Retraction-based numerical methods for continuation, interpolation and time integration on manifolds

Axel Elie Joseph SÉGUIN

EPFL

École
polytechnique
fédérale
de Lausanne

2023

Acknowledgements

The following thesis is not only the output of four years in a mathematics department, it is the culmination of a lifelong journey that was all but a lonely expedition. In the following, I hope that words can convey the profound gratitude I hold for the many people whose direct or indirect contribution was essential in reaching this goal.

I would like start by thanking the members of my PhD defense jury, namely the jury president Maria Colombo and the experts Nicolas Boumal, Bart Vandereycken and Ralf Zimmermann. It has been a real privilege to have such a panel of eminent researchers evaluate my work, and I would like to thank them for their commitment to this demanding task, as well as their valuable suggestions and remarks both at the defense and throughout my PhD years. This extends to my PhD advisor Daniel Kressner, to which I wish to dedicate an additional word of thanks. Learning how to do research on your side has been a great honor. I especially appreciate the time you took to co-write the first article with me. Despite its occasional frustrations, this experience taught me some invaluable lessons both on the scientific and human side. Throughout our years together, you have shown me that life can be approached with positive pragmatism and that there is no such thing as bad weather, only bad clothing. Most of all, I would like to thank you for your trust and for believing in my capacities before I could myself. If it wasn't for you, I would have probably never undertaken the path of a PhD and the opportunity to pursue it under your supervision was one of the greatest gift I ever received.

My thesis was kick-started by another great present: a mathematical problem crafted by the insight of Nicolas Berger. I cannot stress enough how much I appreciate this gift, Nicolas. A part from its sheer beauty, this problem gave me the opportunity to learn many new concepts and served as a beacon in the navigation to new ideas and intuitions. I would like to acknowledge the enlightened support of Michel Rollier, whose vision and charisma has been an inexhaustible source of inspiration and without which this beautiful mathematical adventure would have never been possible.

While taking each of the small steps that brought me closer to this goal, I had the great fortune of being surrounded by many friend and not one, but two teams of passionate, talented and thoughtful colleagues.
My daily work withing the innovation cell, the math group and the software team has been extraordinarily enriching experience from all perspectives. I would like to thank

Acknowledgements

you all for everything you did to support me in the past years and for all the wonderful moments spent together. A special thanks goes to Jonathan, Uri and Dario and for their invaluable guidance.

I will forever cherish also the memories of the time spent within the ANCHP group and the mathematics departments. The academic brotherhood we have acquired goes much beyond a common entry ou our CVs and I would like to thank each of you for the sharing of knowledge, thoughts, life, emotions (and beers) that made these years so enjoyable. Needless to say, the above extends to all my friends inside and outside campus. Gio, Haro, Velo, Gini, Kyara, Giulio, Fabio, Antonio, Gianluca, David, Paride, Léo, Brian, Robin, Jonas, Gil, Julien, Rémi, Edu, Fabienne, Sonia, thanks for your existence and for all your love.

I would like to also pay a warm tribute to the dedication and passion of Vivien Douine, David Burgat and Bertrand Delclos, who, in my in high-school years, introduced me to the joys of science for the first time.

Last but definitely not the least, let me express the final word of thanks to my beloved family, and specifically to my parents. Thank you for your unconditional love and for teaching me by your example the value of hard work. I am deeply grateful for the sacrifices you made to provide my brothers and me with the best possible education. This thesis is dedicated to you.

Abstract

The goal of this thesis is the development and the analysis of numerical methods for problems where the unknown is a curve on a smooth manifold. In particular, the thesis is structured around the three following problems: homotopy continuation, curve interpolation and integration of ordinary differential equations. To accommodate the manifold constraint, all the proposed methods feature as a central ingredient the concept of retractions, as extensively developed in the context of Riemannian optimization methods. A retraction can be thought as a generic device for crafting portions of manifold-constrained curves which are in general computationally cheaper to evaluate than the geodesics defining the Riemannian exponential map. Yet, the axiomatic definition of a retraction reveals to be rich enough for algorithms originally defined on Euclidean spaces to be adapted to the manifold setting using retractions while maintaining properties that are analogous to their Euclidean ancestor. We provide this type of analysis for the methods proposed in the thesis and we showcase the performance of the algorithms with experiments involving matrix manifolds, notably the fixed-rank matrix manifold.

First, we consider a generalization of numerical continuation methods for their application to Riemannian optimization problems. In practice, we propose a retraction-based path-following numerical continuation algorithm for efficiently solving a sequence of Riemannian optimization problems of which the last is the actual problem of interest. Then, we address the problem of Hermite interpolation, whereby a sequence of manifold points are interpolated by a manifold curve whose velocity is prescribed at each interpolation point. For this, we introduce a generalization of the de Casteljau algorithm where suitably chosen retraction curves replace the straight lines of the original algorithm. Lastly, we tackle numerical integration of manifold-constrained ordinary differential equations, in particular for equations evolving on low-rank matrix manifolds encountered in the field of dynamical low-rank approximation. We derive two methods defined using retractions which exhibit second-order convergence of the approximation error with respect to the time integration step.

Keywords. Retraction, numerical continuation, Riemannian optimization, manifold interpolation, fixed-rank manifold.

Résumé

L'objectif de cette thèse est le développement et l'analyse de méthodes numériques pour des problèmes où l'inconnue est une courbe sur une variété différentiable. En particulier, la thèse se structure autour des trois problèmes suivants : la continuation par homotopie, l'interpolation de courbe et l'intégration d'équations différentielles ordinaires. Afin d'accommoder la contrainte de variété, les méthodes proposées se basent sur le concept de *rétraction*, tel qu'amplement développé dans le contexte des méthodes numériques pour l'optimisation riemannienne. Une rétraction peut être interprétée comme un dispositif générique permettant la construction de portions de courbes sur une variété dont le calcul est en général computationnellement moins coûteux que celui pour les géodésiques définissant l'application exponentielle riemannienne. Néanmoins, la définition axiomatique d'une rétraction se révèle suffisamment riche pour adapter des algorithmes initialement conçus sur des espaces euclidiens aux cas d'une variété différentiable en utilisant des rétractions, tout en préservant des propriétés analogues à l'algorithme euclidien d'origine. Ce type d'analyse est fourni pour les méthodes proposées dans cette thèse et les performances des algorithmes sont illustrées avec des expériences sur des variétés matricielles, en particulier la variété des matrices de rang fixé.

En premier lieu, nous généralisons une méthode de continuation numérique pour son application à l'optimisation riemannienne. Concrètement, nous élaborons un algorithme dit de *traçage de courbe* basé sur une rétraction pour résoudre de manière efficace une suite de problèmes d'optimisation riemannienne dont le dernier est celui d'intérêt principal. Par la suite, nous affrontons le problème d'interpolation de Hermite, dans lequel une suite de points sur une variété est interpolée avec une courbe lisse incluse dans la variété et dont la vitesse est prescrite à chaque point d'interpolation. Dans ce but, nous introduisons une généralisation de l'algorithme inventé par de Casteljau, où des segments de courbe de rétraction bien choisis remplacent les segment de droites de l'algorithme original. Enfin, nous traitons l'intégration numérique d'équations différentielles ordinaires contraintes à une variété, en particulier les équations décrivant des dynamiques sur des variétés de matrices de rang bas, telles que considérées dans le cadre de l'*approximation dynamique de rang bas*. Nous proposons deux méthodes exprimées en termes de rétractions qui exhibent une convergence d'ordre deux par rapport au pas d'intégration.

Mots-clés. Rétraction, continuation numérique, optimisation riemannienne, interpolation sur variété différentiable, variété des matrices de rang fixé.

Compendio

L'obiettivo della seguente tesi è lo sviluppo e l'analisi di metodi numerici per la risoluzione di problemi in cui l'incognita è una curva contenuta in una varietà differenziabile. Nello specifico, la tesi si articola intorno ai tre seguenti problemi: la continuazione per omotopia, l'interpolazione di una curva e l'integrazione di equazioni differenziali ordinarie. Per soddisfare il vincolo di varietà, tutti i metodi proposti si avvalgono centralmente del concetto di *retrazione*, come ampiamente usato nel contesto dei metodi numerici per l'ottimizzazione riemanniana. Una retrazione può essere interpretata come un generico dispositivo per costruire porzioni di curve su una varietà, il cui calcolo è in generale computazionalmente meno oneroso di quello per le geodetiche che definiscono la mappa esponenziale riemanniana. Ciò nonostante, la definizione assiomatica di una retrazione risulta sufficientemente ricca per permettere di adattare algoritmi concepiti su spazi euclidei al caso di una varietà differenziabile facendo uso di retrazioni, sempre mantenendo proprietà analoghe all'algoritmo euclideo di origine. Questo tipo di analisi viene fornito per i metodi proposti in questa tesi e le prestazioni degli algoritmi vengono illustrate con esperimenti su varietà matriciali, in particolare la varietà delle matrici di rango fisso.

In primo luogo, viene considerata una generalizzazione di un metodo di continuazione numerica per la sua applicazione all'ottimizzazione riemanniana. In concreto, viene elaborato un algoritmo detto di *tracciamento di curva* basato su una retrazione per risolvere efficacemente una sequenza di problemi di ottimizzazione riemanniana di cui l'ultimo è l'effettivo problema di interesse. In seguito, viene affrontato il problema dell'interpolazione di Hermite, nel quale è richiesto di interpolare una sequenza di punti su una varietà differenziabile con una curva liscia inclusa nella varietà e la cui velocità è prescritta ad ogni punto di interpolazione. A questo scopo, introduciamo una generalizzazione dell'algoritmo di de Casteljau in cui opportuni segmenti di curva di retrazione sostituiscono i segmenti lineari dell'algoritmo d'origine. Infine, viene trattata l'integrazione numerica di equazioni differenziali ordinarie vincolate ad una varietà differenziabile, in particolare equazioni che descrivono dinamiche su varietà di matrici di rango basso come considerate nell'ambito dell'*approssimazione dinamica di rango basso*. Vengono proposti due metodi espressi con retrazioni che esibiscono convergenza dell'errore di ordine due rispetto al passo di integrazione.

Parole chiave. Retrazione, continuazione numerica, ottimizzazione riemanniana, interpolazione su varietà differenziabile, varietà delle matrici di rango fisso.

Contents

Contents

Introduction

One of the defining features of the human mind is its capacity to conceive abstractions. Abstractions are general ideas that do not relate to a specific object or situation but rather unify from a certain viewpoint a collection of objects or situations. The construction of insightful abstractions is at the very essence of the scientific endeavor. Mathematics is a method by which abstractions can be rigorously manipulated and built upon. Then, logical deductions and algorithmic procedure operated at an abstract level can find concrete expression in a multitude of situations of the physical world. This extraordinary compression of reality enabled by well-chosen abstractions certainly carries pragmatic advantages, in addition to a purely aesthetic value to the eyes of some.

The context of the present thesis is the collection of abstractions forming the theory of differential geometry, more specifically the theory of smooth manifolds. This branch of mathematics initiated by Gauss and Riemann in the eighteenth century applies to a multitude of surprisingly diverse mathematical objects, which themselves concretize in a even greater variety of real world entities, as diverse as the leading vibrational modes of a guitar string, the position in three-dimensional space of a robotic manipulator, or a table of numbers representing the preferences of a human population. All these objects share a common feature: they can interpreted as points on a particular differentiable manifold. This enables to consider them all at once from the abstract lens of differential geometry.

Riemannian optimization and retractions

The theory of differentiable manifolds appeared in recent years in many fields of applied mathematics. A prominent example is Riemannian optimization [AMS08, Bou23]. Indeed, mathematical models encountered in contemporary science and engineering frequently involve the minimization of an objective function over an admissible set which admits the structure of a smooth manifold. Riemannian optimization is the now well-established field of research dealing with the development of numerical methods to specifically solve manifold-constrained optimization problems. Riemannian optimization methods are constructed and analyzed for a general Riemannian manifold in terms of precisely defined abstract objects, such as Riemannian metrics and retractions, that materialize differently for each manifold. Providing concrete instances of these objects for the particular manifold at hand is sufficient to automatically benefit the effectiveness of the

Introduction

Riemannian optimization algorithms granted by the general analyses.

The modularity of Riemannian optimization methods as initiated by [AMS08] follows a pattern where algorithmic constructs are decoupled from the particular entities involved in such algorithms and linked by a clear interface. This structure is recurrent in applied mathematics and may be recognized in the language of software engineering as *generic programming*, a popular programming paradigm allowed by many functional and object-oriented programming languages, such as in C++ under the name of *templates*. Riemannian optimization methods are *generic* in that they are defined for a *generic object* (a manifold) for which a set of operations or *abstract methods* used in the algorithm are available (metric, retractions,...). This is precisely the design pattern employed in Manopt, the MATLAB library for Riemannian optimization [BMAS14]. As also put forth in the celebrated book *Design Patterns: Elements of Reusable Object-Oriented Software*[1], this design pattern reduces code duplication. In mathematical theories this translates into reduction of proof duplication. Furthermore, genericity offers a simplified perspective on the algorithmic or mathematical problem at hand, distilled from all the particular details which may hinder understanding and intuition.

An important characterization of Riemannian manifolds prescribes that around each point they can be locally approximated by an Euclidean space. On the other hand, most iterative methods for minimizing a target function on an Euclidean space involve an update rule that is only local. Hence, Riemannian optimization methods are usually derived from pre-existing Euclidean optimization methods based on the above intuition suggesting that the generalization should maintain at least part of the favorable properties of the original method. Generalizing local constructs of flat spaces to manifolds has been a driving force for the development of differential geometry itself, Riemannian optimization, and also the present thesis.

Retractions are the operative tool Riemannian optimization algorithms use to link the local geometry of the manifold to a flat local representation. A retraction is a type of map defined at each point of the manifold from a subset of the tangent space containing the origin to a neighborhood of the point on the manifold. Given any point on the manifold, the defining properties of retractions are sufficient to construct a portion of manifold curve passing through the given point with any prescribed direction. The update rule of most Riemannian optimization algorithms involves traveling along such curves for a short distance. On the one hand, this guarantees the iterates of the method remain constrained to the admissible set of the optimization problem of interest. On the other hand, the definition of retraction proved to be sufficient for establishing convergence guarantees for the Riemannian optimization analogously to their Euclidean ancestor. In this thesis we take advantage of these two aspects and consider retractions as a general manifold curve

---

[1]Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides (1994)

generating device to numerically solve manifold-constrained mathematical problems.

The possibility to reproduce standard results of Euclidean algorithms for their retraction-based Riemannian counterpart seems to have motivated its definition in the first place. In fact, a definition containing the defining feature of what we now call a retraction appeared in the context of numerical integration of ordinary differential equations on manifolds in the work of Shub [Shu86]. In this work, the definition only qualifies a particular type of local chart which can replace to the Riemannian exponential map to generalize the forward Euler method, while still enabling to prove the convergence of the numerical scheme. From a computational perspective, the possibility to use retractions instead of the Riemannian exponential map offers another great benefit which also contributed to the success of retractions in Riemannian optimization. There exist retractions which are much cheaper to compute than the solution to the differential initial value problem that defines an evaluation of the exponential map. As a final historical note, the term retraction seems to have been first employed in [ADM$^+$02] and now earned its place in the vocabulary of applied mathematics, despite a unfortunate naming collision with the much older concept of retraction in topology.

Contributions and structure of the thesis

The present thesis explores possibilities offered by the abstraction of retractions, as popularized and extensively developed in the Riemannian optimization literature, but in a broader context of numerical methods for manifold-constrained problems. Summarizing the preceding discussion, the motivations can be declined as:

- *practical*: the availability of retractions for numerous manifolds of interest makes retraction-based algorithms broadly applicable;
- *computational*: retractions are in general much cheaper to compute than the Riemannian exponential map;
- *theoretical*: the defining properties of retractions are minimal yet rich enough to analyze the Riemannian retraction-based algorithms analogously to their Euclidean counterpart.

From a global perspective, the thesis is constructed around three different manifold-constrained problems for which we propose and analyze retraction-based algorithms. The methods are then tested experimentally on problems where the constraint manifold is a matrix manifold, notably the manifold of fixed-rank matrices.

