = \partial_1 f(t, \varphi(t, t_-, x))y + g(t, y) = A(t)y + g(t, y), \tag{6.2}$$

where $A := (t \mapsto \partial_1 f(t, \varphi(t, t_-, x))) \in C(\mathbb{I}, L(\mathbb{R}^n))$, $g \in C(\mathbb{I} \times \mathbb{R}^n, \mathbb{R}^n)$ and

$$g(t, v) = f(t, v + \varphi(t, t_-, x)) - f(t, \varphi(t, t_-, x)) - \partial_1 f(t, \varphi(t, t_-, x))v,$$

for $t \in \mathbb{I}$ is the nonlinear term. By definition of the derivative we have

$$g(t, x) / |x| \xrightarrow{|x| \to 0} 0$$

for any $t \in \mathbb{I}$. In other words, for any $t \in \mathbb{I}$ and $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta \in (0, 1]$ such that the estimate $\sup \{|g(t, x)| / |x|\,;\, x \in B(0, \delta) \setminus \{0\}\} < \varepsilon$ holds. Due to the uniform continuity of $g|_{\mathbb{I} \times B[0,1]}$, we even obtain that for any $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta \in (0, 1]$ such that

$$\sup \left\{ \frac{|g(t, x)|}{|x|};\, x \in B(0, \delta) \setminus \{0\},\, t \in \mathbb{I} \right\} < \varepsilon. \tag{6.3}$$

We call

$$\dot{y} = \partial_1 f(t, \varphi(t, t_-, x))y \tag{6.4}$$

the linearization of (6.1) along $\varphi(\cdot, t_-, x)$. It is well-known that the associated solution operator Φ of (6.4), interpreted as a linear process, is the linearization of φ along $\varphi(\cdot, t_-, x)$ and that Φ is continuously differentiable in the first argument. Under the differentiability assumption on the norm the growth rates take the form (as introduced in [17, Definition 7])

$$\begin{aligned}
\underline{\lambda}(X, \Phi) &= \min\left\{ \frac{(|\Phi(\cdot, t_-)x|)'(t)}{|\Phi(t, t_-)x|};\ t \in \mathbb{I},\ x \in X \cap \mathcal{S} \right\}, \\
\overline{\lambda}(X, \Phi) &= \max\left\{ \frac{(|\Phi(\cdot, t_-)x|)'(t)}{|\Phi(t, t_-)x|};\ t \in \mathbb{I},\ x \in X \cap \mathcal{S} \right\},
\end{aligned} \tag{6.5}$$

for $X \in \mathrm{Gr}(k, \mathbb{R}^n)$, which can be seen by Lemma 2.12. From this representation and the chain rule $(|\Phi(\cdot, t_-)x|)'(t) = |\cdot|'\,(\Phi(t, t_-)x)\partial_0(\Phi(t, t_-)x)$ for $t \in \mathbb{I}$ we can see that $d_{\mathbb{I}}$ ($\tilde{d}_{\mathbb{I}}$) can be interpreted as some kind of C^1 (semi-)metric for linear solution operators on $\mathbb{I}$.

By classical techniques and Gronwall's lemma one easily establishes that linear right-hand sides $A \in C(\mathbb{I}, L(\mathbb{R}^n))$ map continuously (with respect to $d_{\mathbb{I}}$) to their unique solution operator $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$.

6.1 Lemma. *Let $I \subseteq \mathbb{R}$ be some interval and $\mathbb{I} \subseteq I$ compact. Then the map*

$$\begin{gathered}
(C(I, L(\mathbb{R}^n)), \|\cdot\|_\infty) \to (\mathcal{LP}(\mathbb{I}, \mathbb{R}^n), d_{\mathbb{I}}), \\
A \mapsto \left(\in \left(\left\{\Phi \in C^1(I \times I, L(\mathbb{R}^n));\ \partial_0\Phi(t,s) = A(t)\Phi(t,s),\ \Phi(s,s) = \mathrm{id}_{\mathbb{R}^n}\right\}\right)\right)\big|_{\mathbb{I}\times\mathbb{I}\times\mathbb{R}^n}
\end{gathered}$$

is continuous.

6.1 Linearization of Finite-time ODEs

The next step is to show that Lemma 5.7 applies. Therefore, it remains to establish the necessary order of convergence for the linearization error. To this end, consider Eq. (6.1) and assume that $f \in C^{0,2}(I \times \mathbb{R}^n, \mathbb{R}^n)$, i.e. f is continuous in the first argument and twice continuously differentiable in the second argument. By Taylor's Theorem, the estimate (6.3) on the nonlinear term g improves as follows: there exist $\varepsilon, \delta \in \mathbb{R}_{>0}$ such that

$$\sup\left\{ \frac{|g(t,x)|}{|x|^2};\ x \in B(0,\delta) \setminus \{0\},\ t \in \mathbb{I} \right\} < \varepsilon. \tag{6.6}$$

6.2 Lemma. *Let φ be the solution operator of (6.2) with $f \in C^{0,2}(I \times \mathbb{R}^n, \mathbb{R}^n)$ and Φ be the solution operator of the linearization (6.4) along the reference solution $\varphi(\cdot, t_-, 0) = 0$. Then the function*

$$\begin{aligned} &\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \\ &\eta \mapsto \begin{cases} 0, & \eta = 0, \\ \sup \left\{ \|\varphi(\cdot, t_-, y) - \Phi(\cdot, t_-)y\|_\infty \,;\, y \in B[0, \eta] \right\}, & \text{otherwise,} \end{cases} \end{aligned}$$

is continuous at 0. Moreover, it is of class $\mathcal{O}(\eta^2)$ for $\eta \to 0$.

Proof. Let $\eta \in \mathbb{R}_{>0}$ and $y \in \mathbb{R}^n$, $|y| \leq \eta$. Integrating Eqs. (6.2) and (6.4), we calculate for any $t \in \mathbb{I}$

$$\begin{aligned} |\varphi(t, t_-, y) - \Phi(t, t_-)y| &= \left| \int_{t_-}^{t} A(s) \left(\varphi(s, t_-, y) - \Phi(t, s)y \right) + g(s, \varphi(s, t_-, y)) \, \mathrm{d}s \right| \\ &\leq \int_{t_-}^{t} |A(s)(\varphi(s, t_-, y) - \Phi(s, t_-)y)| \, \mathrm{d}s + \\ &\quad + \int_{t_-}^{t} |g(s, \varphi(s, t_-, y)) - g(s, \Phi(s, t_-)y)| \, \mathrm{d}s + \\ &\quad + \int_{t_-}^{t} |g(s, \Phi(s, t_-)y)| \, \mathrm{d}s \\ &\leq (2 \|A\|_\infty + L_f) \int_{t_-}^{t} |\varphi(s, t_-, y) - \Phi(s, t_-)y| \, \mathrm{d}s + \\ &\quad + (t_+ - t_-)\varepsilon C_\Phi \eta^2, \end{aligned}$$

where $C_\Phi := \|\Phi(\cdot, \cdot)\|_\infty < \infty$ and $\varepsilon \in \mathbb{R}_{>0}$ satisfies Eq. (6.6) with $\delta := C_\Phi \eta$. With the abbreviation

$$C(\eta) := (t_+ - t_-)\varepsilon C_\Phi \eta^2 \mathrm{e}^{(2\|A\|_\infty + L_f)(t_+ - t_-)} \in \mathcal{O}(\eta^2),$$

the uniform estimate from above and Gronwall's lemma we obtain

$$\sup \left\{ \|\varphi(\cdot, t_-, y) - \Phi(\cdot, t_-)y\|_\infty \,;\, y \in B[0, \eta] \right\} \leq C(\eta)$$

and the assertion is proved. $\square$

6.3 Lemma. *Let φ be the solution operator of Eq. (6.2) with $f \in C^{0,2}(I \times \mathbb{R}^n, \mathbb{R}^n)$ and Φ be the solution operator of the linearization (6.4) along the reference solution*

$\varphi(\cdot, t_-, x) = 0$. *Then the function*

$$\begin{aligned}
&\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \\
&\eta \mapsto \begin{cases} 0, & \eta = 0, \\ \sup\left\{\|\partial_0\varphi(\cdot, t_-, y) - \partial_0\Phi(\cdot, t_-)y\|_\infty \,;\, y \in B[0, \eta]\right\}, & \textit{otherwise}, \end{cases}
\end{aligned}$$

is continuous at 0. *Moreover, it is of class* $\mathcal{O}(\eta^2)$ *for* $\eta \to 0$.

Proof. Let $\eta \in \mathbb{R}_{>0}$ and $y \in \mathbb{R}^n$, $|y| \leq \eta$. We calculate

$$\begin{aligned}
|\partial_0\varphi(t, t_-, y) - \partial_0\Phi(t, t_-)y| &= |A(t)(\varphi(t, t_-, y) - \Phi(t, t_-)y) + g(t, \varphi(t, t_-, y))| \\
&\leq \|A\|_\infty |\varphi(t, t_-, y) - \Phi(t, t_-)y| + \\
&\quad + |g(t, \varphi(t, t_-, y)) - g(t, \Phi(t, t_-)y)| + \\
&\quad + |g(t, \Phi(t, t_-)y)| \\
&\leq \|A\|_\infty C(\eta) + (\|A\|_\infty + L_f)C(\eta) + \varepsilon C_\Phi \eta^2 \\
&\leq (2\|A\|_\infty + L_f)C(\eta) + \varepsilon C_\Phi \eta^2,
\end{aligned}$$

where we used the notation from the proof of Lemma 6.2. This proves the assertion. □

In summary, if φ is the solution operator of Eq. (6.2) with $f \in C^{0,2}(I \times \mathbb{R}^n, \mathbb{R}^n)$ and Φ the solution operator of the linearization (6.4) along the reference solution $\varphi(\cdot, t_-, x) = 0$ and $|\cdot|$ is the Euclidean norm, then Lemma 5.7 applies and the function m as defined in Eq. (5.3) is continuous at 0. Hence, Theorem 5.8, Theorem 5.9 and Theorem 5.11 apply.

6.4 Remark. We want to comment on some issues concerning the application of the results in Chapter 5 to ODEs.

1. Concerning Theorem 5.9, note that EMD-subspaces from the starting time-fiber $\{t_-\} \times \mathbb{R}^n$ evolve nonlinearly under φ. Their extensions via φ considered as subsets in the extended state space $\mathbb{I} \times \mathbb{R}^n$ are C^1-manifolds since $\varphi(\cdot, t_-, \cdot)$ is continuously differentiable; see, for instance, [6, Theorem 9.2].

2. Our version of Theorem 5.9 applied to the ODE situation extends some previous work on that topic: in [57, 54, 15] the ODE is first extended to the real line and the desired manifolds are then obtained as the local stable and unstable manifolds of solutions that are hyperbolic on $\mathbb{R}$ in the sense that they admit an exponential dichotomy. Note that our proof, in principle, does not restrict to norms induced by an inner product weighted with a

symmetric, positive definite matrix, as it is done in [15] (cf. also [15, Remark 4(ii)]). In [33] an "intrinsic" proof is presented which is tailored for the case $\mathbb{R}^n = \mathbb{R}^2$. There it is shown, that the domains of attraction and repulsion are not empty under appropriate conditions.

3. In applications, one has to take special care when the underlying vector field is interpolated from a data set which is discrete in space and time. For instance, when using the 3-dimensional local interpolation scheme proposed in [73] one obtains a globally C^1 vector field.

6.2 Robustness & Hyperbolicity Radius

The following result follows from the continuous dependence of the linear solution operator on the linear right-hand side of the ODE established in Lemma 6.1 and was obtained already in [14].

6.5 Theorem (Robustness of EMD, [14, Lemma 3])**.** *Consider*

$$\dot{x} = A(t)x, \tag{6.7}$$

with $A \in C(\mathbb{I}, L(\mathbb{R}^n))$. Suppose the associated solution operator Φ admits an EMD on $\mathbb{I}$. Then there exists $\delta \in \mathbb{R}_{>0}$ such that the solution operator of any $B \in B(A, \delta)$ admits an EMD on $\mathbb{I}$ (with the same extremal projection as A).

The last result is interesting from the following point of view. When imposed on $\mathbb{R}_{\geq 0}$ the inequalities (3.2) in Lemma 3.5 required for an EMD on $\mathbb{R}_{\geq 0}$ correspond to the definition of a so-called *semistrong dichotomy* of (6.7) on $\mathbb{R}_{\geq 0}$; see [108]. [107, 108] yield that semistrong (exponential) dichotomies on $\mathbb{R}_{\geq 0}$ are robust only in the larger class of general exponential dichotomies, but not within semistrong dichotomies. That means that robustness of EMD cannot be deduced from the classical asymptotic analysis but is a pure finite-time result.

Analogously to the robustness investigation for linear processes in Section 3.3 we now address the question of the stability radius for linear ordinary differential equations given on a compact time-interval.

6.6 Definition (Stability radius)**.** Suppose Eq. (6.7) with $A \in C(\mathbb{I}, L(\mathbb{R}^n))$ generates an attracting solution operator on I. Then we define the *stability radius of A* by

$$r(A) := \inf \{ \|B\|_\infty \, ; \, B \in C(\mathbb{I}, L(\mathbb{R}^n)), \, (A + B) \text{ is not attracting on } \mathbb{I} \} .$$

To calculate the stability radius of a given A we make use of an elementary result which can be found in [26] and which specializes to the following result.

6.7 Proposition (cf. [26, Proposition 1, p. 2])**.** *Consider*

$$\dot{x} = A(t)x,$$

with $A \in C(\mathbb{I}, L(\mathbb{R}^n))$*. Suppose the associated solution operator* Φ *is attractive on* $\mathbb{I}$*, i.e.* Φ *admits an EMD on* $\mathbb{I}$ *with the trivial projection* $\mathrm{id}_{\mathbb{R}^n}$ *and* $\overline{\lambda}(\mathbb{R}^n, \Phi) < 0$*. Then for any* $B \in C(\mathbb{I}, L(\mathbb{R}^n))$ *with* $\|B\|_\infty \leq \left|\overline{\lambda}(\mathbb{R}^n, \Phi)\right| =: \delta$ *the solution operator* Ψ *of the perturbed ODE*

$$\dot{y} = (A(t) + B(t))y$$

satisfies $\overline{\lambda}(\mathbb{R}^n, \Psi) \leq \overline{\lambda}(\mathbb{R}^n, \Phi) + \delta = 0$*.*

In other words, $-\overline{\lambda}(\mathbb{R}^n, \Phi) > 0$ is a lower bound on the stability radius around A in $C(\mathbb{I}, L(\mathbb{R}^n))$. The fact that it is also an upper bound follows directly from the well-known correspondence of shifted ODEs and weighted processes (cf. Eq. (3.4)), the definition of spectrum based on the weights/shifts and the fact that for operator norms induced by a vector norm the identity has norm 1. In summary, by a completely analogous argumentation as in the proof of Theorem 3.16 we find that

$$\dot{y} = (A(t) - \overline{\lambda}(\mathbb{R}^n, \Phi)\, \mathrm{id}_{\mathbb{R}^n})y$$

is not attractive. Thus, we obtained that the stability radius of an attractive linear ODE given on $\mathbb{I}$ and the (pseudo-)stability radius of its associated solution operator coincide.

6.8 Theorem (Finite-time Stability Radius)**.** *Let* $A \in C(\mathbb{I}, L(\mathbb{R}^n))$ *and let the associated solution operator* Φ *be attractive. Then*

$$r(A) = -\overline{\lambda}(\mathbb{R}^n, \Phi).$$

Another consequence of EMD-robustness deals with linearizations. The next lemma states that the linearization depends continuously on the initial value of the solution along which we linearize.

6.9 Lemma. *Consider Eq.* (6.1) *and let* φ *denote the associated solution operator. Then the function*

$$\mathbb{R}^n \to C(\mathbb{I}, L(\mathbb{R}^n)), \qquad x \mapsto \partial_1 f(\cdot, \varphi(\cdot, t_-, x)),$$

is continuous.

Proof. Let $t \in \mathbb{I}$, $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}_{>0}$. By the C^1 assumption on f we have in particular that there exists $\delta_1 \in \mathbb{R}_{>0}$ such that

$$|\varphi(t, t_-, x) - \varphi(t, t_-, y)| < \delta_1 \Rightarrow \|\partial_1 f(t, \varphi(t, t_-, x)) - \partial_1 f(t, \varphi(t, t_-, y))\| < \varepsilon.$$

Next note that $y \mapsto \varphi(\cdot, t_-, y) \in C(\mathbb{I}, \mathbb{R}^n)$ is uniformly continuous on any bounded subset $U \subseteq \mathbb{R}^n$ with $x \in U$, i.e. there exists $\delta_2 \in \mathbb{R}_{>0}$ such that for any $y \in U$ we have

$$|x - y| < \delta_2 \quad \Longrightarrow \quad \|\varphi(\cdot, t_-, x) - \varphi(\cdot, t_-, y)\|_\infty < \delta_1.$$

Combining the two continuity observations we obtain that for any $y \in \mathbb{R}^n$ we have

$$|x - y| < \delta_2 \quad \Longrightarrow \quad \|\partial_1 f(\cdot, \varphi(\cdot, t_-, x)) - \partial_1 f(\cdot, \varphi(\cdot, t_-, y))\|_\infty < \varepsilon. \qquad \square$$

Consequently, robustness of EMD carries over to initial values.

6.10 Corollary ([14, Theorem 5]). *Consider Eq. (6.1), let φ be the associated solution operator and $\varphi(\cdot, t_-, x)$ be a hyperbolic/attractive/repulsive reference solution. Then there exists $\delta \in \mathbb{R}_{>0}$ such that for any $y \in B(x, \delta)$ the trajectories $\varphi(\cdot, t_-, y)$ are, respectively, hyperbolic/attractive/repulsive (with respect to the same corresponding subspaces).*

6.3 Finite-time Dynamics on Compact Intervals

One-dimensional differential equations

We consider the following special case of Eq. (6.1)

$$\dot{x} = a(t)x, \tag{6.8}$$

where $a \in C(I, \mathbb{R})$ and $I \subseteq \mathbb{R}$. Let $\mathbb{I} = [t_-, t_+] \subseteq I$ be a compact non-trivial interval and denote the associated solution operator by Φ. Due to the linearity of the right-hand side of Eq. (6.8), to characterize Φ it suffices to consider the solution starting in 1 at time t_-. Obviously, we have $\Phi(\cdot, t_-)1 > 0$ and

$$\begin{aligned}
\underline{\lambda}_1(\Phi) = \underline{\lambda}(\mathbb{R}, \Phi) &= \inf \left\{ \frac{\partial_0 \Phi(t, t_-)1}{\Phi(t, t_-)1};\ t \in \mathbb{I} \right\} = \inf a[\mathbb{I}], \\
\overline{\lambda}_1(\Phi) = \overline{\lambda}(\mathbb{R}, \Phi) &= \sup \left\{ \frac{\partial_0 \Phi(t, t_-)1}{\Phi(t, t_-)1};\ t \in \mathbb{I} \right\} = \sup a[\mathbb{I}],
\end{aligned}$$

and consequently

$$\Sigma(\Phi) = [\inf a[\mathbb{I}], \sup a[\mathbb{I}]],$$

see also [17, Example 21].

Linear time-translation invariant ODEs

We consider the following time-translation invariant differential equation on the time interval $[0, T]$

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

where $A \in L(\mathbb{R}^n)$ and Φ is the associated solution operator. In case A is diagonalizable over $\mathbb{C}$, we have that the generalized eigenspaces are invariant and we have the two possible dynamics: eigenspaces corresponding to real eigenvalues are one-dimensional and we have purely exponential growth/decay, eigenspaces corresponding to pairs of complex conjugate eigenvalues are 2-dimensional and we have exponential growth together with a rotation, that does not affect the Euclidean norm. As is shown in [17, Theorem 22] one can find a norm on $\mathbb{R}^n$ such that $\Sigma^{\mathbb{I}}(\Phi) = \Re[\sigma(A)]$ as follows. For a symmetric positive definite matrix $\Gamma \in \mathbb{R}^{n\times n}$ consider the norm $|\cdot|_\Gamma := \left(x \mapsto \langle x, \Gamma x\rangle^{\frac{1}{2}}\right)$ and choose $\Gamma := T^{-\top}T^{-1}$, where $T \in GL(n, \mathbb{R})$ is the transformation matrix that transforms A to its real Jordan normal form. Then $|\cdot|_\Gamma$ is "aligned" with the eigenspaces in the sense that the eigenspaces are orthogonal to each other in the new coordinates. This yields the intuitively expected result.

In case A is not diagonalizable, this is no longer true. For the general case, there exists a completely independent approach to transient dynamics of time-translation invariant differential equations; see [92].

In order to study the dependence of stable and unstable cones on γ we consider the following time-translation invariant example:

$$\dot{x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} x$$

on the time interval $[0, 1]$. We want to compute for each initial value for which $\gamma \in (-1, 1)$ the associated solution is monotonically increasing or decreasing under the solution operator of

$$\dot{x} = \begin{pmatrix} 1-\gamma & 0 \\ 0 & -1-\gamma \end{pmatrix} x.$$

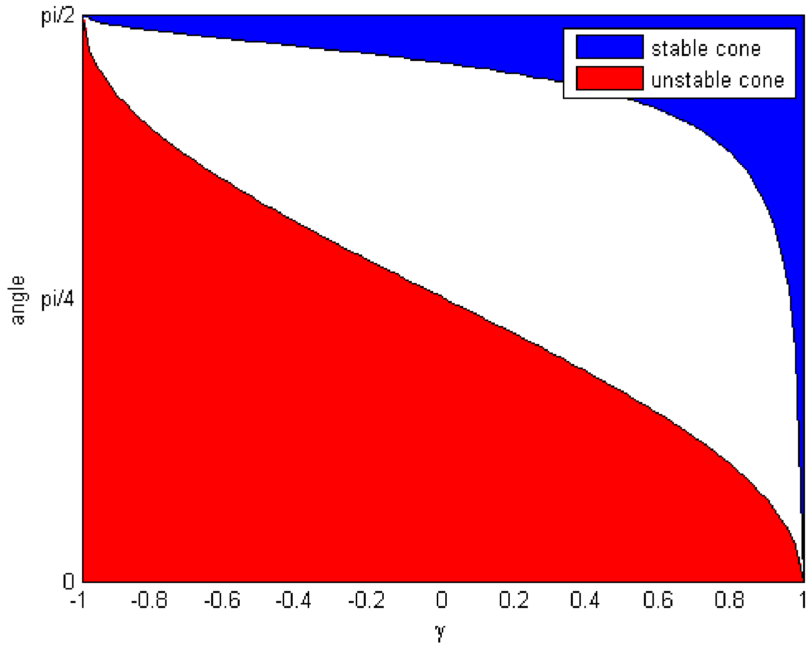


Figure 6.1: Paradigm for dependence of stable and unstable cone on γ

Due to the simple structure of the system and of the solution operator, it suffices to check at the initial and final time point for monotonicity. In other words, the loss of monotonicity can be detected by computing the time derivative of the norm of the trajectory at the initial and the final time point, respectively, depending on whether the normalized initial value is closer to the unstable or the stable direction. The approach is straightforward and a calculation by means of a computer algebra system yields the following angles as functions of the weighting parameter γ: consider α and β as the angles between the abscissa and the boundary of the stable and unstable cones, respectively. The result is:

$$\alpha(\gamma) = \arcsin\left(\mathrm{e}^2\sqrt{-\frac{\gamma-1}{1+\gamma+\mathrm{e}^4-\gamma\mathrm{e}^4}}\right),$$
$$\beta(\gamma) = \arcsin\left(\sqrt{\frac{1-\gamma}{2}}\right),$$

as plotted in Figure 6.1. One can read Figure 6.1 in a vertical way and in a horizontal way. In the vertical way, choose some $\gamma \in (0,1)$ and then read off the opening angles of $V^{\mathrm{s}}(\Phi_\gamma)$ and $V^{\mathrm{u}}(\Phi_\gamma)$ (with respect to the abscissa), and the intermediate segment between the red and the blue lines corresponds to the "neutral" cone $\mathbb{R}^2 \setminus (V^{\mathrm{s}}(\Phi_\gamma) \cup V^{\mathrm{u}}(\Phi_\gamma))$, and the length of the segment to its opening angle. Conversely, in the horizontal way, choose some angle $\zeta \in (0, \pi/2)$ corresponding to a direction of initial values and then read off the lower and upper growth rate of the solution starting in $\begin{pmatrix} \cos\zeta \\ \sin\zeta \end{pmatrix}$. Here, the length of the intermediate segment corresponds to the distance between lower and upper growth rate of the associated trajectory.

7 Finite-time Chaotic Advection

> Nevertheless, chaotic dynamics may occur around lower level structures and the LCS in higher dimensional system may no longer be the only source of chaotic advection.
>
> (Lekien, Shadden & Marsden [74])

> Finally, lower-dimensional LCS can be defined and analyzed in a manner similar to what we pursued for codimension-one LCS in this paper. Normal perturbations to a lower-dimensional material surface, however, span higher-dimensional normal spaces. As a result, several normal repulsion rates and ratios will need to be defined and extremized simultaneously.
>
> (Haller [56])

This chapter is devoted to the study of normal hyperbolicity of special invariant manifolds called *Lagrangian Coherent Structures (LCSs)*. As already mentioned in the introduction, we investigate the impact of a, for instance, solution operator φ on sets/manifolds when evaluated at two time points. This reduces to the investigation of diffeomorphisms/isomorphisms on $\mathbb{R}^n$. In fact, this seems to eliminate the dynamics. However, this approach appears to be suitable and useful in a wide range of applications.

Throughout this chapter, we endow $\mathbb{R}^n$ with the Euclidean norm.

7.1 Finite-time Lyapunov Exponents

We start with an observation, which has been made in [31]. For convenience, we provide a direct and short proof.

7.1 Proposition (cf. [31, Proposition 20 and Theorem 21])**.** *Let* $\mathbb{I} = \{t_-, t_+\}$*,* $t_- < t_+$*,* $\Phi \in \mathcal{LP}(\mathbb{I}, \mathbb{R}^n)$ *and* $0 < \sigma_1 \leq \ldots \leq \sigma_n$ *be the singular values of* $\Phi(t_+, t_-)$*. Then for each* $k \in \{1, \ldots, n\}$ *we have*

$$\underline{\lambda}_{n-k+1}(\Phi) = \frac{1}{t_+ - t_-} \ln \sigma_k, \qquad \overline{\lambda}_k(\Phi) = \frac{1}{t_+ - t_-} \ln \sigma_k.$$

Consequently, $\Sigma(\Phi) = \frac{1}{t_+ - t_-} \ln \sqrt{[\sigma(\Phi(t_+, t_-)^* \Phi(t_+, t_-))]}$ *holds.*

Proof. We calculate

$$\begin{aligned}\overline{\lambda}_k(\Phi) &= \min \overline{\lambda}([\mathrm{Gr}(k, \mathbb{R}^n)], \Phi) = \min_{X \in \mathrm{Gr}(k, \mathbb{R}^n)} \max_{x \in X \setminus \{0\}} \frac{1}{t_+ - t_-} \ln \frac{|\Phi(t_+, t_-)x|}{|x|} \\ &= \frac{1}{t_+ - t_-} \ln \min_{X \in \mathrm{Gr}(k, \mathbb{R}^n)} \max_{x \in X \setminus \{0\}} \frac{|\Phi(t_+, t_-)x|}{|x|} = \frac{1}{t_+ - t_-} \ln \sigma_k,\end{aligned}$$

where the first two equalities hold by definition, the third by the monotonicity of ln and the last by the Courant-Fischer Minimax Theorem [112, Theorem 8.9]. Analogously one can show $\underline{\lambda}_{n-k+1}(\Phi) = \frac{1}{t_+ - t_-} \ln \sigma_k$, where the index $n - k + 1$ is due to the reversed order of the singular values in the analogous "Maximin" Theorem. The representation of the spectrum holds by Theorem 3.10 and by definition of the singular values of $\Phi(t_+, t_-)$. Recall that the singular values of $\Phi(t_+, t_-)$ are defined as the square roots of the eigenvalues of the symmetric positive definite operator $\Phi(t_+, t_-)^* \Phi(t_+, t_-)$. □

The previous proposition motivates the following modification for a solution operator φ of an ODE

$$\dot{x} = f(t, x), \tag{7.1}$$

where $f \in C^{0,1}(I \times \mathbb{R}^n, \mathbb{R}^n)$ and $I \subseteq \mathbb{R}$ is an interval. We introduce the *right Cauchy-Green strain tensor* $C(t, s, x) := \big(\partial_2 \varphi(t, s, x)\big)^* \partial_2 \varphi(t, s, x)$ and the *left Cauchy-Green strain tensor* $B(t, s, x) := \partial_2 \varphi(t, s, x) \big(\partial_2 \varphi(t, s, x)\big)^*$, which are symmetric positive definite linear operators. It is readily verified that $\sigma(C(t, s, x)) = \sigma(B(t, s, x))$. As mentioned before, the roots of their eigenvalues are the singular values of $\partial_2 \varphi(t, s, x)$. We denote the eigenvalues of $C(t, s, x)$ by

$$0 < \lambda_1(t, s, x) \leq \ldots \leq \lambda_n(t, s, x),$$

and an orthonormal eigenbasis of $C(t, s, x)$ by $v_1(t, s, x), \ldots, v_n(t, s, x)$.

For the case when we consider $\varphi(t, s, \cdot)$ for two time-points $t, s \in I$, we name the spectrum of the linearization *finite-time Lyapunov spectrum*.

7.2 Definition (Finite-time Lyapunov spectrum, finite-time Lyapunov exponent, cf. [31]). Consider Eq. (7.1) on some interval $I \subseteq \mathbb{R}$ and let φ denote the associated solution operator. We define

$$\begin{aligned}\Sigma_{\mathrm{FTL}} \colon I \times I \times \mathbb{R}^n &\to 2^{\mathbb{R}}, \\ (t, s, x) &\mapsto \frac{1}{|t-s|} \ln \sqrt{[\sigma\left(C(t,s,x)\right)]}.\end{aligned}$$

and call $\Sigma_{\mathrm{FTL}}(t,s,x)$ the *finite-time Lyapunov spectrum (at x for integration times t and s)*. We refer to its elements as *finite-time Lyapunov exponents (FTLEs) (at x for integration times t and s)*. We define the *maximal finite-time Lyapunov exponent* as

$$\begin{aligned}\sigma \colon I \times I \times \mathbb{R}^n &\to \mathbb{R}, \\ (t,s,x) &\mapsto \sigma(t,s,x) := \max \Sigma_{\mathrm{FTL}}(t,s,x) = \frac{1}{|t-s|} \ln \|\partial_2 \varphi(t,s,x)\| .\end{aligned}$$

For short, we also refer to the maximal finite-time Lyapunov-exponent as *the* finite-time Lyapunov exponent (FTLE).