Numerical continuation. We start by extending with the use of retractions a numerical continuation algorithm that can be applied to Riemannian optimization problems that depend on one scalar parameter. The goal of the algorithm is to track the evolution of a critical point of the optimization problem as the scalar parameter changes. The Riemannian Newton Continuation method, as we denominate it, is a generalization of a predictor-corrector continuation algorithm developed originally for nonlinear equation. In analogy with its Euclidean counterpart, a retraction step produces an estimate for the

solution at the new parameter value and a Riemannian optimization algorithm is invoked to correct the estimate. From a theoretical perspective, we extend the analysis inspired from the Euclidean setting which guarantees the well-posedness of the continuation problem. For illustrating the performance of the algorithm, the central example application is the low-rank matrix-completion problem, tackled from its Riemannian optimization perspective. The developments on this first topic are based on [SK22a].

Interpolation. The second problem of the thesis concerns interpolation on manifolds. In particular, we address the Hermite interpolation problem on a manifold which aims at constructing a smooth manifold curve passing through a set of given interpolation points with a prescribed velocity. We derive a retraction-based interpolation algorithm producing a piecewise solution to the problem that is globally continuously differentiable. The method relies on a generalization of the de Casteljau algorithm where the linear segments are replaced by retraction curves. The building blocks of the algorithm are suitably selected from a novel parametric family of retraction-based curves of which endpoint can be prescribed. Whenever a retraction whose local inverse can be computed efficiently is available, the endpoint retraction curves and, in turn, the retraction-based Hermite interpolation algorithm can be used. This is the case for the fixed-rank manifold, for which the orthographic retraction is available and admits an inverse that is very easy to compute. The well-posedness of the method is proved thanks to the new concept of retraction-convex sets, a generalized notion of convexity for a subset of a manifold involving retraction curves. Furthermore, the classic result characterizing the convergence of the approximation error for polynomial Hermite interpolation is extended to the retraction-based scheme. This work is based on the following submitted preprint [SK22b].

Integration of ODEs. At last, we turn to manifold-constrained ordinary differential equation. We propose two numerical integration methods entirely expressed in terms of retractions which can achieve second-order global error convergence. The first is the accelerated forward Euler scheme, based on incorporating the acceleration of the solution in the update rule thanks to a second-order retraction and a suitable tangent vector. The second algorithm is a two step method echoing an explicit Runge-Kutta method known as the Ralston method. The second step of the method uses the retraction-based Hermite interpolant previously introduced. Hence, we denominate this method the Ralston-Hermite method. The application part focuses on differential equations evolving on the fixed-rank matrix manifold encountered in the context of dynamical low-rank approximation (DLRA). The accuracy and the stability to small-singular values of the solution is assessed via numerical experiments on classical instances of the DLRA literature. A preprint on the contents of this chapter is in preparation [SCK23].

Outline of the thesis.

The first two chapters lay the foundations for the subsequent developments. We report a selection of well-known textbook material in the interest of a self-contained exposition and

with the intent to fix notations. Chapter 1 introduces the fundamentals of differentiable manifolds and Riemannian manifolds while the necessary background on Riemannian optimization tools and in particular retractions is the object of Chapter 2.

In Chapter 3, we delve into retractions by reporting known theoretical results as well as new practical and theoretical possibilities offered by the concept of retraction. The original contributions include a class of retraction curves with prescribed endpoint, the analysis of their well-posedness with the definition of retraction-convexity, and the proof of the existence of retraction-convex set. A digression on the Lipschitz continuity of retractions and the approximating properties of basic retraction curves concludes the chapter.

Chapters 4, 5 and 6 are respectively dedicated to the new retraction-based numerical methods for the three previously announced manifold-constrained problems: continuation, interpolation and numerical integration of ODEs.

1 Manifolds

This chapter is a brief introduction to the basic concepts of differential geometry needed throughout the thesis. We first discuss smooth manifolds, then specialize to Riemannian manifolds. A particular emphasis is given to the case of manifolds embedded into a vector space. The exposition is inspired from the Riemannian optimization references [AMS08, Bou23] and the classic textbooks [Lee13, Lee18]. We conclude the chapter by illustrating these concepts for particular matrix manifolds, thereby collecting well-known material that is used in the following chapters.

1.1   Smooth manifolds

At the most fundamental level, manifolds are topological spaces which can be locally mapped to subsets of $\mathbb{R}^d$. This is achieved with *local charts*.

Definition 1.1. *A* local chart *of a topological space* $\mathcal{M}$ *is a pair* $(\mathcal{U}, \varphi)$*, where* $\mathcal{U} \subseteq \mathcal{M}$ *is an open subset and* $\varphi : \mathcal{U} \to \mathbb{R}^d$ *is a homeomorphism, for some* $d \geq 1$*.*

The local chart is said to contain a point $x \in \mathcal{M}$ if $x \in \mathcal{U}$. Given two local charts $(\mathcal{U}_\alpha, \varphi_\alpha)$ and $(\mathcal{U}_\beta, \varphi_\beta)$ such that $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$, if the *transition map* $\varphi_\alpha \circ \varphi_\beta^{-1}$ is a diffeomorphism from $\varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$ to $\varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$, then the local charts are said to be *smoothly overlapping*.
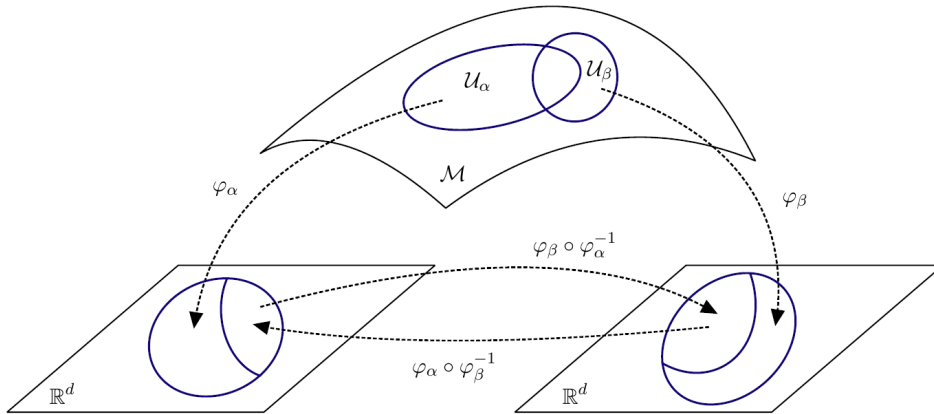
Figure 1.1: Smoothly overlapping local charts and their transition maps.

Two local chart are *compatible* if they are smoothly overlapping or if $\mathcal{U}_\alpha \cap \mathcal{U}_\beta = \emptyset$. A set of compatible local charts $\mathcal{A} = \{(\mathcal{U}_\alpha, \varphi_\alpha)\}$ constitutes a *smooth atlas* provided that $\bigcup_\alpha \mathcal{U}_\alpha = \mathcal{M}$. It is called *maximal* if it is not properly contained in any other smooth atlas. Throughout the whole thesis, unless otherwise stated, we use the term *smooth* to indicate infinitely differentiable. For instance, a diffeomorphism is an infinitely differentiable bijection with an infinitely differentiable inverse.

Definition 1.2 ([AMS08, §3.1.1]). *A d-dimensional smooth manifold is a topological space $\mathcal{M}$ equipped with a maximal smooth atlas $\mathcal{A}$ onto $\mathbb{R}^d$ inducing a second-countable Hausdorff topology.*

The Hausdorff property in Definition 1.2 implies the uniqueness of limit points and, combined with the second-countability, grants the existence of Riemannian metrics [Lee13, Proposition 2.4] and affine connections [Lee13, Proposition 4.12].

Example 1.3 (Linear manifolds). Any finite dimensional vector space $\mathcal{E}$ such as $\mathbb{R}^n$ or $\mathbb{R}^{m \times n}$ can be endowed with a smooth manifold structure. For any choice of basis of $\mathcal{E}$ denoted $\{e_i\}$, the maps $\varphi : \mathcal{E} \to \mathbb{R}^n : v = \sum_{i=1}^d v_i e_i \to (v_1, \ldots, v_n)$ define smoothly overlapping local charts that constitute the linear manifold structure of $\mathcal{E}$.

1.1.1 Smooth maps

One key feature of the definition of a smooth manifold is that it allows the extension of calculus to mappings between manifolds.

Definition 1.4. *A map $F : \mathcal{M} \to \mathcal{N}$ between manifolds $\mathcal{M}$ and $\mathcal{N}$ is a smooth map if for every $x \in \mathcal{M}$ there exist local charts $(\mathcal{U}, \varphi)$ containing x and $(\mathcal{V}, \psi)$ containing $F(x)$ such that $\hat{F} := \psi \circ F \circ \varphi^{-1}$ is smooth.*
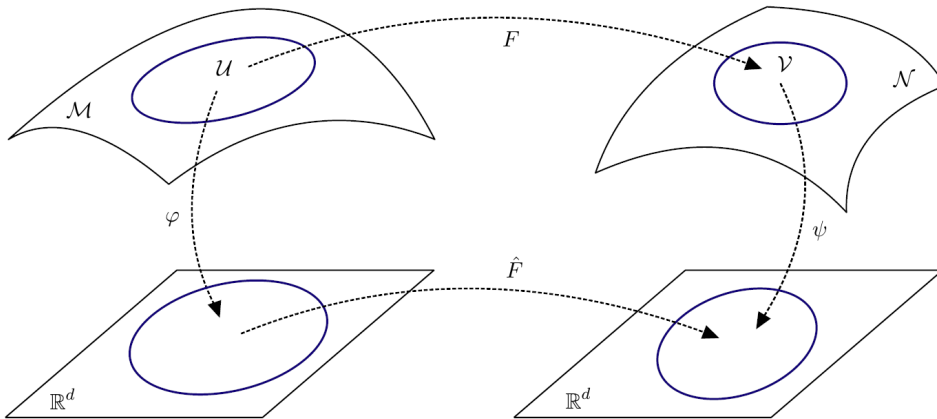
Figure 1.2: Coordinate representation of a map between manifolds.

The function $\hat{F}$ is referred to as the *coordinate representation* of F. We adopt the convention of denoting it with a hat decorating the original symbol. Definition 1.4 gives

meaning to two important type of smooth maps. A *smooth manifold curve* is a smooth map $\gamma : J \subset \mathbb{R} \to \mathcal{M}$ and a *smooth scalar field* is a smooth map $f : \mathcal{M} \to \mathbb{R}$. The set of all smooth scalar fields on $\mathcal{M}$ is denoted $\mathfrak{F}(\mathcal{M})$.

In this work, for any smooth scalar function $g : \mathbb{R} \to \mathbb{R}$ the symbol $\frac{\mathrm{d}}{\mathrm{d}t}$ indicates the usual derivative, i.e. the limit of the incremental ratio $\lim_{h \to 0} \frac{g(t+h)-g(t)}{h} =: \frac{\mathrm{d}g(t)}{\mathrm{d}t}$, also denoted $g'(t)$. The same notation and definition apply for smooth curves $\gamma : \mathbb{R} \to \mathcal{E}$ mapping to a vector space $\mathcal{E}$. For the directional derivative of a smooth function $F : \mathcal{E} \to \mathcal{E}'$ between linear spaces, we employ the notation $\mathrm{D}f(x)\,[v] := \frac{\mathrm{d}}{\mathrm{d}t} F(x+tv)\big|_{t=0}$, for any $x, v \in \mathcal{E}$, and the linear operator $\mathrm{D}f(x) : \mathcal{E} \to \mathcal{E}'$ is the differential of F at x.

1.1.2 The tangent space

The notion of directional derivative for a smooth scalar field on $\mathcal{M}$ is introduced jointly with the notion of direction on a manifold, i.e. tangent vectors.

Definition 1.5 ([AMS08, §3.5.1]). *For any smooth manifold curve $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$, the tangent vector to the curve in $t = 0$ is the linear map $\dot{\gamma}(0) : \mathfrak{F}(\mathcal{M}) \to \mathbb{R}$ defined by $\dot{\gamma}(0)f := \frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}\big|_{t=0}$ for any $f \in \mathfrak{F}(\mathcal{M})$.*
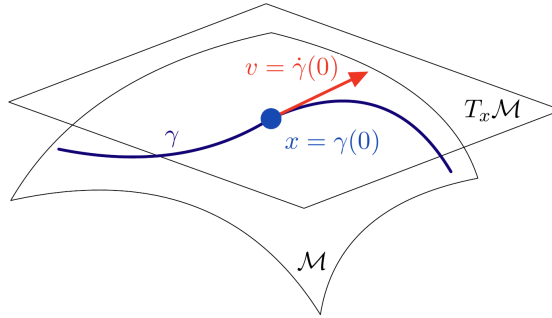
Figure 1.3: A smooth manifold curve γ realizing a tangent vector $v \in T_x\mathcal{M}$.

When it is clear from context that γ is a manifold curve, we also employ the symbol $\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\big|_{t=0}$ to indicate $\dot{\gamma}(0)$. With this definition, two manifold curves γ_1 and γ_2 may share the same tangent vector. In fact, given their coordinate representation $\hat{\gamma}_1$ and $\hat{\gamma}_2$ it holds that $\dot{\gamma}_1(0) = \dot{\gamma}_2(0)$ if and only if $\hat{\gamma}_1'(0) = \hat{\gamma}_2'(0)$ [AMS08, Proposition 3.5.2]. We say γ_1 and γ_2 *realize* the same tangent vector.

Definition 1.6 ([AMS08, Definition 3.5.1]). *A tangent vector v to a manifold $\mathcal{M}$ at x is a linear map $v : \mathfrak{F}(\mathcal{M}) \to \mathbb{R}$ such that there exists a smooth manifold curve realizing v, i.e. $\gamma(0) = x$ an $\dot{\gamma}(0) = v$.*

The collection of all tangent vectors at $x \in \mathcal{M}$ is called the *tangent space* at x and it is denoted $T_x\mathcal{M}$. The tangent space is a vector space of dimension $\dim(\mathcal{M})$ [Lee13, Proposition 3.10].
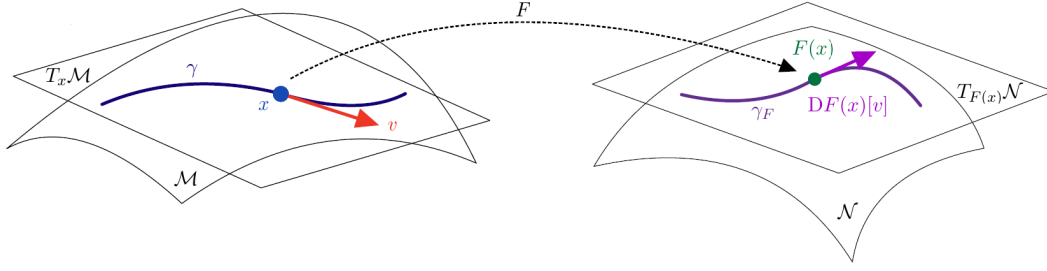
Figure 1.4: The differential of a smooth map F between manifolds $\mathcal{M}$ and $\mathcal{N}$.

Example 1.7 (Tangent space for linear manifolds). For any linear space $\mathcal{E}$ and some $x \in \mathcal{E}$, a tangent vector $v \in T_x\mathcal{E}$ acts on $f \in \mathfrak{F}(\mathcal{E})$ as $vf = \frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}\big|_{t=0} = \mathrm{D}f(x)[\gamma'(0)]$, for some curve γ in $\mathcal{E}$. Since $\mathrm{D}f(x)[\gamma'(0)]$ depends only on $\gamma'(0) =: \bar{v} \in \mathcal{E}$, and not on the particular choice of γ, there is a one-to-one correspondence between any $v \in T_x\mathcal{E}$ and $\bar{v} \in \mathcal{E}$. Therefore, we may identify $T_x\mathcal{E} \simeq \mathcal{E}$ for any vector space $\mathcal{E}$. In particular, for any manifold $\mathcal{M}$ we conflate $T_vT_x\mathcal{M} \simeq T_x\mathcal{M}$ for any $v \in T_x\mathcal{M}$ and $x \in \mathcal{M}$.

For any smooth map between manifolds, the differential is defined as follows.

Definition 1.8. *The differential of a smooth map $F : \mathcal{M} \to \mathcal{N}$ at $x \in \mathcal{M}$ between manifolds $\mathcal{M}$ and $\mathcal{N}$ is the linear map $\mathrm{D}F(x) : T_x\mathcal{M} \to T_{F(x)}\mathcal{N}$ defined by*

$$\mathrm{D}F(x)[v] := \dot{\gamma}_F(0), \quad v \in T_x\mathcal{M},$$

where γ is any smooth curve realizing v and $\gamma_F = F \circ \gamma$ is a smooth manifold curve on $\mathcal{N}$ passing through $F(x)$ in $t = 0$.

1.1.3 The tangent bundle and smooth vector fields

The disjoint union of tangent spaces is denoted

$$T\mathcal{M} = \coprod_{x \in \mathcal{M}} T_x\mathcal{M} := \{(x, v) : x \in \mathcal{M}, \, v \in T_x\mathcal{M}\}$$

and is known as the *tangent bundle*. It admits a natural smooth manifold structure of dimension $2\dim(\mathcal{M})$ with respect to which the projection $\pi_1 : T\mathcal{M} \to \mathcal{M} : (x, v) \mapsto x$ is a smooth map [Lee13, Proposition 3.18]. This projection returns the so called *foot* or *anchor point* of the tangent vector.

Definition 1.9. *A smooth vector field on $\mathcal{M}$ is a smooth map $V : \mathcal{M} \to T\mathcal{M}$ such that $\pi_1(V(x)) = x$ for all $x \in \mathcal{M}$. The set of all smooth vector fields is denoted $\mathfrak{X}(\mathcal{M})$.*

A similar definition hold for a smooth vector field Z along a smooth manifold curve $\gamma : J \subset \mathbb{R} \to \mathcal{M}$. It is a smooth map $Z : J \to T\mathcal{M}$ such that $\pi_1(Z(t)) = \gamma(t), \, \forall t \in J$. The set of smooth vector fields along γ is denoted $\mathfrak{X}(\gamma)$.

1.1.4 Smooth tensor fields

Generalizing the notion of vector field, a tensor field of order k associates at each point of the manifold a tensor of order k, viewed here as a scalar valued map depending linearly on k tangent vectors. The tensor field is considered smooth when it maps any collection of k smooth vector fields to a smooth scalar field.

Definition 1.10 ([Bou23, Definition 10.76]). *A smooth tensor field T of order k on a manifold $\mathcal{M}$ is a map*

$$T : \mathfrak{X}(\mathcal{M})^k \mapsto \mathfrak{F}(\mathcal{M}),$$

that is $\mathfrak{F}(\mathcal{M})$-linear in all its arguments.

The property of $\mathfrak{F}(\mathcal{M})$-linearity means that a tensor field is a pointwise object, in the sense that its value at a given $x \in \mathcal{M}$ depends only on the value of the input vector fields at x. In the same way smooth vector field are maps from $\mathcal{M}$ to the tangent bundle $T\mathcal{M}$ that are smooth in the sense of Definition 1.4, a smooth tensor field is equivalently [Lee13, Proposition 12.19] a smooth map from $\mathcal{M}$ to a tensor bundle of order k defined by,

$$T^k T\mathcal{M} := \left\{ (x, L) : \ x \in \mathcal{M}, \ L : (T_x \mathcal{M})^k \to \mathbb{R} \text{ linear} \right\}. \tag{1.1}$$

A tensor bundle is a particular type of vector bundle [Lee03, §10], in that it associates at each point of the manifold a linear space. Vector bundles, in particular tensor bundles of the form (1.1), can be endowed with a smooth manifold structure [Lee13, Lemma 10.6]. By convention, scalar fields are order zero tensor fields. The differential of a smooth scalar field $f \in \mathfrak{F}(\mathcal{M})$ introduced in Definition 1.8 can be interpreted as a smooth order 1 tensor field as

$$U \in \mathfrak{X}(\mathcal{M}) \mapsto \mathrm{D}f(U) := Uf \in \mathfrak{F}(\mathcal{M}).$$

An example of tensor field of order 2 is the Riemannian metric that we introduce in Section 1.3.

1.2 Embedded submanifolds

All manifolds considered in this thesis are subsets of some finite dimensional vector space $\mathcal{E}$, which we call the *ambient space* or *embedding space*. In this setting, one can define a smooth manifold structure inherited from the linear manifold structure of $\mathcal{E}$ introduced in Example 1.3. This notion is made precise with the concept of *embedded submanifold*, of which we omit the general definition [Lee13, §5] in favor of a definition well-suited for the particular case $\mathcal{M} \subset \mathcal{E}$.

Definition 1.11 ([Bou23, Definition 3.10]). *A nonempty subset $\mathcal{M}$ of a N-dimensional vector space $\mathcal{E}$ is an* embedded submanifold *of $\mathcal{E}$ of dimension d if either*
  1. *$d = N$ and $\mathcal{M}$ is open in $\mathcal{E}$; or*
  2. *$d = N - K$, for some $K \geq 1$ and for every $x \in \mathcal{M}$ there exists a neighborhood $\mathcal{U}$ of x in $\mathcal{E}$ and a smooth function $h : \mathcal{U} \to \mathbb{R}^K$ such that*

(a) $y \in \mathcal{M} \cap \mathcal{U}$ if and only if $h(y) = 0$,

(b) rank $(\mathrm{D}h(x)) = K$. *The integer K is the* codimension *of $\mathcal{M}$.*

An embedded manifold is also a manifold in the sense of Definition 1.2 [Bou23, Proposition 8.32].

A point of the submanifold can be viewed as a point of the ambient space via the *inclusion map* $\iota : \mathcal{M} \to \mathcal{E} : x \to x$. Under the conditions of Definition 1.11, $\mathcal{M}$ admits a manifold structure whose local charts guarantee that ι is a smooth map from $\mathcal{M}$ to $\mathcal{E}$.

Example 1.12 (Stiefel manifold). The set of real column orthogonal matrices of size $n \times k$, is defined as $\mathrm{St}(n, k) := \left\{ X \in \mathbb{R}^{n \times k} : X^{\top} X = I_k \right\}$ and is often called the Stiefel manifold. It is indeed a manifold as it can be seen as an embedded submanifold of $\mathbb{R}^{n \times k}$. It is equal to the zero level set of the smooth function $h : \mathbb{R}^{n \times k} \to \mathbb{S}_k : X \to X^{\top} X - I_k$, mapping to the vector space of $k \times k$ symmetric matrices $\mathbb{S}_k$, isomorphic to $\mathbb{R}^{\dim(\mathbb{S}_k)}$ with $\dim(\mathbb{S}_k) = k(k+1)/2$. The directional derivative of h at $X \in \mathbb{R}^{n \times k}$ along $V \in \mathbb{R}^{n \times k}$ is $\mathrm{D}h(X)[V] = X^{\top} V + V^{\top} X$. As expected, $\mathrm{D}h(X)[V]$ is always symmetric and, for any $X \in \mathrm{St}(n, k)$ and $V = \frac{1}{2} X S$ for any given $S \in \mathbb{S}_k$, then $\mathrm{D}h(X) \left[\frac{1}{2} X S \right] = S$. This shows that for $X \in \mathrm{St}(n, k)$, $\mathrm{D}h(X)$ has constant rank $k(k+1)/2$. Hence, owing to Definition 1.11 the set $\mathrm{St}(n, k)$ is indeed an embedded submanifold of $\mathbb{R}^{n \times k}$ of dimension $nk - \frac{1}{2} k(k+1)$.

1.2.1 Smooth maps on embedded submanifolds

A useful characterization of $\mathfrak{F}(\mathcal{M})$ when $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$ is that smooth scalar fields can be smoothly extended to a neighborhood of $\mathcal{M}$ in the ambient space. This is possible as consequence of a more general result given in [Bou23, Proposition 8.79].

Proposition 1.13. *There exists a neighborhood $\overline{\mathcal{U}}$ of $\mathcal{M}$ in $\mathcal{E}$ such that for any $f \in \mathfrak{F}(\mathcal{M})$ there exists a smooth function $\bar{f} : \overline{\mathcal{U}} \to \mathbb{R}$ such that $\bar{f}\big|_{\mathcal{M}} = f$.*

Any such $\bar{f}$ is called a *smooth extension* of f. Likewise, smooth vector fields also admit smooth extensions [Lee13, Exercise 8.15].

Proposition 1.14. *For every $V \in \mathfrak{X}(\mathcal{M})$ there exists a neighborhood $\overline{\mathcal{V}}$ of $\mathcal{M}$ and a smooth vector field $\overline{V} \in \mathfrak{X}(\overline{\mathcal{V}})$ such that $\overline{V}\big|_{\mathcal{M}} = V$.*

1.2.2 Tangent space on embedded submanifolds

Every smooth manifold curve γ on an embedded submanifold can be interpreted as a smooth curve $\iota \circ \gamma$ in the ambient space $\mathcal{E}$ via the inclusion map ι. The tangent vector to the ambient curve is obtained by ordinary differentiation as $\gamma'(t) = \frac{\mathrm{d}}{\mathrm{d}t} \iota(\gamma(t))$. If for instance $\gamma(0) = x$, the notion of tangent vector $\dot{\gamma}(0) \in T_x \mathcal{M}$ as derivation over $\mathfrak{F}(\mathcal{M})$ can be reconciled with the usual notion of derivative $\gamma'(0)$ for a curve in $\mathcal{E}$. Letting $\bar{f}$ denote a smooth extension of any $f \in \mathfrak{F}(\mathcal{M})$, then it holds that $\dot{\gamma}(0)f = \mathrm{D}\bar{f}(x)[\gamma'(0)]$ [AMS08,

§3.5.7]. This creates a link between a tangent vector of $\mathcal{M}$ and a vector of $\mathcal{E}$ that leads to the identification of the tangent space of an embedded submanifolds with a subspace of $\mathcal{E}$, as summarized in Theorem 1.15 below.

Theorem 1.15 ([Bou23, Theorem 3.15]). *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$ and $x \in \mathcal{M}$. If $\mathcal{M}$ is an open subset of $\mathcal{E}$ then $T_x\mathcal{M} = \mathcal{E}$. Otherwise, $T_x\mathcal{M} = \ker \mathrm{D}h(x)$, for any locally defining function h at x as in Definition 1.11.*

Example 1.16 (Tangent space of the Stiefel manifold). For the defining function of the Stiefel manifold introduced in Example 1.12, we aim at describing its kernel at $X \in \mathrm{St}(n,k)$, i.e. all the matrices $V \in \mathbb{R}^{n \times k}$ such that $\mathrm{D}h(X)[V] = X^\top V + V^\top X = 0$. Observe that this directional derivative vanishes along any $X\Omega$ for some $\Omega \in \mathbb{S}_k^\perp$, the set of skew-symmetric matrices. Indeed $\mathrm{D}h(X)[X\Omega] = \Omega + \Omega^\top = 0$. Furthermore, introducing any matrix $X_\perp \in \mathbb{R}^{n \times (n-k)}$ whose columns span the complement of $\mathrm{span}(X)$, for any $K \in \mathbb{R}^{(n-k) \times k}$ we have $\mathrm{D}h(X)[X_\perp K] = 0$. This shows the kernel of $\mathrm{D}h(X)$ contains a subspace of $\mathbb{R}^{n \times k}$ of dimension $\frac{1}{2}(k-1)k = \dim\left(\mathbb{S}_k^\perp\right)$, and a subspace of dimension $k(n-k)$. Summing the dimensions gives $nk - \frac{1}{2}k(k+1) = \dim(\mathrm{St}(n,k))$ and allows us to conclude

$$T_X\mathrm{St}(n,k) := \left\{ X\Omega + X_\perp K : \Omega \in \mathbb{S}_k^\perp,\ K \in \mathbb{R}^{(n-k) \times k} \right\}. \tag{1.2}$$

1.3 Riemannian manifolds

The objective of defining a notion of distance on a manifold is one motivation for the development of Riemannian manifolds. The starting point is to endow smooth manifolds with an additional structure to measure lengths and angles on each the tangent space.

Definition 1.17. *A* Riemannian metric *on a smooth manifold $\mathcal{M}$ is an inner-product $\langle \cdot, \cdot \rangle_x$ defined on $T_x\mathcal{M}$ for each $x \in \mathcal{M}$ that varies smoothly with x, in the sense that for any smooth vector fields V and W, the function $\langle U, V \rangle$ to be intended as $x \to \langle V(x), W(x) \rangle_x$ defines a smooth scalar field on $\mathcal{M}$.*

The above definition makes of the Riemannian metric a smooth tensor field of order 2 in the sense of Definition 1.10. The Riemannian metric induces on each tangent space a norm $\|v\|_x = \sqrt{\langle v, v \rangle_x}$, $\forall\, v \in T_x\mathcal{M}$. A metric is the only additional structure needed to qualify a manifold of Riemannian.

Definition 1.18. *A* Riemannian manifold *is a smooth manifold equipped with a Riemannian metric.*

Example 1.19 (Euclidean space of matrices). A vector space $\mathcal{E}$ endowed with an inner-product $\langle \cdot, \cdot \rangle$ can be turned into a Riemannian manifold by defining on $T_x\mathcal{E} \simeq \mathcal{E}$ the inner-product of $\mathcal{E}$. With this Riemannian structure, $\mathcal{E}$ is called an *Euclidean space*. For example, the Euclidean space $\mathbb{R}^{m \times n}$ is endowed with the metric $\langle V, W \rangle_X = \langle V, W \rangle := \mathrm{Tr}(V^\top W)$ for all $X \in \mathbb{R}^{m \times n}$ and $V, W \in T_X\mathbb{R}^{m \times n} \simeq \mathbb{R}^{m \times n}$.

Length and distance

The length of a manifold curves is defined next. The definition applies to any piecewise smooth manifold curve [Bou23, Definition 10.2], i.e. a continuous map $\gamma : [a, b] \to \mathcal{M}$ for which there exists a partition $a = a_0 < \cdots < a_N = b$ diving the curve into segments $\gamma\big|_{[a_{i-1}, a_i]}$ which are smooth manifold curves, $\forall\, i = 1, \ldots, N$. For such curve, the tangent vector $\dot{\gamma}(t)$ is well-defined for all $t \in [a, b]$ provided $t \neq a_i$. This is sufficient for the following definition.

Definition 1.20. *The* length *of a piecewise smooth manifold curve* $\gamma : [a, b] \to \mathcal{M}$ *is*

$$L(\gamma) := \int_a^b \|\dot{\gamma}(\tau)\|_{\gamma(\tau)} \, \mathrm{d}\tau.$$

Then, in analogy to Euclidean spaces, the Riemannian distance between to points is the length of the shortest (piecewise smooth) path joining the point. But in contrast, to the Euclidean case, the minimizing path may not be unique. The definition of Riemannian distance is restricted to a connected Riemannian manifold $\mathcal{M}$, for which any two points $x, y \in \mathcal{M}$ can be connected by a piecewise smooth manifold curve [Lee13, Proposition 2.50]. Let Γ_{xy} denote the set of all piecewise smooth manifold curves connecting x and y.

Definition 1.21. *For a connected Riemannian manifold* $\mathcal{M}$, *the* Riemannian distance *between* x *and* y *is given by* $d(x, y) = \inf\limits_{\gamma \in \Gamma_{xy}} \{L(\gamma)\}$.

An important consequence of this definition is that any connected Riemannian manifolds equipped with this distance function becomes a metric spaces [Lee13, Theorem 2.55]. As for any metric space, we can define the metric ball centered in $x \in \mathcal{M}$ of radius $r > 0$ by $B(x, r) := \{y \in \mathcal{M} : d(x, y) < r\}$. Likewise, a Riemannian manifold is said to be *complete* if it is complete as a metric space.