7.3 Remark. In this remark, we want to motivate the reference to Lyapunov spectrum/exponents in the denomination of Σ_{FTL}. Let us first recall the general definition of Lyapunov exponents. Let φ be the solution operator of Eq. (7.1), where f, and hence φ, is C^1 in the last argument (space variable) and defined on the time-set $I = \mathbb{R}_{\geq 0}$. For $x, v \in \mathbb{R}^n$, $v \neq 0$, the *(forward) Lyapunov exponent* is defined as

$$\lambda^+(x,v) := \limsup_{t\to\infty} \frac{1}{t} \log |\partial_2 \varphi(t,0,x)v| . \tag{7.2}$$

It is well-known that $\lambda^+(x,\cdot)$ is constant on one-dimensional subspaces through the origin and that for fixed $x \in \mathbb{R}^n$ the function $\lambda^+(x,\cdot)$ attains at most n different values; see, for instance, [4]. The finite set of those Lyapunov exponents is called the Lyapunov spectrum in x. Assuming the existence of the limit matrix

$$\Lambda := \lim_{t\to\infty} \left((\partial_2 \varphi(t,0,x))^* \partial_2 \varphi(t,0,x) \right)^{\frac{1}{2t}}, \tag{7.3}$$

the Lyapunov spectrum defined above coincides with the logarithm of the eigenvalues of Λ, i.e. with $\ln[\sigma(\Lambda)]$; see, for instance, [97, Theorem B.3]. The motivation for the term finite-time Lyapunov exponents is the similarity of Σ_{FTL} and $\ln[\sigma(\Lambda)]$ when omitting the limit.

By the monotonicity of $\frac{1}{|t-s|} \ln \sqrt{\cdot}$, $t, s \in I$, the FTLE can be characterized by

$$\sigma(t,s,x) = \frac{1}{|t-s|} \ln \sqrt{\max \sigma\left(C(t,s,x)\right)}. \tag{7.4}$$

This is how the FTLE is usually introduced in the literature, cf. [54, 55, 101] for instance. The FTLE is considered as a measure for maximum stretching of nearby trajectories. Analogously, the "minimal" FTLE can be considered as a measure for compression of nearby trajectories (see [58]) due to the following calculation:

$$\begin{aligned}\sigma(s,t,\varphi(t,s,x)) &= \max \Sigma_{\mathrm{FTL}}(s,t,\varphi(t,s,x)) \\ &= \max \frac{1}{|s-t|} \ln \sqrt{\left[\sigma\left(\left(\partial_2\varphi(s,t,\varphi(t,s,x))\right)^* \partial_2\varphi(s,t,\varphi(t,s,x))\right)\right]}\end{aligned}$$

and with $\partial_2\varphi(s,t,\varphi(t,s,x))\partial_2\varphi(t,s,x) = (y \mapsto \varphi(s,t,\varphi(t,s,y)))'(x) = \mathrm{id}_{\mathbb{R}^n}$, it follows that $\partial_2\varphi(s,t,\varphi(t,s,x)) = (\partial_2\varphi(t,s,x))^{-1}$ and thus

$$\begin{aligned}\sigma(s,t,\varphi(t,s,x)) &= \max \frac{1}{|s-t|} \ln \sqrt{\left[\sigma\left(\left((\partial_2\varphi(t,s,x))^{-1}\right)^* (\partial_2\varphi(t,s,x))^{-1}\right)\right]} \\ &= \max \frac{1}{|t-s|} \ln \sqrt{\left[\sigma\left(\left(\partial_2\varphi(t,s,x)(\partial_2\varphi(t,s,x))^*\right)^{-1}\right)\right]} \\ &= \max \frac{1}{|t-s|} \ln \sqrt{\frac{1}{[\sigma\left(B(t,s,x)\right)]}} \\ &= \frac{1}{|t-s|} \ln \sqrt{\frac{1}{\min \sigma\left(B(t,s,x)\right)}} \\ &= -\min \frac{1}{|t-s|} \ln \sqrt{[\sigma(C(t,s,x))]} \\ &= -\min \Sigma_{\mathrm{FTL}}(t,s,x).\end{aligned}$$

By definition, all results related to the finite-time dichotomy spectrum hold for the finite-time Lyapunov spectrum, e.g. the continuous dependence on the linear process established in Proposition 3.11, the continuous dependence on the initial value by Lemma 6.9 and the robustness of hyperbolicity, see Theorem 6.5 and Corollary 6.10.

Next, we investigate the regularity of the FTLE function σ. For a given differential equation (7.1) and for given integration times $t, s \in I$ we refer to the associated mapping $\sigma(t,s,\cdot)$ also as the *FTLE field*. It will play an important role in the investigation of hyperbolic LCSs, as we discuss in the next section.

First, we link the regularity of the right-hand side f of Eq. (7.1) with the regularity of the associated solution operator and its linearization.

7.4 Lemma. *Let $k \in \mathbb{N}_{>0}$ and consider Eq. (7.1) with $f \in C^{0,k}(I \times \mathbb{R}^n, \mathbb{R}^n)$, i.e. continuous in the first argument and k-times continuously differentiable in the second argument. Let φ denote its associated solution operator and $\Phi := \partial_2 \varphi$ be the linearization of φ. Then $\varphi \in C^{1,1,k}(I \times I \times \mathbb{R}^n, \mathbb{R}^n)$ and $\Phi \in C^{1,1,k-1}(I \times I \times \mathbb{R}^n, L(\mathbb{R}^n))$.*

Proof. By classical regularity results as presented in [6, Section 9], the associated solution operator φ is continuously differentiable on the first two arguments (the time variables) and k-times continuously differentiable in the last argument (the space variable). Now the linearization Φ is continuously differentiable in the time variables (see again [6]) and $k-1$-times continuously differentiable in the space variable. □

As a consequence of Lemma 7.4, for regularity considerations regarding σ it suffices to restrict to the maximum function applied to (some vectorized version of) the spectrum of the Cauchy-Green strain tensor C. We start with the special case of simple eigenvalues and present a classical result. We provide a proof for the convenience of the reader. For simplicity, we fix $t, s \in I$ in the next proposition and consider only the dependence on $x \in \mathbb{R}^n$.

7.5 Proposition. *Consider Eq. (7.1) with $f \in C^{0,3}(I \times \mathbb{R}^n, \mathbb{R}^n)$, let φ denote the associated solution operator and $\Phi := \partial_2 \varphi$ be its linearization. Let $t, s \in I$ and $x \in \mathbb{R}^n$ be such that the maximal eigenvalue $\lambda_n(x) := \lambda_n(t, s, x)$ of $C(x) := C(t, s, x)$ is simple. Denote by $v_n(x)$ a normalized eigenvector associated to $\lambda_n(x)$. Then exist a neighborhood U of x and twice continuously differentiable functions $\lambda\colon U \to \mathbb{R}_{>0}$ and $v\colon U \to \mathcal{S}$ such that:*

(a) $\lambda(x) = \lambda_n(x)$ and $v(x) = v_n(x)$;

(b) for any $y \in U$ one has $C(t, s, y)v(y) = \lambda(y)v(y)$.

Proof. Recall that we denote by $\lambda_i(x)$ the i-th repeated eigenvalue of $C(x)$ and by $v_i(x)$ an associated eigenvector. The assertion is shown if we can apply the Implicit Function Theorem. To this end, consider the function

$$f\colon \mathbb{R}^{n+1} \times \mathbb{R}^n \to \mathbb{R}^{n+1}, \qquad (v, \lambda, x) \mapsto \begin{pmatrix} (C(x) - \lambda \,\mathrm{id})v \\ v^\top v - 1 \end{pmatrix},$$

which is obviously C^2. By assumption we have that $f(v_n(x), \lambda_n(x), x) = 0$. Furthermore, we introduce the following block matrix

$$M := \begin{pmatrix} \partial_0 & \partial_1 \end{pmatrix} \otimes f(v_n(x), \lambda_n(x), x) = \begin{pmatrix} C(x) - \lambda_n(x)\,\mathrm{id} & -v_n(x) \\ 2v_n(x)^\top & 0 \end{pmatrix}.$$

Now, if $\det(M) \neq 0$ we are done. The condition can be checked with the following equation:

$$\det(M) = -2 \prod_{i=1}^{n-1} (\lambda_i(x) - \lambda_n(x)). \tag{7.5}$$

It can be derived as follows: it is well-known that the determinant of a matrix equals the product of its eigenvalues. One easily verifies that $\lambda_i(x) - \lambda_n(x)$, $i \in \{1, \ldots, n-1\}$, is an eigenvalue of M with eigenvector $\left(\begin{smallmatrix} v_i(x) \\ 0 \end{smallmatrix}\right)$. Next we derive the remaining two eigenvalues by making the ansatz for eigenvectors $\left(\begin{smallmatrix} v_n(x) \\ z \end{smallmatrix}\right) \in \mathbb{C}^{n+1}$ with $z \in \mathbb{C}$. Consider the linear system

$$M \begin{pmatrix} v_n(x) \\ z \end{pmatrix} = \begin{pmatrix} (C(x) - \lambda_n(x)\,\mathrm{id})v_n(x) - zv_n(x) \\ 2v_n(x)^\top v_n(x) \end{pmatrix} = \begin{pmatrix} -zv_n(x) \\ 2 \end{pmatrix} = \lambda \begin{pmatrix} v_n(x) \\ z \end{pmatrix},$$

from which $z = -\lambda$ and hence $2 = -\lambda^2$ with the two solutions $2\mathrm{i}$ and $-2\mathrm{i}$ and corresponding eigenvectors $\left(\begin{smallmatrix} v_n(x) \\ -2\mathrm{i} \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} v_n(x) \\ 2\mathrm{i} \end{smallmatrix}\right)$, respectively, follows. This proves Eq. (7.5), from which we read off that $\det(M) \neq 0$ if and only if $\lambda_n(x)$ is simple. □

Note that Proposition 7.5 naturally applies to any other eigenvalue and its associated eigenvector. Obstacles to the direct application of more general classical results as in [67, Section II-6.3] are the multi-parameter dependence and the finite order of differentiability of C. Observe that the spectrum $\sigma(C)$ cannot be represented globally by C^2 eigenvalue functions due to possible coincidence of eigenvalues. Consequently, $\max \sigma(C)$ is in general not the composition of two differentiable functions. At this point, the observation that $\max \circ\, \sigma \colon GL(n, \mathbb{R}) \subset L(\mathbb{R}^n) \to \mathbb{R}$ is a *spectral function* is very useful.

7.6 Remark. Let us briefly recall the notion of *spectral functions*. A function

$$F \colon \mathrm{Sym}_n \to \mathbb{R}$$

where Sym_n denotes the set of symmetric operators in $L(\mathbb{R}^n)$, is called *spectral* if it is invariant under orthogonal similarity transformations, i.e. for all $A \in \mathrm{Sym}_n$ and $U \in O(n)$ we have $F(A) = F(U^\top AU)$. Due to this invariance it is sufficient to define F on the diagonal matrices, since by the well-known spectral theorem for self-adjoint matrices every $A \in \mathrm{Sym}_n$ can be diagonalized by an orthogonal matrix. Thus, we can equivalently define F by some $f \colon \mathbb{R}^n \to \mathbb{R}$, acting on n-dimensional vectors which give rise to diagonal matrices via the diag-operator. In summary, any spectral function $F \colon \mathrm{Sym}_n \to \mathbb{R}$ can be equivalently decomposed as $F = f \circ \vec{\sigma}$. Here f is symmetric in the sense that it is invariant under permutations of the arguments and $\vec{\sigma}(\cdot)$ denotes some vectorization of the spectrum. We do not specify the vectorization rule, since this is not important and second could confuse the reader. For instance, if the spectrum is vectorized in decreasing order, then the max-function applied to $\vec{\sigma}$ becomes the projection on the first coordinate, which is linear and hence indefinitely differentiable. Points of non-differentiability are then hidden in the vectorization rule. Therefore, for regularity investigations a specification of the vectorization rule is misleading.

The following theorem states the regularity of the FTLE in the general case, i.e. without assumptions on the space dimension n and without assumptions on the vector field f except for necessary regularity. It states assertions on the regularity in [101, 74] more precisely.

7.7 Theorem (Regularity of the FTLE)**.** *Consider Eq.* (7.1) *with* $f \in C^{0,3}(I \times \mathbb{R}^n, \mathbb{R}^n)$*. Then for any* $t, s \in I$ *and* $x \in \mathbb{R}^n$ *the functions* $\lambda_n(\cdot, t, x), \lambda_n(t, \cdot, x) \colon I \to \mathbb{R}$ *are continuously differentiable except on an (at most) countable set and the function* $\lambda_n(t, s, \cdot) \colon \mathbb{R}^n \to \mathbb{R}$ *is twice continuously differentiable on an open and dense set.*

Proof. By Lemma 7.4 we have that $C \in C^{1,1,2}(I \times I \times \mathbb{R}^n)$. By Proposition 7.5 λ_n is (twice with respect to the last argument) continuously differentiable at $C(t, s, x)$, $t, s \in I$, $x \in \mathbb{R}^n$, if the maximal eigenvalue of $C(t, s, x)$ is simple. Thus, $\max \colon \mathbb{R}^n \to \mathbb{R}$ is (twice, respectively) continuously differentiable on the dense open set $\mathbb{R}^n \setminus M$, where

$$M := \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n; \bigvee_{\substack{i,j \in \{1,\dots,n\} \\ i \neq j}} x_i = x_j = \max\{x_1, \dots, x_n\} \right\},$$

is the closed nowhere dense critical set. In summary, the only obstruction for differentiability is the preimage N of M under the continuous function $\vec{\sigma} \circ C$, i.e. $N := (\vec{\sigma} \circ C)^{-1}[M] \subseteq I \times I \times \mathbb{R}^n$.

Concerning the first two arguments, fix $x \in \mathbb{R}^n$ and $s \in I$ and consider

$$(\vec{\sigma} \circ C(\cdot, s, x))^{-1}[M], (\vec{\sigma} \circ C(s, \cdot, x))^{-1}[M] \subset I.$$

We discuss only the first case since the second is completely analogous. The only critical points in $(\vec{\sigma} \circ C(\cdot, s, x))^{-1}[M]$ for differentiability are those where the two top eigenvalues cross transversally, i.e. at a nonvanishing angle. We show that the critical points occur at most finitely often in any compact interval $J \subseteq I$. First, Theorem II-6.8 in [67] ensures the existence of n continuously differentiable functions $\mu_1, \dots, \mu_n \colon I \to \mathbb{R}_{>0}$ that represent the repeated (unsorted) eigenvalues of $C(\cdot, s, x)$. Next, we show that the set of points t where some pair of eigenvalue functions, not necessarily the two top ones, intersect transversally, is finite on any compact $J \subseteq I$. Since this set is a superset of $(\vec{\sigma} \circ C(\cdot, s, x))^{-1}[M]$ we will be done. Without loss of generality consider $h_{12} := \mu_1 - \mu_2$, then h_{12} is continuously differentiable. A transversal crossing of μ_1 and μ_2 then corresponds to roots t of h_{12}, i.e. $h_{12}(t) = 0$, such that $h_{12}'(t) \neq 0$. We call such t *transversal roots* for short. Due to the local monotonicity of h_{12} at some transversal root t the function h_{12} is locally injective and hence every transversal root is isolated from the others. Hence, the set of transversal roots of h_{12} is a discrete and closed subset of the compact set J, thus it is finite. Analogously we can proceed with any other combination h_{ij}, $i, j \in \{1, \dots, n\}$, $i \neq j$. Since there are only finitely many such combinations, the union over all transversal roots of all combinations is finite on any compact $J \subseteq I$.

Concerning the last argument, fix $t, s \in I$ and consider $(\vec{\sigma} \circ C(t, s, \cdot))^{-1}[M]$. As before, we discuss different combinations of possible crossings separately. Note that the above argument with transversal crossings does not apply here directly due to the multi-parameter dependence and the absence of differentiable eigenvalue functions. We cover M with the closed nowhere dense sets

$$M_{ij} := \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n;\ x_i = x_j = \max\{x_1, \dots, x_n\} \right\},$$

where $i, j \in \{1, \dots, n\}$, $i \neq j$. Now define the closed sets

$$N_{ij} := (\lambda \circ C(t, s, \cdot))^{-1}[M_{ij}] \subseteq \mathbb{R}^n$$

for $i, j \in \{1, \dots, n\}$, $i \neq j$, such that

$$N = \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n} N_{ij} \subseteq \mathbb{R}^n.$$

The analogues to the transversal crossings in the 1-parametric case above are exactly those points that lie on the boundary of some N_{ij}. This is for the following reason: choose some $x \in \operatorname{int} N_{ij}$, if it exists, that does not belong to the boundary of some other N_{km}. Then there exists some neighborhood of x on which the two top eigenvalues μ_i and μ_j coincide and the maximum is locally uniquely determined by μ_i. The boundary of the closed sets N_{ij} coincides with the boundary of their respective open complement $\mathbb{R}^n \setminus N_{ij}$, hence ∂N_{ij} as the boundary of an open set is closed and nowhere dense. So is the finite union of ∂N_{ij} with $i, j \in \{1, \dots, n\}$, $i \neq j$. Thus, its complement

$$\mathbb{R}^n \setminus \bigcap_{\substack{i,j\in\{1,\dots,n\}\\ i\neq j}} \partial N_{ij}$$

is an open dense set in $\mathbb{R}^n$. $\square$

7.8 Remark. The possibility of crossing eigenvalues can be excluded by considering special problem classes. For instance, in [101], where finite-time ODEs in domains of $\mathbb{R}^2$ are considered, the separation of the eigenvalues is assumed directly. Another such problem class consists of ODEs with divergence-free vector fields in $\mathbb{R}^2$, or similarly, hyperbolic planar Hamiltonian systems; see e.g. [57].

7.2 Hyperbolic Lagrangian Coherent Structures

In this section, we resume the variational approach to hyperbolic Lagrangian Coherent Structures (LCSs) presented in [56] and extend it towards hyperbolic LCSs of higher codimensions. The extension is rather motivated by its mathematical possibility than by physical necessity. So far, to the best of the author's knowledge, the only indications for a potential interest in such structures are quoted in the beginning of this chapter.

In the following, we consider a C^3-process $\varphi \in \mathcal{P}(\mathbb{I}, \mathbb{R}^n)$ on $\mathbb{I} = \{t_-, t_+\}$, which we identify with the orientation-preserving C^3-diffeomorphism

$$\varphi(t_+, t_-, \cdot)\colon \mathbb{R}^n \to \mathbb{R}^n,$$

again denoted by φ. The diffeomorphism φ can be thought of as the restriction of the solution operator of Eq. (7.1) to $\{t_+\} \times \{t_-\} \times \mathbb{R}^n$, where $f \in C^{0,3}(I \times \mathbb{R}^n, \mathbb{R}^n)$. As before, we have

$$\Phi := \partial\varphi \in C^2(\mathbb{R}^n, GL(n, \mathbb{R})) \quad \text{and} \quad C := \Phi^*\Phi \in C^2(\mathbb{R}^n, \mathrm{Sym}_n \cap GL(n, \mathbb{R}))$$

with the pointwise inverse function

$$C^{-1} := \left(x \mapsto (C(x))^{-1}\right) \in C^2(\mathbb{R}^n, \mathrm{Sym}_n \cap GL(n, \mathbb{R})).$$

Throughout this section, $\cdot^*$ and $\cdot^{-1}$ are meant pointwise whenever there is no argument. Since adjunction and inversion commute, we abbreviate the composition of both operations by $\cdot^{-*}$. In this sense, we have $C^{-1} = \Phi^{-1}\Phi^{-*}$. Recall that we parametrize the spectrum of C by the continuous functions $0 < \lambda_1 \leq \ldots \leq \lambda_n$ and an orthonormal eigenbasis of C by $v_1, \ldots, v_n$, where v_i, $i \in \{1, \ldots, n\}$, is an eigenvector to the eigenvalue λ_i. By Proposition 7.5 λ_i and v_i can be chosen locally twice continuously differentiable if λ_i is simple.

7.2.1 Hyperbolic LCSs of arbitrary dimension

In the following, $\widetilde{\mathcal{M}}$ denotes an orientable k-dimensional embedded differentiable submanifold, $k \in \{1, \ldots, n-1\}$, and $\mathcal{M} \subseteq \widetilde{\mathcal{M}} \subset \mathbb{R}^n$ denotes a compact, connected subset of $\widetilde{\mathcal{M}}$ with boundary and non-empty interior relative to $\widetilde{\mathcal{M}}$. We endow $\mathbb{R}^n$ with its standard (Riemannian) metric. The interior of $\mathcal{M}$ relative to $\widetilde{\mathcal{M}}$ is a submanifold itself and at these interior points notions like tangent space and normal space are well-defined and coincide with the notions defined with respect to $\widetilde{\mathcal{M}}$. For points on the boundary of $\mathcal{M}$ relative to $\widetilde{\mathcal{M}}$, when we refer to its tangent space and normal space, we define these notions with respect to $\widetilde{\mathcal{M}}$. In this sense, we consider the tangent bundle and the normal bundle to be well-defined on (interior and boundary of) $\mathcal{M}$.

We denote by $T\mathcal{M}$ and $T^{\perp}\mathcal{M}$ the tangent and the normal bundle, respectively, and by $T_x\mathcal{M}$ and $T_x^{\perp}\mathcal{M}$ the tangent and the normal space at $x \in \mathcal{M}$, respectively, i.e. for any $x \in \mathcal{M}$ we have $\mathbb{R}^n = T_x\mathcal{M} \oplus T_x^{\perp}\mathcal{M}$. Due to the orientability and the differentiability assumption on $\mathcal{M}$ the tangent and the normal bundle are continuously differentiable vector bundles of rank k and $n-k$, respectively. That means, the mappings $\mathcal{M} \ni x \mapsto T_x\mathcal{M} \in \mathrm{Gr}(k, \mathbb{R}^n)$ and $\mathcal{M} \ni x \mapsto T_x^{\perp}\mathcal{M} \in \mathrm{Gr}(n-k, \mathbb{R}^n)$ are continuously differentiable.

Since φ is a diffeomorphism we have that $\varphi[\mathcal{M}] \subset \mathbb{R}^n$ is a C^1-submanifold. It is well-known that $\Phi = \partial\varphi$ is a vector bundle isomorphism and we have that $\Phi[T_x\mathcal{M}] = T_{\varphi(x)}\varphi[\mathcal{M}]$. Due to the following calculation, we find that Φ^{-*} is a vector bundle isomorphism between the respective normal bundles: consider a tangent vector $e \in T_x\mathcal{M}$ and a normal vector $n \in T_x^{\perp}\mathcal{M}$, then we have

$$\left\langle \Phi(x)e, \Phi(x)^{-*}n \right\rangle = \langle e, n \rangle = 0.$$

7.9 Remark. From a more abstract point of view it is not surprising that Φ^{-*} is the vector bundle isomorphism between the normal spaces of $\mathcal{M}$ and $\varphi[\mathcal{M}]$. The (pointwise) adjoint operator of Φ, Φ considered as a linear relation in $\mathbb{R}^n \times \mathbb{R}^n$, is defined as $\Phi^* = -(\Phi^{-1})^{\perp}$. Hence, we observe that $\Phi^{-*} = -\Phi^{\perp}$, saying that up to the minus sign Φ^{-*} is the orthogonal subspace in the direct sum of the Hilbert spaces $\mathbb{R}^n$ and $\mathbb{R}^n$. The inner product in $\mathbb{R}^n \oplus \mathbb{R}^n$ is just the sum of the inner products with respect to the arguments and the images, respectively, i.e.

$$\langle (u,v),(x,y)\rangle_{\mathbb{R}^n\oplus\mathbb{R}^n} := \langle u,x\rangle_{\mathbb{R}^n} + \langle v,y\rangle_{\mathbb{R}^n}.$$

Now, $\Phi^{-*} = -\Phi^{\perp}$ says that for any $u, x \in \mathbb{R}^n$ the image $-\Phi^{-*}u$ satisfies

$$\left\langle (u,-\Phi^{-*}u),(x,\Phi x)\right\rangle_{\mathbb{R}^n\oplus\mathbb{R}^n} = \langle u,x\rangle_{\mathbb{R}^n} - \left\langle \Phi^{-*}u,\Phi x\right\rangle_{\mathbb{R}^n} = 0,$$

or equivalently

$$\langle u,x\rangle_{\mathbb{R}^n} = \left\langle \Phi^{-*}u,\Phi x\right\rangle_{\mathbb{R}^n}.$$

Roughly speaking, Φ^{-*} preserves orthogonality with respect to Φ. Thus, we get $\Phi(x)^{-*}[T_x^{\perp}M] \perp \Phi(x)[T_xM] = T_{\varphi(x)}\varphi[\mathcal{M}]$ and consequently $\Phi(x)^{-*}[T_x^{\perp}M] = T_{\varphi(x)}^{\perp}\varphi[\mathcal{M}]$.

One of the essential ideas in [56] is to introduce hyperbolic LCSs by a description and a comparison of growth of normal and tangential perturbations, respectively. Taking the definitions and representations of *repulsion rate* and *repulsion ratio* in [56] as a motivation, we generalize these notions in a straightforward manner to our context. Note that these notions have an asymptotic predecessor in the *generalized Lyapunov type numbers* in [41].

7.10 Definition (Repulsion rate, repulsion ratio)**.** We define

$$\rho\colon \mathcal{M}\to\mathbb{R},\quad x\mapsto \min_{n\in T_x^{\perp}\mathcal{M}\cap\mathcal{S}} \left\langle n, C(x)^{-1}n\right\rangle^{-1/2} = \left\|\Phi(x)^{-*}|_{T_x^{\perp}\mathcal{M}}\right\|^{-1},$$
$$\nu\colon \mathcal{M}\to\mathbb{R},\quad x\mapsto \frac{\rho(x)}{\max_{v\in T_x\mathcal{M}\cap\mathcal{S}}|\Phi(x)v|} = \left\|\Phi(x)|_{T_x\mathcal{M}}\right\|^{-1}\left\|\Phi(x)^{-*}|_{T_x^{\perp}\mathcal{M}}\right\|^{-1},$$

and call ρ *repulsion rate* and ν *repulsion ratio*.

7.11 Remark. (i) Note that the last definition is well-defined since Φ is invertible. To see that the representation of the repulsion rate in terms of the operator norm of the restriction of $\Phi(x)^{-*}$ holds, observe the following calculation: let $T \in GL(n,\mathbb{R})$, $X \in \mathrm{Gr}(k,\mathbb{R}^n)$ be a k-dimensional subspace and $V \in L(\mathbb{R}^k,\mathbb{R}^n)$ with $V^*V = \mathrm{id}_{\mathbb{R}^k}$ and $V[\mathbb{R}^k] = X$. We consider $\mathbb{R}^n = \mathbb{R}^k \oplus \mathbb{R}^{n-k}$ as well as $\mathbb{R}^n = X \oplus X^{\perp}$. We extend V to an operator in $\mathbb{R}^n$, which we again denote by V, by mapping $\mathbb{R}^k$ to V and $\mathbb{R}^{n-k}$ to $X^{\perp}$, such that $V^*V = \mathrm{id}_{\mathbb{R}^n}$. In particular, we have

$V^*|_{\operatorname{im} V} V = \mathrm{id}_{\mathbb{R}^k}$ by construction and both V and V^* are invertible. Now, we calculate

$$\begin{aligned}
\min_{v \in X \cap \mathcal{S}} |Tv| = \min_{v \in X \cap \mathcal{S}} \langle Tv, Tv \rangle^{1/2} &= \min_{x \in (\mathbb{R}^k \oplus \{0\}) \cap \mathcal{S}} \langle TVx, TVx \rangle^{1/2} \\
&= \min_{x \in (\mathbb{R}^k \oplus \{0\}) \cap \mathcal{S}} \langle x, V^* T^* T V x \rangle^{1/2} \\
&= \frac{1}{\max_{x \in (\mathbb{R}^k \oplus \{0\}) \cap \mathcal{S}} \langle x, (V^* T^* T V)^{-1} x \rangle^{1/2}} \\
&= \frac{1}{\max_{x \in (\mathbb{R}^k \oplus \{0\}) \cap \mathcal{S}} \langle x, V^* (T^* T)^{-1} V x \rangle^{1/2}} \\
&= \frac{1}{\max_{x \in (\mathbb{R}^k \oplus \{0\}) \cap \mathcal{S}} \left\langle \left(T^{-1}\right)^* Vx, \left(T^{-1}\right)^* Vx \right\rangle^{1/2}} = \frac{1}{\left\| \left(T^{-1}\right)^* |_X \right\|},
\end{aligned}$$

where the equality of the second and the third line is due to the self-adjointness of $V^* T^* T V$ and the spectral theorem.

(ii) Observe the double parameter-dependence in Definition 7.10: the operators themselves depend on x as well as the subspaces to which the operators are restricted. The operator norm of the restriction can be reduced to the operator norm $\|\Phi(y) V(y)\|$, see also (i), where $V(\cdot)$ is C^1 by the smoothness assumption on $\mathcal{M}$. From this we directly read off the continuity of the repulsion rate and the growth ratio.

7.12 Definition (cf. [56, Def. 3]). $\mathcal{M}$ is called *normally repelling* if there exists $c \in \mathbb{R}_{>1}$ such that $\rho[\mathcal{M}] \geq c$ and $\nu[\mathcal{M}] \geq c$ hold.

7.13 Remark. In this work, we deal explicitly with *normally repelling* manifolds only. As is done in [56], we define the *normally attracting* counterparts as those which are normally repelling "in backward time". That aims at identifying submanifolds $\mathcal{N}$, considered as subsets of the t_+-fiber of the extended state space, which are normally repelling under φ^{-1}. To simplify the presentation, we do not consider $\mathcal{M}$ together with $\varphi[\mathcal{M}]$ (or $\mathcal{N}$ together with $\varphi^{-1}[\mathcal{N}]$) as subsets in the extended state space $\{t_-, t_+\} \times \mathbb{R}^n$. Instead, we work only in the initial time-fibers, i.e. $\{t_-\} \times \mathbb{R}^n$ for φ and $\{t_+\} \times \mathbb{R}^n$ for φ^{-1}. In this sense, we will sometimes loosely refer to a normally repelling or a normally attracting manifold as *normally hyperbolic*, and it is clear from the particular characteristic if the manifold is situated in the earlier or in the latter time-fiber.