1.3.1 Connections and covariant differentiation

Another central construction of Riemannian geometry is a generalization for the notion of directional derivative of smooth vectors fields defined on a manifold. The generalization is not immediate as measuring the variation of a vector field along a certain direction implies comparing tangent vectors belonging to different tangent spaces. A priori, there is no canonical way for doing this and it needs to be specified. Indirectly, this is the purpose of a *connection*.

Definition 1.22. *A* connection *on a smooth manifold* $\mathcal{M}$ *is a mapping*

$$\nabla \,:\, \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$$
$$(U, V) \mapsto \nabla_U V$$

satisfying the following properties for any $U, V, W \in \mathfrak{X}(\mathcal{M})$, $f, g \in \mathfrak{F}(\mathcal{M})$ *and* $a, b \in \mathbb{R}$:

1. $\mathfrak{F}(\mathcal{M})$-*linearity in the first argument:* $\nabla_{fU+gV}W = f\nabla_U W + g\nabla_V W$,
2. $\mathbb{R}$-*linearity in the second argument:* $\nabla_U(aV + bW) = a\nabla_U W + b\nabla_V W$,
3. *Leibniz' product rule:* $\nabla_U(fV) = Uf + f\nabla_U V$.

The vector field $\nabla_U V$ is called the *covariant derivative* of V along U. Its value at $x \in \mathcal{M}$, that we denote $(\nabla_U V)(x)$, depends only on $U(x) = u$ [Lee18, Proposition 4.5]. Hence, the notation $\nabla_u V$ is legitimate and to be understood as $(\nabla_U V)(x)$ for any smooth vector field such that $U(x) = u$. We allow ourselves to use the notation $\nabla_u V$ also when $u \in T_x\mathcal{M}$ to actually refer to the derivative along $(x, u) \in T\mathcal{M}$.

For a smooth manifold $\mathcal{M}$, there exist infinitely many connections satisfying the properties of Definition 1.22 [Lee18, Lemma 4.10]. If $\mathcal{M}$ is a Riemannian manifold, the following theorem shows how to filter out a unique connection.

Theorem 1.23 ([AMS08, Theorem 5.3.1])**.** *On a Riemannian manifold $\mathcal{M}$ there exists a unique connection ∇ known as the* Riemannian connection *which verifies, for any $U, V, W \in \mathfrak{X}(\mathcal{M})$:*
1. *Symmetry:* $\nabla_U V - \nabla_V U = [U, V]$, *the Lie-Bracket[1] of the vector fields,*
2. *Compatibility with the metric:* $U\langle V, W \rangle = \langle \nabla_U V, W \rangle + \langle V, \nabla_U W \rangle$.

Example 1.24 (the Euclidean connection)**.** In an Euclidean space $\mathcal{E}$, the Riemannian connection is called the Euclidean connection and coincides with the usual directional derivative of vector fields [Lee18, Proposition 5.12]. It is denoted $\mathrm{D}_U V$ and it is given by

$$(\mathrm{D}_U V)(x) := \mathrm{D}V(x)\,[U(x)] = \left.\frac{\mathrm{d}}{\mathrm{d}t}V(x + tU(x))\right|_{t=0}, \quad \forall U, V \in \mathfrak{X}(\mathcal{E}).$$

For the derivative of smooth vector fields along a manifold curve, a separate construction has to be given. In fact, for example, curves that self-intersect admit smooth vector fields which are not the restriction of a manifold vector field to the curve. Nevertheless, every connection induces a natural way to define derivative of vector fields along curves. For conciseness the following definition, is restricted to the case of the Riemannian connection.

Theorem 1.25 ([Bou23, Theorem 8.67])**.** *Let $\gamma : J \subset \mathbb{R} \to \mathcal{M}$ be smooth curve on a manifold $\mathcal{M}$ equipped with the Riemannian connection ∇. There exists a unique mapping $\frac{\mathrm{D}}{\mathrm{d}t} : \mathfrak{X}(\gamma) \to \mathfrak{X}(\gamma)$ called the* induced covariant derivative *which satisfies the following properties for all $Y, Z \in \mathfrak{X}(\gamma)$, $U \in \mathfrak{X}(\mathcal{M})$, $g \in \mathfrak{F}(J)$ and $a, b \in \mathbb{R}$:*
1. $\mathbb{R}$-*linearity:* $\frac{\mathrm{D}}{\mathrm{d}t}(aY + bZ) = a\frac{\mathrm{D}}{\mathrm{d}t}Y + b\frac{\mathrm{D}}{\mathrm{d}t}Z$;
2. *Leibniz product rule:* $\frac{\mathrm{D}}{\mathrm{d}t}(gZ) = g'Z + g\frac{\mathrm{D}}{\mathrm{d}t}Z$;
3. *Chain rule:* $\frac{\mathrm{D}}{\mathrm{d}t}(U \circ \gamma)(t) = (\nabla_{\dot{\gamma}(t)}U)(\gamma(t))$, *for all $t \in J$;*
4. *Compatibility with the metric:* $\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z \rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t}Y, Z \right\rangle + \left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z \right\rangle$.

For a sufficiently smooth vector field, induced covariant differentiation may be repeated several times. We use $\frac{\mathrm{D}^k}{\mathrm{d}t^k}$ to indicate the covariant differentiation of order k.

---

[1] The Lie-Bracket of $U, V \in \mathfrak{X}(\mathcal{M})$ is a vector field defined as $[U, V]f = U(Vf) - V(Uf)$, $\forall f \in \mathfrak{F}(\mathcal{M})$

For any smooth manifold curve γ, the map $t \to \dot{\gamma}(t)$ defines a smooth vector field along the curve called the velocity field. Hence, the *acceleration* of γ is defined as $\ddot{\gamma} := \frac{D}{dt}\dot{\gamma} \in \mathfrak{X}(\gamma)$. A vector field along a curve is said to be *parallel* if the induced covariant derivative vanishes identically. Those curves whose velocity field is parallel are of particular interest and they are the object of the next section.

A connection on a manifold $\mathcal{M}$ not only determines the covariant differentiation of smooth vector fields but also of general smooth tensor fields as introduced with Definition 1.10.

Definition 1.26 ([Bou23, Definition 10.77]). *On a manifold $\mathcal{M}$ equipped with a connection ∇, the covariant derivative of a smooth tensor field T of order k is a tensor field of order $k+1$ defined by*

$$\nabla T(U_1, \ldots, U_k, W) = WT(U_1, \ldots U_k) - \sum_{j=1}^{k} T(U_1, \ldots, \nabla_W U_j, \ldots, U_k).$$

for any $U_1, \ldots, U_k, W \in \mathfrak{X}(\mathcal{M})$.

Just as covariant derivative of manifold vector fields induces a unique induced covariant derivative of vector fields along curves, the so-defined covariant derivative of tensor fields determines a unique induced covariant derivative for tensor fields along curves [Bou23, Theorem 10.80]. Induced covariant differentiation of tensor fields is also denoted $\frac{D}{dt}$ as it admits properties that are analogous to those stated for the vector fields counterpart in Theorem 1.25, except for the compatibility with the metric which does not apply. Concerning the chain rule, the notation $\nabla_{\dot{\gamma}(t)}T$ for a smooth order-k tensor field T restricted on a smooth curve γ is to be intended as follows. For any smooth vector fields $U_1, \ldots, U_k \in \mathfrak{X}(\gamma)$, we denote

$$\nabla_{\dot{\gamma}(t)}T\big(U_1(\gamma(t)), \ldots, U_k(\gamma(t))\big) := \nabla T\big(U_1(\gamma(t)), \ldots, U_k(\gamma(t)), \dot{\gamma}(t)\big), \quad \forall\, t \in J.$$

As a consequence of Definition 1.26, we can compute the derivative along the curve γ of the tensor field with input vector fields $U_1, \ldots, U_k$, for every $t \in J$, as

$$\begin{aligned}
\frac{d}{dt}T\big(U_1(t), \ldots, U_k(t)\big) &= \dot{\gamma}(t)T\big(U_1(t), \ldots, U_k(t)\big) \\
&= \nabla T\big(U_1(\gamma(t)), \ldots, U_k(\gamma(t)), \dot{\gamma}(t)\big) \\
&\quad + \sum_{j=1}^{k} T\Big(U_1\big(\gamma(t)\big), \ldots, \frac{D}{dt}U_j\big(\gamma(t)\big), \ldots, U_k\big(\gamma(t)\big)\Big).
\end{aligned} \tag{1.3}$$

1.3.2 Geodesics and the exponential map

In an Euclidean space, linear segments are the unique curves with zero acceleration. This observation motivates generalizing the concept of straight lines on a manifold using the definition of acceleration for a manifold curve defined in the previous section.

Definition 1.27. *A geodesic on $\mathcal{M}$ is a smooth manifold curve $\gamma : J \subset \mathbb{R} \to \mathcal{M}$ with identically zero covariant acceleration, i.e. $\ddot{\gamma}(t) = 0, \quad \forall\, t \in J.$*

Via the solution of an initial value problem associated to the zero acceleration condition, it is always possible to obtain a geodesic $\gamma_{x,v}$ passing at time $t = 0$ through a given $x \in \mathcal{M}$ with prescribed velocity $v \in T_x\mathcal{M}$ on an interval J containing zero [Lee13, Corollary 4.28]. It is unique by Picard-Lindelöf theorem, up to the extent of the parametrization interval J. We let $\gamma_{x,v}$ denote the *maximal* geodesic, i.e. when J is taken as large as possible. This means that there does not exist a geodesic $\tilde{\gamma}_{x,v}$ defined on $\tilde{J}$ such that $J \subsetneq \tilde{J}$ and $\tilde{\gamma}_{x,v}\big|_J = \gamma_{x,v}$. We denote by

$$\mathcal{O} := \{(x, v) \in T\mathcal{M} : \gamma_{x,v} \text{ is defined on an interval containing } [0, 1]\}$$

the subset of the tangent bundle for which the associated maximal geodesic is defined at least up to $t = 1$.

Definition 1.28 ([Bou23, Definition 10.16]). *The exponential map of Riemannian manifold $\mathcal{M}$ equipped with the connection ∇ is the map $\mathrm{Exp} : \mathcal{O} \to \mathcal{M}$ defined*

$$\mathrm{Exp}(x, v) = \mathrm{Exp}_x(v) = \gamma_{x,v}(1),$$

where Exp_x denotes the restriction on $T_x\mathcal{M}$, also called the exponential map at x.
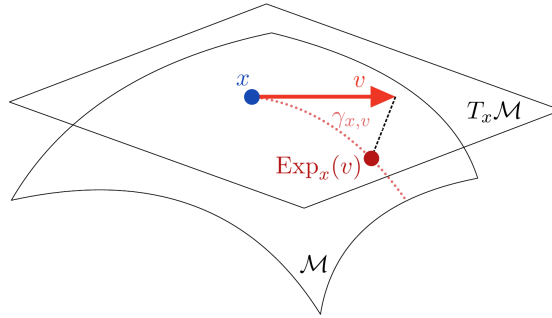
Figure 1.5: The Riemannian exponential map.

The exponential map can be used to parametrize the geodesic $\gamma_{x,v}$ as it holds that $\gamma_{x,v}(t) = \mathrm{Exp}(tv)$, for all $t \in [0, 1]$. For later reference, we collect properties of the exponential map in the following proposition.

Proposition 1.29 ([Lee18, Proposition 5.19]). *The exponential map of a Riemannian manifold $\mathcal{M}$ satisfies the following properties.*

1. *$\mathcal{O}$ is an open subset of $T\mathcal{M}$ which contains $(x, 0)$ for all $x \in \mathcal{M}$.*
2. *The exponential map is a smooth map from $T\mathcal{M}$ to $\mathcal{M}$.*
3. *For a geodesic $\gamma_{x,v}$ such that $(x, v) \in \mathcal{O}$, it holds that $\gamma_{x,v}(t) = \mathrm{Exp}(tv), \forall\, t \in [0, 1]$.*
4. *The differential of Exp_x in $0 \in T_x\mathcal{M}$ denoted $\mathrm{DExp}_x(0)$ is the identity map on $T_x\mathcal{M}$ under the identification $T_0 T_x\mathcal{M} \simeq T_x\mathcal{M}$.*

Example 1.30 (Stiefel manifold exponential map). The exponential map for the Stiefel manifold with the embedded metric (1.5) defined in Section 1.4 and the associated Riemannian connection admits an expression involving the matrix exponential[2]. For $X \in \mathrm{St}(n, k)$ and $V = X\Omega + X_\perp K \in T_X \mathrm{St}(n, k)$ as in (1.2), denote $S = V^\top V$. Then the exponential map is given by [EAS99, §2.2.2]

$$\mathrm{Exp}_X(V) = \begin{pmatrix} X & V \end{pmatrix} \exp\left(t \begin{pmatrix} \Omega & -S \\ I_{k \times k} & \Omega \end{pmatrix}\right) \begin{pmatrix} I_{k \times k} \\ 0_{k \times k} \end{pmatrix} e^{-t\Omega}, \tag{1.4}$$

By the properties of the exponential map stated in Proposition 1.29, the inverse function theorem guarantees the exponential map at any x is locally invertible. A quantitative indication on the size of the domain for which Exp_x always admits an inverse is given by the notion of injectivity radius at x.

Definition 1.31 ([Bou23, Definition 10.19]). *For a Riemannian manifold $\mathcal{M}$, the injectivity radius at $x \in \mathcal{M}$ is*

$$\mathrm{inj}(x) := \sup\left(r > 0: \ \mathrm{Exp}_x \ is \ a \ diffeomorphism \ on \ B_T(0, r)\right),$$

where $B_T(0, r) \subset T_x \mathcal{M}$ denote that open ball of radius r centered at the origin of the tangent space at x.

By the inverse function theorem, $\mathrm{inj}(x) > 0$ for every $x \in \mathcal{M}$. Restricting the domain of exponential at x to the ball $B_T(0, \mathrm{inj}(x))$ guarantees by construction that Exp_x is a bijection onto $\mathrm{Exp}_x(B_T(0, \mathrm{inj}(x)))$. With this restriction of the domain, the inverse of the exponential map is known as the logarithmic map. The following definition combines [Bou23, Definition 10.20] with [Bou23, Corollary 10.25].

Definition 1.32. *On a Riemannian manifold $\mathcal{M}$, the Riemannian logarithm or logarithmic map at $x \in \mathcal{M}$ is defined for every $y \in \mathrm{Exp}_x(B_T(0, \mathrm{inj}(x)))$ as $\mathrm{Log}_x(y) = v$, where v is the unique tangent vector in $B_T(0, \mathrm{inj}(x)) \subset T_x \mathcal{M}$ such that $\mathrm{Exp}_x(v) = y$.*

The so-defined logarithm is a map $\mathrm{Log} : \mathcal{M} \times \mathcal{M} \to T\mathcal{M}$ that is smooth on the following set [Bou23, Corollary 10.25]

$$\{(x, y) \in \mathcal{M} \times \mathcal{M} : d(x, y) < \mathrm{inj}(x)\}.$$

A very useful property of the logarithmic map is its connections to the Riemannian distance function.

Proposition 1.33 ([Bou23, Proposition 10.22]). *For any $x, y \in \mathcal{M}$ such that $d(x, y) < \mathrm{inj}(x)$, a geodesic joining x and y is parametrized as $\gamma(t) = \mathrm{Exp}_x(t\mathrm{Log}_x(y))$ for $t \in [0, 1]$ and it is the unique curve such that $d(x, y) = L(\gamma)$. In particular, $d(x, y) = \|\mathrm{Log}_x(y)\|_x$.*

---

[2]The matrix exponential is defined for any $A \in \mathbb{R}^{n \times n}$ by $\exp(A) = e^A := \sum_{k=0}^{+\infty} A^k / k!$.

The geodesic γ in Proposition 1.33 is unique up to reparametrization and is called the *length-minimizing geodesic*. These length-minimizing geodesics generalize to Riemannian manifolds the Euclidean notion of linear segments from the perspective of length-minimizing paths between two points.

1.3.3 Parallel transport

In the same way as the zero-acceleration condition of geodesics coupled with an initial condition leads to the definition of the exponential map, parallel vector fields Z along any smooth curve γ can be constructed by integrating the vanishing induced covariant derivative condition $\frac{D}{dt}Z(t) = 0$, $\forall t \in J$. This construction is always possible. In fact, for any $t_0 \in J$ and $Z_0 \in T_{\gamma(t_0)}\mathcal{M}$ there exists a unique parallel vector field $Z \in \mathfrak{X}(\gamma)$ such that $Z(t_0) = Z_0$ [Lee18, Theorem 4.32]. In turn, this leads to the definition of a map to transport tangent vectors along the curve.

Definition 1.34 ([Bou23, Definition 10.35]). *Given a smooth curve γ on $\mathcal{M}$, the parallel transport along γ from t_0 to t_1 is the linear map*

$$\mathrm{P}^{\gamma}_{t_1 \leftarrow t_0} : T_{\gamma(t_0)}\mathcal{M} \to T_{\gamma(t_1)}\mathcal{M}$$

defined for any $v \in T_{\gamma(t_0)}\mathcal{M}$ by $\mathrm{P}^{\gamma}_{t_1 \leftarrow t_0}v = Z(t_1)$, where $Z \in \mathfrak{X}(\mathcal{M})$ is the only parallel vector field along γ such that $Z(t_0) = v$.
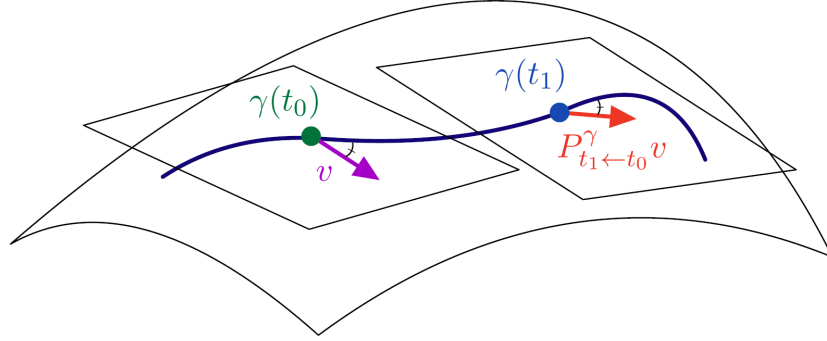
Figure 1.6: Parallel transport along a smooth manifold curve γ.

With the Riemannian connection, parallel transport is an isometry between tangent spaces [Bou23, Proposition 10.36].

1.4 Riemannian submanifolds

Let us now specialize to the case where $\mathcal{M}$ is an embedded submanifold of an Euclidean space $\mathcal{E}$, a vector space whose inner-product $\langle \cdot, \cdot \rangle$ is used as Riemannian metric, see Example 1.19. In light of Theorem 1.15, we can interpret tangent vectors of $T_x\mathcal{M}$ as elements of $T_x\mathcal{E} \simeq \mathcal{E}$. This allows the definition of a natural Riemannian metric for $\mathcal{M}$

by restricting the inner-product of $\mathcal{E}$ to each $T_x\mathcal{M}$, i.e. $\langle u, v \rangle_x = \langle u, v \rangle$, $\forall u, v \in T_x\mathcal{M}$, $x \in \mathcal{M}$. This is the so-called *induced metric* on $\mathcal{M}$.

Definition 1.35. *A* Riemannian submanifold *is an embedded manifold $\mathcal{M}$ of an Euclidean space $\mathcal{E}$ endowed with the induced metric.*

Example 1.36 (Stiefel manifold embedded metric). The embedded metric on the Stiefel manifold is the restriction of the Euclidean metric of $\mathbb{R}^{n \times k}$ introduced in Example 1.19 on each tangent space of $\mathrm{St}(n, k)$. For every $X \in \mathrm{St}(n, k)$ and $V_1, V_2 \in T_X\mathrm{St}(n, k)$ written as in (1.2), i.e. $V_i = X\Omega_i + X_\perp K_i$, for $i = 1, 2$, we have

$$\langle V_1, V_2 \rangle = \mathrm{Tr}(V_1^\top V_2) = \mathrm{Tr}(\Omega_1^\top \Omega_2) + \mathrm{Tr}(K_1^\top K_2). \tag{1.5}$$

Note that this metric weighs the subspace associated to the anti-symmetric part Ω_i twice as much as the other directions. For this reason, the embedded metric is often replaced by the so-called *canonical metric* [EAS99, §2.4.1] which is a non-induced metric defined by $\langle V_1, V_2 \rangle_X = \frac{1}{2}\mathrm{Tr}(\Omega_1^\top \Omega_2) + \mathrm{Tr}(K_1^\top K_2)$.

1.4.1 Projections on Riemannian submanifolds

For each $x \in \mathcal{M}$, the identification of $T_x\mathcal{M}$ with a vector subspace of $\mathcal{E}$ allows the definition of the orthogonal projection from $\mathcal{E}$ to $T_x\mathcal{M}$ that we denote $\Pi(x)$. Then, at every x, we can write any $v \in \mathcal{E}$ as the sum of a tangential component $\Pi(x)v \in T_x\mathcal{M}$ and normal component $(I - \Pi(x))v =: \Pi(x)^\perp v \in T_x\mathcal{M}^\perp$. The orthogonal complement of $T_x\mathcal{M}$ as a subspace in $\mathcal{E}$ is called the *normal space* at x denoted $N_x\mathcal{M}$.

Example 1.37 (Stiefel manifold tangent space projection). For any $X \in \mathrm{St}(n, k)$, the orthogonal projection onto $T_X\mathrm{St}(n, k)$ is given by

$$\Pi(X)V = X\,\mathrm{Skew}(X^\top V) + (I - XX^\top)V$$

where $\mathrm{Skew}(A) := \frac{1}{2}\left(A - A^\top\right)$ for any square matrix A. In fact, it is a linear idempotent and self-adjoint map whose image spans $T_X\mathrm{St}(n, k)$.

The tangent space projection at $\Pi(x)$ returns the closest tangent vector in $T_x\mathcal{M}$ to a given ambient vector. The closest manifold point to an ambient space point is provided by the metric projection from $\mathcal{E}$ to $\mathcal{M}$, which is defined as

$$\Pi_\mathcal{M}(y) := \operatorname*{arg\,min}_{x \in \mathcal{M}} \|y - x\|.$$

When $\mathcal{M}$ is closed in $\mathcal{E}$, this mapping is single-valued on a dense subset of $\mathcal{E}$ [Bou23, Theorem 5.53] or if $\mathcal{M}$ is the boundary of a convex set. For embedded manifolds which are not closed subsets of the ambient space, the metric projection can still be defined, but it is single-valued only in a neighborhood of $\mathcal{M}$ [AM12, Lemma 3.1].

1.4.2 The tangential connection

For a Riemannian submanifold $\mathcal{M}$ of an Euclidean space $\mathcal{E}$, it is possible to define a connection by projecting onto the tangent space the ambient space Euclidean connection defined in Example 1.24. This is the so-called *tangential connection* that is denoted $\nabla^\top$ and defined by

$$(\nabla_U^\top V)(x) = \Pi(x)(\mathrm{D}_U \overline{V})(x), \quad \forall\, x \in \mathcal{M},\, U, V \in \mathfrak{X}(\mathcal{M}),$$

where $\overline{V}$ is any smooth extensions of V. The tangential connection is the Riemannian connection of a Riemannian submanifold [Lee13, Proposition 5.12]. Since this is the only connection we consider for Riemannian submanifold, from now on we simply denote it ∇. For any smooth curve $\gamma : J \subset \mathbb{R} \to \mathcal{M}$, the induced covariant derivative uniquely determined by the tangential connection is given for any $Z \in \mathfrak{X}(\gamma)$ by [Bou23, Proposition 5.31]

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \Pi(\gamma(t)) Z'(t), \quad \forall\, t \in J,$$

where $Z'(t)$ is obtained from interpreting Z as a curve in $\mathcal{E}$. Hence, we may write the acceleration of curve on a Riemannian submanifold as

$$\ddot{\gamma}(t) = \Pi(\gamma(t))\gamma''(t).$$

It is customary to the refer to $\ddot{\gamma}$ as the *intrinsic* acceleration, obtained as the tangential part of the *extrinsic* or Euclidean acceleration $\gamma''(t)$. Therefore, on a Riemannian submanifolds equipped with the Riemannian connection, geodesics are the manifold curve whose extrinsic acceleration is orthogonal to the manifold.

1.5 Manifolds of interest

In the following sections we gather definitions and well-known results about some Riemannian manifolds considered in the applications sections of the thesis.

1.5.1 The sphere

In order to have the possibility to visually appreciate the output of some algorithm presented in the thesis, we apply them to a very simple manifold: the sphere. While illustrative example restrict to the case of the sphere in $\mathbb{R}^3$, let us introduce the manifolds structure of the sphere in $\mathbb{R}^n$. For a complete presentation, the reader is referred to [Bou23, §7.2].
Let us consider $\mathbb{R}^n$ with it Euclidean structure given by the inner-product $\langle u, v \rangle = u^\top v$, $\forall\, u, v \in \mathbb{R}^n$. The sphere in $\mathbb{R}^n$ is defined as

$$S^{n-1} := \{ x \in \mathbb{R}^n : \langle x, x \rangle = 1 \}.$$
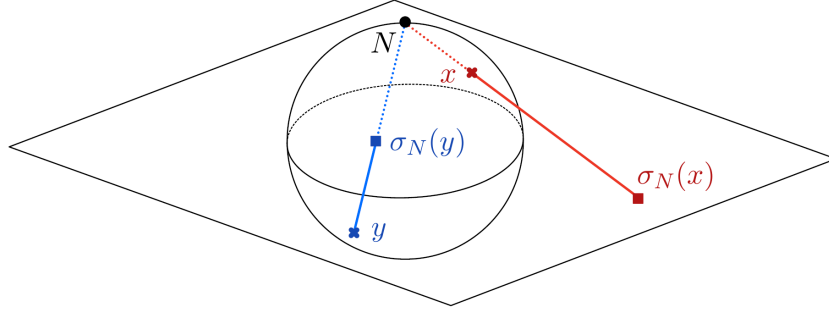
Figure 1.7: Stereographic projection through the north pole N of a northern hemisphere point $x \in S^2$ and of a southern hemisphere point $y \in S^2$.

The function $h(x) = \langle x, x \rangle - 1$ is smooth and the sphere coincides with the zero level set. The differential of the defining function is $\mathrm{D}h(x)[v] = 2\langle x, v \rangle$ and is surjective on $\mathbb{R}$ for any $x \in S^{n-1}$. This shows S^{n-1} is an embedded submanifold of $\mathbb{R}^n$ of dimension $n-1$ in the sense of Definition 1.11.

Alternatively, S^{n-1} can be argued to be a smooth manifold in the sense of Definition 1.2 by constructing explicitly a smooth atlas [Lee13, Exercise 1-7]. Let us indicate by $N = (0, \ldots, 0, 1) \in S^{n-1}$ the *north pole* of the sphere. The *stereographic projection* through the north pole of S^{n-1} is $\sigma_N : S^{n-1} \setminus \{N\} \to \mathbb{R}^{n-1}$ defined by

$$\sigma_N\big((x_1, \ldots, x_n)\big) = \frac{(x_1, \ldots, x_{n-1})}{1 - x_n}.$$

This mapping associates to any point of $x \in S^{n-1} \setminus \{N\}$ the intersection between the plane $x_n = 0$ and the line joining N and x, see Figure 1.7. Analogously, the south pole of the sphere is $S := -N$, and the associated stereographic projection is $\sigma_S(x) = -\sigma_N(-x)$, defined for any $x \in S^{n-1} \setminus \{S\}$. Both σ_N and σ_S are bijective and the transition map $\sigma_S \circ \sigma_N^{-1}$ is smooth. Since $S^{n-1} \setminus \{N\}$ and $S^{n-1} \setminus \{S\}$ form an open cover for S^{n-1}, then $\{\sigma_N, \sigma_S\}$ is a smooth atlas. The manifold structure which properly contains this atlas is the same manifold structure implied by Definition 1.11 and the above defining function h.

Following Proposition 1.15, the tangent space at $x \in S^{n-1}$ can be expressed as

$$T_x S^{n-1} = \ker \mathrm{D}h(x) = \{v \in \mathbb{R}^n : \langle x, v \rangle = 0\}.$$

Hence, the projection onto the tangent space is $\Pi(x)w = w - \langle x, w \rangle x$, for every $w \in \mathbb{R}^n$. The sphere is the boundary of a closed convex subset of $\mathbb{R}^n$, so the closest point projection is always uniquely defined. It coincides with the normalization

$$\Pi_{S^{n-1}}(y) := \frac{y}{\|y\|}. \tag{1.6}$$

The sphere is endowed with the embedded metric of $\mathbb{R}^n$ and the associated Riemannian metric. This yields a Riemannian submanifold structure for S^{n-1} which admits an

expression for the exponential map involving trigonometric functions [Bou23, Example 5.37] which is a constant-speed parametrization of a great circle of the sphere

$$\mathrm{Exp}_x(v) = \cos(t\,\|v\|)x + \frac{\sin(t\,\|v\|)}{\|v\|}v. \tag{1.7}$$

For any $x \in S^{n-1}$, by appropriately restricting the domain of the exponential map on the open disk of radius π around the origin of the tangent space, the exponential map is bijective and its inverse coincides with the logarithmic map. It is defined for any every point on the sphere except the antipodal point and admits the following expression [Bou23, Example 10.21]

$$\mathrm{Log}_x(y) = \arccos(y^\top x)\frac{\Pi(x)y}{\|\Pi(x)y\|}, \quad \forall\, y \in S^{n-1} \setminus \{-x\}. \tag{1.8}$$

1.5.2 Symmetric positive definite matrices

Symmetric positive definite matrices often arise in applications, notably as covariance matrices. Here, we present a commonly encountered manifold structure for this set of matrices which yields favorable properties, notably completeness of the manifold. The material in this section can be found in the reference textbook on positive definite matrices [Bha07, §6] and in [Moa05, §2.3].

The set of symmetric positive definite matrices are a subset of the vector space of symmetric matrices $\mathbb{S}_n$ defined by

$$\mathbb{S}_n^+ := \left\{ X \in \mathbb{S}_n \;:\; v^\top X v > 0, \quad \forall\, v \in \mathbb{R}^n \text{ if } v \neq 0 \right\}.$$

In particular, it is a convex cone in the sense $\alpha X + \beta Y \in \mathbb{S}_n^+$ for any $X, Y \in \mathbb{S}_n^+$ provided $\alpha, \beta > 0$. Furthermore, it can be characterized by the inequalities $\det(X_{1:k,1:k}) > 0$, for all $k = 1, \ldots, n$. By continuity of the functions $X \to \det(X_{1:k,1:k})$, we obtain $\mathbb{S}_n^+$ as the finite intersection of such open subsets. Hence, $\mathbb{S}_n^+$ is open in $\mathbb{S}_n$ and therefore it is an embedded submanifold of $\mathbb{S}_n$ of dimension $n(n+1)/2 = \dim(\mathbb{S}_n)$. The tangent space at $X \in \mathbb{S}_n^+$ can be identified with the ambient vector space:

$$T_X \mathbb{S}_n^+ \simeq \mathbb{S}_n.$$

Under the embedded metric, $\mathbb{S}_n^+$ would not be a complete metric space since it is open in $\mathbb{S}_n$. An alternative metric can be defined with the intent to preserve lengths under congruence transformations. In fact, consider any $n \times n$ invertible matrix X, the congruent matrix $X^\top Y X \in \mathbb{S}_n^+$ if $Y \in \mathbb{S}_n^+$. Therefore, it is required that for any smooth curve γ on $\mathbb{S}_n^+$ the metric implies length-invariance under congruence transformation, i.e. $L(t \to \gamma(t)) = L(t \to X^\top \gamma(t)X)$. This is the case for any metric such that

$$\left\langle X^\top V X, X^\top W X \right\rangle_{X^\top Y X} = \langle V, W \rangle_Y, \quad \forall\, V, W \in \mathbb{S}_n. \tag{1.9}$$

23

A frequently used Riemannian metric verifying this property has the following expression

$$\langle V, W \rangle_X = \text{Tr}\left(X^{-1}VX^{-1}W\right). \tag{1.10}$$

Due to property (1.9), this metric is often called the *bi-invariant metric* of $\mathbb{S}_n^+$. The induced norm $\|V\|_X$, by cyclic permutation of the trace operator, can be written as

$$\|V\|_X = \left\|X^{-1/2}VX^{-1/2}\right\|_F.$$

With the bi-invariant metric (1.10) and the associated Riemannian connection, any $X, Y \in \mathbb{S}_n^+$ can be joined by a unique geodesic which can be parametrized by [Bha07, Theorem 6.1.6]

$$\gamma_{XY}(t) = X^{1/2}\exp(t\log(X^{-1/2}YX^{-1/2}))X^{1/2}, \tag{1.11}$$

where exp and log are the matrix exponential and logarithm[3]. Differentiating (1.11) in $t = 0$ tell us the Riemannian logarithmic map and exponential map on $\mathbb{S}_n^+$ with the bi-invariant metric are, $\forall\, X, Y \in \mathbb{S}_n^+$ and $V \in T_X\mathbb{S}_n^+$.

$$\begin{aligned}
\text{Log}_X(Y) &:= X^{1/2}\log(X^{-1/2}YX^{-1/2})X^{1/2}, \\
\text{Exp}_X(V) &:= X^{1/2}\exp(X^{-1/2}VX^{-1/2}))X^{1/2}.
\end{aligned} \tag{1.12}$$

In turn, the Riemannian distance function reads as

$$d(X, Y) = \|\log(X^{-\frac{1}{2}}YX^{-\frac{1}{2}})\|_F. \tag{1.13}$$

The completeness of $\mathbb{S}_n^+$ as metric space with this distance follows from the Hopf-Rinow Theorem [Lee13, Theorem 6.19] by observing the exponential map is well-defined on the whole tangent bundle.