7.14 Lemma. *For any $x \in \mathcal{M}$ the following estimates hold:*

$$\sqrt{\lambda_1(x)} \leq \rho(x) \leq \sqrt{\lambda_{k+1}(x)} \qquad \textit{and} \qquad \nu(x) \leq \sqrt{\frac{\lambda_{k+1}(x)}{\lambda_k(x)}}.$$

Proof. First, we observe that the ordered singular values of Φ and Φ^{-*} are

$$0 < \sqrt{\lambda_1} \leq \ldots \leq \sqrt{\lambda_n} \qquad \text{and} \qquad 0 < \sqrt{\frac{1}{\lambda_n}} \leq \ldots \leq \sqrt{\frac{1}{\lambda_1}},$$

respectively. The estimates now follow from the Courant-Fischer Minimax Theorem [112, Theorem 8.9], or more precisely from the following estimates: for $X \in \mathrm{Gr}(k, \mathbb{R}^n)$ and $Y \in \mathrm{Gr}(n-k, \mathbb{R}^n)$ we have

$$\left\|\Phi|_X\right\| \geq \sqrt{\lambda_k}, \qquad \qquad \left\|\Phi^{-*}|_Y\right\| \geq \frac{1}{\sqrt{\lambda_{k+1}}}. \qquad \square$$

In the following, we want to specify which perturbations of $\mathcal{M}$ we are going to consider. Since $\mathcal{M}$ is a k-dimensional Riemannian submanifold of $\mathbb{R}^n$, we can parametrize $\mathcal{M}$ locally in normal coordinates. Let $y \in \mathcal{M}$, $U \subset \mathbb{R}^k$ a neighborhood of 0 and $F_y \colon U \subseteq \mathbb{R}^k \to \mathcal{M} \subset \mathbb{R}^n$ be a local normal parametrization around $y = F(0)$, then with the definition $e_{i+1}(y) := \partial_i F(0)$, $i \in \{0, \ldots, k-1\}$, the set $\{e_1(y), \ldots, e_k(y)\}$ is an orthonormal basis of $T_y\mathcal{M}$. We extend this basis by $\{e_{k+1}(y), \ldots, e_n(y)\}$, an orthonormal basis of $T_y^\perp\mathcal{M}$, which is orthonormal to the basis of $T_y\mathcal{M}$ in the tangent space of the ambient space $T_y\mathbb{R}^n = \mathbb{R}^n$. A *family of C^1-normal perturbations of $\mathcal{M}$* is a (linearly) parametrized family of C^1-submanifolds $\mathcal{M}_\varepsilon$, $\varepsilon \in [-\theta, \theta]$, $\theta \in \mathbb{R}_{>0}$, $\mathcal{M}_0 = \mathcal{M}$, such that for each $y \in \mathcal{M}$ there exists a neighborhood of 0 in $\mathbb{R}^k$ and a bounded continuously differentiable function $\alpha \colon U \subseteq \mathbb{R}^k \to \mathbb{R}^{n-k}$ called *perturbation profile* such that locally the submanifolds can be parametrized by

$$x_y \colon (-\theta, \theta) \times U \subset \mathbb{R}^{k+1} \to \mathbb{R}^n,$$

$$\begin{aligned}(\varepsilon, u) \mapsto x_y(\varepsilon, u) &= F(u) + \varepsilon \begin{pmatrix} | & & | \\ e_{k+1}(F(u)) & \cdots & e_n(F(u)) \\ | & & | \end{pmatrix} \alpha(u) \\ &= F(u) + \varepsilon N(F(u))\alpha(u) \in \mathcal{M}_\varepsilon,\end{aligned} \tag{7.6}$$

where $N(F(u)) := \begin{pmatrix} e_{k+1}(F(u)) & \cdots & e_n(F(u)) \end{pmatrix}$ and $\partial F(u)^* N(F(u)) = 0$. Furthermore, observe that the parametrization depends linearly on the perturbation parameter, i.e. in particular $\varepsilon \mapsto x_y(\varepsilon, 0)$ is C^∞. In this sense, it would be more precise to speak of a C^∞-family of C^1-normal perturbations of $\mathcal{M}$. Note also that for the case $k = n-1$ exactly this specific type of perturbations is studied in [56].

7.15 Remark. This type of perturbation can be considered as the linearization (with respect to ε) of an arbitrary C^1 normal perturbation, and hence as sufficiently general. Here α should depend in a C^2 manner on ε.

For notational convenience, we will omit the dependence of the local parametrization on y and denote by x an associated local normal parametrization under consideration; recall that with this notation one has $x(0,0) = y$.

In our next definition we are going to accentuate those normally repelling manifolds $\mathcal{M}$ that satisfy a certain optimality condition of the repulsion rate with respect to all families of C^1-normal perturbations of itself. Namely, we will require that for any $y \in \mathcal{M}$ and the associated perturbation parametrization x one has that $\rho \circ x(\cdot, 0) = \big(\varepsilon \mapsto \rho(x(\varepsilon, 0))\big)$ has a local (nondegenerate) maximum at $\varepsilon = 0$ for arbitrary admissible perturbation profiles.

7.16 Definition (Repelling WLCS and LCS (in the broad sense))**.** Assume $\mathcal{M}$ is normally repelling. Then $\mathcal{M}$ is called a *(k-dimensional) repelling Weak LCS (WLCS) in the broad sense* if for any family of C^1-normal perturbations of $\mathcal{M}$ and any $y \in \mathcal{M}$ we have that there exists a neighborhood $U \subseteq [-\theta, \theta]$ of 0 such that for any $\varepsilon \in U$ we have $\rho(x(\varepsilon, 0)) \leq \rho(x(0,0))$. $\mathcal{M}$ is called a *(k-dimensional) repelling LCS in the broad sense* if for any family of C^1-normal perturbations of $\mathcal{M}$ and any $y \in \mathcal{M}$ we have that there exists a neighborhood $U \subseteq [-\theta, \theta]$ of 0 such that for any $\varepsilon \in U \setminus \{0\}$ we have $\rho(x(\varepsilon, 0)) < \rho(x(0,0))$.

This can be considered as a generalization of the physical definition of hyperbolic LCSs in [56]. Note that the above definition does not incorporate any derivatives but expresses formally the role LCSs should play, as "locally the strongest repelling or attracting" manifolds, see [56, Definition 1]. As with normal attraction, we define *attracting WLCSs/LCSs* as repelling WLCSs/LCSs in backward time and Remark 7.13 applies analogously. Furthermore, *hyperbolic WLCSs/LCSs* are attracting or repelling WLCSs/LCSs.

To strengthen Definition 7.16 towards a definition that leads to computable mathematical criteria, we discuss the regularity of the repulsion rate function with respect to the perturbation parameter first. To this end, for some $y \in \mathcal{M}$ we consider the function

$$\rho \circ x(\cdot, 0)\colon [-\theta, \theta] \to \mathbb{R},$$
$$\varepsilon \mapsto \left\| \Phi(x(\varepsilon,0))^{-*}|_{T_x^\perp \mathcal{M}_\varepsilon} \right\|^{-1} = \left(\max \sigma(V^*(\varepsilon) C(x(\varepsilon,0))^{-1} V(\varepsilon))\right)^{-1/2},$$

where $V\colon (-\theta, \theta) \to L(\mathbb{R}^{n-k}, \mathbb{R}^n)$ is C^∞ with $V^*V = \mathrm{id}_{\mathbb{R}^{n-k}}$ and $V(\varepsilon)[\mathbb{R}^{n-k}] = T^\perp_{x(\varepsilon,0)} \mathcal{M}_\varepsilon$. As in Proposition 7.5 (note that ε is a scalar variable here), we have that $\rho \circ x(\cdot, 0)$ is continuous and piecewise twice continuously differentiable. In the general case, it can happen that $\rho \circ x(\cdot, 0)$ is not differentiable in 0. However, in this case it is continuously differentiable in a punctured neighborhood of 0. In the following we restrict our considerations to *well-behaved* manifolds $\mathcal{M}$ in the sense that $\rho \circ x(\cdot, 0)$ is twice continuously differentiable at 0 for any $y \in \mathcal{M}$ and for any local perturbation profile α. Note that normally hyperbolic hypersurfaces $\mathcal{M}$ are always well-behaved due to the separation of the two largest eigenvalues as a consequence of Lemma 7.14 and the twice continuous differentiability of ρ. For the general case, a sufficient criterion for well-behavior is the fact that in any $y \in \mathcal{M}$ we have that the maximal eigenvalue of $V^*C(y)^{-1}V$ is simple, where V is the respective subspace spanning matrix.

7.17 Definition (Repelling WLCS and LCS (in the strict sense))**.** Assume $\mathcal{M}$ is well-behaved and normally repelling. Then $\mathcal{M}$ is called a *(k-dimensional) repelling Weak LCS* if for any family of C^1-normal perturbations of $\mathcal{M}$ and any $y \in \mathcal{M}$ we have $\rho(x(\cdot, 0))'(0) = 0$. $\mathcal{M}$ is called a *(k-dimensional) repelling LCS* if $\mathcal{M}$ is a repelling Weak LCS and additionally $\rho(x(\cdot, 0))''(0) < 0$ for any family of C^1-normal perturbations of $\mathcal{M}$ and any $y \in \mathcal{M}$.

The next theorem states necessary and sufficient conditions for $\mathcal{M}$ to be a repelling Weak LCS.

7.18 Theorem. *Let* $\dim \mathcal{M} = k$ *and* $\lambda_{k+1}\colon \mathbb{R}^n \to \mathbb{R}_{>0}$ *be continuously differentiable in all points* $y \in \mathcal{M}$*. Then* $\mathcal{M}$ *is a repelling WLCS if and only if for any* $y \in \mathcal{M}$ *the following conditions hold:*

(i) $\lambda_k(y) \neq \lambda_{k+1}(y) > 1$;

(ii) $\operatorname{span}\{v_{k+1}(y), \dots, v_n(y)\} = T_y^\perp \mathcal{M}$;

(iii) $\partial_{v_i(y)} \lambda_{k+1}(y) = 0$ *for any* $i \in \{k+1, \dots, n\}$.

Proof. Our proof is essentially an adaptation of the one in [56, Theorem 7] to the general case, sometimes revealing the underlying nature of the arguments.

"⇒": By the assumptions in Definition 7.12 there exists $c \in \mathbb{R}_{>1}$ such that with Lemma 7.14 we obtain

$$1 < c \leq \rho(y) \leq \sqrt{\lambda_{k+1}(y)} \qquad \text{and} \qquad 1 < c \leq \nu(y) \leq \frac{\rho(y)}{\sqrt{\lambda_k(y)}} \tag{7.7}$$

for all $y \in \mathcal{M}$, from which (i) follows. In order to compute the (derivatives of the) repulsion rate for points in $\mathcal{M}_\varepsilon$ (in $\mathcal{M}$, respectively) we need to derive a first order approximation of the normal space $T^\perp_{x(\varepsilon,0)}\mathcal{M}_\varepsilon$ in $x(\varepsilon, 0)$. As before, we denote the reference point $F(0) = x(0, 0)$ by y. First, we derive a representation for the tangent space $T_{x(\varepsilon,0)}\mathcal{M}_\varepsilon$. To this end, we differentiate its parametrization Eq. (7.6) in the coordinate point 0 and obtain

$$\begin{aligned} \partial x(\varepsilon,\cdot)(0) &= \partial F(0) + \varepsilon\big(x \mapsto (N \circ F)(x)\alpha(x)\big)'(0) \\ &= \partial F(0) + \varepsilon\left(\partial(N \circ F)(0)(\alpha(0)) + N(y)\partial\alpha(0)\right) \in \mathbb{R}^{n\times k}. \end{aligned} \tag{7.8}$$

By the C^∞-differentiability assumption on the perturbation parameter we know that the normal space $T^\perp_{x(\varepsilon,0)}\mathcal{M}_\varepsilon$ in $x(\varepsilon, 0)$, spanned by $N_\varepsilon(x(\varepsilon, 0)) \in \mathbb{R}^{n\times(n-k)}$, or more precisely on the coordinate point 0, is C^∞ with respect to ε and hence has a local representation as

$$N_\varepsilon(x(\varepsilon,0)) = N(y) + \varepsilon K + \varepsilon^2 L + \mathcal{O}(\varepsilon^3), \tag{7.9}$$

where $N(y) \in \mathbb{R}^{n\times(n-k)}$ spans $T^\perp_y\mathcal{M}$ and $K \in \mathbb{R}^{n\times(n-k)}$ needs to be determined; we compute L in the special case of $\dim \mathcal{M} = n - 1$ in the next section. This is done by plugging Eqs. (7.8) and (7.9) into the characterizing equalities

$$N_\varepsilon(x(\varepsilon,0))^\top N_\varepsilon(x(\varepsilon,0)) = I_{n-k}, \qquad \text{and} \qquad N_\varepsilon(x(\varepsilon,0))_\varepsilon^\top \partial x(\varepsilon,\cdot)(0) = 0.$$

These equalities determine the subspace uniquely. However, from the theory of the Grassmann manifold it is clear that this system of equations is solved by an equivalence class of matrices, generated by one representative and $(n - k) \times (n - k)$-dimensional invertible matrices, multiplied from the right. Expanding the two characterizations and comparing coefficients of powers of ε we obtain $K = -\partial F(0)\partial\alpha(0)^*$. In the following argument, only the image of K plays a role, see Eq. (a) below. Obviously, right-multiplication with invertible matrices does not change the image of K. Hence, there are no problems with non-uniqueness of the representation of K.

Next we show that the fact that $\rho(x(\cdot, 0))'(0) = 0$ for any local perturbation profile α is equivalent to (ii) and (iii). To this end, we compute the derivative explicitly, making use of our well-behavior assumption in the following way: we represent

$$\rho(x(\varepsilon,0)) = \left\langle N_\varepsilon(x(\varepsilon,0))v(\varepsilon), C(x(\varepsilon,0))^{-1}N_\varepsilon(x(\varepsilon,0))v(\varepsilon)\right\rangle^{-1/2}, \tag{7.10}$$

where $v \in C^2([-\theta,\theta],\mathcal{S})$ is a twice continuously differentiable function with values on the unit circle in $\mathbb{R}^{n-k}$. For each $\varepsilon \in [-\theta,\theta]$ we have that $v(\varepsilon)$ is a normalized eigenvector of $N_\varepsilon(x(\varepsilon,0))^*C(x(\varepsilon,0))^{-1}N_\varepsilon(x(\varepsilon,0))$, corresponding to its maximal eigenvalue, i.e. in particular for $\mu = \frac{1}{\rho(y)^2} = \|\Phi(y)^{-*}N(y)\|$ holds

$$N(y)^*C(y)^{-1}N(y)v(0) = \mu v(0). \tag{7.11}$$

Due to the normalization for all $\varepsilon \in [-\theta,\theta]$ we have that $\langle v'(\varepsilon), v(\varepsilon)\rangle = 0$ and consequently, by the eigenvector property

$$\left\langle N_\varepsilon(x(\varepsilon,0))v'(\varepsilon), C(x(\varepsilon,0))^{-1}N_\varepsilon(x(\varepsilon,0))v(\varepsilon)\right\rangle = 0.$$

With these observations at hand, we calculate

$$\begin{aligned}
0 &= \rho(x(\cdot,0))'(0)\\
&= \left(\varepsilon \mapsto \left\langle N_\varepsilon(x(\varepsilon,0))v(\varepsilon), C(x(\varepsilon,0))^{-1}N_\varepsilon(x(\varepsilon,0))v(\varepsilon)\right\rangle^{-1/2}\right)'(0)\\
&= -\frac{1}{2}\rho(y)^3\left(\left\langle Kv(0)+N(y)v'(0), C(y)^{-1}N(y)v(0)\right\rangle + \right.\\
&\qquad + \left\langle N(y)v(0), (C^{-1}\circ x(\cdot,0))'(0)N(y)v(0)\right\rangle +\\
&\qquad \left. + \left\langle N(y)v(0), C(y)^{-1}(Kv(0)+N(y)v'(0))\right\rangle\right).
\end{aligned}$$

We compute

$$(C^{-1}\circ x(\cdot,0))'(0) = (\partial C^{-1})(y)x(\cdot,0)'(0) = \partial C^{-1}(y)N(y)\alpha(0),$$

and together with the symmetry of C^{-1} and of the inner product we obtain

$$\begin{aligned}
0 = -\frac{1}{2}\rho(y)^3\Big(&-2\left\langle \partial F(0)\partial\alpha(0)^*v(0), C(y)^{-1}N(y)v(0)\right\rangle +\\
&+ \left\langle N(y)v(0), \left(\partial C^{-1}(y)(N(y)\alpha(0))\right)N(y)v(0)\right\rangle\Big).
\end{aligned} \tag{7.12}$$

Since $\rho(y) > 0$, Eq. (7.12) holds for any continuously differentiable function α, or more precisely for any $\alpha(0) \in \mathbb{R}^{n-k}$ and $\partial\alpha(0) \in L(\mathbb{R}^k,\mathbb{R}^{n-k})$, if and only if the following hold:

$$C(y)^{-1}N(y)v(0) \in \partial F(0)[\mathbb{R}^k]^\perp = T_y^\perp\mathcal{M}, \tag{a}$$

$$\left\langle N(y)v(0), \left(\partial C^{-1}(y)\left[T_y^\perp\mathcal{M}\right]\right)N(y)v(0)\right\rangle = 0. \tag{b}$$

Additionally, we have

$$N(y)v(0) \in T_y^{\perp}\mathcal{M}. \tag{c}$$

From Eqs. (c) and (a) we see that for $\tau \in \mathbb{R}$

$$C(y)^{-1}N(y)v(0) - \tau N(y)v(0) \in T_y^{\perp}\mathcal{M},$$

and from Eq. (7.11) we conclude that for $\tau = \mu = \frac{1}{\rho(y)^2}$

$$C(y)^{-1}N(y)v(0) - \mu N(y)v(0) \in \ker N(y)^* = (\operatorname{im} N(y))^{\perp} = T_y\mathcal{M}.$$

Here, the first equality holds by a well-known consequence of the Projection Theorem. In summary, we obtain that $N(y)v(0)$ is an eigenvector of $C(y)^{-1}$ corresponding to the eigenvalue $\mu = \frac{1}{\rho(y)^2}$ and from Eq. (7.7) we see that $\mu = \frac{1}{\lambda_{k+1}(y)}$ and $N(y)v(0) = v_{k+1}(y)$. Since $\lambda_k(y) < \lambda_{k+1}(y)$ there exists only one subspace V such that the operator norm of Φ^{-*} restricted to V attains μ. Thus, we obtain that

$$N(y)v(0) = v_{k+1}(y) \quad \text{and} \quad T_y^{\perp}\mathcal{M} = \operatorname{span}\left\{v_{k+1}(y), \dots, v_n(y)\right\}.$$

Note that the conclusion that $N(y)v(0)$ is an eigenvector is simpler in the codimension-1 case; cf. [56]. Conversely, if $T_y^{\perp}\mathcal{M} = \operatorname{span}\{v_{k+1}(y), \dots, v_n(y)\}$ then (a) is obvious. As a consequence, we have that $T_y\mathcal{M} = \operatorname{span}\{v_1(y), \dots, v_k(y)\}$ and without loss of generality we may assume that $e_i(y) = v_i(y)$, $i \in \{1, \dots, n\}$.

To see (iii) we manipulate the resulting eigenvalue problem as follows. We consider

$$C(y)^{-1}v_{k+1}(y) = \frac{1}{\lambda_{k+1}(y)}v_{k+1}(y),$$

and differentiation at y and application to $v_i(y)$, $i \in \{k+1, \dots, n\}$, yields

$$\begin{aligned}\left(\partial C^{-1}(y)v_i(y)\right)v_{k+1}(y) + C(y)^{-1}(\partial v_{k+1}(y))v_i(y) &= \\ = -\frac{1}{\lambda_{k+1}(y)^2}v_{k+1}(y)\partial\lambda_{k+1}(y)v_i(y) &+ \frac{1}{\lambda_{k+1}(y)}\partial v_{k+1}(y)v_i(y).\end{aligned}$$

By composing the operator with $v_{k+1}(y)^*$ and making use of its eigenvector property, of $v_{k+1}(y)^* \partial v_{k+1}(y) = 0$ and Eq. (b) we obtain

$$\begin{aligned} 0 &= \left\langle v_{k+1}(y), \left(\partial C^{-1}(y) v_i(y)\right) v_{k+1}(y) \right\rangle \\ &= v_{k+1}(y)^* \left(\partial C^{-1}(y) v_i(y)\right) v_{k+1}(y) \\ &= -\frac{1}{\lambda_{k+1}(y)^2} (\partial \lambda_{k+1}(y)) v_i(y) \\ &= -\frac{1}{\lambda_{k+1}(y)^2} \partial_{v_i(y)} \lambda_{k+1}(y). \end{aligned}$$

Due to linearity of $\partial C^{-1}(y)$ and the positivity of λ_{k+1} we have that Eq. (b) is satisfied if and only if $\partial_{v_i(y)} \lambda_{k+1}(y) = 0$ for any $i \in \{k+1, \ldots, n\}$, i.e. (iii) holds.

"$\Leftarrow$": By the equivalence shown in the first part of the proof it remains to show that $\mathcal{M}$ is normally repelling. From assumption (i) and (ii) we have that the continuous functions ρ and ν obey the estimates

$$\rho = \sqrt{\lambda_{k+1}} > 1, \qquad \nu = \sqrt{\frac{\lambda_{k+1}}{\lambda_k}} > 1,$$

on the compactum $\mathcal{M}$ and hence do so uniformly for some $c \in \mathbb{R}_{>1}$. □

7.19 Remark. Unfortunately, in the general case we are not able to derive an equivalent but easier computable condition to the non-degeneracy of the second derivative of $\rho \circ x(\cdot, 0)$, i.e. a complete characterization for repelling LCSs. However, the case $k = n-1$, which we consider next, suggests that the characterizing condition might be $\partial^2_{v_i(y)} \lambda_{k+1}(y) < 0$ for $i \in \{k+1, \ldots, n\}$; cf. Remark 7.24. Intuitively, the conditions on the second order directional derivatives of λ_{k+1} in the normal directions seem to be sufficient for the non-degeneracy of $(\rho \circ x(\cdot, 0))''(0)$.

7.2.2 Hyperbolic LCSs of codimension-1

Next, we specialize our results to the (traditional) hypersurface case considered in [56]. Here, the normal bundle is a C^1 line bundle and we denote the continuously differentiable normal vector field on $\mathcal{M}$ by

$$n_0 \colon \mathcal{M} \to T^{\perp}\mathcal{M}, \qquad x \mapsto n_0(x) \in T_x^{\perp}\mathcal{M},$$

where $|n_0(x)| = \langle n_0(x), n_0(x)\rangle^{\frac{1}{2}} = 1$ for any $x \in \mathcal{M}$. The normal vector field on the image manifold $\varphi[\mathcal{M}]$ then takes the form

$$n_1(\varphi(x)) := \frac{\Phi(x)^{-*}n_0(x)}{|\Phi(x)^{-*}n_0(x)|}$$

for any point $\varphi(x) \in \varphi[\mathcal{M}]$, $x \in \mathcal{M}$. To motivate the definition of the repulsion rate, we consider

$$\begin{aligned}
\rho(x) &= \frac{1}{\langle n_0(x), C(x)^{-1}n_0(x)\rangle^{1/2}} \\
&= \frac{1}{\langle \Phi(x)^{-*}n_0(x), \Phi(x)^{-*}n_0(x)\rangle^{1/2}} \\
&= \frac{1}{|\Phi(x)^{-*}n_0(x)|} = \left\langle \frac{\Phi(x)^{-*}n_0(x)}{|\Phi(x)^{-*}n_0(x)|}, \Phi(x)n_0(x) \right\rangle \\
&= \langle n_1(\varphi(x)), \Phi(x)n_0(x)\rangle .
\end{aligned}$$

The last expression corresponds to the projection of $\Phi(x)n_0(x)$ onto the normal vector $n_1(\varphi(x))$; see also [56, Figure 10].

In the situation considered here, Theorem 7.18 can be improved to the following result.

7.20 Theorem (cf. [56, Theorem 7] and [39])**.** *Let* $\dim \mathcal{M} = n - 1$. *Then the following equivalences hold:*

1. *$\mathcal{M}$ is a repelling WLCS if and only if for any $y \in \mathcal{M}$ the following conditions hold:*
   (i) $\lambda_{n-1}(y) \neq \lambda_n(y) > 1$;
   (ii) $v_n(y) \perp T_y\mathcal{M}$;
   (iii) $\partial_{v_n(y)}\lambda_n(y) = 0$.
2. *$\mathcal{M}$ is a repelling LCS if and only if the following conditions hold:*
   (i) $\mathcal{M}$ is a repelling WLCS;
   (ii) for any $y \in \mathcal{M}$ either the matrix $L(y)$ as given in Eq. (7.13) is positive definite or, in case that $v_1, \dots, v_n$ are continuously differentiable at y, the inequality $\partial^2_{v_n(y)}\lambda_n(y) < 0$ holds.

Proof. The assertion 1. corresponds to the general statements in Theorem 7.18. It remains to show the equivalence between the definiteness of the second derivative of the repulsion rate and assertion 2(ii). As before we can assume without loss of generality that

$$e_i(y) = v_i(y), \quad i \in \{1, \ldots, n\},$$

where $\operatorname{span}\{v_1(y), \ldots, v_{n-1}(y)\} = T_y\mathcal{M}$ and $v_n(y) \perp T_y\mathcal{M}$. Recall that we have

$$\rho(y) = \left\langle v_n(y), C(y)^{-1} v_n(y) \right\rangle^{-1/2} = \sqrt{\lambda_n(y)}.$$

Since N, N_ε, K, L are now vectors, we denote them (in accordance with the notation in [56]) by $n, n_\varepsilon, \beta, \gamma$, respectively. We have for β and γ, i.e. L in Eq. (7.9),

$$\begin{aligned}
\beta &= -\sum_{i=1}^{n-1} \partial_{i-1}\alpha(0) v_i(y), \\
\gamma &= -\frac{1}{2}\left(\sum_{i=1}^{n-1} \partial_{i-1}\alpha(0)^2\right) v_n(y) - \alpha(0) \sum_{i=1}^{n-1} \langle \partial_{i-1}(v_n \circ F)(0), \beta \rangle \, v_i(y)
\end{aligned}$$

and furthermore, for $\varepsilon \in (-\theta, \theta)$, that

$$\begin{aligned}
\rho(x(\cdot,0))'(\varepsilon) = -\frac{1}{2}\rho(x(\varepsilon,0))^3 \Big(2 \left\langle \beta + 2\varepsilon\gamma, C(x(\varepsilon,0))^{-1} n_\varepsilon(x(\varepsilon,0)) \right\rangle + \\
+ \alpha(0) \left\langle n_\varepsilon(x(\varepsilon,0)), \left(\partial C^{-1}(x(\varepsilon,0)) v_n(y)\right) n_\varepsilon(x(\varepsilon,0)) \right\rangle \Big).
\end{aligned}$$

Due to the stationarity assumption $\rho(x(\cdot,0))'(0) = 0$ and the fact that $-\frac{1}{2}\rho(y)^3 = -\frac{(\lambda_n(y))^{3/2}}{2}$ the product rule of differentiation yields

$$\begin{aligned}
\rho(x(\cdot,0))''(0) = -\frac{(\lambda_n(y))^{3/2}}{2} \Big(\varepsilon \mapsto 2 \left\langle \beta + 2\varepsilon\gamma, C(x(\varepsilon,0))^{-1} n_\varepsilon(x(\varepsilon,0)) \right\rangle + \\
+ \alpha(0) \left\langle n_\varepsilon(x(\varepsilon,0)), \left(\partial C^{-1}(x(\varepsilon,0)) v_n(y)\right) n_\varepsilon(x(\varepsilon,0)) \right\rangle \Big)'(0).
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\frac{-2\rho(x(\cdot,0))''(0)}{\lambda_n(y)^{3/2}} &= 2\alpha(0) \left\langle \beta, \left(\partial C^{-1}(y) v_n(y)\right) v_n(y) \right\rangle + 2 \left\langle \beta, C(y)^{-1}\beta \right\rangle + \\
&\quad + 4 \left\langle \gamma, C(y)^{-1} v_n(y) \right\rangle + \\
&\quad + \alpha(0) \left\langle \beta, \left(\partial C^{-1}(y) v_n(y)\right) v_n(y) \right\rangle + \\
&\quad + \alpha(0)^2 \left\langle v_n(y), \partial^2 C^{-1}(y)(v_n(y), v_n(y)) v_n(y) \right\rangle + \\
&\quad + \alpha(0) \left\langle v_n(y), \left(\partial C^{-1}(y) v_n(y)\right) \beta \right\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \alpha(0)^2 \left\langle v_n(y), \partial^2 C^{-1}(y)(v_n(y), v_n(y))v_n(y) \right\rangle + \\
&\quad - 4\sum_{i=1}^{n-1} \alpha(0)\partial_{i-1}\alpha(0) \left\langle v_i(y), \left(\partial C^{-1}(y)v_n(y)\right) v_n(y) \right\rangle + \\
&\quad + 2\sum_{i=1}^{n-1} (\partial_{i-1}\alpha(0))^2 \left(\frac{1}{\lambda_i(y)} - \frac{1}{\lambda_n(y)} \right) \\
&= \left\langle \begin{pmatrix} \alpha(0) \\ \partial_0\alpha(0) \\ \vdots \\ \partial_{n-2}\alpha(0) \end{pmatrix}, L(y) \begin{pmatrix} \alpha(0) \\ \partial_0\alpha(0) \\ \vdots \\ \partial_{n-2}\alpha(0) \end{pmatrix} \right\rangle,
\end{aligned}
$$

where we have used the symmetry of $\partial C^{-1}(y)v_n(y)$ and

$$
L = \begin{pmatrix}
\left\langle v_n, \partial^2 C^{-1}(v_n, v_n)v_n \right\rangle & -2\left\langle v_1, (\partial C^{-1}v_n)v_n \right\rangle & \cdots & -2\left\langle v_{n-1}, (\partial C^{-1}v_n)v_n \right\rangle \\
-2\left\langle v_1, (\partial C^{-1}v_n)v_n \right\rangle & 2\frac{\lambda_n - \lambda_1}{\lambda_1\lambda_n} & & \\
\vdots & & \ddots & \\
-2\left\langle v_{n-1}, (\partial C^{-1}v_n)v_n \right\rangle & & & 2\frac{\lambda_n - \lambda_{n-1}}{\lambda_{n-1}\lambda_n}
\end{pmatrix}. \tag{7.13}
$$

Clearly, we have $\rho(x(\cdot, 0))''(0) < 0$ in every point $y = x(0, 0) \in \mathcal{M}$ and for any perturbation profile α if and only if the matrix $L(y)$ is positive definite in every $y \in \mathcal{M}$.

Next we consider the case that $v_1, \dots, v_n$ are differentiable at $y \in \mathcal{M}$. We need to prove the equivalence between the positive definiteness of $L(y)$ and the inequality $\partial^2_{v_n(y)}\lambda_n(y) < 0$. To this end, we first derive an equivalent representation of the matrix $L(y)$ under the additional differentiability assumption on the eigenvector functions. The calculations follow essentially those in [56, p. 583–584].