With this geometry, the parallel transport along a geodesic γ from $\gamma(0) = X$ to $\gamma(1) = Y$ has the following expression [SH15, Equation 3.4] for every $V \in \mathbb{S}_n$

$$\text{P}_{1\leftarrow0}^{\gamma}V = EVE^\top, \quad \text{with } E = (YX^{-1})^{1/2}. \tag{1.14}$$

1.5.3 Fixed-rank matrices

The set of $m \times n$ matrices of rank $k \leq \min\{m, n\}$ is defined as

$$\mathcal{M}_k := \left\{X \in \mathbb{R}^{m\times n} : \dim(\text{span}(X)) = k\right\}.$$

It is an embedded submanifold of $\mathbb{R}^{m\times n}$ of dimension $(m+n-k)k$ [Lee03, Example 8.14]. It is often encountered in applications to approximate numerically low-rank matrices. Among many equivalent parametrizations of this manifold, in this work we adopt the

---

[3]The (principal) matrix logarithm is uniquely defined for any non-singular $B \in \mathbb{R}^{n\times n}$ with no eigenvalues on the negative real axis. Then $\log(B) =: A$ is the unique matrix such that $e^A = B$ and whose eigenvalues lie in the strip $\{z \in \mathbb{C} : -\pi < \text{Im}(z) < \pi\}$, see [Hig08, Theorem 1.31].

conventions used in the MATLAB Riemannian optimization library Manopt [BMAS14] to represent rank-k matrices and tangent vectors. The reader is referred to [Van13, UV20] for an extensive overview. Any point $X \in \mathcal{M}_k$ is associated with its economy-sized SVD

$$X = U\Sigma V^\top \tag{1.15}$$

where $U \in \mathrm{St}(m,k)$, $V \in \mathrm{St}(n,k)$ and $\Sigma = \mathrm{diag}(\sigma_i)$ is the diagonal matrix with the non-zero singular values in descending order $\sigma_1 \geq \cdots \geq \sigma_k > 0$. The tangent space at X is parametrized as

$$T_X\mathcal{M}_k = \left\{ UMV^\top + U_pV^\top + UV_p^\top \in \mathbb{R}^{m\times n} : \begin{array}{ll} M \in \mathbb{R}^{k\times k}, & \\ U_p \in \mathbb{R}^{m\times k}, & U_p^\top U = 0, \\ V_p \in \mathbb{R}^{n\times k}, & V_p^\top V = 0 \end{array} \right\}. \tag{1.16}$$

The manifold of rank-k matrices is endowed with the Riemannian submanifold structure of $\mathbb{R}^{m\times n}$, i.e. with the metric $\langle Z, W \rangle = \mathrm{Tr}(Z^\top W)$, $\forall W, Z \in T_X\mathcal{M}_k$. Then, the orthogonal projection onto the tangent space at X has the following expression

$$\Pi(X)W = P_U W P_V + P_U W P_V^\perp + P_U^\perp W P_V. \tag{1.17}$$

where $P_U = UU^\top$, $P_V = VV^\top$, $P_U^\perp = I - P_U$, $P_V^\perp = I - P_V$ are orthogonal projections onto the spaces spanned by the columns of U and V, and their orthogonal complements, respectively.

A notable property of $\mathcal{M}_k$ under this Riemannian submanifold structure is that it is not complete. This is because it is not closed in $\mathbb{R}^{m\times n}$. For instance, consider the sequence of matrices of rank-k defined by

$$A_i = U\mathrm{diag}(\sigma_1, \ldots, \sigma_{k-1}, \sigma_k/i)V^\top, \quad \forall i \in \mathbb{N}$$

for any column orthogonal U and V and $\sigma_1, \ldots, \sigma_k > 0$. This defines a Cauchy sequence contained $\mathcal{M}_k$ but whose limit lies in $\mathcal{M}_{k-1}$. Nevertheless, under the Euclidean structure of $\mathbb{R}^{m\times n}$, the closest point projection onto $\mathcal{M}_k$ of a given $A \in \mathbb{R}^{m\times n}$ can be uniquely defined provided $\sigma_k(A) > \sigma_{k+1}(A)$. In fact, it is the solution to the problem

$$\arg\min_{Y \in \mathcal{M}_k} \|A - Y\|_F$$

which notoriously can be computed with the rank-k truncated SVD of the matrix A [HJ13, §7.4.2]

$$\Pi_{\mathcal{M}_r}(A) := \sum_{i=1}^{k} \sigma_i u_i v_i^\top, \tag{1.18}$$

where σ_i are the k largest singular values and u_i, v_i the associated singular vectors.

To mitigate storage requirements when m and n are large and $k \ll \min\{m,n\}$, it is convenient to store elements of $\mathcal{M}_k$ with the triplet (U, Σ, V). Likewise, tangent

vectors can be identified with the triplet (M, U_p, V_p) and the triplet of the anchor point. Metric projection, tangent space projection and Riemannian metric can be efficiently implemented to operate with these formats.

A part from numerically integrating the geodesic equation in Definition 1.27 on the fixed-rank manifold by some standard procedure, to the best of our knowledge, there is no specifically tailored strategy to compute the exponential map of $\mathcal{M}_k$ under this embedded geometry. However, a different Riemannian structure on $\mathcal{M}_k$ enables to leverage the exponential map (1.4) obtained for the Stiefel manifold to construct the exponential map for $\mathcal{M}_k$ under this alternative geometry [AAM14, §6].

2 | Retractions for Riemannian optimization

In this chapter, we introduce the basics of Riemannian optimization with particular emphasis on retractions. As an illustrative example and for later reference, we describe some standard Riemannian optimization algorithms as well as some retractions for the manifolds introduced in the previous chapter that are relevant for the thesis. The reader is referred to [AMS08, Bou23] for a complete overview.

2.1 Riemannian optimization fundamentals

Given a smooth manifold $\mathcal{M}$ and a smooth scalar field $f \in \mathfrak{F}(\mathcal{M})$, consider the following constrained optimization problem

$$\min_{x \in \mathcal{M}} f(x). \tag{2.1}$$

When $\mathcal{M}$ is an embedded submanifold of an Euclidean space $\mathcal{E}$ that can be conveniently described using defining functions, the problem could be addressed with methods from general constrained optimization, e.g., the method of Lagrange multipliers. For certain $\mathcal{M}$ however, this approach cannot be pursued due to the complexity of the defining functions. Additionally, the smoothness of the constraint set offers some simplifications which can be leveraged to develop more specific methods. For instance, a necessary optimality condition for (2.1) can be formulated as follows.

Proposition 2.1 ([Bou23, Proposition 4.5])**.** *If $x \in \mathcal{M}$ is a local minimizer of a scalar field $f \in \mathcal{M}$, then x is a* critical point *of f, i.e. a point for which*

$$vf = \mathrm{D}f(x)[v] = 0, \quad \forall\, v \in T_x\mathcal{M}. \tag{2.2}$$

2.1.1 The Riemannian gradient

When the manifold is endowed with a Riemannian metric, the criticality condition (2.2) can be made even more explicit. By Riesz representation theorem, the inner-product on each tangent space grants the existence of the Riemannian gradient.

Definition 2.2 ([Bou23, Definition 3.58])**.** *Let $\mathcal{M}$ denote a Riemannian manifold with a metric $\langle \cdot, \cdot \rangle_x$ for every $x \in \mathcal{M}$. The* Riemannian gradient *of a scalar field $f \in \mathfrak{F}(\mathcal{M})$*

at x is the unique vector $\mathrm{grad}f(x) \in T_x\mathcal{M}$ which satisfies

$$\mathrm{D}f(x)\,[v] = \langle \mathrm{grad}f(x), v \rangle_x\,, \quad \forall\, v \in T_x\mathcal{M}.$$

The Riemannian gradient of a smooth function is a smooth vector field on $\mathcal{M}$ [Bou23, Proposition 3.70]. This same definition holds when $\mathcal{M} = \mathcal{E}$ an Euclidean space endowed the standard inner-product. The so-defined Riemannian gradient of $f \in \mathfrak{F}(\mathcal{E})$ coincides with the usual gradient of f defined as the vector of partial derivatives. For this particular case, we reserve the term Euclidean gradient and use the symbol ∇f.

The necessary optimality condition stated in Proposition 2.1 can now be expressed in terms of the Riemannian gradient, recovering the usual notion of critical point for a scalar field on an Euclidean space.

Proposition 2.3 ([Bou23, Proposition 4.5]). *For a smooth scalar field $f \in \mathfrak{F}(\mathcal{M})$ on a Riemannian manifold $\mathcal{M}$, the point x is a critical point if and only if $\mathrm{grad}f(x) = 0$.*

In the case where $\mathcal{M}$ is a Riemannian submanifold of an Euclidean space $\mathcal{E}$, the possibility to smoothly extend scalar fields to the ambient space stated in Proposition 1.13 allows us to relate the Riemannian gradient with the Euclidean gradient of any smooth extension. Letting $\bar{f}$ denote any smooth extension for $f \in \mathfrak{F}(\mathcal{M})$, then for $x \in \mathcal{M}$ and any $v \in T_x\mathcal{M}$

$$\mathrm{D}f(x)\,[v] = \mathrm{D}\bar{f}(x)\,[v]\,,$$

where the v on the right-hand side is interpreted as an element of $T_x\mathcal{E} \simeq \mathcal{E}$. Hence, since the Riemannian metric is the induced metric, from the definition of Riemannian and Euclidean gradient it must hold that for all $v \in T_x\mathcal{M}$

$$\begin{aligned}
\langle \mathrm{grad}f(x), v \rangle &= \left\langle \nabla\bar{f}(x), v \right\rangle, \\
&= \left\langle \Pi(x)\nabla\bar{f}(x), v \right\rangle + \left\langle \Pi(x)^{\perp}\nabla\bar{f}(x), v \right\rangle \\
&= \left\langle \Pi(x)\nabla\bar{f}(x), v \right\rangle,
\end{aligned}$$

where $\Pi(x)$ and $\Pi(x)^{\perp}$ are respectively the orthogonal projection from $\mathcal{E}$ to $T_x\mathcal{M}$ and its complement. This shows that the Riemannian gradient is the tangential component of the Euclidean gradient of a smooth extension. This is independent of the choice of smooth extension as summarized by the following.

Proposition 2.4 ([Bou23, Proposition 3.61]). *For a Riemannian submanifold $\mathcal{M}$ to an Euclidean space $\mathcal{E}$, the Riemannian gradient of a scalar field $f \in \mathfrak{F}(\mathcal{M})$ is given by*

$$\mathrm{grad}f(x) = \Pi(x)\nabla\bar{f}(x),$$

for any smooth extension $\bar{f}$ of f in a neighborhood of $\mathcal{M}$ in $\mathcal{E}$.

In turn, Proposition 2.4 yields a practical way of computing Riemannian gradients on Riemannian submanifolds: choose any smooth extension of the scalar field, compute the Euclidean gradient and project it to the tangent space.
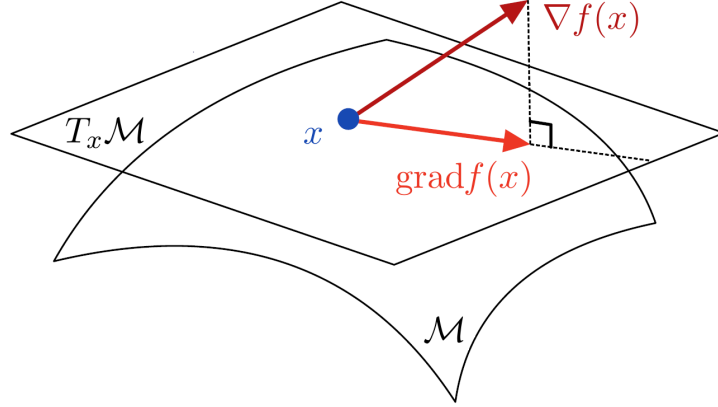
Figure 2.1: The Riemannian gradient as orthogonal projection of the Euclidean gradient.

2.1.2 The Riemannian Hessian

Having identified a critical point of a scalar field $f \in \mathfrak{F}(\mathcal{M})$ is not sufficient to claim local optimality for problem (2.1). To distinguish locally minimizing critical points from local maxima and saddle points, it can be sufficient to consider the second-order variation of the objective function. The extension of this concept for manifold scalar fields can be done endowing the manifold with a Riemannian structure.

Definition 2.5 ([Bou23, Definition 5.14]). *For a Riemannian manifold $\mathcal{M}$ endowed with its Riemannian connection, the* Riemannian Hessian *at $x \in \mathcal{M}$ of a smooth scalar field $f \in \mathfrak{F}(\mathcal{M})$ is the linear operator $\mathrm{Hess} f(x) : T_x\mathcal{M} \to T_x\mathcal{M}$ defined by*

$$\mathrm{Hess} f(x)\,[v] := \nabla_v \mathrm{grad} f(x), \quad \forall\, v \in T_x\mathcal{M}.$$

The Riemannian Hessian is smoothly varying in x in the sense that for any smooth vector field $V \in \mathfrak{X}(\mathcal{M})$, the map $x \mapsto (x, \mathrm{Hess} f(x)\,[V(x)])$ is a smooth vector field. Alternatively, the smoothness can be understood in the sense the map $U, V \mapsto \langle U, \mathrm{Hess}(f)\,[V] \rangle$ for every $U, V \in \mathfrak{X}(\mathcal{M})$ defines a smooth tensor field as per Definition 1.10. The Riemannian Hessian is defined with the Riemannian connection so that the compatibility with the Riemannian metric implies the Riemannian Hessian is a self-adjoint operator with respect to the metric [Bou23, Proposition 5.15]. This guarantees that the positive semi-definiteness of the Riemannian Hessian translates into a condition on its eigenvalues.

Proposition 2.6 ([Bou23, Proposition 6.3]). *A local minimizer $x \in \mathcal{M}$ of a scalar field $f \in \mathfrak{F}(\mathcal{M})$ satisfies $\mathrm{grad} f(x) = 0$ and $\mathrm{Hess} f(x) \succeq 0$.*

This condition is still not sufficient for local optimality: if the Hessian is not full rank, the critical point may be a local maxima or a saddle point of order greater than two. A sufficient condition reads as follows.