For $i \in \{1, \dots, n-1\}$ and $y \in \mathcal{M}$ differentiate the two equations

$$
\langle v_i(y), v_n(y) \rangle = 0 \qquad \text{and} \qquad \left\langle v_i(y), C(y)^{-1}v_n(y) \right\rangle = 0,
$$

to obtain $\partial v_i(y)^* v_n(y) + v_i(y)^* \partial v_n(y) = 0$ and

$$
\partial v_i(y)^* C(y)^{-1} v_n(y) + v_i(y)^* \left(\partial C^{-1}(y)v_n(y) + C^{-1}(y)\partial v_n(y) \right) = 0.
$$

This together with the eigenvector property of v_n and v_i and the symmetry of $C(y)^{-1}$ we get

$$
v_i(y)^* \left(\partial C^{-1}(y)v_n(y) \right) = -\frac{\lambda_n(y) - \lambda_i(y)}{\lambda_i(y)\lambda_n(y)} v_i(y)^* \partial v_n(y).
$$

Application to $v_n(y)$ and reordering the terms yields

$$\left\langle v_i(y), \left(\partial C^{-1}(y) v_n(y)\right) v_n(y) \right\rangle = -\tfrac{\lambda_n(y)-\lambda_i(y)}{\lambda_i(y)\lambda_n(y)} \left\langle v_i(y), (\partial v_n(y)) v_n(y) \right\rangle, \tag{7.14}$$

such that L takes the form

$$L = \begin{pmatrix} \left\langle v_n, \partial^2 C^{-1}(v_n, v_n) v_n \right\rangle & 2\frac{\lambda_n-\lambda_1}{\lambda_1\lambda_n}\langle v_1, (\partial v_n) v_n\rangle & \cdots & 2\frac{\lambda_n-\lambda_{n-1}}{\lambda_{n-1}\lambda_n}\langle v_{n-1}, (\partial v_n) v_n\rangle \\ 2\frac{\lambda_n-\lambda_1}{\lambda_1\lambda_n}\langle v_1, (\partial v_n) v_n\rangle & 2\frac{\lambda_n-\lambda_1}{\lambda_1\lambda_n} & & \\ \vdots & & \ddots & \\ 2\frac{\lambda_n-\lambda_{n-1}}{\lambda_{n-1}\lambda_n}\langle v_{n-1}, (\partial v_n) v_n\rangle & & & 2\frac{\lambda_n-\lambda_{n-1}}{\lambda_{n-1}\lambda_n} \end{pmatrix},$$

as in [56, Eq. (31)] with the correction in [39]. Next, we consider

$$C(y)^{-1} v_n(y) = \frac{1}{\lambda_n(y)} v_n(y),$$

and differentiation at y and composition with $v_n(y)^*$ yields the equation

$$v_n(y)^* \left(\partial C^{-1}(y)(\cdot)\right) v_n(y) = -\frac{1}{\lambda_n(y)^2} \partial\lambda_n(y)$$

in $L(\mathbb{R}^n, \mathbb{R})$. By differentiating at y and evaluating at $(v_n(y), v_n(y)) \in \mathbb{R}^n \times \mathbb{R}^n$ we obtain

$$\begin{aligned} \left\langle v_n(y), \partial^2 C^{-1}(y)(v_n(y), v_n(y)) v_n(y) \right\rangle &= \\ = -2 \left\langle v_n(y), \left(\partial C^{-1}(y) v_n(y)\right) ((\partial v_n(y)) v_n(y)) \right\rangle &+ \\ - \frac{1}{\lambda_n(y)^2} (\partial^2 \lambda_n(y))(v_n(y), v_n(y)), & \end{aligned}$$

where we have used that $(\partial\lambda_n(y)) v_n(y) = 0$ and the symmetry of $\partial C^{-1}(y) v_n(y)$. Since $v_n(y)^*(\partial v_n(y)) v_n(y) = 0$, we have that

$$(\partial v_n(y)) v_n(y) = \sum_{i=1}^{n-1} \langle v_i(y), (\partial v_n(y)) v_n(y) \rangle \, v_i(y).$$

Using Eq. (7.14) we obtain

$$\begin{aligned} \left\langle v_n(y), \partial C^{-1}(y)((\partial v_n(y)) v_n(y)) \right\rangle &= \\ = \sum_{i=1}^{n-1} \langle v_i(y), (\partial v_n(y)) v_n(y) \rangle \left\langle v_n(y), \left(\partial C^{-1}(y) v_n(y)\right) v_i(y) \right\rangle & \\ = -\sum_{i=1}^{n-1} \frac{\lambda_n(y) - \lambda_i(y)}{\lambda_i(y)\lambda_n(y)} \langle v_i(y), (\partial v_n(y)) v_n(y) \rangle^2, & \end{aligned}$$

and finally

$$\left\langle v_n(y), \left(\partial^2 C^{-1}(y)(v_n(y), v_n(y))\right) v_n(y)\right\rangle =$$
$$= 2\sum_{i=1}^{n-1} \frac{\lambda_n(y) - \lambda_i(y)}{\lambda_i(y)\lambda_n(y)} \langle v_i(y), (\partial v_n(y))v_n(y)\rangle^2 +$$
$$- \frac{1}{\lambda_n(y)^2}(\partial^2\lambda_n(y))(v_n(y), v_n(y)). \quad (7.15)$$

Obviously, $L(y)$ is positive definite if and only if $JL(y)J$ is positive definite, where $J = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}$ is the (orthogonal) inverted identity matrix. However, Sylvester's criterion applied to $JL(y)J$ corresponds to taking the leading principal minors of $L(y)$ from the lower right corner instead of taking them from the upper left corner. Since all "lower" leading principal minors of order lower than n are positive in any case (they just correspond to the product of the diagonal entries), by Sylvester's criterion, $L(y)$ is positive definite if and only if $\det L(y) > 0$. Now we make use of the block structure of $L(y)$ as follows: consider

$$L = \begin{pmatrix} a & b^\top \\ b & D \end{pmatrix},$$

where

$$a = \left\langle v_n, \partial^2 C^{-1}(v_n, v_n)v_n\right\rangle, \qquad b = \begin{pmatrix} 2\frac{\lambda_n-\lambda_1}{\lambda_1\lambda_n}\langle v_1, (\partial v_n)v_n\rangle \\ \vdots \\ 2\frac{\lambda_n-\lambda_{n-1}}{\lambda_{n-1}\lambda_n}\langle v_{n-1}, (\partial v_n)v_n\rangle \end{pmatrix},$$
$$D = \operatorname{diag}\left(2\frac{\lambda_n-\lambda_1}{\lambda_1\lambda_n}, \dots, 2\frac{\lambda_n-\lambda_{n-1}}{\lambda_{n-1}\lambda_n}\right),$$

and for $y \in \mathcal{M}$ we have $a(y) \in \mathbb{R}$, $b(y) \in \mathbb{R}^{n-1}$ and $D(y) \in \mathbb{R}^{(n-1)\times(n-1)}$. Since D is invertible, we have that

$$\det L = \det D \cdot \det(a - b^\top D^{-1}b) = \det D(a - b^\top D^{-1}b),$$

cf. [112, Theorem 2.2]. With $\det D = \prod_{i=1}^{n-1} 2\frac{\lambda_n-\lambda_i}{\lambda_i\lambda_n} > 0$ we obtain that $\det L > 0$ if and only if

$$a - b^\top D^{-1}b = \left\langle v_n, (\partial^2 C^{-1})(v_n, v_n)v_n\right\rangle - 2\sum_{i=1}^{n-1}\frac{\lambda_n-\lambda_i}{\lambda_i\lambda_n}\langle v_i, (\partial v_n)v_n\rangle^2$$
$$= -\frac{1}{\lambda_n^2}(\partial^2\lambda_n)(v_n, v_n) > 0,$$

or equivalently $\partial^2_{v_n(y)}\lambda_n(y) < 0$. Here, we have used Eq. (7.15). □

7.21 Remark. 1. By Sylvester's criterion a necessary condition for $L(y)$ to be positive definite is the inequality

$$\left\langle v_n(y), \left(\partial^2 C^{-1}(y)(v_n(y), v_n(y))\right) v_n(y) \right\rangle > 0;$$

see [56, Proposition 8].

2. The second alternative in Theorem 7.20 2/(ii) is a significant simplification of the first alternative. Note that for $n = 2$, the differentiability of the two eigenvectors follows by the simplicity of the eigenvalues by Theorem 7.20 1/(i) and Proposition 7.5. However, in higher dimensions the differentiability requirement is in general restricting.

3. The second alternative in Theorem 7.20 2/(ii) has been observed for 2-dimensional systems in [39].

7.3 Regularity and Robustness of Hyperbolic LCSs

So far, we assumed that we have a manifold $\mathcal{M}$ with properties as described in Section 7.2.1. We derived characterizing properties in Theorem 7.18 and Theorem 7.20 for $\mathcal{M}$ to be a hyperbolic LCS. Now, we want to go the other way and use those properties to define (or to extract) a candidate set $\mathcal{M}$ for a hyperbolic LCS, which is has submanifold structure. To this end, we restrict our considerations first to the codimension-1 case and to the open set

$$H := \{y \in \mathbb{R}^n;\ \lambda_{n-1}(y) \neq \lambda_n(y) > 1\},$$

on which λ_n and v_n are twice continuously differentiable by Proposition 7.5. Now, define M as the 0-level set of the continuously differentiable function

$$H \ni y \mapsto \partial_{v_n(y)}\lambda_n(y) = \partial\lambda_n(y)v_n(y) = \mathcal{L}_{v_n(y)}\lambda_n(y), \tag{7.16}$$

where $\mathcal{L}_{v_n(y)}\lambda_n(y)$ denotes the Lie derivative of λ_n in the direction of the vector field v_n. By the Implicit Function Theorem, there exists a special local parametrization of M around $y \in M$, where one component is expressed as a C^1-function depending on the other $n-1$ components, if the derivative

$$(x \mapsto \partial\lambda_n(x)v_n(x))'(y) = \partial^2\lambda_n(y)v_n(y) + \partial\lambda_n(y)\partial v_n(y) \neq 0. \tag{7.17}$$

Clearly,

$$\partial^2\lambda_n(y)(v_n(y), v_n(y)) + \partial\lambda_n(y)\partial v_n(y)v_n(y) \neq 0 \tag{7.18}$$

is sufficient for Eq. (7.17) to hold. Conversely, suppose that the inequality (7.17) holds for all $y \in M$, then M is locally an embedded submanifold and we can speak of the tangent space T_yM of M in $y \in M$. Now define

$$\mathcal{M} := \{y \in M;\ v_n(y) \perp T_yM\}\,. \tag{7.19}$$

Since the derivative in Eq. (7.17), considered as a gradient, is orthogonal to the 0-level set, we obtain that on $\mathcal{M}$ Eqs. (7.18) and (7.17) are equivalent.

However, we would like to point out that, in general, it seems to be unclear whether Eq. (7.19) leads to a submanifold of codimension-1. While there are analytical examples in [56, Section 8], hyperbolic LCSs may typically not exist as embedded submanifolds in general finite-time processes considered in applications.

The above discussion for hyperbolic codimension-1 LCSs can be partially transferred to the general case, developed in Section 7.2.1. To this end, we restrict our considerations to

$$H := \{y \in \mathbb{R}^n;\ \lambda_k(y) \neq \lambda_{k+1}(y) \neq \lambda_{k+2},\ \lambda_{k+1} > 1\}\,.$$

Then λ_{k+1} is continuously differentiable on H. For $i \in \{k+1, \ldots, n\}$ introduce the 0-level set

$$H_i := (\mathcal{L}_{v_i}\lambda_{k+1})^{-1}[\{0\}].$$

By requiring corresponding versions of Eq. (7.17) for each H_i, we ensure that locally the H_i are embedded codimension-1 submanifolds and their intersection

$$M := \bigcap_{i=k+1}^{n} H_i$$

is typically a k-dimensional submanifold. We define

$$\mathcal{M} := \{y \in M;\ v_i(y) \perp T_yM,\ i \in \{k+1, \ldots, n\}\}\,.$$

Again, in general, it seems to be difficult to affirm that $\mathcal{M}$ thus defined is a submanifold and that $\mathcal{M}$ is well-behaved. These difficulties in the extraction of (exact) LCSs based on the variational approach prompted the relaxation of the orthogonality requirement in [38].

As for robustness, assume that $\mathcal{M}$ is a codimension-1 hyperbolic LCS, on which Eq. (7.18) is satisfied. Then $\mathcal{M}$ is, by [56, Theorem 11], robust under continuously differentiable perturbations of φ in the following, peculiar sense: hyperbolic

LCSs do not persist under perturbations along with all four properties stated in Theorem 7.20 (1)-(2). However, they persist as *hyperbolic Quasi-LCSs*, see [56, Def. 10], i.e. as manifolds that do not satisfy necessarily the alignment requirement Theorem 7.20 1/(ii). Note, that a perturbation of the, initial or final, time parameter leads in particular to a continuously differentiable perturbation of the flow map φ. Consequently, the exact position of hyperbolic LCSs depends crucially on the choice of the time parameters and hyperbolic LCSs might get lost under small variations. Instead, there existence may be indicated for nearby parameters by hyperbolic Quasi-LCSs. The issue of missing robustness in the strict sense, but existing robustness in the aforementioned quasi-sense is also addressed in [38].

7.4 Generalized Extrema and Hyperbolic LCSs

In this section, we discuss the relation between hyperbolic LCSs and the differential geometric notion of *ridges* and their antonym, *valleys* (also referred to as *troughs* or *courses*), with umbrella term *crease*. As is discussed in [35], there exist several mathematical definitions of creases; for a textbook on ridges we refer the reader to [34] and for a historical review see [71].

Let us first recall that for $f \in C^1(\mathbb{R}^n, \mathbb{R})$ a point $x \in \mathbb{R}^n$ is said to be a *critical point* for f if $\partial f(x) = 0 \in L(\mathbb{R}^n, \mathbb{R})$, or equivalently $\operatorname{grad} f(x) = 0 \in \mathbb{R}^n$. Critical points are further classified if $f \in C^2(\mathbb{R}^n, \mathbb{R})$ as follows. The function f is said to have a *minimum/maximum* at x if $\partial^2 f(x)$ is positive/negative definite. We represent the second derivative by the Hessian matrix $H(x) := \partial \otimes \operatorname{grad} f(x)$, where $\otimes$ denotes the formal Kronecker product of $\partial = \begin{pmatrix} \partial_0 & \dots & \partial_{n-1} \end{pmatrix}$ and $\operatorname{grad} = \partial^\top$. Any such point is called an *extreme point* of f. This definition of extreme points can be generalized to "conditional" extremality. Let $k \in \{1, \dots, n\}$, $L \in \operatorname{Gr}(k, \mathbb{R}^n)$ and $A \in \operatorname{St}^*(k, \mathbb{R}^n)$ with $\pi(A) = L$. The function f has a *generalized minimum/-maximum of type $n-k$ at x with respect to L* if

(i) $A^\top \operatorname{grad} f(x) = 0 \in \mathbb{R}^k$, or equivalently, $\operatorname{grad} f(x) \perp L$, and

(ii) $A^\top H(x) A$ is positive/negative definite.

Such points x are called *generalized extreme points of type $n-k$ for f with respect to L*.

Now we are in the position to define *FTLE ridges*, thereby modifying the definition in [56].

7.22 Definition (FTLE ridge, cf. [56, Def. 12])**.** A k-dimensional manifold $\mathcal{M}$ as in Section 7.2.1 is called an *FTLE ridge*, if each $x \in \mathcal{M}$ is a generalized maximum of type k for λ_{k+1} with respect to span $\{v_{k+1}(x), \dots, v_n(x)\}$.

In fact, FTLE ridges are not ridges in the usual, geometric sense, which is introduced as follows. Let again $f \in C^2(\mathbb{R}^n, \mathbb{R})$, $x \in \mathbb{R}^n$ and consider the Hessian matrix $H(x)$ of f at x. Denote by $\kappa_1 \leq \dots \leq \kappa_n$ the ordered eigenvalues of H and by $v_1, \dots, v_n$ corresponding eigenvectors. A point $x \in \mathbb{R}^n$ is called a *ridge point of type $n-k$* if $\kappa_k < 0$ and x is a generalized maximum point of type $n-k$ for f with respect to span $\{v_1, \dots, v_k\}$. A point $x \in \mathbb{R}^n$ is called a *valley point of type $n-k$* if $\kappa_{n-k+1} > 0$ and x is a generalized minimum point of type $n-k$ for f with respect to span $\{v_{n-k+1}, \dots, v_n\}$. We call an embedded, differentiable submanifold of dimension $n-k$ in $\mathbb{R}^n$ a *(height) ridge* (*(height) valley*) if each of its points is a ridge (valley, respectively) point of type $n-k$. Finally, *second-derivative ridges* as defined in [101, 74] are height ridges $\mathcal{M}$ of codimension-1, such that in each point $x \in \mathcal{M}$ one has that grad $\lambda_n(x) \in T_x\mathcal{M}$ and $\lambda_n(x) > \lambda_{n-1}(x)$.

Historically, the connection between hyperbolic LCSs and ridges developed as follows. [54, 55] proposed that a repelling LCS should appear as a locally maximizing curve, or ridge in the higher-dimensional case, of the FTLE field. Indeed, ridges of the FTLE field have been found to be accurate indicators of LCSs in a number of applications; see [88]. Despite an example in [55] showing that FTLE maximizing curves do not necessarily correspond to coherent trajectory patterns, [101, 74] define hyperbolic LCSs as a family of second-derivative ridges of the FTLE field over time. While the extraction of these ridges has been performed visually, see [100], the differential-geometric notion of a second-derivative ridge was used to derive estimates on the flux across the LCS. [99] initiated many applications of that concept. However, recent results show that the focus on second-derivative ridges is too restrictive. First, [56] shows that the estimate on the flux formula holds only under additional restricting assumptions. Second, [83] proves that second-derivative ridges in 2-dimensional systems are necessarily straight lines. Our next aim is to advocate a change of focus from ridges to generalized extrema, based on the results in Section 7.2.2.

To this end, let us consider the codimension-1 case, for which we obtain the following trivial characterization of repelling LCSs as certain FTLE ridges. This result shows that hyperbolic LCS are in fact of generalized-extremum-type, rather than of ridge-type, due to the missing reference to the Hessian of the FTLE field.

7.23 Theorem (Characterization of Hyperbolic LCSs by FTLE Ridges)**.** *Assume that $\mathcal{M}$ is of codimension-1 and the functions $v_1, \dots, v_n$ are continuously differentiable* *on $\mathcal{M}$. Then $\mathcal{M}$ is a repelling LCS if and only if $\mathcal{M}$ is an FTLE ridge such that for each $x \in \mathcal{M}$ the following conditions are satisfied:*

$$\lambda_{n-1}(x) \neq \lambda_n(x) > 1, \tag{7.20}$$

$$v_n(x) \perp T_x\mathcal{M}. \tag{7.21}$$

Proof. The FTLE ridge assumption together with conditions (7.20)–(7.21) are just re-statements of the conditions listed in Theorem 7.20. □

7.24 Remark. Recall that, in the general case of Section 7.2.1, we were not able to find an analytically computable criterion that distinguishes for repelling WLCS from repelling LCS. However, Theorem 7.23 suggests that this criterion is the second-derivative condition in the definition of generalized extreme points of type k, thus justifying Remark 7.19.

7.5 Embeddings of Hyperbolic LCS

In this section, we want to formally introduce hyperbolic LCSs embedded in another manifold. We leave a rigorous investigation as well as a (physical) motivation for future research. One possible application is the introduction of the variational LCS-approach to finite-time flows on manifolds; another one is the derivation of substructures on LCSs defined in $\mathbb{R}^n$. Concerning the second application, we would like to point out that the concept of codimension-1 LCSs as introduced in [56] essentially gives dynamical information on normal perturbations of some hypersurface under investigation. From the point of view of applications, that is probably the most important one, since $(n-1)$-dimensional manifolds locally divide the space into two halves and are therefore considered to act as transport barriers. However, the methods presented so far do not give any information on the dynamical behavior "within" the hyperbolic LCS, i.e. information on tangential perturbations.

The first of the aforementioned cases corresponds to the study of a diffeomorphism

$$\varphi\colon \mathcal{N} \to \mathcal{N},$$

while the second corresponds to the study of a diffeomorphism

$$\varphi\colon \mathcal{N} \subset \mathbb{R}^N \to \varphi[\mathcal{N}] \subset \mathbb{R}^N.$$

Here $\mathcal{N}$ denotes an n-dimensional differentiable manifold.

Let $\mathcal{M}$ with $\dim \mathcal{M} = k$ be a differentiable submanifold of another differentiable manifold $\mathcal{N}$ with $\dim \mathcal{N} = n$, which is possibly embedded in $\mathbb{R}^N$. We define the normal bundle of $\mathcal{M}$ as the (fiber-wise) orthogonal complement of the tangent bundle $T\mathcal{M}$ in $T\mathcal{N}$, i.e. we have

$$T\mathcal{N} = T\mathcal{M} \oplus T^{\perp}\mathcal{M}.$$

Now repulsion rate and repulsion ratio are defined literally as in Definition 7.10. In this setting Definition 7.12 makes sense as well. Finally, we have to restrict the families of normal perturbations to those that lead to manifolds $\mathcal{M}_\varepsilon$ which stay in $\mathcal{N}$. This can be achieved by using special Riemannian normal coordinates $F\colon D(F) \subseteq \mathbb{R}^n \to \mathcal{N}$ of $\mathcal{N}$ in each point $x \in \mathcal{M} \subset \mathcal{N}$. By means of the exponential map at $x \in \mathcal{M}$ and an appropriate choice of an orthonormal basis in $\mathbb{R}^n$ one can ensure that the resulting coordinates are aligned with tangent and normal space of $\mathcal{M}$ at x, i.e. $F(0) = x$ and

$$\operatorname{span}\{\partial_0 F(0), \dots, \partial_{k-1} F(0)\} = T_x\mathcal{M}, \quad \operatorname{span}\{\partial_k F(0), \dots, \partial_{n-1} F(0)\} = T_x^{\perp}\mathcal{M}.$$

Let $\alpha\colon U \subseteq \mathbb{R}^k \to \mathbb{R}^{n-k}$ be bounded and continuously differentiable, where U is a neighborhood of the origin. With the local parametrization F and the perturbation profile α we consider perturbations of the form

$$(-\theta, \theta) \times U \subset \mathbb{R}^{k+1} \to \mathcal{N}, \qquad (\varepsilon, u) \mapsto F(u, \varepsilon\alpha(u)).$$

Obviously, all perturbations stay in $\mathcal{N}$ and $(\varepsilon \mapsto F(0, \varepsilon\alpha(0)))'(0) \in T_x^{\perp}\mathcal{M}$, i.e. we indeed have a normal perturbation. In this setting, Definition 7.17 is meaningful and due to the similarity one can hope that Theorem 7.18 and Theorem 7.20 hold analogously. In the general context $\varphi\colon \mathcal{N} \to \mathcal{N}$, where $\mathcal{N}$ is an abstract differentiable manifold, the directional derivatives in 1(iii) of the mentioned theorems have to be replaced by the Lie derivatives of the eigenvalue functions in the directions of the eigenvector fields.

Next, we want to discuss what can be expected when applying the aforementioned procedure to the setting discussed in Section 7.2. More precisely, suppose we consider a diffeomorphism $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$, where we have identified a, for instance, codimension-1 manifold $\mathcal{M}_1$ as a repelling LCS. Assuming the characterizations in Theorem 7.18 and Theorem 7.20 to hold analogously and restricting φ to $\mathcal{M}_1$, we get that the linearization of φ and hence the Cauchy-Green strain tensor are restricted to the direct sum of eigenspaces corresponding to the lower $n-1$ eigenvalues of C. Studying the restricted dynamics therefore means analyzing structures based on the lower eigenvalues, too. When applied iteratively, one might be able to get dynamical information from the whole family of finite-time Lyapunov spectra $\Sigma_{\mathrm{FTL}}(t_+, t_-, \cdot)$.

One can draw parallels between the construction of spectral cones and manifolds in Section 3.4 on the one hand and the iterative embeddings of hyperbolic LCS on the other hand: it is the task of finding a "dynamical skeleton" which captures the most important dynamical features. However, for both notions the mathematical and the physical relevance seem to be still unclear; see also the discussion in the last paragraphs of Section 3.4.

8 Conclusions

In this thesis we proposed an abstract framework for evolutionary processes on compact time-sets, which covers different special cases from the literature. By relying on the elementary notion of logarithmic difference quotient and by introducing a natural topology on the linear finite-time processes, we were able to generalize and to extend several existing approaches and results, such as continuity of growth rates and of the spectrum. From that we easily obtained robustness of finite-time hyperbolicity including sharp perturbation bounds. In the same framework, we obtained finite-time analogues to classical linearization theorems by means of continuity arguments rather than artificially extending the processes to unbounded time-sets and invoking asymptotic arguments. Thus, we obtained a complete description locally around a reference trajectory, realizing and making explicit use of the inherent non-uniqueness of EMD-projections and local stable manifolds, for instance. An open problem in this context is the establishment of a state space decomposition into spectral subspaces/cones.

Furthermore, we described and implemented a general numerical approach to the computation of growth rates. Future research should aim at finding sufficient conditions for uniqueness of extremal subspaces, which in turn will make the growth rate functions differentiable at their optimizers and the numerical approach via the Newton algorithm feasible. A comparison with existing alternative optimization codes as well as the implementation of a trust-region algorithm for the computation of growth rates are further research targets. With powerful computational methods at hand, applications to real-world problems will hopefully lead to new insights and to new theoretical challenges. In particular, they would give the opportunity to find counter-examples to the state space decomposition conjecture; recall the discussion at the end of Section 3.4.

As we mentioned in this work, hyperbolic LCSs are a very popular tool in the analysis of oceanic, atmospheric and biological systems. So far, only codimension-1 manifolds were considered in applications. In higher-dimensional systems, also hyperbolic LCSs of higher codimension may play an important role in understanding the dynamics better. However, a physical motivation and an example demonstrating the relevance of these structures are still to be found.

Another interesting research direction was started only recently. It is about *predicting* the formation of certain patterns in the close future, based on extracted information over bounded time-sets; see [79, 84]. Once a theory for finite-time predictions exists, the next step would be to establish a theory of finite-time *control*. Also, one can study finite-time processes on manifolds, static or dynamic, which could open up new fields of application and again feed-back open problems to theoretical research.

A Differential Geometry

In this section, we recall notions from differential geometry and fix the notation. For textbooks on differential geometry we refer the reader to [70, 85, 21, 1, 18, 80]. Note, however, that our notation differs at some points from the usual notation; for more details and proofs see also [66].

Throughout this chapter let M, N denote two sets. We define

$$\mathrm{Map}(M, N) := \{A \subseteq M \times N;\ A \text{ right-unique}\}\,.$$

For $f \in \mathrm{Map}(M, N)$ we set $D(f) := [N]f$, i.e. the domain of f.

A.1 Differentiable Manifolds

A.1 Definition (Atlas). Let $n \in \mathbb{N}$ and $r \in \mathbb{N}_0 \cup \{\infty\}$. We call a set $\mathcal{A} \subseteq \mathrm{Map}(M, \mathbb{R}^n)$ an *n-dimensional C^r-Atlas of M*, if

(a) $\bigcup_{\varphi \in \mathcal{A}} D(\varphi) = M$,

(b) any $\varphi \in \mathcal{A}$ is injective,

(c) for any $\varphi, \psi \in \mathcal{A}$ one has: $\varphi[D(\psi)]$ is open and the so-called *change of coordinates*

$$\psi \circ \varphi^{-1} \colon \varphi[D(\psi)] \subseteq \mathbb{R}^n \to \mathbb{R}^n$$

is r-times continuously differentiable.

We refer to elements $\varphi \in \mathcal{A}$ of the atlas as *charts* and to $(M, \mathcal{A})$ as an *n-dimensional C^r-manifold*.

Let $n \in \mathbb{N}, r \in \mathbb{N}_0 \cup \{\infty\}$ and $\mathcal{A}$ be an n-dimensional C^r-atlas of M. Then by

$$\mathcal{T}_{\mathcal{A}} := \{U \subseteq M;\ \forall \varphi \in \mathcal{A} \colon \varphi[U] \text{ open}\}$$

a topology on M is given. With respect to this topology any chart is a homeomorphism on its image. For $p \in M$, we denote by $\mathcal{U}_p(\mathcal{T}_\mathcal{A})$ the filter of neighborhoods (with respect to $\mathcal{T}_\mathcal{A}$) around p.

In the following, we restrict to C^∞-manifolds, which can be justified by Whitney's theorem, see [21], and refer to these as differentiable manifolds for short. Even shorter, we refer to n-dimensional differentiable manifolds as n-manifolds.

A.2 Definition (Differentiability)**.** Let $m, n \in \mathbb{N}$, $(M, \mathcal{A})$ be an m-manifold, $(N, \mathcal{B})$ be an n-manifold, $p \in M$, $U \in \mathcal{U}_p(\mathcal{T}_\mathcal{A})$ and $f\colon U \subseteq M \to N$ be a function. We say that f *is differentiable in* p, if for any $\varphi \in \mathcal{A}$ with $p \in D(\varphi)$ and any $\psi \in \mathcal{B}$ with $f(p) \in D(\psi)$ one has that

$$\psi \circ f \circ \varphi^{-1}\colon \varphi[f^{-1}[D(\psi)]] \subseteq \mathbb{R}^m \to \mathbb{R}^n$$

is indefinitely differentiable in $\varphi(p)$. We call a function $f\colon M \to N$ *differentiable*, if f is differentiable in any $p \in M$. We call $f\colon M \to N$ a *diffeomorphism*, if f is bijective and both f, f^{-1} are differentiable.

One can show that differentiability implies continuity and that, in order to check for differentiability at a point, it suffices to test with an appropriate map $\varphi \in \mathcal{A}$ and a $\psi \in \mathcal{B}$. Furthermore, compositions of differentiable functions between differentiable manifolds are differentiable. For two in $p \in M$ differentiable functions $f, g\colon M \to \mathbb{R}$ we have that $f + g$, fg, and λf, $\lambda \in \mathbb{R}$, are differentiable in p.

In the following, let $(M, \mathcal{A})$ be an n-manifold, $p \in M$ and $\varphi \in \mathcal{A}$ with $p \in D(\varphi)$. For $j \in \{1, \ldots, n\}$ we denote by $\pi_j\colon \mathbb{R}^n \to \mathbb{R}$ the canonical projection on the j-th coordinate. Additionally, let $\varphi_j := \pi_j \circ \varphi$ for $j \in \{1, \ldots, n\}$. Let $U \in \mathcal{U}_p(\mathcal{T}_\mathcal{A})$ and $f\colon U \subseteq M \to \mathbb{R}$ be differentiable in p. Then we define

$$\partial_{\varphi_j} f(p) := \partial_j (f \circ \varphi^{-1})(\varphi(p)) \quad (j \in \{1, \ldots, n\}).$$

Furthermore, we define an equivalence relation on functions that are differentiable in p as follows: let $U, V \in \mathcal{U}_p(\mathcal{T}_\mathcal{A})$ and $f\colon U \subseteq M \to \mathbb{R}, g\colon V \subseteq M \to \mathbb{R}$ be differentiable in p. Then $f \sim_p g$, if there exists $W \in \mathcal{U}_p(\mathcal{T}_\mathcal{A})$ with $W \subseteq U \cap V$, such that $f\big|_W = g\big|_W$. The set

$$\mathcal{F}_p(M) := \left\{f\colon U \to \mathbb{R};\ U \in \mathcal{U}_p(\mathcal{T}_\mathcal{A}), f \text{ differentiable in } p\right\}\big/{\sim_p}$$

is called the set of *(function) germs in* p. We denote the elements of $\mathcal{F}_p(M)$ by $[f]_{\sim_p}$, where f is as in the definition. We define addition, multiplication and point evaluation on $\mathcal{F}_p(M)$ pointwise using representatives. In particular, for $\varphi \in \mathcal{A}$ with $p \in D(\varphi)$ we have $[\varphi_j]_{\sim_p} \in \mathcal{F}_p(M)$ for each $j \in \{1, \ldots, n\}$.

A.3 Definition (Derivation, tangent space)**.** A map $X\colon \mathcal{F}_p(M) \to \mathbb{R}$ is called *derivation on* $\mathcal{F}_p(M)$, if X is linear and for any $[f]_{\sim_p}, [g]_{\sim_p} \in \mathcal{F}_p(M)$ one has

$$\begin{aligned} X([f]_{\sim_p}[g]_{\sim_p}) &= [f]_{\sim_p}(p)X([g]_{\sim_p}) + X([f]_{\sim_p})[g]_{\sim_p}(p) \\ &= f(p)X([g]_{\sim_p}) + X([f]_{\sim_p})g(p). \end{aligned}$$

The set

$$T_pM := \big\{X\colon \mathcal{F}_p(M) \to \mathbb{R};\ X \text{ derivation on } \mathcal{F}_p(M)\big\}$$

is called the *tangent space of M at p*.

The tangent space T_pM is a vector space and for each $j \in \{1, \dots, n\}$ we have that

$$\partial_{\varphi_j}\big|_p\colon \mathcal{F}_p(M) \to \mathbb{R}, \qquad\qquad [f]_{\sim_p} \mapsto \partial_{\varphi_j} f(p),$$

is a derivation on $\mathcal{F}_p(M)$. Moreover, $\left(\partial_{\varphi_j}\big|_p\right)_{j\in\{1,\dots,n\}}$ form the *canonical basis* of T_pM with respect to the coordinates induced by φ. With respect to the canonical basis a derivation $X \in T_pM$ can be represented as

$$X = \sum_{j=1}^{n} X([\varphi_j]_{\sim_p})\partial_{\varphi_j}\big|_p.$$

In $\mathbb{R}^n$, considered as an n-manifold with the atlas generated by the global $\mathrm{id}_{\mathbb{R}^n}$-chart, derivations with a coefficient vector $v = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^\top$ with respect to the canonical basis can be identified with $\partial_v\big|_p$, the directional derivative at p with direction $v \in \mathbb{R}^n$.

A.4 Definition (Tangent bundle)**.** We denote by

$$TM := \biguplus_{p\in M} T_pM := \bigcup_{p\in M} \{p\} \times T_pM$$

the *total space of M* and

$$\pi_M\colon TM \to M, \qquad\qquad (p, X) \mapsto p$$

the *bundle projection*. The triple (TM, π_M, M) is referred to as the *tangent bundle of M*.

We identify an element $(p, X) \in TM$ with the derivation X. If M is an n-manifold, then TM is a $2n$-manifold.

A.5 Definition (Vector field)**.** A differentiable function $X\colon M \to TM$ with $\pi_M \circ X = \mathrm{id}_M$ is called a *vector field on* M. The set of all vector fields on M is denoted by $\mathcal{X}(M)$.

We denote by $\mathcal{F}(M)$ the set of all differentiable real-valued functions on M. This together with the pointwise addition and multiplication forms a commutative unital ring. Moreover, $\mathcal{F}(M)$ and $\mathcal{X}(M)$ are vector spaces over $\mathbb{R}$. Additionally, we can consider $\mathcal{X}(M)$ as an $\mathcal{F}(M)$-module with the following "scalar"-multiplication:

$$(fX)(p) := f(p)X(p),$$

where $X \in \mathcal{X}(M)$ and $f \in \mathcal{F}(M)$.

We can consider a given vector field $X \in \mathcal{X}(M)$ equivalently as

$$\tilde{X}\colon \mathcal{F}(M) \to \mathcal{F}(M), \qquad\qquad f \mapsto (p \mapsto X([f]_{\sim_p})).$$

$\tilde{X}$ is sometimes referred to as the *extension of* X. However, we use X synonymously for X itself and its extension $\tilde{X}$.

A.6 Definition (Lie bracket)**.** For two vector fields $X, Y \in \mathcal{X}(M)$ we define the *Lie bracket* $[X, Y]$ by

$$[X, Y](f) := XYf - YXf,$$

where $f \in \mathcal{F}(M)$.

The Lie bracket $[X, Y]$ of two vector fields $X, Y \in \mathcal{X}(M)$ is a vector field again, i.e. $[X, Y] \in \mathcal{X}(M)$. Consequently, the Lie bracket can be established as a function $\mathcal{L} := [\cdot, \cdot]\colon \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$, and one speaks of $\mathcal{L}_X Y := \mathcal{L}(X, Y)$ as the Lie derivative of Y in the direction of X.

A.7 Definition (Differential/push-forward)**.** Let $F\colon M \to N$ be a differentiable function. We call

$$DF\colon TM \to TN, \qquad X \mapsto DF(X) = \left([f]_{\sim_{F(\pi_M(X))}} \mapsto X\left([f \circ F]_{\sim_{\pi_M(X)}}\right)\right)$$

the *push-forward of* f. For clarity, if we apply DF to some $X \in T_pM$ we will also write $DF\big|_p X$.

For the differential, we have the following chain rule: $D(G \circ F)\big|_p = DG\big|_{F(p)} \circ DF\big|_p$.

A.8 Definition (Immersion, submersion, embedding, submanifold)**.** Let $(M, \mathcal{A})$ be an m-manifold, $(N, \mathcal{B})$ an n-manifold and $F\colon M \to N$ be differentiable. If for any $\varphi \in \mathcal{A}$, $\psi \in \mathcal{B}$ and $p \in D(\varphi) \cap F^{-1}[D(\psi)]$ one has that $\partial(\psi \circ F \circ \varphi^{-1})(\varphi(p))$ is injective/surjective, then F is called an *immersion/submersion*. If F is an immersion and $F\colon M \to F[M]$ is a homeomorphism onto its image, then F is called *embedding*. If $M \subseteq N$ and $\mathrm{id}\colon M \to N$ is an embedding, then M is called an *(embedded) submanifold* of N.

For immersions we have the following property.

A.9 Lemma. *Let $F\colon M \to N$ be an immersion. Then DF is injective.*

The image $F[M] \subseteq N$ of an m-manifold M under an embedding F is a submanifold of the n-manifold N. Necessarily, we have in this case $m \leq n$. In the submersion case, we have the following result, also known as the *submersion theorem*.

A.10 Proposition. *Let M be an m-manifold, N an n-manifold and $F\colon M \to N$ be a submersion. If $y \in F[M]$, then $F^{-1}[\{y\}] \subset M$ is a submanifold of M of dimension $m - n$.*

A.2 Riemannian Manifolds

In the following, let M denote an n-manifold with (maximal) atlas $\mathcal{A}$.

A.11 Definition ((Riemannian) Metric, Riemannian manifold)**.** We call a family $g := \{g_p;\ p \in M\}$ of symmetric, positive-definite $\mathbb{R}$-bilinear forms $g_p\colon T_pM \times T_pM \to \mathbb{R}$, $p \in M$, a *Riemannian metric on M* if for all $\varphi \in \mathcal{A}$ and any $i, j \in \mathbb{N}_{\leq n}$ the functions

$$g_{\varphi,ij}\colon D(\varphi) \to \mathbb{R}, \qquad\qquad p \mapsto g_p(\partial_{\varphi_i}\big|_p, \partial_{\varphi_j}\big|_p) \tag{A.1}$$

are differentiable. (M, g) is called an n-dimensional *Riemannian manifold*.

A.12 Example. We consider $\mathbb{R}^n$ as an n-manifold and identify its tangent spaces with $\mathbb{R}^n$. Then $g = (g_p)_{p \in \mathbb{R}^n}$, with $g_p = \langle \cdot, \cdot \rangle$ the Euclidean scalar product, is a Riemannian metric on $\mathbb{R}^n$.

A.13 Definition (Affine connection)**.** An *affine connection on M* is a mapping

$$\nabla\colon \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$$

satisfying for all $f \in \mathcal{F}(M)$, $X, Y, Z \in \mathcal{X}(M)$:

(a) $\nabla(X, Y+Z) = \nabla(X,Y) + \nabla(X,Z)$ and $\nabla(X+Y, Z) = \nabla(X,Z) + \nabla(Y,Z)$ (biadditivity);

(b) $\nabla(fX, Y) = f\nabla(X,Y)$ ($\mathcal{F}(M)$-linearity in the first argument);

(c) $\nabla(X, fY) = (Xf)Y + f\nabla(X,Y)$.

We call $\nabla(X,Y)$ the *covariant derivative of Y along X with respect to ∇*. An affine connection ∇ is called *torsion-free* or *symmetric* if additionally for all $X, Y \in \mathcal{X}(M)$ the equation $\nabla(X,Y) - \nabla(Y,X) = [X,Y]$ holds.

A.14 Example. Let $U \subseteq \mathbb{R}^n$ be open, $X, Y \in \mathcal{X}(U)$ with the canonical representation $X = \sum_{i=1}^n f_i\partial_i$ and $Y = \sum_{j=1}^n g_j\partial_j$. Then ∇ defined as

$$\nabla(X,Y) := \sum_{j=1}^n X(g_j)\partial_j = \sum_{i,j=1}^n f_i(\partial_i g_j)\partial_j$$

is an affine connection on U.

The *gradient* of some $f \in \mathcal{F}(M)$ with respect to a Riemannian metric g on M, denoted by $\operatorname{grad}_g f$, is defined as the unique vector field satisfying $g(\operatorname{grad}_g f, X) = \nabla(X, f) := Xf$ for any $X \in \mathcal{X}(M)$. Obviously, $\nabla f := \nabla(\cdot, f)$ as a covector/1-form and $\operatorname{grad}_g f \in \mathcal{X}(M)$ are different objects. However, in $\mathbb{R}^n$ endowed with the standard structure, i.e. with the atlas generated by the chart $\mathrm{id}_{\mathbb{R}^n}$, they can be identified by transposition.

A.15 Definition (Curve in M, regular)**.** Let $a, b \in \mathbb{R}$ with $a < b$. A differentiable function $c\colon (a,b) \to M$ is called a *curve in M*. We define $\dot{c}\colon (a,b) \to TM$ by

$$\dot{c}(t)\left([f]_{\sim_{c(t)}}\right) := Dc\big|_t\left(\partial\big|_t\right)\left([f]_{\sim_{c(t)}}\right) = (f \circ c)'(t),$$

where $t \in (a,b)$ and $[f]_{\sim_{c(t)}} \in \mathcal{F}_{c(t)}(M)$. By Definition A.7 we have $\dot{c}(t) \in T_{c(t)}M$. The curve c is called *regular* if $\dot{c}(t) \neq 0$ for all $t \in (a,b)$. If (M,g) is a Riemannian manifold then c is said to be *evenly parametrized* if for any $t \in (a,b)$ one has

$$g_{c(t)}(\dot{c}(t), \dot{c}(t)) = 1.$$

A.16 Theorem. *Let (M, g) be a Riemannian manifold. Then there exists a unique torsion-free affine connection ∇ on M, satisfying the* Ricci-identity*:*

$$\bigwedge_{X,Y,Z \in \mathcal{X}(M)} Xg(Y,Z) = g(\nabla(X,Y),Z) + g(Y,\nabla(X,Z)). \tag{A.2}$$

∇ is called Riemannian connection*. Furthermore, ∇ satisfies the* Koszul-formula*: for $X, Y, Z \in \mathcal{X}(M)$ one has*

$$\begin{aligned} g(\nabla(X,Y),Z)) = &\frac{1}{2}(Xg(Y,Z) + Yg(X,Z) - Zg(X,Y)) + \\ &+ \frac{1}{2}(g(Z,[X,Y]) + g(Y,[Z,X]) - g(X,[Y,Z])). \end{aligned} \tag{A.3}$$

The Riemannian connection induces a notion of parallelism as follows.

A.17 Definition (parallel)**.** Let (M, g) be a Riemannian n-manifold and ∇ the associated Riemannian connection. Then $Y \in \mathcal{X}(M)$ is said to be *parallel*, if $\nabla(X,Y) = 0$ holds for any $X \in \mathcal{X}(M)$.

A.18 Definition (Vector field along a curve)**.** Let M, N be differentiable manifolds and $c\colon I \subseteq \mathbb{R} \to M$ be a regular curve. A differentiable function $X\colon I \to TM$ satisfying $\pi_M \circ X = c$ is called a *(differentiable) vector field along c*. We denote by $\mathcal{X}_c(M)$ the set of vector fields along c, and $\mathcal{X}_c(M)$ is canonically an $\mathcal{F}(I)$-module: for $f \in \mathcal{F}(I)$ and $X \in \mathcal{X}_c(M)$ we have that fX with $(fX)(t) := f(t)X(t)$, $t \in I$, is an element of $\mathcal{X}_c(M)$.

A.19 Example. By Definition A.15 we have that $\dot{c}$ is a vector field along c.

Next, we want to define what it means that a vector field is parallel along a curve. To this end, we observe the following. Let (M, g) be a Riemannian manifold with affine connection $\nabla\colon \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ and $c\colon I := (a,b) \subseteq \mathbb{R} \to M$ be a regular curve. Then $\nabla(\cdot, Y)$ is a tensor field of type $(1,1)$ for each $Y \in \mathcal{X}(M)$. By a representation result for tensor fields there exists for any $t \in I$ a unique mapping $(\nabla(\cdot,Y))_{c(t)}\colon T_{c(t)} \to T_{c(t)}$ such that

$$\nabla(X,Y)(c(t)) = (\nabla(\cdot,Y))_{c(t)}(X(c(t))).$$

We introduce the notation

$$\nabla(\dot{c}(t),Y) := (\nabla(\cdot,Y))_{c(t)}(\dot{c}(t)) \in T_{c(t)}M$$

for $t \in I$ as well as $\nabla(\dot{c},Y) := (t \mapsto \nabla(\dot{c}(t),Y)) \in \mathcal{X}_c(M)$. Thus, the affine connection ∇ induces a covariant derivative along vector fields along regular curves on $\mathcal{X}(M)$ and we can establish for any $Y \in \mathcal{X}(M)$ the following function:

$$\nabla(\cdot, Y)\colon \mathcal{X}_c(M) \to \mathcal{X}_c(M), \quad X \mapsto \Big(t \mapsto (\nabla(\cdot, Y))_{c(t)}(X(c(t)))\Big) =: \nabla(X, Y).$$

Our next aim is to define an analogous function for $Y \in \mathcal{X}_c(M)$.

A.20 Definition (Covariant derivative of vector fields along curves, cf. [24, Def. 10.1.10]**).** Let (M, g) be a Riemannian manifold with affine connection ∇ and $c\colon I := (a, b) \subseteq \mathbb{R} \to M$ be a regular curve. An *associated covariant derivative along* c is a mapping

$$\underset{c}{\nabla}\colon \mathcal{X}_c(M) \to \mathcal{X}_c(M),$$

satisfying

(a) $\underset{c}{\nabla}$ is additive,

(b) for any $f \in \mathcal{F}(I)$ and $X \in \mathcal{X}_c(M)$ we have $\underset{c}{\nabla}(fX) = \dot{f}X + f\underset{c}{\nabla}X$,

(c) for any $Y \in \mathcal{X}(M)$ and $t \in I$ we have $\underset{c}{\nabla}(Y \circ c)(t) = \nabla(\dot{c}(t), Y)$.

A.21 Lemma. *Let (M, g) be a Riemannian n-manifold and c be a regular curve in M. For each affine connection ∇ on M there exists a unique covariant derivative along c associated to ∇.*

Proof. Uniqueness: First, we prove local uniqueness. To this end, let $\varphi \in \mathcal{A}$ and $\partial_1, \ldots, \partial_n$ be the local basis vector fields (with respect to φ) on $D(\varphi)$ and $X \in \mathcal{X}_c(M)$. Then X and $\dot{c}$ can be represented for $t \in I$ as

$$X(t) = \sum_{i=1}^{n} x_i(t)(\partial_i \circ c)(t), \qquad \dot{c}(t) = \sum_{j=1}^{n} \dot{c}_j(t)(\partial_j \circ c)(t).$$

Note that $\partial_i \in \mathcal{X}(D(\varphi))$ can be extended to a differentiable vector field on M. By definition we have

$$\begin{aligned}
\underset{c}{\nabla} X &= \underset{c}{\nabla}\left(\sum_{i=1}^{n} x_i(\partial_i \circ c)\right) = \sum_{i=1}^{n}\Big(\dot{x}_i(\partial_i \circ c) + x_i \underset{c}{\nabla}(\partial_i \circ c)\Big) \\
&= \sum_{i=1}^{n}\left(\dot{x}_i(\partial_i \circ c) + x_i \nabla(\dot{c}, \partial_i)\right) \\
&= \sum_{i=1}^{n}\left(\dot{x}_i(\partial_i \circ c) + x_i \nabla\left(\sum_{j=1}^{n} \dot{c}_j(\partial_j \circ c), \partial_i\right)\right) \\
&= \sum_{i=1}^{n}\left(\dot{x}_i(\partial_i \circ c) + \sum_{j,k=1}^{n} x_i \dot{c}_j \left(\Gamma_{ji}^{k} \circ c\right)(\partial_k \circ c)\right) \\
&= \sum_{k=1}^{n}\left(\dot{x}_k(\partial_k \circ c)\right) + \sum_{i,j,k=1}^{n}\Big(x_i \dot{c}_j \left(\Gamma_{ji}^{k} \circ c\right)(\partial_k \circ c)\Big) \\
&= \sum_{k=1}^{n}\left(\dot{x}_k + \sum_{i,j=1}^{n} x_i \dot{c}_j \left(\Gamma_{ji}^{k} \circ c\right)\right)(\partial_k \circ c) \in \mathcal{X}_c(M),
\end{aligned} \tag{A.4}$$

with obvious evaluation at $t \in I$. Here, Γ_{ji}^{k} denote the Christoffel symbols with respect to φ. Hence, the covariant derivative along c is uniquely determined by the local coefficients of $\dot{c}$ and X as well as by the Christoffel symbols. Uniqueness of $\underset{c}{\nabla}$ now follows from their uniqueness.

Existence: By the above formula we can define a covariant derivative along c locally, and the 3 requirements are easily checked. From uniqueness follows that in regions of overlapping domains of charts $D(\varphi) \cap D(\psi)$ for two charts $\varphi, \psi \in \mathcal{A}$ the two sets of Christoffel symbols define the same covariant derivative along c, such that the definitions can be matched together along the whole curve c. $\square$

A.22 Definition (Parallelism along curves, geodesics)**.** Let (M, g) be a Riemannian manifold with affine connection ∇ and $c\colon I := (a, b) \subseteq \mathbb{R} \to M$ be a regular curve. Then $Y \in \mathcal{X}(M)$ and $X \in \mathcal{X}_c(M)$ are said to be *parallel along c* if $\nabla(\dot{c}, Y) = 0$ and $\underset{c}{\nabla} X = 0$, respectively. A regular curve c in M is called *geodesic* if $\dot{c}$ is parallel along c, i.e. $\underset{c}{\nabla}\dot{c} = 0$.

A.23 Theorem (Existence and uniqueness of parallel vector fields)**.** *Let M be an n-manifold with affine connection ∇, $c\colon I \subset \mathbb{R} \to M$ be a regular curve in M, $t_0 \in I$ and $V \in T_{c(t_0)}M$. Then there exists a unique vector field $X \in \mathcal{X}_c$ along c such that $X(t_0) = V$ and X is parallel along c.*

For convenience, we give the full proof.

Proof. First, let $c[I] \subseteq D(\varphi)$ for some $\varphi \in \mathcal{A}$. By the defining formula (A.4) we have that $X = \sum_{i=1}^{n} x_{i,\varphi}(\partial_{\varphi_i} \circ c) \in \mathcal{X}_c$ is parallel along c if and only if

$$\sum_{k=1}^{n} \left(\dot{x}_{k,\varphi} + \sum_{i,j=1}^{n} x_{i,\varphi} \dot{c}_{j,\varphi} \left(\Gamma_{ji}^{k} \circ c \right) \right) (\partial_{\varphi_k} \circ c) = 0.$$

Due to the linear independence of $(\partial_{\varphi_k} \circ c)_{k\in\{1,\dots,n\}}$, this is equivalent to vanishing coefficients, i.e. for $k \in \{1,\dots,n\}$ we have

$$\dot{x}_{k,\varphi}(t) + \sum_{i,j=1}^{n} \dot{c}_{j,\varphi}(t) \Gamma_{ji}^{k}(c(t)) x_{i,\varphi}(t) = 0,$$

or rewritten in a system

$$\begin{pmatrix} \dot{x}_{1,\varphi} \\ \vdots \\ \dot{x}_{n,\varphi} \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{n} \dot{c}_{j,\varphi}(\Gamma_{j1}^{1} \circ c) & \dots & \sum_{j=1}^{n} \dot{c}_{j,\varphi}(\Gamma_{jn}^{1} \circ c) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} \dot{c}_{j,\varphi}(\Gamma_{j1}^{n} \circ c) & \dots & \sum_{j=1}^{n} \dot{c}_{j,\varphi}(\Gamma_{jn}^{n} \circ c) \end{pmatrix} \begin{pmatrix} x_{1,\varphi} \\ \vdots \\ x_{n,\varphi} \end{pmatrix} = 0. \tag{A.5}$$

This is a first-order linear homogeneous system of differential equations with C^∞-coefficients, which, considered as an initial value problem with initial value

$$\begin{pmatrix} x_{1,\varphi}(t_0) \\ \vdots \\ x_{n,\varphi}(t_0) \end{pmatrix} = \begin{pmatrix} V([\varphi_1]_{\sim_{c(t_0)}}) \\ \vdots \\ V([\varphi_n]_{\sim_{c(t_0)}}) \end{pmatrix} \tag{A.6}$$

is uniquely solvable. By setting

$$X_\varphi(t) := \sum_{k=1}^{n} x_{k,\varphi}(t) \partial_{\varphi_k}|_{c(t)}$$

for $t \in I$ we obtain a vector field, which is parallel along c and satisfies the initial value condition. It remains to show independence of the local chart. Having this, we can omit the initial requirement $c[I] \subseteq D(\varphi)$ and obtain the vector field for all $t \in I$ by sticking together the pieces.

Without loss of generality assume $t_0 \in c^{-1}[D(\varphi) \cap D(\psi)]$, $\psi \in \mathcal{A}$, and solve Eq. (A.5) with ψ instead of φ with the corresponding initial value as in Eq. (A.6). For $t \in I$ with $c(t) \in D(\psi)$ we set

$$X_\psi(t) := \sum_{k=1}^{n} x_{k,\psi}(t) \partial_{\psi_k}|_{c(t)}.$$

We want to show that $X_\psi(t) = X_\varphi(t)$ holds for all $t \in I$ with $c(t) \in D(\varphi) \cap D(\psi)$. Observe that the coefficients $x_{i,\varphi}$ are, by Eq. (A.4), the unique solutions of

$$\sum_{i=1}^{n} \Big(\dot{x}_{i,\varphi}(\partial_{\varphi_i} \circ c) + x_{i,\varphi} \underset{c}{\nabla} (\partial_{i,\varphi} \circ c) \Big) = 0 \tag{A.7}$$

with initial value (A.6). For $t \in [D(\varphi) \cap D(\psi)]c$ we have

$$\begin{aligned} X_\psi(t) &= \sum_{j=1}^{n} x_{j,\psi}(t) \sum_{k=1}^{n} \partial_j(\varphi_k \circ \psi^{-1})(\psi(c(t)))\partial_{\varphi_k}\big|_{c(t)} \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} x_{j,\psi}(t)\partial_j(\varphi_k \circ \psi^{-1})(\psi(c(t))) \right) \partial_{\varphi_k}\big|_{c(t)}. \end{aligned}$$

Now, for $k \in \{1, \dots, n\}$ we set

$$\chi_k(t) := \sum_{j=1}^{n} x_{j,\psi}(t)\partial_j(\varphi_k \circ \psi^{-1})(\psi(c(t))).$$

If we can show that $(\chi_1, \dots, \chi_n)$ also satisfy Eq. (A.7) with initial value (A.6), then by the uniqueness of solutions follows $\chi_k(t) = x_{k,\varphi}(t)$ and consequently $X_\psi(t) = X_\varphi(t)$ for all $t \in c^{-1}[D(\varphi) \cap D(\psi)]$. The initial value condition is satisfied by

$$\chi_k(t_0) = \sum_{j=1}^{n} V([\psi_j]_{\sim_{c(t_0)}})\partial_j(\varphi_k \circ \psi^{-1})(\psi(c(t_0))) = V([\varphi_k]_{\sim_{c(t_0)}}) = x_{k,\varphi}(t_0).$$

To check the second condition we calculate

$$\begin{aligned} &\sum_{i=1}^{n} \dot{\chi}_i(\partial_{\varphi_i} \circ c) + \chi_i \underset{c}{\nabla} (\partial_{\varphi_i} \circ c) = \\ &\quad= \sum_{i=1}^{n}\sum_{j=1}^{n} \Big(\dot{x}_{j,\psi}\partial_j(\varphi_i \circ \psi^{-1})(\psi \circ c) + x_{j,\psi}(\partial_j(\varphi_i \circ \psi^{-1})(\psi \circ c))' \Big) \partial_{\varphi_i} \circ c \\ &\qquad + \sum_{i=1}^{n}\sum_{j=1}^{n} x_{j,\psi}\partial_j(\varphi_i \circ \psi^{-1})(\psi \circ c) \underset{c}{\nabla} (\partial_{\varphi_i} \circ c) \\ &\quad= \sum_{j=1}^{n} \dot{x}_{j,\psi}(\partial_{\psi_j} \circ c) \\ &\qquad + \sum_{j=1}^{n} x_{j,\psi} \sum_{i=1}^{n} \Big((\partial_j(\varphi_i \circ \psi^{-1})(\psi \circ c))'\partial_{\varphi_i} \circ c + \partial_j(\varphi_i \circ \psi^{-1})(\psi \circ c) \underset{c}{\nabla} (\partial_{\varphi_i} \circ c) \Big) \\ &\quad= \sum_{j=1}^{n} \dot{x}_{j,\psi}(\partial_{\psi_j} \circ c) + x_{j,\psi} \underset{c}{\nabla} \left(\sum_{i=1}^{n} \partial_j(\varphi_i \circ \psi^{-1})(\psi \circ c)\partial_{\varphi_i} \circ c \right) \end{aligned}$$

$$= \sum_{j=1}^{n} \dot{x}_{j,\psi}(\partial_{\psi_j} \circ c) + x_{j,\psi} \underset{c}{\nabla} \left(\partial_{\psi_j} \circ c\right)$$
$$= 0.$$

This finishes the proof. □

The vector field X in the previous theorem is also said to be *parallel transported along* c.

A.24 Theorem. *Let (M, g) be a Riemannian n-manifold with Riemannian connection ∇ and c be a regular curve in M. Then the following statements hold:*

(a) For any $p \in M$ and $X \in T_pM$ with $g_p(X, X) = 1$ there exist $\varepsilon \in \mathbb{R}_{>0}$ and a unique, evenly parametrized geodesic $c\colon (-\varepsilon, \varepsilon) \to M$ such that $c(0) = p$ and $\dot{c}(0) = X$. We denote this curve by $c_X^{(p)}\colon (-\varepsilon, \varepsilon) \to M$.

(b) Let $p \in M$ and $\mathcal{S}(T_pM) = \left\{X \in T_pM;\ g_p(X, X) = 1\right\}$ be the unit sphere in T_pM. Then there exists $\varepsilon \in \mathbb{R}_{>0}$ such that the mapping

$$\exp_p\colon (-\varepsilon, \varepsilon)[\mathcal{S}(T_pM)] \subseteq T_pM \to M, \qquad tV \mapsto \exp_p(tV) := c_V^{(p)}(t),$$

is injective and differentiable. It is called the exponential map in p.

(c) Let $p \in M$. Then there exist $\varepsilon \in \mathbb{R}_{>0}$ and an open neighborhood $U \subseteq M$ of p, such that $\exp_p\colon B(0, \varepsilon) \subseteq T_pM \to U \subseteq M$ is a diffeomorphism. Hence, the exponential map defines a local parametrization around p and its inverse $\exp_p^{-1}\colon U \subseteq M \to T_pM \cong \mathbb{R}^n$ defines a chart around p, where $\exp_p^{-1}(p) = 0$. The associated coordinates are referred to as (Riemannian) normal coordinates.

Observe that the normal coordinates can be further aligned to tangent and normal spaces of embedded submanifolds , for instance, if necessary, by composing $A \circ \exp_p^{-1}$, where $A \in GL(n, \mathbb{R})$.

B Grassmann Manifolds

In the following we give a self-consistent introduction to Grassmann manifolds and their representation by matrix manifolds. The presentation is based on [42, 2] and all results are taken from these references if not stated otherwise.

Throughout, let $n \in \mathbb{N}_{>0}$ and $k \in \{1, \dots, n\}$. The object of investigation is the set of all k-dimensional linear subspaces $V \subseteq \mathbb{R}^n$, i.e.

$$\mathrm{Gr}(k, \mathbb{R}^n) := \{V \subseteq \mathbb{R}^n;\ V \text{ is a linear subspace},\ \dim V = k\}\,.$$

The objective is to establish $\mathrm{Gr}(k, \mathbb{R}^n)$ as a Riemannian (differentiable) manifold.

B.1 Topological & Metric Structure

In this section, we follow [42]. We consider $A \in \mathbb{R}^{n\times k}$ as

$$A = (a_{ij})_{i\in\{1,\dots,n\},j\in\{1,\dots,k\}} = \begin{pmatrix} a_1 & \cdots & a_k \end{pmatrix},$$

where $a_i = (a_{ji})_{j\in\{1,\dots,n\}} \in \mathbb{R}^n$, $i \in \{1, \dots, k\}$. We endow $\mathbb{R}^{n\times k}$ with the norm induced by the Frobenius inner product, i.e.

$$\|A\| = \sqrt{\mathrm{tr}(A^\top A)} = \left|\begin{pmatrix} |a_1|_2 & \cdots & |a_k|_2 \end{pmatrix}^\top\right|_2 = \sqrt{\sum_{i,j} a_{ij}^2}. \tag{B.1}$$

We define the *(non-compact) Stiefel manifold* as

$$\begin{aligned}\mathrm{St}(k, \mathbb{R}^n) &:= \left\{A \in \mathbb{R}^{n\times k};\ \mathrm{rk}(A) = k\right\}\\ &= \left\{A = \begin{pmatrix} a_1 & \cdots & a_k \end{pmatrix} \in \mathbb{R}^{n\times k};\ a_1, \dots, a_k \text{ are linearly independent}\right\}.\end{aligned}$$

$\mathrm{St}(k, \mathbb{R}^n)$ is often referred to as the set of all *k-frames* (of linearly independent vectors) in $\mathbb{R}^n$. Note that $\mathrm{St}(k, \mathbb{R}^n)$ is an open subset of $\mathbb{R}^{n\times k}$ (as the preimage of the open set $\mathbb{R} \setminus \{0\}$ under the continuous map $A \mapsto \det(A^\top A)$). We define the *compact Stiefel manifold* as

$$\mathrm{St}^*(k, \mathbb{R}^n) := \left\{ A \in \mathrm{St}(k, \mathbb{R}^n);\ A^\top A = I_k \right\}.$$

Clearly, $\mathrm{St}^*(k, \mathbb{R}^n)$ is a bounded (by $\sqrt{k}$) and closed (as the preimage of the closed set $\{I_k\}$ under the continuous map $A \mapsto A^\top A$) and hence compact subset of $\mathrm{St}(k, \mathbb{R}^n)$ and $\mathrm{St}^*(k, \mathbb{R}^n) \hookrightarrow \mathrm{St}(k, \mathbb{R}^n)$. It is often referred to as the set of all *orthonormal k-frames* of linearly independent vectors in $\mathbb{R}^n$.