Proposition 2.7 ([Bou23, Proposition 6.5]). *If $x \in \mathcal{M}$ is a critical point of the scalar field $f \in \mathfrak{F}(\mathcal{M})$ and $\mathrm{Hess} f(x) \succ 0$ then x is a strict local minimizer.*

Analogously to Proposition 2.4 concerning the Riemannian gradient, let us discuss one way of computing the Riemannian Hessian of a given objective function for the case of a Riemannian submanifold $\mathcal{M}$ in an Euclidean space $\mathcal{E}$. On the one hand, as detailed in Section 1.4.2, the Riemannian connection consists of projecting onto the tangent space the Euclidean derivative of the vector field smoothly extended in the ambient space. On the other hand, according to Proposition 2.4, the Riemannian gradient will coincide with the tangential projection of the Euclidean gradient. Hence, for a Riemannian submanifold the covariant derivative of the Riemannian gradient features the differential of tangent space projection. From every $x \in \mathcal{M}$, the differential of the tangent space projection is a linear map from $T_x\mathcal{M}$ to $T_x\mathcal{E}$ and we denote it $\mathrm{D}\Pi(x)[v]$.

Proposition 2.8 ([AMT13, §3]). *The Riemannian Hessian of a scalar field f defined on a Riemannian submanifold $\mathcal{M}$ of an Euclidean space can be computed as*

$$\mathrm{Hess} f(x)[v] = \Pi(x)\nabla^2 \bar{f}(x)\,[v] + \Pi(x)\mathrm{D}\Pi(x)[v]\Pi(x)^\perp \nabla \bar{f}(x), \quad \forall\, x \in \mathcal{M},\, v \in T_x\mathcal{M},$$

where $\bar{f}$ is any smooth extension of f to $\mathcal{E}$ and $\nabla \bar{f}$ and $\nabla^2 \bar{f}$ indicate respectively its Euclidean gradient and Hessian.

2.1.3   Gradient descent along geodesics

Analogously to the Euclidean case, when $\mathrm{grad} f(x) \neq 0$ the Riemannian gradient produces the direction of steepest ascent for the function f in the sense

$$\frac{\mathrm{grad} f(x)}{\|\mathrm{grad} f(x)\|_x} = \arg\max_{\|v\|_x = 1} \mathrm{D}f(x)\,[v].$$

Therefore, following the opposite direction along a geodesic produces a greedy minimization algorithm for problem (2.1). In fact, provided the objective function is continuously differentiable, a sufficiently small step along the geodesic passing trough the current iterate in the direction of the negated Riemannian gradient is guaranteed to decrease the objective function without leaving the constraint manifold [RW12, Proposition 1]. With the tools introduced so far, the iteration can be written using the exponential map of $\mathcal{M}$ as

$$x_{k+1} = \mathrm{Exp}_{x_k}(-\alpha_k\, \mathrm{grad} f(x_k)), \quad \forall\, k \geq 0, \tag{2.3}$$

for some suitably chosen step sizes $\alpha_k > 0$ and an initial guess $x_0 \in \mathcal{M}$. This first appearance of this kind of algorithm dates back to [Lue72, Gab82].
The exponential map (see Definition 1.28) is the solution to an initial value problem which can be a demanding problem to solve in general. Although sometimes an analytical expression is available, see Example 1.30, it may still be too onerous to compute when it involves matrix functions for matrices of large dimension. The objective to alleviate the cost of computing the Riemannian exponential map in a procedure like (2.3) motivates the developments of retractions.

2.2   Retractions

The defining properties of retractions mirror essential features of the exponential map.

Definition 2.9 ([AMS08, Definition 4.4.1]). *A retraction is a smooth map*

$$(x, v) \in T\mathcal{M} \mapsto R(x, v) \in \mathcal{M}$$

defined in a neighborhood of $\{(x, 0) \in T\mathcal{M} : x \in \mathcal{M}\}$. Denoting R_x the restriction of R onto $T_x\mathcal{M}$ for some $x \in \mathcal{M}$, retractions satisfy

1. *$R_x(0_x) = x$, where 0_x is the zero vector of $T_x\mathcal{M}$,*
2. *$\mathrm{D}R_x(0_x) = I_{T_x\mathcal{M}}$, the identity map on the tangent space at x.*

Concrete examples of retractions for the manifolds considered in the thesis are provided in Section 2.4. As expected from the intent in defining the retraction, the exponential map on a Riemannian manifold is indeed a retraction, owing to Proposition 1.29. The second property of Definition 2.9 is referred to as *local rigidity* and has several implications. First and foremost, it guarantees that for any $x \in \mathcal{M}$ and $v \in T_x\mathcal{M}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} R_x(tv)\big|_{t=0} = v = \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Exp}_x(tv)\big|_{t=0}.$$

This justifies replacing the exponential map in (2.3) with any retraction while maintaining the guarantee to reduce the objective function at each step for sufficiently small steps [RW12, Proposition 1]. The algorithm so-obtained is called Riemannian gradient descent, see Section 2.3.

Note that the definition of retraction does not require $\mathcal{M}$ to be endowed with a Riemannian structure. Yet, on Riemannian manifolds some retractions may satisfy a property related to the acceleration of certain curves defined using retractions.

Definition 2.10. *A retraction on a Riemannian manifold $\mathcal{M}$ is said to be a second-order retraction if for any $x \in \mathcal{M}$ and $v \in T_x\mathcal{M}$ the manifold curve $\sigma_{x,v}(t) = R_x(tv)$ has zero covariant acceleration in $t = 0$, that is $\ddot{\sigma}_{x,v}(0) = 0$.*

Two general ways of constructing second-order retractions are discussed in Section 2.2.2 and provided explicitly for specific manifolds in Section 2.4. When $\mathcal{M}$ is a Riemannian submanifold, the second-order condition for a retraction can be interpreted as requiring the extrinsic acceleration $\sigma''_{x,v}(0)$ to be normal to the manifold at x.

Furthermore, when the curve $\sigma_{x,v}$ is interpreted as a curve in $\mathcal{E}$ it can be compared with the geodesic $\gamma_{x,v}(t) = \mathrm{Exp}_x(tv)$, also interpreted in the ambient space. Then, the local rigidity condition and the second-order condition imply that the retraction curve matches the Taylor expansion of the geodesic up to first and second order respectively, as summarized by the following result. This is a particular case of [AM12, Proposition 2.2-2.3] in the setting like ours where both the manifold and the retraction are assumed to be infinitely differentiable.
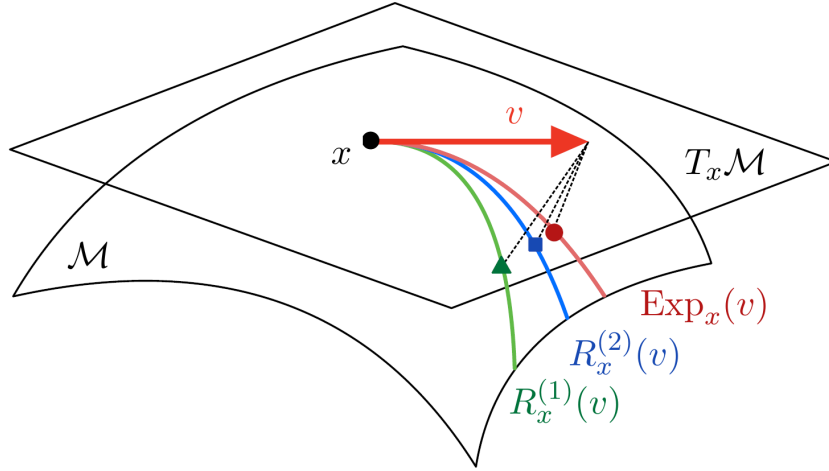
Figure 2.2: Pictorial comparison between a (first-order) retraction $R^{(1)}$, a second-order retraction $R^{(2)}$ and the Riemannian exponential map Exp.

Proposition 2.11. *For $\mathcal{M}$ an embedded submanifold of an Euclidean space $\mathcal{E}$, let R denote a smooth map from an open neighborhood of $\{(x,0) \in T\mathcal{M} : x \in \mathcal{M}\}$ to $\mathcal{M}$. For any $(x,v) \in T\mathcal{M}$, consider the geodesic curve $\gamma_{x,v}(t) = \mathrm{Exp}_x(tv)$ and the smooth curve $\sigma_{x,v}(t) = R_x(tv)$, which can always be both defined around $t = 0$. Then:*

1. R is a retraction if and only if

$$\sigma_{x,v}(t) = \gamma_{x,v}(t) + O(t^2), \quad \text{as } t \to 0,$$

2. R is a second-order retraction if and only if

$$\sigma_{x,v}(t) = \gamma_{x,v}(t) + O(t^3), \quad \text{as } t \to 0.$$

A useful consequence of this result is that it provides a way to recognize retractions on Riemannian submanifolds based on how close retraction curve approximate geodesics. Vice-versa, retractions may be constructed by truncating the Taylor series of the exponential map, provided this leads to a point on the manifold.

2.2.1   Inverse retraction

Another important consequence of the local rigidity condition in the definition of retraction is that retractions are locally invertible. The following property [Bou23, Corollary 10.27] conveniently characterizes the subset of the retraction's domain over which the inverse retraction can be defined.

Proposition 2.12. *For any retraction R there exists a continuous function $\Delta : \mathcal{M} \to (0, \infty]$ defining an open subset $\mathcal{D} := \{(x,v) \in T\mathcal{M} : \|v\|_x < \Delta(x)\}$ of the tangent bundle*

such that

$$E : \mathcal{D} \subset T\mathcal{M} \to \mathcal{M} \times \mathcal{M}$$
$$(x, v) \mapsto (x, R_x(v))$$

is a diffeomorphism.

Accordingly, if $v \in T_x\mathcal{M}$ is such that $\|v\|_x < \Delta(x)$, then the retraction is well-defined. Likewise, whenever $x, y \in \mathcal{M}$ are such that $(x, y) \in E(\mathcal{D})$, we say that the inverse retraction $R_x^{-1}(y)$ is well-defined and is equal to the unique $v \in T_x\mathcal{M}$ such that $E(x, v) = (x, y)$. Moreover, with this definition the inverse retraction is smooth jointly in x and y. This definition of inverse retraction is analogous to the one given for the Riemannian logarithm, a particular inverse of the exponential map, see Section 1.3.2. In fact, the inverse retraction for the exponential retraction coincides with the Riemannian logarithm. In this case, the function $\Delta(x)$ can be taken to be the injectivity radius given by Definition 1.31.

2.2.2 Projective retractions for embedded submanifolds

We now turn to the case of a Riemannian manifold $\mathcal{M}$ embedded into an Euclidean space $\mathcal{E}$ to present two general constructions leading to a well-defined retraction. In both approaches, the retraction in $(x, v) \in T\mathcal{M}$ is computed by first perturbing the point x into ambient space as $x + v \in \mathcal{E}$ and then suitably projecting this ambient point back to the manifold. This class of retractions goes by the name of projective-like retractions as introduced in [AM12].
The *projective retraction*, that we denote R^π, maps the ambient point $x + v$ to the closest point on the manifold via the closest-point projection $\Pi_\mathcal{M}$:

$$R_x^\pi(v) := \Pi_\mathcal{M}(x + v) = \underset{y \in \mathcal{M}}{\arg\min} \|x + v - y\|. \tag{2.4}$$

The smoothness and well-posedness of the metric projection $\Pi_\mathcal{M}$ in a neighborhood of the manifold [AM12, Lemma 3.1] guarantees the projective retraction fulfills Definition 2.9. The *orthographic retraction* R^o projects the perturbed point $x + v$ back onto the manifold along vectors from the normal space of the starting point, see Figure 2.3. Formally this reads:

$$R_x^o(v) := \underset{y \in (x + v + N_x\mathcal{M}) \cap \mathcal{M}}{\arg\min} \|x + v - y\|. \tag{2.5}$$

A convenient strategy to compute the solution of this optimization is available for the the sphere, the fixed-rank manifold and the Stiefel manifold, see Section 2.4. The major advantage of the orthographic retraction is that the inverse retraction is trivial. As illustrated by Figure 2.3, it is sufficient to project the ambient space difference onto the tangent space as

$$(R_x^o)^{-1}(y) = \Pi(x)(y - x). \tag{2.6}$$

Owing to a general result about projection-like retraction [AM12, Theorem 4.9], if $\mathcal{M}$ is a
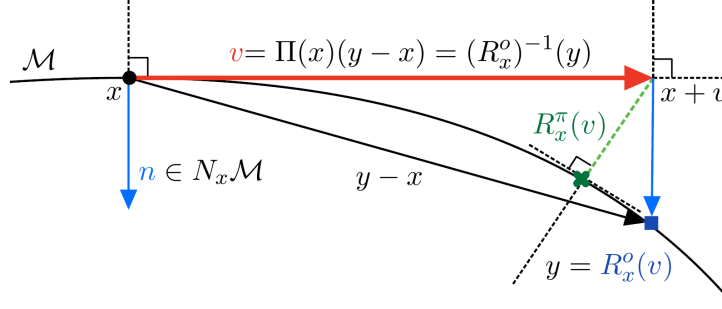
Figure 2.3: The projective retraction R^π, the orthographic retraction R^o and the inverse orthographic retraction $(R^o)^{-1}$.

Riemannian submanifold, both the projective retraction and the orthographic retraction are second-order retractions.

2.3   Riemannian optimization algorithms

Let us briefly introduce some standard Riemannian optimization algorithms used throughout the thesis.

2.3.1   Retraction-based line search methods

Most optimization methods considered in this work are part of the class of retraction-based *line search* methods. The recurrent iterative procedure consists in identifying a direction along which the objective function decreases and performing a step along this direction using a retraction. The gradient descent method along geodesics mentioned in Section 2.1.3 is one such method, the one that chooses the steepest descent direction and the most natural retraction, the exponential map. More in general, at each iteration a certain *descent direction* $v_k \in T_{x_k}\mathcal{M}$, i.e. satisfying $\langle \mathrm{grad}f(x_k), v_k \rangle_{x_k} < 0$, is scaled by a step size $\alpha_k > 0$ and is followed using a retraction R to produce $x_{k+1} = R_{x_k}(\alpha_k v_k)$. Provided the so-called Armijo-Goldstein sufficient decrease condition

$$f(R_{x_k}(\alpha_k v_k)) < f(x_k) + \alpha_k \beta \langle \mathrm{grad}f(x_k), v_k \rangle_{x_k} \tag{2.7}$$

can be guaranteed at each step and the objective function is lower bounded on $\mathcal{M}$ and with compact level-sets, the so-defined general descent method is guaranteed to converge to a critical point of the objective function for any initial condition [AMS08, Corollary 4.3.2]. When this is the case, the method is said to be *globally convergent*. The speed at which the iterates approach the limit is characterized by the *order of local convergence*. We say the method has local convergence of order one if a measure of distance to the local minimizer at iteration k, denoted Δ_k, is reduced asymptotically in a geometric way, i.e.

$$\Delta_{k+1} = c\Delta_k, \quad \forall\, k \geq K,$$

for some $c \in (0,1)$ and an integer $K > 0$. Commonly used measures of distance include the Riemannian distance, the gap of objective function values or the norm of the gradient. Superlinear convergence refers to the case where the constant c can be taken arbitrarily small asymptotically. Equivalently, a method is superlinearly convergent if there exist $\theta > 0$ such that

$$\Delta_{k+1} = c\Delta_k^{1+\theta}, \quad \forall k \geq K.$$

If $\theta = 1$ the method is said to be quadratically convergent or of second-order.

The simplest example of line search method is the Riemannian gradient descent (RGD) method, outlined in Algorithm 2.1. It is a globally convergent method [AMS08, Corollary 4.3.2] and under suitable assumption on the objective function exhibits first-order convergence in terms of objective function values [AMS08, Theorem 4.5.6].

Algorithm 2.1 Riemannian gradient descent (RGD)

Input: $x_0 \in \mathcal{M}$, $f \in \mathfrak{F}(\mathcal{M})$, $\varepsilon_{\text{tol}} > 0$, $k_{\max} \in \mathbb{N}$, $\beta \in (0,1)$.