We consider the map

$$\pi\colon \mathrm{St}(k, \mathbb{R}^n) \to \mathrm{Gr}(k, \mathbb{R}^n), \qquad A = \begin{pmatrix} a_1 & \cdots & a_k \end{pmatrix} \mapsto \mathrm{span}\{a_1, \dots, a_k\}, \tag{B.2}$$

and endow $\mathrm{Gr}(k, \mathbb{R}^n)$ with the final topology with respect to π, which we call the *Grassmann topology*, i.e. $U \subseteq \mathrm{Gr}(k, \mathbb{R}^n)$ is open if and only if $\pi^{-1}[U]$ is open in $\mathrm{St}(k, \mathbb{R}^n)$. Let $f \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^n)$ be injective and $A \in \mathrm{St}(k, \mathbb{R}^n)$ be the matrix representation with respect to the canonical bases in $\mathbb{R}^k$ and $\mathbb{R}^n$. Then $\pi(A)$ can be equivalently considered as $f[\mathbb{R}^k]$, i.e. the image of f.

B.1 Lemma. *Consider* $\mathrm{Gr}(k, \mathbb{R}^n)$ *with the Grassmann topology.*

(i) The map π is surjective and for any $A \in \mathrm{St}(k, \mathbb{R}^n)$ we have

$$\pi^{-1}[\pi(A)] = \{AP;\ P \in GL(k, \mathbb{R})\}. \tag{B.3}$$

(ii) The map π is continuous and open.

Proof. Obviously, π is surjective and continuous. Furthermore, $A, B \in \mathrm{St}(k, \mathbb{R}^n)$ span the same subspace, i.e. $\pi(A) = \pi(B)$ if and only if there exists a $P \in GL(k, \mathbb{R})$ such that $B = AP$. To see the sufficiency note that by assumption we have $A = BP$ and $B = AQ = BPQ$ for some $P, Q \in \mathbb{R}^{k \times k}$, from which we directly read off that $P = Q^{-1} \in GL(k, \mathbb{R})$. On the other hand, the necessity is obvious. For an open subset $V \subseteq \mathrm{St}(k, \mathbb{R}^n)$ we have $\pi^{-1}[\pi[V]] = \bigcup_{P \in GL(k,\mathbb{R})} [V]P$, i.e. $\pi^{-1}[\pi[V]]$ is a union of open subsets. Thus, $\pi[V] \subseteq \mathrm{Gr}(k, \mathbb{R}^n)$ is open by the definition of the Grassmann topology. $\square$

Eq. (B.3) identifies $\mathrm{Gr}(k, \mathbb{R}^n)$ with the quotient space

$$\mathrm{St}(k, \mathbb{R}^n)/GL(k, \mathbb{R}) := \{A[GL(k, \mathbb{R})];\ A \in \mathrm{St}(k, \mathbb{R}^n)\}, \tag{B.4}$$

which can be endowed with the structure of a quotient manifold; see [3, Section 3.4]. However, we define a topological and differentiable structure on $\mathrm{Gr}(k, \mathbb{R}^n)$ directly.

We consider the restriction of π to $\mathrm{St}^*(k, \mathbb{R}^n)$, i.e. $\overline{\pi} := \pi|_{\mathrm{St}^*(k,\mathbb{R}^n)}$, and the "Gram-Schmidt orthonormalization map", i.e. $GS(A)$ is the matrix obtained by the application of the Gram-Schmidt orthonormalization method to the columns $a_1, \ldots, a_n$ of A. It is well-known that the Gram-Schmidt method defines a continuous map. Thus, we have $\overline{\pi} \circ GS = \pi$. Clearly, $\overline{\pi}$ is surjective and continuous.

Next, we show a characterization of the continuity of functions f defined on a topological space endowed with the final topology with respect to some other function and mapping to another topological space.

B.2 Lemma. *Consider* $\mathrm{Gr}(k, \mathbb{R}^n)$ *with the Grassmann topology. Let* Ω *be a topological space and*

$$f\colon \mathrm{Gr}(k, \mathbb{R}^n) \to \Omega.$$

Then the following statements are equivalent:

(*a*) *f is continuous;* (*b*) *$f \circ \pi$ is continuous;* (*c*) *$f \circ \overline{\pi}$ is continuous.*

Proof. (a) $\Leftrightarrow$ (b) follows from the definition of final topology with respect to π.

(b) $\Rightarrow$ (c) is obvious since $\overline{\pi}$ is the restriction of π.

(c) $\Rightarrow$ (b) follows from the representation $f \circ \pi = f \circ \overline{\pi} \circ GS$ and the continuity of GS. □

One can show that the set

$$\Delta := \left\{(A, B) \in \mathrm{St}(k, \mathbb{R}^n)^2;\ \exists\, P \in GL(k, \mathbb{R}): \quad AP = B\right\}$$

is closed in $\mathrm{St}(k, \mathbb{R}^n) \times \mathrm{St}(k, \mathbb{R}^n)$ by realizing that $(A, B) \in \Delta$ if and only if $\mathrm{rk}\begin{pmatrix} A & B \end{pmatrix} = k$. This condition can be characterized by the requirement that all (continuous) $k+1$ minors of $\begin{pmatrix} A & B \end{pmatrix}$ vanish. The set of matrices satisfying this condition is hence closed. Since $\mathrm{Gr}(k, \mathbb{R}^n) = \overline{\pi}[\mathrm{St}^*(k, \mathbb{R}^n)]$, $\overline{\pi}$ is continuous and $\mathrm{St}^*(k, \mathbb{R}^n)$ compact, we have that $\mathrm{Gr}(k, \mathbb{R}^n)$ is compact, too, and it remains to prove that the Grassmann topology is Hausdorff. To this end, let $L, M \in \mathrm{Gr}(k, \mathbb{R}^n)$ with $L \neq M$, $A \in \pi^{-1}[\{L\}]$, $B \in \pi^{-1}[\{M\}]$. By Lemma B.1 we have $(A, B) \notin \Delta$ and since Δ is closed, there exist neighborhoods U_A and U_B of A and B, respectively, such that $(U_A \times U_B) \cap \Delta = \varnothing$. By Lemma B.1 $\pi[U_A]$ and $\pi[U_B]$ are disjoint open sets containing $\pi(A) = L$ and $\pi(B) = M$. Summarizing, we obtain the following result.

B.3 Proposition. $\mathrm{Gr}(k,\mathbb{R}^n)$ *is a compact space with respect to the Grassmann topology.*

Often, $\mathrm{Gr}(k,\mathbb{R}^n)$ is endowed and investigated with the so-called *gap metric*, which we introduce next.

B.4 Definition (Gap metric)**.** We define

$$\Theta\colon \mathrm{Gr}(k,\mathbb{R}^n)\times\mathrm{Gr}(k,\mathbb{R}^n)\to\mathbb{R}_{\geq 0}, \qquad (L,M)\mapsto \|\Pi_L-\Pi_M\|, \tag{B.5}$$

where Π_X denotes the orthogonal projection on a subspace X and $\|\cdot\|$ denotes the operator norm. Θ is called the *gap metric* on $\mathrm{Gr}(k,\mathbb{R}^n)$.

In the literature, $\Theta(L,M)$ is also referred to as the *opening* between the subspaces L and M and was originally introduced in [72]. By the Projection Theorem for Hilbert spaces, there is a one-to-one correspondence between (closed) subspaces U and orthogonal projections Π_U with $\operatorname{im}\Pi_U = U$ and $\ker\Pi_U = U^\perp$. Therefore, it is easy to see that Θ is indeed a metric on $\mathrm{Gr}(k,\mathbb{R}^n)$. Next we show some further properties of the gap metric.

B.5 Proposition ([48, Theorems 13.1.1 & 13.1.2])**.** *For* $L,M\in\mathrm{Gr}(k,\mathbb{R}^n)$

(i) one has

$$\Theta(L,M)=\max\left(\sup_{x\in\mathcal{S}\cap L} d(x,M),\ \sup_{x\in\mathcal{S}\cap M} d(x,L)\right), \tag{B.6}$$

where $d(x,M):=\inf\{|x-y|\,;\,y\in M\}$;

(ii) $\Theta(L,M)=\Theta(L^\perp,M^\perp)\leq 1$;

(iii) $\Theta(L,M)<1 \quad\Leftrightarrow\quad L\cap M^\perp = L^\perp\cap M=\{0\} \quad\Leftrightarrow\quad \mathbb{R}^n = L\oplus M^\perp = L^\perp\oplus M$.

Proof. (i) Let $x\in L\cap\mathcal{S}$. Then $\Pi_M x\in M$ and it follows

$$|x-\Pi_M x| = |(\Pi_L-\Pi_M)x| \leq \|\Pi_L-\Pi_M\|.$$

Therefore,

$$\sup_{x\in\mathcal{S}\cap L} d(x,M)\leq\|\Pi_L-\Pi_M\|.$$

Analogously we have

$$\sup_{x\in\mathcal{S}\cap M} d(x,L)\leq\|\Pi_L-\Pi_M\|$$

and we obtain

$$\max\left(\sup_{x\in\mathcal{S}\cap L} d(x,M),\ \sup_{x\in\mathcal{S}\cap M} d(x,L)\right) \le \|\Pi_L - \Pi_M\| = \Theta(L,M).$$

To derive the complementary inequality, observe that by the Projection Theorem in Hilbert spaces we have

$$\begin{aligned}
\rho_L &:= \sup_{x\in\mathcal{S}\cap L} d(x,M) = \sup_{x\in\mathcal{S}\cap L} |x-\Pi_M x| = \sup_{x\in\mathcal{S}\cap L} |(I-\Pi_M)x|\,,\\
\rho_M &:= \sup_{x\in\mathcal{S}\cap M} d(x,L) = \sup_{x\in\mathcal{S}\cap M} |x-\Pi_L x| = \sup_{x\in\mathcal{S}\cap M} |(I-\Pi_L)x|\,.
\end{aligned}$$

Consequently, we have for every $x \in \mathbb{R}^n$ that

$$|(I-\Pi_L)\Pi_M x| \le \rho_M\,|\Pi_M x|\,, \qquad \text{and} \qquad |(I-\Pi_M)\Pi_L x| \le \rho_L\,|\Pi_L x|\,. \tag{B.7}$$

We estimate

$$\begin{aligned}
|\Pi_M(I-\Pi_L)x|^2 &= \langle \Pi_M(I-\Pi_L)x, \Pi_M(I-\Pi_L)x\rangle\\
&= \langle \Pi_M(I-\Pi_L)x, (I-\Pi_L)x\rangle\\
&= \langle \Pi_M(I-\Pi_L)x, (I-\Pi_L)(I-\Pi_L)x\rangle\\
&= \langle (I-\Pi_L)\Pi_M(I-\Pi_L)x, (I-\Pi_L)x\rangle\\
&\le |((I-\Pi_L)\Pi_M(I-\Pi_L)x)|\,|(I-\Pi_L)x|\,,
\end{aligned}$$

thus by Eq. (B.7)

$$|\Pi_M(I-\Pi_L)x|^2 \le \rho_M\,|\Pi_M(I-\Pi_L)x|\,|(I-\Pi_L)x|$$

and hence

$$|\Pi_M(I-\Pi_L)x| \le \rho_M\,|(I-\Pi_L)x|\,. \tag{B.8}$$

On the other hand, using

$$\Pi_M - \Pi_L = \Pi_M(I-\Pi_L) - (I-\Pi_M)\Pi_L$$

and the orthogonality of Π_M, we obtain by Eqs. (B.8) and (B.7)

$$\begin{aligned}
|(\Pi_M-\Pi_L)x|^2 &= \langle (\Pi_M(I-\Pi_L) - (I-\Pi_M)\Pi_L)\,x, (\Pi_M(I-\Pi_L) - (I-\Pi_M)\Pi_L)\,x\rangle\\
&= |\Pi_M(I-\Pi_L)x|^2 + |(I-\Pi_M)\Pi_L x|^2\\
&\le \rho_M^2\,|(I-\Pi_L)x|^2 + \rho_L^2\,|\Pi_L x|^2\\
&\le \max\left\{\rho_M^2, \rho_L^2\right\}\,|x|^2\,.
\end{aligned}$$

Thus, we obtain

$$\|\Pi_L - \Pi_M\| \leq \max\{\rho_M, \rho_L\},$$

which finishes the proof of (a).

(ii) The equality $\Theta(L,M) = \Theta(L^\perp, M^\perp)$ follows directly from the definition, since $\Pi_{L^\perp} = I - \Pi_L$. The estimate $\Theta(L,M) \leq 1$ holds by the observation that for $x \in \mathcal{S} \cap L$ we have $d(x,M) = |\Pi_{M^\perp} x| = |x| - |\Pi_M x| \leq |x| = 1$.

(iii) First note that the second equivalence is clear by the definition of decompositions. The first one follows from the following observation.

B.6 Claim. Let $M \in \mathrm{Gr}(k, \mathbb{R}^n)$ and $x \in \mathcal{S}$. Then $x \in M^\perp$ if and only if $d(x,M) = 1$.

Proof of claim. The proof relies on the relation $|x|^2 = |\Pi_M x|^2 + |\Pi_{M^\perp} x|^2$. From the relation we directly read off that

$$1 = |x|^2 = |\Pi_{M^\perp} x|^2 = |(I - \Pi_M)x|^2 = |x - \Pi_M x|^2 = d(x,M).$$

Conversely, from $d(x,M) = |\Pi_{M^\perp} x|^2 = |x|^2 = 1$ follows $|\Pi_M x| = 0$ and hence $x \in M^\perp$. □

Due to the continuity of the norm and the compactness of $\mathcal{S} \cap L$ the value of $\Theta(L,M)$ is attained, so from $\Theta(L,M) < 1$ follows $L \cap M^\perp = L^\perp \cap M = \{0\}$. On the other hand, from $\Theta(L,M) = 1$, which we (w.l.o.g.) assume to be attained by $\sup_{x \in \mathcal{S} \cap L} d(x,M) = 1$, follows that there exists some $z \in \mathcal{S} \cap L$ such that $d(z,M) = 1$ and by the claim $z \in M^\perp$ holds. □

As a consequence, we get that for any $L \in \mathrm{Gr}(k, \mathbb{R}^n)$

$$B[L,1] = \mathrm{Gr}(k,\mathbb{R}^n), \quad \text{and} \quad B(L,1) = \left\{ M \in \mathrm{Gr}(k,\mathbb{R}^n);\ \mathbb{R}^n = L \oplus M^\perp \right\}.$$

Our next aims are to prove that the gap metric on $\mathrm{Gr}(k, \mathbb{R}^n)$ is equivalent to the Hausdorff metric on the sections of two subspaces with the unit sphere (called the spherical gap), and second to show the equivalence of the Grassmann topology to the topology induced by the gap metric/Hausdorff metric.

B.7 Definition (Spherical gap metric, cf. also [67])**.** We define

$$\widetilde{\Theta} \colon \mathrm{Gr}(k,\mathbb{R}^n) \times \mathrm{Gr}(k,\mathbb{R}^n) \to \mathbb{R}_{\geq 0},$$
$$(L,M) \mapsto d_{\mathrm{H}}(\mathcal{S} \cap L, \mathcal{S} \cap M) = \max\left(\sup_{x \in \mathcal{S} \cap L} d(x, \mathcal{S} \cap M), \sup_{x \in \mathcal{S} \cap M} d(x, \mathcal{S} \cap L) \right),$$

where d_{H} denotes the Hausdorff metric defined on closed subsets of $\mathbb{R}^n$. $\widetilde{\Theta}$ is called the *spherical gap metric* on $\mathrm{Gr}(k, \mathbb{R}^n)$.

B.8 Lemma ([49]). *For any $L, M \in \mathrm{Gr}(k, \mathbb{R}^n)$ holds*

$$\Theta(L, M) \leq \widetilde{\Theta}(L, M) \leq 2\Theta(L, M).$$

Proof. We follow [67, p. 199]. The first inequality follows trivially from $d(x, M) \leq d(x, M \cap \mathcal{S})$, $x \in \mathbb{R}^n$. To show the second inequality it suffices to prove $d(x, M \cap \mathcal{S}) \leq 2d(x, M)$ for $x \in \mathcal{S}$. By the definition of infimum, for any $\varepsilon \in \mathbb{R}_{>0}$ there exists $y \in M \setminus \{0\}$ such that $|x - y| < d(x, M) + \varepsilon$. Define $y_0 := \frac{y}{|y|} \in \mathcal{S} \cap M$. Then the estimate $d(x, M \cap \mathcal{S}) \leq |x - y_0| \leq |x - y| + |y - y_0|$ holds. Furthermore, since y and y_0 are parallel, we have

$$|y - y_0| = \left| y - \frac{y}{|y|} \right| = ||y| - 1| = ||y| - |x|| \leq |y - x| \, .$$

In summary, we have $d(x, M \cap \mathcal{S}) < 2d(x, M) + 2\varepsilon$, and since ε is arbitrary, the prove is finished. $\square$

B.9 Corollary. *The topologies on $\mathrm{Gr}(k, \mathbb{R}^n)$ induced by Θ and $\widetilde{\Theta}$, respectively, coincide.*

B.10 Remark. For a sequence $(X_n)_{n \in \mathbb{N}} \in (\mathrm{Gr}(k, \mathbb{R}^n))^{\mathbb{N}}$ converging to $X \in \mathrm{Gr}(k, \mathbb{R}^n)$ with respect to Θ, i.e. $\|\Pi_{X_n} - \Pi_X\| \xrightarrow{n \to \infty} 0$, we have by the definition of Hausdorff metric that for any $x \in X$ there exists a sequence $(x_n)_{n \in \mathbb{N}} \in (\mathcal{S})^{\mathbb{N}}$ with $x_n \in X_n$ for each $n \in \mathbb{N}$ such that $x_n \xrightarrow{n \to \infty} x$, and hence with respect to any norm.

B.11 Remark. When Θ is introduced by Eq. (B.6) in the set of subspaces of general Banach spaces, Θ does not necessarily satisfy the triangle inequality and can therefore not be used to define a metric/topology. However, in Hilbert space one can show equality to the expression given in Eq. (B.5), which we have used for the definition and which obviously satisfies the conditions for a metric. On the other hand, $\widetilde{\Theta}$ defines a proper metric even in the general Banach space case; see, for instance, [67, IV.2.1] and the references therein.

Next we address our second aim, to show that the Grassmann topology coincides with the topology induced by the gap metric Θ (or equivalently $\widetilde{\Theta}$). To this end, we need to prove that the identity map from the topological space $\mathrm{Gr}(k, \mathbb{R}^n)$ with the Grassmann topology to the topological space $\mathrm{Gr}(k, \mathbb{R}^n)$ with the gap topology is continuous together with its inverse. Since $\mathrm{Gr}(k, \mathbb{R}^n)$ with the Grassmann topology is compact and the metric space $\mathrm{Gr}(k, \mathbb{R}^n)$ with the metric Θ is Hausdorff, the continuity of the inverse follows directly from the continuity of the identity map. Furthermore, by Lemma B.2 it suffices to show that $\overline{\pi}\colon \mathrm{St}^*(k, \mathbb{R}^n) \to (\mathrm{Gr}(k, \mathbb{R}^n), \Theta)$ is continuous.

B.12 Proposition. *The Grassmann topology in* $\mathrm{Gr}(k, \mathbb{R}^n)$ *coincides with the topologies induced by* Θ *and* $\widetilde{\Theta}$*, respectively.*

Proof. Let $A, B \in \mathrm{St}^*(k, \mathbb{R}^n)$ and $\overline{\pi}(A) = L$, $\overline{\pi}(B) = M$, i.e. $L, M \in \mathrm{Gr}(k, \mathbb{R}^n)$. It is well-known that the orthogonal projections onto L and M have the form $\sum_{i=1}^k \langle a_i, \cdot\rangle a_i$ and $\sum_{i=1}^k \langle b_i, \cdot\rangle b_i$, respectively. Now we estimate

$$
\begin{aligned}
\Theta(\overline{\pi}(A), \overline{\pi}(B)) &= \Theta(L, M) = \|\Pi_L - \Pi_M\| \\
&= \left\| \sum_{i=1}^k \langle a_i, \cdot\rangle a_i - \sum_{i=1}^k \langle b_i, \cdot\rangle b_i \right\| \\
&= \left\| \sum_{i=1}^k \langle a_i, \cdot\rangle (a_i - b_i) + \sum_{i=1}^k \langle a_i - b_i, \cdot\rangle b_i \right\| \\
&\le \sum_{i=1}^k |a_i|\,|a_i - b_i| + |a_i - b_i|\,|b_i| \\
&\le 2k \sum_{i=1}^k |a_i - b_i| = 2k\langle (1)_{i\in\{1,\dots,k\}}, (|a_i - b_i|)_{i\in\{1,\dots,k\}}\rangle \\
&\le 2k\sqrt{k}\,\|A - B\|,
\end{aligned}
$$

which implies the Lipschitz-continuity of $\overline{\pi}$. ☐

B.13 Corollary. *The metric space* $(\mathrm{Gr}(k, \mathbb{R}^n), \Theta)$ *is complete.*

B.2 Differentiable Structure

This section merges ideas of [42, 2]. To introduce the differentiable structure, we make use of *local affine cross sections*. Let $A \in \mathrm{St}(k, \mathbb{R}^n)$ and define

$$
\begin{aligned}
S_A &:= \left\{ B \in \mathrm{St}(k, \mathbb{R}^n);\ A^\top (B - A) = 0 \right\} = \left\{ B \in \mathrm{St}(k, \mathbb{R}^n);\ A^\top B = A^\top A \right\} \\
&\subseteq \left\{ B \in \mathrm{St}(k, \mathbb{R}^n);\ \det(A^\top B) \neq 0 \right\} =: T_A \subset \mathrm{St}(k, \mathbb{R}^n),
\end{aligned}
$$

orthogonal to the fiber $A[GL(k,\mathbb{R})]$ crossing through A. For $B \in \mathrm{St}(k,\mathbb{R}^n)$ the equivalence class $B[GL(k,\mathbb{R})] = \pi^{-1}[\pi(B)]$ intersects the cross section S_A if and only if $B \in T_A$, and then in $B(A^\top B)^{-1}A^\top A$. This can be seen by plugging BP with $P \in GL(k,\mathbb{R})$ in the definition of S_A:

$$A^\top(BP - A) = 0 \quad \Leftrightarrow \quad A^\top BP = A^\top A \quad \Leftrightarrow \quad P = (A^\top B)^{-1}A^\top A.$$

Since for each $A \in \mathrm{St}(k,\mathbb{R}^n)$ we have $A \in T_A$ and T_A is open, $(T_A)_{A\in\mathrm{St}(k,\mathbb{R}^n)}$ is an open covering of $\mathrm{St}(k,\mathbb{R}^n)$. On the other hand, for $B \in \mathrm{St}(k,\mathbb{R}^n)$ the set $\{A \in \mathrm{St}(k,\mathbb{R}^n);\ B \in T_A\}$ equals T_B and is therefore open as well. Let

$$U_A := \pi[T_A] = \left\{\pi(B);\ B \in \mathrm{St}(k,\mathbb{R}^n),\ \det(A^\top B) \neq 0\right\} \subset \mathrm{Gr}(k,\mathbb{R}^n)$$

be the set of subspaces whose representing fibers $B[GL(k,\mathbb{R})]$ intersect the cross section S_A. We call the mapping

$$\sigma_A \colon U_A \to S_A, \qquad L = \pi(B) \mapsto B(A^\top B)^{-1}A^\top A \tag{B.9}$$

the *cross section mapping*.

B.14 Lemma. *The following statements hold:*

(i) $(U_A)_{A\in\mathrm{St}(k,\mathbb{R}^n)}$ *is an open covering of* $\mathrm{Gr}(k,\mathbb{R}^n)$.

(ii) For each $A \in \mathrm{St}(k,\mathbb{R}^n)$ *one has* $\pi^{-1}[U_A] = T_A$.

(iii) For each $A \in \mathrm{St}(k,\mathbb{R}^n)$ *one has that* $\pi\big|_{T_A} : T_A \subset \mathrm{St}(k,\mathbb{R}^n) \to U_A$ *is continuous and open.*

(iv) For each $A \in \mathrm{St}(k,\mathbb{R}^n)$ *one has that* σ_A *is continuous,* $\pi \circ \sigma_A = \mathrm{id}_{U_A}$ *and* $\sigma_{AP}(L) = \sigma_A(L)P$ *for* $P \in GL(k,\mathbb{R})$ *and* $L \in U_A$.

(v) For each $B \in \mathrm{St}(k,\mathbb{R}^n)$ *one has that* $T_B \ni A \mapsto \sigma_A(\pi(B))$ *is differentiable.*

Proof. (i) The covering property is clear by the surjectivity of π and the fact that $(T_A)_{A\in\mathrm{St}(k,\mathbb{R}^n)}$ is an open covering of $\mathrm{St}(k,\mathbb{R}^n)$. The fact that for each $A \in \mathrm{St}(k,\mathbb{R}^n)$ the set U_A is open follows from the fact that both T_A and π are open.

(ii) We have by definition that $\pi^{-1}[U_A] = \pi^{-1}[\pi[T_A]] = \bigcup_{P\in GL(k,\mathbb{R})}[T_A]P = T_A$, since T_A is invariant under right-multiplication with elements from $GL(k,\mathbb{R})$.

(iii) This is clear by the continuous embedding of T_A in $\mathrm{St}(k,\mathbb{R}^n)$ via the identity.

(iv) In order to prove continuity of σ_A, by Lemma B.2 it is equivalent to show the continuity of $\sigma_A \circ \pi \colon T_A \subset \mathrm{St}(k, \mathbb{R}^n) \to S_A$, which is obvious by the representation in Eq. (B.9). To see that $\pi \circ \sigma_A = \mathrm{id}$ observe that for $B \in T_A$ we have $(A^\top B)^{-1} A^\top A \in GL(k, \mathbb{R})$. Let $B \in T_A$ and $L := \pi(B) \in U_A$, then $\pi(\sigma_A(L)) = \pi(B(A^\top B)^{-1} A^\top A) = \pi(B) = L$. We show the homogeneity property by calculating

$$\begin{aligned}\sigma_{AP}(\pi(B)) &= B((AP)^\top B)^{-1}(AP)^\top AP = B(P^\top A^\top B)^{-1}(AP)^\top AP \\ &= B(A^\top B)^{-1}(P^\top)^{-1} P^\top A^\top AP = B(A^\top B)^{-1} A^\top AP \\ &= \sigma_A(\pi(B))P.\end{aligned}$$

(v) This is clear with the fact that W_B is open and with the representation in Eq. (B.9). □

Next, choose for each $A \in \mathrm{St}(k, \mathbb{R}^n)$ an $A_\perp \in \mathrm{St}^*(n-k, n)$ such that $A^\top A_\perp = 0 \in \mathbb{R}^{k \times (n-k)}$ and $A_\perp^\top A = 0 \in \mathbb{R}^{(n-k) \times k}$ in the following way. Let $A = U \begin{pmatrix} \Sigma \\ 0_{(n-k)\times k} \end{pmatrix} V^\top$ be a singular value decomposition (SVD) with $U \in O(n)$, $V \in O(k)$ and $\Sigma \in \mathbb{R}^{k \times k}$ a diagonal matrix with positive diagonal entries. Define $A_\perp := U \begin{pmatrix} 0_{k \times (n-k)} \\ W \end{pmatrix}$ for some $W \in \mathbb{R}^{(n-k)\times(n-k)}$ satisfying $WW^\top = I_{n-k}$. We then verify

$$\begin{aligned}A^\top A_\perp &= V \begin{pmatrix} \Sigma^\top & 0_{k\times(n-k)} \end{pmatrix} U^\top U \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix} \\ &= V \begin{pmatrix} \Sigma^\top & 0_{k\times(n-k)} \end{pmatrix} \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix} = 0_{k\times(n-k)},\end{aligned} \tag{B.10}$$

and consequently $A_\perp^\top A = 0_{(n-k)\times k}$. This justifies the notation $A_\perp$ since we have $\pi(A)^\perp = \pi(A_\perp)$ due to the full rank of both matrices and their orthogonality from Eq. (B.10). For preparatory reasons, we first consider

$$A_\perp A_\perp^\top = U \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix} \begin{pmatrix} 0_{k\times(n-k)} & W^\top \end{pmatrix} U^\top = U \begin{pmatrix} 0_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & I_{n-k} \end{pmatrix} U^\top$$

and secondly

$$\begin{aligned}A(A^\top A)^{-1} A^\top &= U \begin{pmatrix} \Sigma \\ 0_{(n-k)\times k} \end{pmatrix} V^\top \left(V \Sigma \Sigma V^\top \right)^{-1} V \begin{pmatrix} \Sigma & 0_{(n-k)\times k} \end{pmatrix} U^\top \\ &= U \begin{pmatrix} \Sigma \\ 0_{(n-k)\times k} \end{pmatrix} \Sigma^{-1} \Sigma^{-1} \begin{pmatrix} \Sigma & 0_{(n-k)\times k} \end{pmatrix} U^\top\end{aligned}$$

$$
\begin{aligned}
&= U \begin{pmatrix} I_k \\ 0_{(n-k)\times k} \end{pmatrix} \begin{pmatrix} I_k & 0_{(n-k)\times k} \end{pmatrix} U^\top \\
&= U \begin{pmatrix} I_k & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & 0_{(n-k)\times(n-k)} \end{pmatrix} U^\top \\
&= U \left(I_n - \begin{pmatrix} 0_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & I_{n-k} \end{pmatrix} \right) U^\top \\
&= I_n - A_\perp A_\perp^\top.
\end{aligned} \tag{B.11}
$$

Note that $(A^\top A)^{-1}A^\top$ is the Moore-Penrose inverse to A since A has full rank. It is well-known that for $B \in \mathrm{St}^*(k,\mathbb{R}^n)$ the matrix $BB^\top$ represents the orthogonal projection onto $\pi(B)$. Thus, we derived in Eq. (B.11) that $A(A^\top A)^{-1}A^\top$ represents the orthogonal projection onto $\pi(A)$ and equivalently $I_n - A(A^\top A)^{-1}A^\top$ represents the orthogonal projection onto $\pi(A)^\perp$.

With respect to $A \in \mathrm{St}(k,\mathbb{R}^n)$ we define the following family of functions

$$
\begin{aligned}
\varphi_A \colon U_A &\to \mathbb{R}^{(n-k)\times k}, \\
L = \pi(B) &\mapsto A_\perp^\top \sigma_A(L) = A_\perp^\top B(A^\top B)^{-1}A^\top A.
\end{aligned} \tag{B.12}
$$

B.15 Theorem. *$\{\varphi_A;\ A \in \mathrm{St}(k,\mathbb{R}^n)\}$ define a structure of differentiable manifold of dimension $k(n-k)$ in $\mathrm{Gr}(k,\mathbb{R}^n)$ with parametrization*

$$
p_A \colon \mathbb{R}^{(n-k)\times k} \to U_A, \qquad K \mapsto \pi(A + A_\perp K)
$$

for $A \in \mathrm{St}(k,\mathbb{R}^n)$.

Proof. We need to prove that (a) the domains of the charts cover $\mathrm{Gr}(k,\mathbb{R}^n)$, (b) all charts are injective, (c) the image of any chart is mapped by any other chart to an open subset of $\mathbb{R}^{(n-k)\times k}$ and (d) coordinate changes are smooth.

(a) By Lemma B.14(a) the domains of the charts φ_A, $A \in \mathrm{St}(k,\mathbb{R}^n)$, are a covering for $\mathrm{Gr}(k,\mathbb{R}^n)$.

(b) To prove the invertibility of φ_A for an arbitrary $A \in \mathrm{St}(k,\mathbb{R}^n)$ we need to check if the parametrization p_A is its inverse to obtain surjectivity. To this end, let $K \in \mathbb{R}^{(n-k)\times k}$ and we calculate

$$
\begin{aligned}
\varphi_A(p_A(K)) &= A_\perp^\top (A + A_\perp K)(A^\top(A + A_\perp K))^{-1}A^\top A \\
&= \left(A_\perp^\top A + A_\perp^\top A_\perp K\right)\left(A^\top A + A^\top A_\perp K\right)^{-1} A^\top A \\
&= K\left(A^\top A\right)^{-1} A^\top A = K.
\end{aligned}
$$

To prove injectivity, let $L \in U_A$, $B \in \pi^{-1}[\{L\}]$ and by making use of Eq. (B.11) we calculate

$$\begin{aligned}p_A(\varphi_A(L)) &= \pi(A + A_\perp A_\perp^\top B(A^\top B)^{-1}A^\top A)\\&= \pi(A + B(A^\top B)^{-1}A^\top A - A(A^\top A)^{-1}A^\top B(A^\top B)^{-1}A^\top A)\\&= \pi(B(A^\top B)^{-1}A^\top A) = \pi(B) = L.\end{aligned}$$

(c) Clearly, φ_A and p_A are continuous for every $A \in \mathrm{St}(k, \mathbb{R}^n)$, i.e. φ_A is a homeomorphism. For $A \in \mathrm{St}(k, \mathbb{R}^n)$ the domain of φ_A is open, thus $U_A \cap U_B$ is open and consequently $\varphi_B[U_A \cap U_B] = p_B^{-1}[U_A \cap U_B] \subseteq \mathbb{R}^{(n-k)\times k}$ is open due to continuity of p_B for any $B \in \mathrm{St}(k, \mathbb{R}^n)$.

(d) Let $A, B \in \mathrm{St}(k, \mathbb{R}^n)$ be such that $U_A \cap U_B \neq \varnothing$. Then for $K \in \varphi_A[U_B]$ we have

$$\begin{aligned}\varphi_B \circ p_A(K) &= B_\perp^\top(A + A_\perp K)(B^\top(A + A_\perp K))^{-1}B^\top B\\&= \left(B_\perp^\top A + B_\perp^\top A_\perp K\right)\left(B^\top A + B^\top A_\perp K\right)^{-1} B^\top B,\end{aligned}$$

which is smooth as the composition of smooth functions (multiplication and inversion). □

From Eq. (B.12) we directly read off the differentiability of π.

B.16 Corollary. *Let $\mathrm{Gr}(k, \mathbb{R}^n)$ be endowed with the differentiable structure from Theorem B.15. Then $\pi\colon \mathrm{St}(k, \mathbb{R}^n) \to \mathrm{Gr}(k, \mathbb{R}^n)$ as defined in Eq.* (B.2) *is differentiable.*

Furthermore, with the charts at hand we obtain that σ_A is an immersion for any $A \in \mathrm{St}(k, \mathbb{R}^n)$.

B.17 Lemma. *For any $A \in \mathrm{St}(k, \mathbb{R}^n)$ the function $\sigma_A\colon U_A \to S_A$ is an immersion.*

Proof. First, differentiability with respect to the charts is obvious. We need to show that $\partial(\mathrm{id} \circ \sigma_A \circ \varphi_A^{-1})(\varphi_A(L)) \in L(\mathbb{R}^{(n-k)\times k}, \mathbb{R}^{n\times k})$ is injective for any $L \in U_A$. By Theorem B.15 we have that $\varphi_A^{-1} = p_A$ and that $\sigma_A \circ p_A = (K \mapsto A + A_\perp K)$, which has the constant derivative $A_\perp$. By construction, $A_\perp$ has full rank and the injectivity is proved. □

B.3 (Fiber) Bundle Structure

In this section, we follow [42]. It is known that $GL(k, \mathbb{R})$ forms a real Lie-group, see for instance [40, Example 2.4]. Next we establish $\pi\colon \operatorname{St}(k, \mathbb{R}^n) \to \operatorname{Gr}(k, \mathbb{R}^n)$ as a principal $GL(k, \mathbb{R})$-bundle, i.e. a differentiable bundle with a differentiable right $GL(k, \mathbb{R})$-action. This differentiable right $GL(k, \mathbb{R})$-action $\cdot$ on $\operatorname{St}(k, \mathbb{R}^n)$ is the right-multiplication with matrices $P \in GL(k, \mathbb{R})$

$$\operatorname{St}(k, \mathbb{R}^n) \times GL(k, \mathbb{R}) \to \operatorname{St}(k, \mathbb{R}^n), \qquad (A, P) \mapsto AP. \tag{B.13}$$

Also, if $U \subset \operatorname{Gr}(k, \mathbb{R}^n)$ is open, we consider the $GL(k, \mathbb{R})$-action on $U \times GL(k, \mathbb{R})$ defined by

$$(U \times GL(k, \mathbb{R})) \times GL(k, \mathbb{R}) \to U \times GL(k, \mathbb{R}), \quad ((L, P), Q) \mapsto (L, PQ). \tag{B.14}$$

It is immediate that both are differentiable and effective, i.e. $AP = A$ implies $P = I$ and $(L, PQ) = (L, P)$ implies $Q = I$ with $A \in \operatorname{St}(k, \mathbb{R}^n)$, $P, Q \in GL(k, \mathbb{R})$ and $L \in \operatorname{Gr}(k, \mathbb{R}^n)$.

B.18 Proposition. *The function π together with the $GL(k, \mathbb{R})$-actions defined in Eqs. (B.13) and (B.14) is a principal $GL(k, \mathbb{R})$-bundle.*

Proof. We need to show that

(i) π is a fiber bundle, i.e. the "local triviality" is satisfied: for every $L \in \operatorname{Gr}(k, \mathbb{R}^n)$ there is an open neighborhood U of L in $\operatorname{Gr}(k, \mathbb{R}^n)$ and a diffeomorphism $\psi\colon \pi^{-1}[U] \to U \times GL(k, \mathbb{R})$ such that the diagram

$$\begin{tikzcd} \pi^{-1}[U] \arrow[r, "\psi"] \arrow[d, "\pi"'] & U \times GL(k, \mathbb{R}) \arrow[dl, "\operatorname{proj}_0"] \\ U & \end{tikzcd}$$

commutes; here, proj_0 denotes the projection onto the first component;

(ii) for every $L \in \operatorname{Gr}(k, \mathbb{R}^n)$ its fiber $\pi^{-1}[\{L\}]$ is diffeomorphic to $GL(k, n)$, and

(iii) the function ψ in a) is equivariant.

Concerning (i) let $L \in \operatorname{Gr}(k, \mathbb{R}^n)$ and $A \in \pi^{-1}[\{L\}]$. Then U_A is an open neighborhood of L and with

$$\begin{aligned} \psi_A\colon T_A &\to U_A \times GL(k, \mathbb{R}), & B &\mapsto (\pi(B), (A^\top A)^{-1} A^\top B), \\ \theta_A\colon U_A \times GL(k, \mathbb{R}) &\to T_A, & (L, P) &\mapsto \sigma_A(L) P, \end{aligned}$$

the diagram commutes, as is easily checked. In particular, ψ_A and θ_A are differentiable (by the differentiability of π and σ_A), inverse functions of each other and consequently ψ_A is a diffeomorphism. Requirement (ii) is satisfied by Lemma B.1(i). Finally, ψ_A is equivariant by the following calculation: let $B \in T_A$ and $P \in GL(k, \mathbb{R})$, then

$$\psi_A(BP) = (\pi(BP), (A^\top A)^{-1}A^\top BP) = (\pi(B), (A^\top A)^{-1}A^\top BP) = \psi_A(B)P. \quad \square$$

As a consequence, we have that π is a surjective submersion.

B.4 Riemannian Structure

The Stiefel manifold can be endowed with the structure of a Riemannian manifold by the Riemannian metric

$$\langle X, Y \rangle_A := \operatorname{tr}((A^\top A)^{-1} X^\top Y), \tag{B.15}$$

where $A \in \operatorname{St}(k, \mathbb{R}^n)$ and $X, Y \in T_A \operatorname{St}(k, \mathbb{R}^n)$. Note that this is different from the Euclidean Riemannian structure we will also consider later. Since the Grassmann manifold can be equivalently considered as a quotient manifold with respect to the Stiefel manifold modulo $GL(k, \mathbb{R})$ (see Eq. (B.4)), it can inherit the Riemannian structure from the Stiefel manifold by turning it into a Riemannian quotient manifold. Since subspaces are represented by matrices that span the subspaces, the aim of this section is to find representations for notions connected with the tangent bundle of the Grassmann manifold by objects connected with the tangent bundle of the Stiefel manifold. To this end, the two notions of *vertical bundle* and *horizontal bundle* are introduced as follows.

Let $A \in \operatorname{St}(k, \mathbb{R}^n)$. Since $\operatorname{St}(k, \mathbb{R}^n)$ is open in $\mathbb{R}^{n \times k}$ one has that $T_A \operatorname{St}(k, \mathbb{R}^n) = T_A \mathbb{R}^{n \times k} = \mathbb{R}^{n \times k}$. Next we decompose the tangent bundle $T \operatorname{St}(k, \mathbb{R}^n)$ into two subbundles, the aforementioned vertical and horizontal bundle. First, observe that by Lemma B.1 π is surjective and as a consequence of the fiber bundle property π is a submersion. Hence, by Proposition A.10 each fiber is a submanifold of dimension k^2. The well-defined tangent space to the fiber $\pi^{-1}[\{\pi(A)\}]$ is a subspace of $T_A \operatorname{St}(k, \mathbb{R}^n)$ and is called *vertical space* at A, denoted by V_A, i.e.

$$V_A := T_A(\pi^{-1}[\{\pi(A)\}]) = T_A(A[GL(k, \mathbb{R})]) \cong A[\mathbb{R}^{k \times k}],$$

since $GL(k,\mathbb{R})$ is open in $\mathbb{R}^{k\times k}$. Equivalently, the vertical space can be represented as

$$V_A = \ker D\pi\big|_A,$$

since π is constant on fibers. The *horizontal space* H_A at A is, in the case of Riemannian manifolds considered here, the orthogonal complement in $T\,\mathrm{St}(k,\mathbb{R}^n)$ to V_A. This yields the tangent space to the local cross section through A, i.e.

$$H_A := V_A^\perp = \left\{Y \in T\,\mathrm{St}(k,\mathbb{R}^n);\ A^\top Y = 0\right\} \cong A_\perp[\mathbb{R}^{(n-k)\times k}] \cong T_A S_A,$$

i.e. all matrices "orthogonal" to A. We denote by $V\,\mathrm{St}(k,\mathbb{R}^n)$ and $H\,\mathrm{St}(kn)$ the vertical and horizontal bundle associated to $\mathrm{St}(k,\mathbb{R}^n)$, respectively.

B.19 Lemma. *For any $A \in \mathrm{St}(k,\mathbb{R}^n)$ one has $T_A\,\mathrm{St}(k,\mathbb{R}^n) = V_A \oplus H_A$.*

Proof. This is easily verified making use of the full rank of A and $A_\perp$ and their orthogonality. □

For short, we have the decomposition of the tangent bundle

$$T\,\mathrm{St}(k,\mathbb{R}^n) = V\,\mathrm{St}(k,\mathbb{R}^n) \oplus H\,\mathrm{St}(k,\mathbb{R}^n),$$

which is meant fiberwise and abbreviates the statement of Lemma B.19. By Theorem B.15 we have that $\mathrm{Gr}(k,\mathbb{R}^n)$ is a differentiable manifold of dimension $k(n-k)$. Hence, the tangent space at an arbitrary point $L \in \mathrm{Gr}(k,\mathbb{R}^n)$ is a vector space of the same dimension. By Lemma B.17 σ_A is an immersion and due to Lemma A.9 the differential $D\sigma_A\colon TU_A \subset T\,\mathrm{Gr}(k,\mathbb{R}^n) \to TS_A \subset H_A \subset T\,\mathrm{St}(k,\mathbb{R}^n)$ is injective. Furthermore, the tangent spaces have equal dimension so that $D\sigma_A$ is a vector space isomorphism between the tangent space $T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$ and the horizontal space H_A of the Stiefel manifold.

We denote by π_{Gr} and π_{St} the tangent bundle projections of $\mathrm{Gr}(k,\mathbb{R}^n)$ and $\mathrm{St}(k,\mathbb{R}^n)$, respectively.

B.20 Definition (Horizontal lift)**.** We define

$$\{(A,X) \in \mathrm{St}(k,\mathbb{R}^n) \times T\,\mathrm{Gr}(k,\mathbb{R}^n);\ \pi(A) = \pi_{\mathrm{Gr}}(X)\} \to H\,\mathrm{St}(k,\mathbb{R}^n),$$
$$(A,X) \mapsto \overline{X}_A := D\sigma_A\big|_{\pi(A)} X,$$

and call $\overline{X}_A$ the *horizontal lift of $X \in T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$ at A.*

By what has been said before, $\left(X \mapsto \overline{X}_A\right) : T_{\pi(A)} \operatorname{Gr}(k, \mathbb{R}^n) \to H_A$ is an isomorphism, i.e. $\overline{Y}_A$ is the *only* horizontal vector that represents $Y \in T_{\pi(A)} \operatorname{Gr}(k, \mathbb{R}^n)$ in a sense specified in Proposition B.21(ii). In the next proposition we collect some simple properties of the horizontal lift. We denote by $\partial_A f(x)$ the directional derivative of f in the direction A at x.

B.21 Proposition. *Let $A \in \operatorname{St}(k, \mathbb{R}^n)$, $f \in \mathcal{F}_{\pi(A)}(\operatorname{Gr}(k, \mathbb{R}^n))$, $X \in \mathcal{X}(\operatorname{Gr}(k, \mathbb{R}^n))$, $Y \in T_{\pi(A)} \operatorname{Gr}(k, \mathbb{R}^n)$, $Z \in U_A$ and $P \in GL(k, \mathbb{R})$. Then the following statements hold:*

(i) $\overline{X \circ \pi} := (A \mapsto \overline{X(\pi(A))}_A) \in \mathcal{X}(\operatorname{St}(k, \mathbb{R}^n))$, *i.e. the horizontal lift of a vector field on* $\operatorname{Gr}(k, \mathbb{R}^n)$ *is a vector field on* $\operatorname{St}(k, \mathbb{R}^n)$.

(ii) $D\pi\big|_A \overline{Y}_A = Y$, *or, equivalently,* $\left(D\pi\big|_A\right)\Big|_{H_A}$ *is the inverse to* $\left(X \mapsto \overline{X}_A\right)$. *By definition, that means that* $\overline{Y}_A$ *is* π*-related to* Y.

(iii) $\overline{X}_{AP} = \overline{X}_A P$.

(iv) $A^\top \overline{Y}_{AP} = 0$.

(v) $Yf \cong \partial_{\overline{Y}_A}(f \circ \pi)(A)$.

Proof. (i) This holds by the definition of the push-forward and by Lemma B.14(v).

(ii) By Lemma B.14(iv) we have that

$$\mathrm{id}_{U_A} = \pi \circ \sigma_A.$$

Differentiating both sides yields in particular

$$Y = D(\pi \circ \sigma_A)Y = D\pi\big|_A D\sigma_A\big|_{\pi(A)} Y = D\pi\big|_A \overline{Y}_A.$$

(iii) This follows from the homogeneity property of σ_A proved in Lemma B.14(iv).

(iv) This is a paraphrase of the orthogonality of A and H_A.

(v) First observe that $f \circ \pi \in \mathcal{F}(\operatorname{St}(k, \mathbb{R}^n))$. In submanifolds of $\mathbb{R}^l$ derivations (tangent vectors) and directional derivatives can be identified:

$$\partial_{\overline{Y}_A}(f \circ \pi)(A) \cong \overline{Y}_A(f \circ \pi) = D\pi\big|_A \overline{Y}_A(f) \overset{(ii)}{=} Yf. \qquad \square$$

Next we define a Riemannian metric on $\mathrm{Gr}(k, \mathbb{R}^n)$, which is in fact induced by the horizontal lift and the Riemannian metric on $\mathrm{St}(k, \mathbb{R}^n)$ given in Eq. (B.15). It is natural in the following sense: it is the only Riemannian metric that turns π into a *Riemannian submersion*, i.e. a submersion whose restriction of the differential to the horizontal bundle is an isometry; see Eq. (B.16) in connection with Proposition B.21(ii).

B.22 Proposition (Riemannian metric)**.** *For $L \in \mathrm{Gr}(k, \mathbb{R}^n)$, $A \in \pi^{-1}[\{L\}] \subset \mathrm{St}(k, \mathbb{R}^n)$ and $X, Y \in T_L \mathrm{Gr}(k, \mathbb{R}^n)$ we define*

$$\langle X, Y \rangle_L = \langle X, Y \rangle_{\pi(A)} := \mathrm{tr}((A^\top A)^{-1} \overline{X}_A^\top \overline{Y}_A) = \langle \overline{X}_A, \overline{Y}_A \rangle_A . \tag{B.16}$$

Then the family $(\langle \cdot, \cdot \rangle_L)_{L \in \mathrm{Gr}(k, \mathbb{R}^n)}$ is a Riemannian metric on $\mathrm{Gr}(k, \mathbb{R}^n)$.

Proof. The inner product nature of $\langle \cdot, \cdot \rangle_L$ for every $L \in \mathrm{Gr}(k, \mathbb{R}^n)$ is obvious as well as the differentiability, so it remains to check whether the definition is independent of the choice of the representing matrix. Let therefore $A, B \in \pi^{-1}[\{L\}]$ and $P \in GL(k, \mathbb{R})$ such that $B = AP$. Then

$$\begin{aligned}
\langle \overline{X}_{AP}, \overline{Y}_{AP} \rangle_{AP} &= \mathrm{tr}(((AP)^\top AP)^{-1} \overline{X}_{AP}^\top \overline{Y}_{AP}) \\
&= \mathrm{tr}((P^\top A^\top AP)^{-1} P^\top \overline{X}_A^\top \overline{Y}_A P) \\
&= \mathrm{tr}(P^{-1} (A^\top A)^{-1} \overline{X}_A^\top \overline{Y}_A P) \\
&= \mathrm{tr}((A^\top A)^{-1} \overline{X}_A^\top \overline{Y}_A) \\
&= \langle \overline{X}_A, \overline{Y}_A \rangle_A
\end{aligned}$$

by the similarity-invariance of the trace. □

B.23 Proposition (Lie bracket)**.** *Let $X, Y \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n))$ and $A \in \mathrm{St}(k, \mathbb{R}^n)$. Then*

$$\overline{[X, Y]_{\mathrm{Gr}}(\pi(A))}_A = \Pi_{\pi(A_\perp)} [\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A), \tag{B.17}$$

where $\Pi_{\pi(A_\perp)} = A_\perp A_\perp^\top = I_n - A(A^\top A)^{-1} A^\top$ (recall Eq. (B.11)) and

$$[\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A) = \partial_{\overline{X(\pi(A))}_A} (\overline{Y \circ \pi})(A) - \partial_{\overline{Y(\pi(A))}_A} (\overline{X \circ \pi})(A)$$

is the Lie bracket for vector fields on $\mathrm{St}(k, \mathbb{R}^n)$.

Proof. By [85, Lemma 1.22] we have that $[\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A) \in T_A \mathrm{St}(k, \mathbb{R}^n)$ is π-related to $[X, Y]_{\mathrm{Gr}}(\pi(A))$. The projection $\Pi_{\pi(A_\perp)}$ makes it horizontal, such that

$$\Pi_{\pi(A_\perp)} [\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A) \in H_A$$

and is π-related to $[X,Y]_{\mathrm{Gr}}(\pi(A))$. These are the characterizing properties of $\overline{[X,Y]_{\mathrm{Gr}}(\pi(A))}_A$ and Eq. (B.17) is proved. $\square$

In the following we endow $\mathrm{Gr}(k,\mathbb{R}^n)$ with the Riemannian structure induced by the Riemannian metric given in Eq. (B.16).

The *gradient* of a function $f \in \mathcal{F}(\mathrm{Gr}(k,\mathbb{R}^n)$ with respect to the Riemannian metric, denoted by $\mathrm{grad}_{\mathrm{Gr}} f$, is the vector field satisfying $\langle \mathrm{grad}_{\mathrm{Gr}} f, X\rangle = Xf$ for any $X \in \mathcal{X}(\mathrm{Gr}(k,\mathbb{R}^n))$. Recall that in $\mathrm{St}(k,\mathbb{R}^n)$ the Euclidean, i.e. induced by $\mathbb{R}^{n\times k}$ as in Eq. (B.1), Riemannian metric is given by $\langle A,B\rangle_{\mathrm{St}} = \mathrm{tr}(A^\top B)$. Thus, the Euclidean gradient for $g \in \mathcal{F}(\mathrm{St}(k,\mathbb{R}^n))$ with

$$\mathrm{grad}_{\mathrm{St}}\, g(A) = (\partial_{ij} g(A))_{i,j} =: \partial \otimes g(A) \in \mathbb{R}^{n\times k} \tag{B.18}$$

is characterized by

$$X(f\circ\pi) \cong \partial_X(f\circ\pi)(A) = \mathrm{tr}(X^\top \mathrm{grad}_{\mathrm{St}}(f\circ\pi)(A))$$

for $A \in \mathrm{St}(k,\mathbb{R}^n)$, $X \in T_A\,\mathrm{St}(k,\mathbb{R}^n)$ and $f \in \mathcal{F}(\mathrm{Gr}(k,\mathbb{R}^n))$. Alternatively, we can define another gradient with respect to the metric defined in Eq. (B.15) and denoted by $\mathrm{grad}_A f$ accordingly. Note that for $f \in \mathcal{F}(\mathrm{Gr}(k,\mathbb{R}^n))$ we have that $f\circ\pi$ is constant on fibers and hence it follows that

$$\langle \mathrm{grad}_A(f\circ\pi)(A), X\rangle_A = \partial_X(f\circ\pi)(A) = 0$$

for all $X \in V_A$. Consequently, $\mathrm{grad}_A(f\circ\pi)(A) \in H_A$.

B.24 Lemma (Gradient)**.** *Let* $A \in \mathrm{St}(k,\mathbb{R}^n)$ *and* $f \in \mathcal{F}(\mathrm{Gr}(k,\mathbb{R}^n))$. *Then*

$$\overline{\mathrm{grad}_{\mathrm{Gr}} f(\pi(A))}_A = \mathrm{grad}_A(f\circ\pi)(A) = \mathrm{grad}_{\mathrm{St}}(f\circ\pi)(A)A^\top A. \tag{B.19}$$

Proof. Since $\mathrm{grad}_A(f\circ\pi)(A) \in H_A$ it suffices to consider horizontal tangent vectors, that can be uniquely represented as the horizontal lift $\overline{X}_A$ of a tangent vector $X \in T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$. Consider the following equalities:

$$\begin{aligned}
\overline{X}_A([f\circ\pi]_{\sim_A}) &= \langle \overline{X}_A, \mathrm{grad}_A(f\circ\pi)(A)\rangle_A & \text{(B.20)}\\
&= \mathrm{tr}\big((A^\top A)^{-1}\overline{X}_A^\top \mathrm{grad}_A(f\circ\pi)(A)\big), & \text{(B.21)}\\
\overline{X}_A([f\circ\pi]_{\sim_A}) &= \langle \overline{X}_A, \mathrm{grad}_{\mathrm{St}}(f\circ\pi)(A)\rangle_{\mathrm{St}} &\\
&= \mathrm{tr}\big(\overline{X}_A^\top \mathrm{grad}_{\mathrm{St}}(f\circ\pi)(A)\big), & \text{(B.22)}\\
\overline{X}_A([f\circ\pi]_{\sim_A}) &= X([f]_{\sim_{\pi(A)}}) = \langle X, \mathrm{grad}_{\mathrm{Gr}} f(\pi(A))\rangle_{\pi(A)} &\\
&= \left\langle \overline{X}_A, \overline{\mathrm{grad}_{\mathrm{Gr}} f(\pi(A))}_A \right\rangle_A. & \text{(B.23)}
\end{aligned}$$

Now, from Eqs. (B.21) and (B.22) we read off the second claimed equality, and from Eqs. (B.20) and (B.23) the first one. □

If M is a submanifold of a Euclidean space, then the Riemannian connection $\nabla(X, Y)$, $X \in T_pM$ and $Y \in \mathcal{X}(M)$, consists in taking the directional derivative of Y in the direction of X in the ambient Euclidean space and projecting the result into T_pM. The projection step can be omitted in the case of $\mathrm{St}(k, \mathbb{R}^n)$ since it is an open subset of $\mathbb{R}^{n\times k}$. In other words, in $\mathrm{St}(k, \mathbb{R}^n)$ the Riemannian connection ∇_{St} is the directional derivative. On the other hand, in $\mathrm{Gr}(k, \mathbb{R}^n)$ the Riemannian metric gives rise to a unique Riemannian connection ∇_{Gr}. The next result shows that both Riemannian connections are related to each other through a horizontal projection.

B.25 Proposition (Riemannian connection)**.** *Let $A \in \mathrm{St}(k, \mathbb{R}^n)$ and $X, Y \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n))$. Then*

$$\overline{\nabla_{\mathrm{Gr}}(X(\pi(A)), Y)}_A = \Pi_{\pi(A_\perp)} \nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y\circ\pi}\big) \tag{B.24}$$

with

$$\nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y\circ\pi}\big) = \partial_{\overline{X(\pi(A))}_A} \overline{Y\circ\pi}(A). \tag{B.25}$$

Proof. Obviously, both sides of Eq. (B.24) are elements of H_A. Eq. (B.24) holds if both sides have the same scalar product with every horizontal tangent vector, which can be represented by $\overline{Z}_A$ for $Z \in T_{\pi(A)}\mathrm{Gr}(k, \mathbb{R}^n)$. Since the vertical component of $\nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y\circ\pi}\big)$ is orthogonal to the horizontal space and hence there is no contribution to the scalar product, it suffices to prove for each $Z \in T_{\pi(A)}\mathrm{Gr}(k, \mathbb{R}^n)$ that

$$\left\langle \nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y\circ\pi}\big), \overline{Z}_A \right\rangle_A = \langle \nabla_{\mathrm{Gr}}(X(\pi(A)), Y), Z\rangle_{\pi(A)}. \tag{B.26}$$

This follows by expanding both sides in the Koszul formula Eq. (A.3), where it suffices to consider terms of the following two types. Let $X, Y, Z \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n))$ and denote by $\overline{X\circ\pi} := (A \mapsto \overline{X\circ\pi}_A) \in \mathcal{X}(\mathrm{St}(k, \mathbb{R}^n))$ the horizontal vector field and consider the following as functions on $\mathrm{St}(k, \mathbb{R}^n)$:

$$\begin{aligned}
\overline{X\circ\pi}\,\langle\overline{Y\circ\pi}, \overline{Z\circ\pi}\rangle &= \left(A \mapsto \overline{X\circ\pi}_A\left([\langle\overline{Y\circ\pi}, \overline{Z\circ\pi}\rangle]_{\sim_A}\right)\right)\\
&= \left(A \mapsto \overline{X\circ\pi}_A\left(\left[\langle Y\circ\pi, Z\circ\pi\rangle_{\pi(\cdot)}\right]_{\sim_A}\right)\right)\\
&= \left(A \mapsto \overline{X\circ\pi}_A\left([\langle Y, Z\rangle_{\cdot}\circ\pi]_{\sim_A}\right)\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left(A \mapsto D\pi\big|_A \overline{X \circ \pi}_A \left([\langle Y, Z \rangle]_{\sim_{\pi(A)}} \right) \right) \\
&= \left(A \mapsto X(\pi(A)) \left([\langle Y, Z \rangle]_{\sim_{\pi(A)}} \right) \right) \\
&= (X \langle Y, Z \rangle) \circ \pi, \\
\left\langle \overline{X \circ \pi}, \left[\overline{Y \circ \pi}, \overline{Z \circ \pi} \right] \right\rangle &= \left(A \mapsto \left\langle \overline{X \circ \pi}_A, \left[\overline{Y \circ \pi}, \overline{Z \circ \pi} \right]_{\mathrm{St}} (A) \right\rangle_A \right) \\
&= \left(A \mapsto \left\langle \overline{X \circ \pi}_A, \overline{[Y, Z]_{\mathrm{Gr}} (\pi(A))}_A \right\rangle_A \right) \\
&= \left(A \mapsto \langle X(\pi(A)), [Y, Z]_{\mathrm{Gr}}(\pi(A)) \rangle_{\pi(A)} \right) \\
&= (\langle X, [Y, Z] \rangle) \circ \pi.
\end{aligned}
$$

In summary, corresponding terms in the expansion in the Koszul formula of both sides of Eq. (B.26) equal each other and the proof is finished. □

In the following we consider regular curves on $\mathrm{Gr}(k, \mathbb{R}^n)$. We do not state the domains explicitly, but assume implicitly that they contain 0.

Let $t \mapsto C(t) \in \mathrm{Gr}(k, \mathbb{R}^n)$ be a regular curve on $\mathrm{Gr}(k, \mathbb{R}^n)$. Recall that a vector field $X \in \mathcal{X}_C(\mathrm{Gr}(k, \mathbb{R}^n))$ along C is said to be parallel transported along C if $\underset{C}{\nabla} X = 0$. A (regular) curve $A\colon t \mapsto c(t) \in \mathrm{St}(k, \mathbb{R}^n)$ on $\mathrm{St}(k, \mathbb{R}^n)$ is called *horizontal* if $\dot{A}(t) \in H_{A(t)}$ for any $t \in D(A)$.

Let $t \mapsto C(t) \in \mathrm{Gr}(k, \mathbb{R}^n)$ be a regular curve on $\mathrm{Gr}(k, \mathbb{R}^n)$, $\pi(A_0) = C(0) \in \mathrm{Gr}(k, \mathbb{R}^n)$, $A_0 \in \mathrm{St}(k, \mathbb{R}^n)$. Then there exists a unique horizontal curve $A\colon t \mapsto A(t)$ such that $A(0) = A_0$ and $\pi(A(t)) = C(t)$ for all $t \in D(C)$. The reason for that is that locally the image of C is a submanifold in $\mathrm{Gr}(k, \mathbb{R}^n)$. For simplicity assume that this holds globally. The preimage of C under π is by Proposition A.10 a submanifold of $\mathrm{St}(k, \mathbb{R}^n)$, on which we can define a horizontal vector field that is constant on fibers, i.e. for $B \in \pi^{-1}[C[D(C)]] \subseteq \mathrm{St}(k, \mathbb{R}^n)$ define

$$X(B) := \overline{\dot{C}(\pi(B))}_B.$$

For each A_0 there exists a unique integral curve A through A_0, i.e. a solution to $A(0) = A_0$ and $\dot{A}(t) = X(A(t))$, satisfying $\pi(A(t)) = C(t)$ for all $t \in D(C)$. By definition of X we have that A is horizontal and the projection property follows from uniqueness. The curve A is called the *horizontal lift of C through A_0*.

B.26 Proposition (Parallel transport)**.** *Let C be a regular curve on $\mathrm{Gr}(k, \mathbb{R}^n)$, $X \in \mathcal{X}_C(\mathrm{Gr}(k, \mathbb{R}^n))$, A be a horizontal lift of C. Then*

$$\overline{X}_A := \left(t \mapsto \overline{X(t)}_{A(t)}\right) \in \mathcal{X}_C(\mathrm{St}(k, \mathbb{R}^n))$$

is the horizontal lift of X along A and X is parallel transported along C if and only if

$$\dot{\overline{X}}_A(t) + A(t)\left(A(t)^\top A(t)\right)^{-1} \dot{A}(t)^\top \overline{X}_A(t) = 0, \tag{B.27}$$

where

$$\dot{\overline{X}}_A(t) = (s \mapsto \overline{X(s)}_{A(s)})'(t).$$

Proof. Let $t \in D(C)$. We have that $\dot{A}(t) = \overline{\dot{C}(t)}_{A(t)}$ and

$$\begin{aligned}
\left(\underset{C}{\nabla} X\right)(t) &= \Pi_{\pi(A(t)_\perp)} \nabla_{\mathrm{St}}\left(\overline{\dot{C}(t)}_{A(t)}, \overline{X}_A\right) \\
&= \Pi_{\pi(A(t)_\perp)} \nabla_{\mathrm{St}}\left(\overline{\dot{C}(t)}_{A(t)}, \overline{X}_A\right) \\
&= \Pi_{\pi(A(t)_\perp)} \partial_{\overline{\dot{C}(t)}_{A(t)}} \overline{X}_A(A(t)) \\
&= \Pi_{\pi(A(t)_\perp)} \partial_{\dot{A}(t)} \overline{X}_A(A(t)) \\
&= \Pi_{\pi(A(t)_\perp)} \dot{\overline{X}}_A(t)
\end{aligned}$$

by Proposition B.25. Thus, X is parallel transported along C if and only if

$$0 = \left(\underset{C}{\nabla} X\right)(t) = \Pi_{\pi(A(t)_\perp)} \dot{\overline{X}}_A(t),$$

i.e. $\dot{\overline{X}}_A(t) \in V_{A(t)}$. Consequently it is of the form

$$\dot{\overline{X}}_A(t) = A(t)M(t) \tag{B.28}$$

for some $M\colon t \mapsto M(t) \in \mathbb{R}^{k \times k}$. By Proposition B.21(iv) we have

$$A(t)^\top \overline{X}_A(t) = 0.$$

Differentiation of the last equation yields

$$\dot{A}(t)^\top \overline{X}_A(t) + A(t)^\top \dot{\overline{X}}_A(t) = 0.$$

Plugging in Eq. (B.28) and solving for $M(t)$ we obtain

$$M(t) = -\left(A(t)^\top A(t)\right)^{-1} \dot{A}(t)^\top \overline{X}_A(t),$$

and the proof is finished. ∎

Next, we consider geodesics $C\colon t \mapsto C(t) \in \mathrm{Gr}(k, \mathbb{R}^n)$ with some reference point $C(0) = C_0$ and initial direction $\dot{C}_0 \in T_{C(0)} \mathrm{Gr}(k, \mathbb{R}^n)$. The unique geodesic C is characterized by $\underset{C}{\nabla} \dot{C} = 0$ and $\dot{C}(0) = \dot{C}_0$. With this notation the following holds.

B.27 Proposition (Geodesics)**.** *Let $A_0 \in \pi^{-1}[\{C_0\}] \subset \mathrm{St}(k, \mathbb{R}^n)$, $\overline{\dot{C}_0}_{A_0} =: \dot{A}_0$ and*

$$\dot{A}_0 (A_0^\top A_0)^{-1/2} = U \Sigma V^\top$$

be a thin singular value decomposition (SVD), i.e. $U \in \mathrm{St}^*(k, \mathbb{R}^n)$, $V \in \mathrm{St}^*(k,k)$ and $\Sigma \in \mathbb{R}^{k\times k}$ is diagonal with nonnegative entries; see, for instance, [51, Section 2.5.4]. Then

$$C(t) = \pi(A_0(A_0^\top A_0)^{-1/2} V \cos(t\Sigma) + U \sin(t\Sigma)).$$

We define $\exp(\dot{C}_0) := C(1)$.

Proof. Let $t \in D(C)$ and A be the unique horizontal lift of C through A_0, so that $\dot{A}(t) = \overline{\dot{C}(t)}_{A(t)}$. Then by (B.27) we have

$$\ddot{A}(t) + A(t) \left(A(t)^\top A(t) \right)^{-1} \dot{A}(t)^\top \dot{A}(t) = 0. \tag{B.29}$$

Since A is horizontal, one has

$$A(t)^\top \dot{A}(t) = 0 = \dot{A}(t)^\top A(t). \tag{B.30}$$

Thus, we have

$$\dot{\widehat{(A^\top A)}}\,(t) = \dot{A}(t)^\top A(t) + A(t)^\top \dot{A}(t) = 0,$$

saying that $t \mapsto A(t)^\top A(t)$ is a constant function. Differentiation of Eq. (B.30) yields

$$\dot{A}(t)^\top \dot{A}(t) + A(t)^\top \ddot{A}(t) = 0,$$

or equivalently

$$\dot{A}(t)^\top \dot{A}(t) = -A(t)^\top \ddot{A}(t). \tag{B.31}$$

By plugging Eq. (B.31) into Eq. (B.29) we get

$$\ddot{A}(t) - A(t) \left(A(t)^\top A(t) \right)^{-1} A(t)^\top \ddot{A}(t) = \Pi_{\pi(A(t))^\perp} \ddot{A}(t) = 0,$$

saying that $\ddot{A}(t) \in V_{A(t)}$ and consequently it is of the form

$$\ddot{A}(t) = A(t)M(t) \tag{B.32}$$

for some $M\colon t \mapsto M(t) \in \mathbb{R}^{k\times k}$. With Eqs. (B.32) and (B.31) we obtain that

$$\begin{aligned}
\dot{\widehat{(\dot{A}^\top A)}}\,(t) &= \ddot{A}(t)^\top \dot{A}(t) + \dot{A}(t)^\top \ddot{A}(t) \\
&= (A(t)M(t))^\top \dot{A}(t) + \dot{A}(t)^\top A(t)M(t) \\
&= M(t)^\top A(t)^\top \dot{A}(t) + \dot{A}(t)^\top A(t)M(t) = 0,
\end{aligned}$$

i.e. $t \mapsto \dot{A}(t)^\top \dot{A}(t)$ is constant, too. Now consider the thin SVD

$$\dot{A}_0(A_0^\top A_0)^{-1/2} = U\Sigma V^\top \tag{B.33}$$

as in the statement of the proposition. Eq. (B.29) right-multiplied with $(A_0^\top A_0)^{-1/2}$ yields

$$\ddot{A}(t)(A_0^\top A_0)^{-1/2} + A(t)(A_0^\top A_0)^{-1/2}((A_0^\top A_0)^{-1/2}\dot{A}^\top)(\dot{A}(A_0^\top A_0)^{-1/2}) = 0.$$

Plugging Eq. (B.33) into the last equation we obtain

$$\ddot{A}(t)(A_0^\top A_0)^{-1/2} + A(t)(A_0^\top A_0)^{-1/2}V\Sigma^2V^\top = 0,$$

and after right-multiplication with V

$$\ddot{A}(t)(A_0^\top A_0)^{-1/2}V + A(t)(A_0^\top A_0)^{-1/2}V\Sigma^2 = 0.$$

With the abbreviation $B(t) := A(t)(A_0^\top A_0)^{-1/2}V$, the last equation reads

$$\ddot{B}(t) + B(t)\Sigma^2 = 0,$$

from which the solution

$$B(t) = B(0)\cos(t\Sigma) + \dot{B}(0)\Sigma^{-1}\sin(t\Sigma),$$

or equivalently

$$A(t)(A_0^\top A_0)^{-1/2}V = A_0(A_0^\top A_0)^{-1/2}V\cos(t\Sigma) + \dot{A}_0(A_0^\top A_0)^{-1/2}V\Sigma^{-1}\sin(t\Sigma),$$

can be read off directly. The assertion now follows from

$$C(t) = \pi(A(t)) = \pi(A(t)(A_0^\top A_0)^{-1/2}V),$$

since $(A_0^\top A_0)^{-1/2}V \in GL(k,\mathbb{R})$. $\square$

Proposition B.27 shows that the Grassmann manifold with the Riemannian structure given by the Riemannian metric in (B.16) is complete, i.e. the geodesics are defined on $\mathbb{R}$. By [110] the Grassmann manifold is connected and hence, by the Hopf-Rinow-theorem (see, for instance, [24, Theorem 10.4.16]), a complete metric space with respect to the Riemannian distance function, also referred to as the geodesic distance. However, the completeness of $\mathrm{Gr}(k, \mathbb{R}^n)$ with respect to the Riemannian distance function can be obtained alternatively with the notion of principal angles; for an introduction see, for instance, [51, 12.4.3]. The distance function induced by the Riemannian metric (B.16) is the 2-norm of the vector of principal angles denoted by θ. On the other hand, the gap metric Θ corresponds to the sin of the largest principal angle, i.e. $\sin|\theta|_\infty$. Now, one easily verifies the (strong) equivalence of the gap metric and the geodesic distance, i.e.

$$\sin|\theta|_\infty \leq |\theta|_2 \leq \frac{2\sqrt{k}}{\pi} \sin|\theta|_\infty\,,$$

from which the completeness of $\mathrm{Gr}(k, \mathbb{R}^n)$ with respect to the geodesic distance function follows. For other definitions of distance functions in terms of principal angles see [36, 4.3].

Bibliography

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer, New York, 2. edition, 1988.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre. Riemannian Geometry of Grassmann Manifolds with a View on Algorithmic Computation. *Acta Applicandae Mathematicae*, 80(2):199–220, 2004.

[3] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

[4] L. Ya. Adrianova. *Introduction to linear systems of differential equations*, volume 146 of *Translations of mathematical monographs*. AMS, Providence, 1995.

[5] B. B. Aldridge, G. Haller, P. K. Sorger, and D. A. Lauffenburger. Direct Lyapunov exponent analysis enables parametric study of transient signalling governing cell behaviour. *IEE Proceedings - Systems Biology*, 153(6):425–432, 2006.

[6] H. Amann. *Ordinary differential equations: an introduction to nonlinear analysis*, volume 13 of *De Gruyter studies in mathematics*. Walter de Gruyter, 1990.

[7] H. Aref. Stirring by chaotic advection. *Journal of Fluid Mechanics*, 143:1–21, 1984.

[8] H. Aref. The development of chaotic advection. *Physics of Fluids*, 14(4):1315–1325, 2002.

[9] E Aurell, G Boffetta, A Crisanti, G Paladin, and A Vulpiani. Predictability in the large: an extension of the concept of Lyapunov exponent. *Journal of Physics A: Mathematical and General*, 30(1):1, 1997.

[10] L. Barreira and Y. B. Pesin. *Nonuniform hyperbolicity: dynamics of systems with nonzero Lyapunov exponents*. Encyclopedia of mathematics and its applications. Cambridge University Press, 2007.

[11] L. Barreira and C. Valls. *Stability of Nonautonomous Differential Equations*, volume 1926 of *Lecture Notes in Mathematics*. Springer, 2008.

[12] A. E. Bashirov, E. M. Kurpınar, and A. Özyapıcı. Multiplicative calculus and its applications. *Journal of Mathematical Analysis and Applications*, 337(1):36–48, 2008.

[13] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. Lyapunov Characteristic Exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. Part 1: Theory. *Meccanica*, 15:9–20, 1980.

[14] A. Berger. More on finite-time hyperbolicity. *Boletín de la Sociedad Española de Matemática Aplicada*, 51:25–32, 2010.

[15] A. Berger. On finite-time hyperbolicity. *Communications on Pure and Applied Analysis*, 10(3):963–981, 2011.

[16] A. Berger, T. S. Doan, and S. Siegmund. A Remark on Finite-Time Hyperbolicity. *PAMM*, 8(1):10917–10918, 2008.

[17] A. Berger, T. S. Doan, and S. Siegmund. A definition of spectrum for differential equations on finite time. *Journal of Differential Equations*, 246(3):1098–1118, 2009.

[18] M. Berger and B. Gostiaux. *Differential geometry : manifolds, curves, and surfaces*, volume 115 of *Graduate texts in mathematics*. Springer, 1988.

[19] P. Bohl. Über Differentialungleichungen. *Journal für die Reine und Angewandte Mathematik*, 144:284–318, 1913.

[20] M. Bohner and A. Peterson. *Dynamic Equations on Time Scales: An Introduction with Applications*. Birkhäuser, 2001.

[21] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*, volume 120 of *Pure and Applied Mathematics*. Elsevier, 1986.

[22] D. Chelidze and J. P. Cusumano. Phase space warping: nonlinear time-series analysis for slowly drifting systems. *Philosophical Transactions of the Royal Society A*, 364(1846):2495–2513, 2006.

[23] F. Colonius, P. E. Kloeden, and M. Rasmussen. Morse spectrum for nonautonomous differential equations. *Stochastics and Dynamics*, 8(3):351–363, 2008.

[24] L. Conlon. *Differentiable Manifolds*. Birkhäuser Advanced Texts. Birkhäuser, 2. edition, 2001.

[25] A. R. Conn, N. I. M. Gould, and P. L. Toint. *Trust-Region Methods*. MPS-SIAM Series on Optimization. SIAM, 2000.

[26] W. A. Coppel. *Dichotomies in Stability Theory*, volume 629 of *Lecture Notes in Mathematics*. Springer-Verlag Berlin Heidelberg New York, 1978.

[27] C. M. Dafermos. An invariance principle for compact processes. *Journal of Differential Equations*, 9:239–252, 1971. erratum: ibid. 10 (1971), 179-180.

[28] Yu. L. Daletskij and M. G. Krejn. *Stability of solutions of differential equations in Banach space.* , volume 43 of *Translations of Mathematical Monographs*. American Mathematical Society (AMS), Providence, R.I., 1974.

[29] M. Dellnitz, G. Froyland, C. Horenkamp, and K. Padberg. On the Approximation of Transport Phenomena – a Dynamical Systems Approach. *GAMM-Mitteilungen*, 32(1):47–60, 2009.

[30] P. Deuflhard. *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*. Springer Series in Computational Mathematics. Springer, 2011.

[31] T. S. Doan, D. Karrasch, T. Y. Nguyen, and S. Siegmund. A unified approach to finite-time hyperbolicity which extends finite-time Lyapunov exponents. *Journal of Differential Equations*, 252(10):5535–5554, 2012.

[32] T. S. Doan, K. J. Palmer, and S. Siegmund. Transient Spectral Theory, Stable and Unstable Cones and Gershgorin's Theorem for Finite-Time Differential Equations. *Journal of Differential Equations*, 250(11):4177–4199, 2011.

[33] L. H. Duc and S. Siegmund. Hyperbolicity and Invariant Manifolds for Planar Nonautonomous Systems on Finite Time Intervals. *International Journal of Bifurcation and Chaos*, 18(3):641–674, 2008.

[34] D. Eberly. *Ridges in Image and Data Analysis*, volume 7 of *Computational Imaging and Vision*. Kluwer Academic Publishers, 1996.

[35] D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach. Ridges for image analysis. *J. Math. Imaging Vis.*, 4:353–373, 1994.

[36] A. Edelman, T. A. Arias, and S. T. Smith. The Geometry of Algorithms with Orthogonality Constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, 1998.

[37] A. Edelman and R. A. Lippert. Nonlinear eigenvalue problems with orthogonality constraints. In Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors, *Templates for the solution of Algebraic Eigenvalue Problems: A Practical Guide*, chapter 8.3, pages 67–81. SIAM, 2000.

[38] M. Farazmand and G. Haller. Computing Lagrangian coherent structures from their variational theory. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(1):013128, 2012.

[39] M. Farazmand and G. Haller. Erratum and addendum to "A variational theory of hyperbolic Lagrangian coherent structures [Physica D 240 (2011) 574–598]". *Physica D*, 241(4):439–441, 2012.

[40] H. D. Fegan. *Introduction to compact Lie groups*, volume 13 of *Series in Pure Mathematics*. World Scientific, Singapore, 1991.

[41] N. Fenichel. Persistence and Smoothness of Invariant Manifolds for Flows. *Indiana Univ. Math. J.*, 21(3):193–226, 1971.

[42] J. Ferrer, M. I. Garcia, and F. Puerta. Differentiable families of subspaces. *Linear Algebra and its Applications*, 199:229–252, 1994.

[43] G. Froyland, S. Llyod, and N. Santitissadeekorn. Coherent sets for nonautonomous dynamical systems. *Physica D*, 239(16):1527–1541, 2010.

[44] G. Froyland and K. Padberg. Almost-invariant sets and invariant manifolds - Connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D*, 238(16):1507–1523, 2009.

[45] G. Froyland, K. Padberg, M. H. England, and A. M. Treguier. Detection of Coherent Oceanic Structures via Transfer Operators. *Physical Review Letters*, 98(22):224503–1–224503–4, 2007.

[46] G. Froyland and K. Padberg-Gehle. Finite-time entropy: A probabilistic approach for measuring nonlinear stretching. *Physica D*, 241(19):1612–1628, 2012.

[47] P. Giesl and M. Rasmussen. Areas of attraction for nonautonomous differential equations on finite time intervals. *Journal of Mathematical Analysis and Applications*, 390(1):27–46, 2012. doi: 10.1016/j.jmaa.2011.12.051.

[48] I. C. Gohberg, P. Lancaster, and L. Rodman. *Invariant subspaces of matrices with applications*. Classics in applied mathematics. SIAM Society of Industrial and Applied Mathematics, 2006.

[49] I. C. Gohberg and A. S. Markus. Two theorems on the opening between subspaces of a Banach space. *Uspekhi Mat. Nauk*, 89:135–140, 1959. in Russian.

[50] I. Goldhirsch, P.-L. Sulem, and S. A. Orszag. Stability and Lyapunov stability of dynamical systems: A differential approach and a numerical method. *Physica D*, 27(3):311–337, 1987.

[51] G.H. Golub and C.F.V. Loan. *Matrix computations*. Johns Hopkins studies in the mathematical sciences. Johns Hopkins University Press, 3. edition, 1996.

[52] N. I. M. Gould. An introduction to algorithms for continuous optimization. Technical report, Oxford University Computing Laboratory, 2006.

[53] G. Haller. Finding finite-time invariant manifolds in two-dimensional velocity fields. *Chaos*, 10(1):99–108, 2000.

[54] G. Haller. Distinguished material surfaces and coherent structures in three-dimensional fluid flows. *Physica D*, 149(4):248–277, 2001.

[55] G. Haller. Lagrangian coherent structures from approximate velocity data. *Physics of Fluids*, 14(6):1851–1861, 2002.

[56] G. Haller. A variational theory of hyperbolic Lagrangian Coherent Structures. *Physica D*, 240(7):574–598, 2011.

[57] G. Haller and A. C. Poje. Finite time transport in aperiodic flows. *Physica D*, 119(3):352–380, 1998.

[58] G. Haller and T. Sapsis. Lagrangian coherent structures and the smallest finite-time Lyapunov exponent. *Chaos*, 21(2):023115, 2011.

[59] G. Haller and G. Yuan. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D*, 147(3-4):352–370, 2000.

[60] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture notes in mathematics*. Springer, Berlin, 1981.

[61] R. A. Johnson, K. J. Palmer, and G. R. Sell. Ergodic Properties of Linear Dynamical Systems. *SIAM Journal on Mathematical Analysis*, 18(1):1–33, 1987.

[62] B. Joseph and B. Legras. Relation between Kinematic Boundaries, Stirring, and Barriers for the Antarctic Polar Vortex. *Journal of the Atmospheric Sciences*, 59(7):1198–1212, 2002.

[63] N. Ju and S. Wiggins. On Roughness of Exponential Dichotomy. *Journal of Mathematical Analysis and Applications*, 262(1):39–49, 2001.

[64] D. Karrasch. Comment on "A variational theory of hyperbolic Lagrangian Coherent Structures, Physica D 240 (2011) 574–598". *Physica D*, 241(17):1470–1473, 2012.

[65] D. Karrasch. Linearization of hyperbolic finite-time processes. *Journal of Differential Equations*, 254(1):256–282, 2013.

[66] D. Karrasch and S. Trostorff. Differentialgeometrie. Unpublished manuscript, 2011.

[67] T. Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer-Verlag Berlin Heidelberg New York, corrected printing of the second edition, 1995. Reprint of the 1980 edition, Originally published in Grundlehren der mathematischen Wissenschaften; Bd. 132.

[68] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*, volume 54 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1995.

[69] P. E. Kloeden and M. Rasmussen. *Nonautonomous Dynamical Systems*, volume 176 of *Mathematical Surveys and Monographs*. AMS, 2011.

[70] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume 1, 2 of *Interscience Tracts in Pure and Applied Mathematics, 15*. Interscience, 1963.

[71] J. J. Koenderink and A. J. van Doorn. Local features of smooth shapes: ridges and courses. In B. C. Vemuri, editor, *Geometric Methods in Computer Vision II*, volume 2031, pages 2–13. SPIE, 1993.

[72] M. G. Krein and M. A. Krasnoselski. Fundamental theorems concerning the extension of Hermitian operators and some of their applications to the theory of orthogonal polynomials and the moment problem. *Uspekhi Mat. Nauk*, 2(3), 1947. in Russian.

[73] F. Lekien and J. E. Marsden. Tricubic interpolation in three dimensions. *International Journal for Numerical Methods in Engineering*, 63(3):455–471, 2005.

[74] F. Lekien, S. C. Shadden, and J. E. Marsden. Lagrangian coherent structures in n-dimensional systems. *J. Math. Phys.*, 48(6):065404. 1–19, 2007.

[75] A. M. Lyapunov. *Problème Général de la Stabilité du Mouvement*, volume 17 of *Annals of Mathematics Studies*. Princeton University Press, 1949.

[76] J. A. J. Madrid and A. M. Mancho. Distinguished trajectories in time dependent vector fields. *Chaos*, 19(1):1–18, 2009.

[77] R. H. Martin. Conditional stability and separation of solutions to differential equations. *Journal of Differential Equations*, 13(1):81–105, 1973.

[78] J. L. Massera and J. J. Schäffer. *Linear differential equations and function spaces*, volume 21 of *Pure and applied mathematics*. Academic Press, New York, 1966.

[79] I. Mezic, S. Loire, V. A. Fonoberov, and P. Hogan. A New Mixing Diagnostic and Gulf Oil Spill Movement. *Science*, 330(6003):486–489, 2010.

[80] P. W. Michor. *Topics in differential geometry*, volume 93 of *Graduate studies in mathematics*. AMS, 2008.

[81] J. S. Muldowney. Dichotomies and Asymptotic Behaviour for Linear Differential Systems. *Transactions of the American Mathematical Society*, 283(2):465–484, 1984.

[82] R. Naulin and M. Pinto. Roughness of (h, k)-Dichotomies. *Journal of Differential Equations*, 118(1):20–35, 1995.

[83] G. Norgard and P.-T. Bremer. Second derivative ridges are straight lines and the implications for computing Lagrangian Coherent Structures. *Physica D*, 2012. doi: 10.1016/j.physd.2012.05.006.

[84] M. J. Olascoaga and G. Haller. Forecasting sudden changes in environmental pollution patterns. *Proceedings of the National Academy of Sciences*, 109(13):4738–4743, 2012.

[85] B. O'Neill. *Semi-Riemannian geometry. With applications to relativity.*, volume 103 of *Pure and Applied Mathematics*. Academic Press, 1983.

[86] V. I. Oseledec. A multiplicative ergodic theorem. *Trudy Moskov. Mat. Obsc.*, 19:197–231, 1968.

[87] J. M. Ottino. *The Kinematics of Mixing: Stretching, Chaos, and Transport*. Cambridge University Press, Cambridge, England, 1989.

[88] T. Peacock and J. Dabiri. Introduction to Focus Issue: Lagrangian Coherent Structures. *Chaos*, 20(1):017501, 2010.

[89] Y. B. Pesin. *Lectures on Partial Hyperbolicity and Stable Ergodicity*. Zürich Lectures in Advanced Mathematics. EMS, 2004.

[90] R. T. Pierrehumbert. Large-scale horizontal mixing in planetary atmospheres. *Physics of Fluids A: Fluid Dynamics*, 3(5):1250–1260, 1991.

[91] R. T. Pierrehumbert and H. Yang. Global Chaotic Mixing on Isentropic Surfaces. *Journal of the Atmospheric Sciences*, 50(15):2462–2480, 1993.

[92] E. Plischke. *Transient Effects of Linear Dynamical Systems*. PhD thesis, University Bremen, 2005.

[93] V. A. Pliss and G. R. Sell. Robustness of Exponential Dichotomies in Infinite-Dimensional Dynamical Systems. *Journal of Dynamics and Differential Equations*, 11(3):471–513, 1999.

[94] L. H. Popescu. Exponential dichotomy roughness on Banach spaces. *Journal of Mathematical Analysis and Applications*, 314(2):436–454, 2006.

[95] M. Rasmussen. Finite-time attractivity and bifurcation for nonautonomous differential equations. *Differential Equations and Dynamical Systems*, 18(1):57–78, 2010.

[96] K. Rateitschak and O. Wolkenhauer. Thresholds in transient dynamics of signal transduction pathways. *Journal of Theoretical Biology*, 264(2):334–346, 2010.

[97] D. Ruelle. Ergodic theory of differentiable dynamical systems. *Publications Mathématiques de L'IHÉS*, 50:27–58, 1979.

[98] R. J. Sacker and G. R. Sell. A Spectral Theory for Linear Differential Systems. *Journal of Differential Equations*, 27(3):320–358, 1978.

[99] S. C. Shadden. *A dynamical systems approach to unsteady systems*. PhD thesis, California Institute of Technology, Pasadena, California, 2006.

[100] S. C. Shadden. Lagrangian Coherent Structures. In R. Grigoriev, editor, *Transport and Mixing in Laminar Flows: From Microfluidics to Oceanic Currents*. WILEY-VCH Verlag, 2011.

[101] S. C. Shadden, F. Lekien, and J. E. Marsden. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D*, 212(3-4):271–304, 2005.

[102] S. Siegmund. *Spektral-Theorie, glatte Faserungen und Normalformen für Differentialgleichungen vom Carathéodory-Typ*. PhD thesis, Augsburg, 1999.

[103] S. Siegmund. Dichotomy Spectrum for Nonautonomous Differential Equations. *Journal of Dynamics and Differential Equations*, 14(1):243–258, 2002.

[104] S. Siegmund. Normal Forms for Nonautonomous Differential Equations. *Journal of Differential Equations*, 178(2):541–573, 2002.

[105] S. T. Smith. *Optimization techniques on Riemannian manifolds*, pages 113–136. Hamiltonian and Gradient Flows, Algorithms and Control. Fields Institute Communications, AMS, 1994.

[106] C. Udrişte. *Convex Functions and Optimization Methods on Riemannian Manifolds*. Kluwer Acad. Publ., Dordrecht, 1994.

[107] R. E. Vinograd. Exact bounds for exponential dichotomy roughness I. Strong dichotomy. *Journal of Differential Equations*, 71(1):63–71, 1988.

[108] R. E. Vinograd. Exact bounds for exponential dichotomy roughness III. Semistrong dichotomy. *Journal of Differential Equations*, 91(2):245–267, 1991.

[109] S. Wiggins. *Chaotic Transport in Dynamical Systems*. Interdisciplinary Applied Mathematics. Springer, 1992.

[110] Y. Wong. Differential Geometry of Grassmann Manifolds. *Proceedings of the National Academy of Sciences*, 57(3):589–594, 1967.

[111] T. J. Ypma. Historical Development of the Newton-Raphson Method. *SIAM Review*, 37(4):531–551, 1995.

[112] F. Zhang. *Matrix Theory: Basic Results and Techniques*. Universitext. Springer, 2. edition, 2011.

Index

Affirmation / Versicherung

Hereby I affirm that I wrote the present thesis without any inadmissible help by a third party and without using any other means than indicated. Thoughts that were taken directly or indirectly from other sources are indicated as such. This thesis has not been presented to any other examination board in this or a similar form, neither in Germany nor in any other country.

I have written this dissertation at Dresden University of Technology under the scientific supervision of Prof. Dr. rer. nat. habil. Stefan Siegmund.

There have been no prior attempts to obtain a PhD at any university.

I accept the requirements for obtaining a PhD (Promotionsordnung) of the Faculty of Science of the TU Dresden, issued February 23, 2011 with the changes in effect since June 15, 2011.

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Dissertation habe ich an der Technischen Universität Dresden unter der wissenschaftlichen Betreuung von Prof. Dr. rer. nat. habil. Stefan Siegmund angefertigt.

Es wurden zuvor keine Promotionsvorhaben unternommen.

Ich erkenne die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften der TU Dresden vom 23. Februar 2011 mit der letzten Änderung durch den Fakultätsratsbeschluss vom 15.06.2011 an.