1: $k = 0$;
2: **while** $\| \operatorname{grad} f(x_k) \|_{x_k} > \varepsilon_{\text{tol}}$ & $k < k_{\max}$ **do**
3: $\quad v_k = -\operatorname{grad} f(x_k)$;
4: $\quad$ Choose step size α_k so that (2.7) is verified;
5: $\quad x_{k+1} = R_{x_k}(\alpha_k v_k)$;
6: $\quad k \leftarrow k + 1$;
7: **end while**
8: **return** x_k;

One practical way to meet the sufficient decrease requirement is via the so-called *backtracking line search*. An initial step size is reduced geometrically until the condition (2.7) is verified, as detailed in Algorithm 2.2. The procedure is guaranteed to terminate at each step under the assumption the objective function is such that the pullback $\hat{f} = f \circ R_x^{-1}$ has Lipschitz continuous gradient, uniformly for all $x \in \mathcal{M}$ [Bou23, Proposition 4.8].

Algorithm 2.2 Backtracking line search

Input: $x \in \mathcal{M}$, $f \in \mathfrak{F}(\mathcal{M})$, $v \in T_x\mathcal{M}$, $\alpha_0 > 0$, $\beta \in (0,1)$, $\tau \in (0,1)$.

1: **if** $\langle \operatorname{grad} f(x), v \rangle \geq 0$ **then**
2: $\quad$ **return error**;
3: **end if**
4: **while** $f(R_x(\alpha_m v)) \geq f(x) + \alpha_m \beta \langle \operatorname{grad} f(x), v \rangle_x$ **do**
5: $\quad \alpha_{m+1} = \tau \alpha_m$;
6: $\quad m \leftarrow m + 1$;
7: **end while**
8: **return** α_m;

2.3.2 Riemannian Newton and Riemannian trust-region

The enhanced local information about the objective function provided by the Riemannian Hessian function allows for the the generalization of second-order optimization schemes to manifolds. Let us first introduce the generalization of the Newton method for scalar optimization. The procedure is outlined in Algorithm 2.3. The update direction v_k at iteration k is taken as the solution of the following generalization of the Newton equation

$$\mathrm{Hess}f(x)\,[v_k] = -\,\mathrm{grad}f(x_k).$$

The direction v_k is of descent only when the Riemannian Hessian is positive definite, i.e. when the objective function is convex. Therefore, in general the Newton method is not guaranteed to converge to a critical point. However, if initialized sufficiently close to a local minimizer with a positive definite Riemannian Hessian, the Newton method converges to it at least quadratically in terms of Riemannian distance [AMS08, Theorem 6.3.2].

Algorithm 2.3 Riemannian Newton (RN)

Input: $x_0 \in \mathcal{M}$, $f \in \mathfrak{F}(\mathcal{M})$, $\varepsilon_{\mathrm{tol}} > 0$, $k_{\max} \in \mathbb{N}$.

1: $k = 0$;
2: **while** $\|\,\mathrm{grad}f(x_k)\|_{x_k} > \varepsilon_{\mathrm{tol}}$ & $k < k_{\max}$ **do**
3:     Solve $\mathrm{Hess}f(x_k)\,[v_k] = -\,\mathrm{grad}f(x_k)$ for $v_k \in T_{x_k}\mathcal{M}$;
4:     $x_{k+1} = R_{x_k}(v_k)$;
5:     $k = k + 1$;
6: **end while**
7: **return** x_k;

The trust-region method provides remedy to the unreliable global convergence properties of the Newton method in the general non-convex case, while maintaining the second-order local convergence. As detailed in Algorithm 2.4, the mechanism allowing this best-of-both-worlds behavior involves a varying size radius indicating the size of the region around the current iterate where the approximation of the objective function provided by the Hessian is considered valid. The next iterate is sought inside this *trust-region* with the goal to reduce a second-order model of the objective function. For this reason, the trust-region algorithm is not a line search method. Together with an acceptance-rejection criterion for the next candidate iterate, which also adjusts the size of the trust-region, the method is guaranteed to reach a critical point [AMS08, Theorem 7.4.2]. Asymptotically, if the critical point is a local minimizer, the trust-region method mimics the Newton method and so converges quadratically in terms of Riemannian distance [AMS08, Theorem 7.4.11].

Algorithm 2.4 Riemannian trust-region (RTR)

Input: $x_0 \in \mathcal{M}$, $f \in \mathfrak{F}(\mathcal{M})$, $\varepsilon_{\text{tol}} > 0$, $k_{\max} \in \mathbb{N}$, $\bar{\Delta} > 0$, $\Delta_0 \in (0, \bar{\Delta})$, $\rho' \in [0, 1/4)$.

1: $k = 0$;
2: **while** $\| \text{grad} f(x_k) \|_{x_k} > \varepsilon_{\text{tol}} \, \& \, k < k_{\max}$ **do**
3:     Choose v_k as an approximate solution of

$$\min_{v \in T_{x_k}\mathcal{M}} m_{x_k}(v) := f(x_k) + \langle \text{grad} f(x_k), v \rangle_{x_k} + \frac{1}{2} \langle v, \text{Hess} f(x_k)[v] \rangle_{x_k}$$

$$\text{subject to } \|v\|_{x_k} \le \Delta_k;$$

4:     $\rho_k = \frac{f(x_k) - f(R_{x_k}(v_k))}{m_{x_k}(0) - m_{x_k}(v_k)}$;
5:     **if** $\rho_k < 1/4$ **then**
6:         $\Delta_{k+1} = \frac{1}{4}\Delta_k$;
7:     **else if** $\rho_k > 3/4$ and $\|v_k\|_{x_k} = \Delta_k$ **then**
8:         $\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$;
9:     **else**
10:        $\Delta_{k+1} = \Delta_k$;
11:    **end if**
12:    $k = k + 1$;
13:    **if** $\rho_k > \rho'$ **then**
14:       $x_{k+1} = R_{x_k}(v_k)$;
15:    **else**
16:       $x_{k+1} = x_k$;
17:    **end if**
18: **end while**
19: **return** x_k;

2.3.3 Riemannian Conjugate Gradient and Riemannian BFGS

Classical optimization methods such as the conjugate gradient descent or quasi-Newton methods, combine the gradients and the update directions of the previous iterates to craft improved descent directions. When generalizing such optimization methods to the manifold setting, one must be careful to move vectors to a common tangent space before performing the combinations. The notion of parallel transport introduced in Definition 1.34 would serve the purpose. However, as the exponential map and the weaker definition of retraction, the essential features of parallel transport can be collected into the more general notion of *vector transport*. It is specifically crafted to operate in combination with a retraction.

Definition 2.13 ([AMS08, Definition 8.1.1]). *On a differentiable manifold $\mathcal{M}$, a* vector transport $\mathcal{T}$ *associated to a retraction R is a mapping*

$$\mathcal{T} : T\mathcal{M} \oplus T\mathcal{M} \to T\mathcal{M}$$
$$(u, v) \mapsto \mathcal{T}_u(v)$$

where $T\mathcal{M} \oplus T\mathcal{M} = \{(u, v) : u, v \in T_x\mathcal{M}, x \in \mathcal{M}\}$ such that
  1. *(Associated retraction) $\mathcal{T}_u(v) \in T_{R_x(u)}\mathcal{M}$,*
  2. *(Consistency) $\mathcal{T}_{0_x}(v) = v$,*
  3. *(Linearity) $\mathcal{T}_u(av + bw) = a\mathcal{T}_u(v) + b\mathcal{T}_u(w)$.*

If furthermore, the manifold is equipped with a Riemannian metric, the vector transport is said to be *isometric* if it preserves the metric, that is,

$$\langle v, w \rangle_x = \langle \mathcal{T}_u(v), \mathcal{T}_u(w) \rangle_{R_x(u)}, \quad \forall v, u, w \in T_x\mathcal{M}.$$

As expected from the intent of defining vector transport, parallel transport along a retraction curve $t \to R_x(tv)$ from $t = 0$ to $t = 1$, provided the retraction curve itself is well-defined, is always a vector transport for any retraction. In particular, parallel transport along a retraction curve is an isometric vector transport. A general construction for a vector transport on an embedded submanifold $\mathcal{M}$ to an Euclidean space $\mathcal{E}$ leverages the usual identification of tangent vectors of $\mathcal{M}$ with vectors in $\mathcal{E}$. Given a retraction R on $\mathcal{M}$, projecting a tangent vector $w \in T_x\mathcal{M}$ to the tangent space of $R_x(v)$ defines a vector transport:

$$\mathcal{T}^\pi_v(w) = \Pi(R_x(v))w, \quad \forall v, w \in T_x\mathcal{M}. \tag{2.8}$$

Another commonly used vector transport is the so-called differentiated retraction vector transport, given by

$$\mathcal{T}^\delta_v(w) = \mathrm{D}R_x(v)[w], \quad \forall v, w \in T_x\mathcal{M}.$$

With this concept of vector transport at hand, the Riemannian conjugate gradient (RCG) and the Riemannian BFGS (RBFGS) methods generalizing their Euclidean counterpart to a manifold setting are given in Algorithm 2.5 and Algorithm 2.6. Note that that

choice of β_{k+1} at line 6 of RCG must be consistent throughout an execution and the vector transport used in the RBFGS method needs to be invertible. Regarding convergence properties, the RCG method is a globally convergent method [RW12, Proposition 15]. The RBFGS scheme is also globally convergent if the objective function is convex [RW12, Proposition 10] and locally superlinearly convergent [RW12, Corollary 13]. For a general objective function, the method can be modified to update the linear operator H_k at iteration k only if $\langle y_k, s_k \rangle > 0$. This guarantees updated H_{k+1} is positive definite if H_k also is [RW12, Lemma 9]. With this modification, the algorithm goes by the name of cautious RBFGS and is globally convergent [HAG18, Theorem 4.2] and locally superlinearly convergent [HAG18, Corollary 5.2].

Algorithm 2.5 Riemannian conjugate gradient (CG)

Input: $x_0 \in \mathcal{M}$, $f \in \mathfrak{F}(\mathcal{M})$, $\varepsilon_{\text{tol}} > 0$, $k_{\max} \in \mathbb{N}$, $\beta \in (0,1)$.

1: $k = 0$;
2: $v_0 = -\operatorname{grad}f(x_k)$
3: **while** $\|\operatorname{grad}f(x_k)\|_{x_k} > \varepsilon_{\text{tol}} \ \& \ k < k_{\max}$ **do**
4: $\quad$ Choose step size α_k so that (2.7) is verified;
5: $\quad$ $x_{k+1} = R_{x_k}(\alpha_k v_k)$;
6: $\quad$ Compute β_{k+1} as

$$\beta_{k+1}^{\text{FR}} = \frac{\langle \operatorname{grad}f(x_{k+1}), \operatorname{grad}f(x_{k+1}) \rangle_{x_{k+1}}}{\langle \operatorname{grad}f(x_k), \operatorname{grad}f(x_k) \rangle_{x_k}} \qquad \text{(Fletcher-Reeves)}$$

$\quad$ or

$$\beta_{k+1}^{\text{PR}} = \frac{\langle \operatorname{grad}f(x_{k+1}), \operatorname{grad}f(x_{k+1}) - \mathcal{T}_{\alpha_k v_k}(\operatorname{grad}f(x_k)) \rangle_{x_{k+1}}}{\langle \operatorname{grad}f(x_k), \operatorname{grad}f(x_k) \rangle_{x_k}}; \qquad \text{(Polak–Ribière)}$$

7: $\quad$ $v_{k+1} = -\operatorname{grad}f(x_{k+1}) + \beta_{k+1}\mathcal{T}_{\alpha_k v_k}(v_k)$
8: $\quad$ $k \leftarrow k + 1$;
9: **end while**
10: **return** x_k;

Algorithm 2.6 Riemannian BFGS (RBFGS)

Input: $x_0 \in \mathcal{M}$, $f \in \mathfrak{F}(\mathcal{M})$, $\varepsilon_{\text{tol}} > 0$, $k_{\max} \in \mathbb{N}$, $H_0 : T_{x_0}\mathcal{M} \to T_{x_0}\mathcal{M}$ linear.

 1: $k = 0$;
 2: **while** $\| \operatorname{grad} f(x_k) \|_{x_k} > \varepsilon_{\text{tol}}$ & $k < k_{\max}$ **do**
 3:     Solve $H_k v_k = -\operatorname{grad} f(x_k)$ for $v_k \in T_{x_k}\mathcal{M}$;
 4:     Choose step size α_k so that (2.7) is verified;
 5:     $x_{k+1} = R_{x_k}(\alpha_k v_k)$;
 6:     $y_k = \nabla f(x_{k+1}) - \mathcal{T}_{\alpha_k v_k} \nabla f(x_k)$;
 7:     $s_k = \mathcal{T}_{\alpha_k v_k}(\alpha_k v_k)$;
 8:     Define the linear operator $H_{k+1} : T_{x_{k+1}}\mathcal{M} \longmapsto T_{x_{k+1}}\mathcal{M}$ by

$$
H_{k+1}w = \tilde{H}_k w - \frac{\langle s_k, \tilde{H}_k w \rangle_{x_{k+1}}}{\langle s_k, \tilde{H}_k s_k \rangle_{x_{k+1}}} \tilde{H}_k s_k + \frac{\langle y_k, w \rangle_{x_{k+1}}}{\langle y_k, s_k \rangle_{x_{k+1}}} y_k, \quad \forall w \in T_{x_{k+1}}\mathcal{M},
$$

       where $\tilde{H}_k = \mathcal{T}_{\alpha_k v_k} \circ H_k \circ (\mathcal{T}_{\alpha_k v_k})^{-1}$;
 9:     $k = k + 1$;
10: **end while**
11: **return** x_k

2.4   Retractions for manifolds of interest

The following section reviews the definition of well-known retractions for the manifolds of interest considered in the thesis.

2.4.1   Sphere

Before introducing retractions on the sphere S^{n-1} presented in Section 1.5.1, we point out that for the sphere the analytical expression for the exponential map (1.7) under the embedded geometry is simple enough to lend itself as a valid candidate for a retraction in practical algorithms. With the expression for the associated Riemannian logarithm (1.8), the inverse retraction of the exponential retraction can be computed easily.

The metric projection retraction (2.4) is the most frequently used retraction on the sphere. From the expression of the metric projection (1.6) we get that for every $x \in S^{n-1}$ it is defined as

$$
R_x^\pi(v) = \frac{x+v}{\|x+v\|}, \quad \forall v \in T_x S^{n-1}. \tag{2.9}
$$

This retraction is defined on the whole tangent space since the metric projection on S^{n-1} is always uniquely defined, but from the point x, only the open half sphere $H_x^+ := \{y \in S^{n-1} : y^\top x > 0\}$ can be reached. Yet, on this hemisphere the retraction is bijective and admits the following inverse

$$
(R_x^\pi)^{-1}(y) = \frac{y}{x^\top y} - x, \quad \forall y \in H_x^+.
$$

The derivation of the orthographic retraction (2.4) in the case of the sphere uses that the normal space at $x \in S^{n-1}$ only contains scalar multiples of x. Then, for some $v \in T_x S^{n-1}$, we impose $R_x^o(v) = x + v - \alpha x$, where $\alpha > 0$ is the smallest real root of

$$\|x + v - \alpha x\|^2 = 1 \quad \Leftrightarrow \quad \alpha = 1 \pm \sqrt{1 - v^\top v}$$

We evince that the orthographic retraction at $x \in S^{n-1}$ is well-defined only if $\|v\| \leq 1$ and has the expression

$$R_x^o(v) = x\sqrt{1 - v^\top v} + v.$$

The restriction on the domain of the orthographic retraction makes it less appealing than the metric projection retraction, although the orthographic retraction at x maps to a slightly larger region as it is able to reach the closed half hemisphere $\overline{H_x^+} =: \left\{ y \in S^{n-1} : x^\top y \geq 0 \right\}$. The formula for the inverse orthographic retraction (2.6) is general and holds as such on the sphere; its domain is restricted to the closed hemisphere $\overline{H_x^+}$.

2.4.2   Stiefel manifold

For the Stiefel manifold presented throughout Chapter 1, we introduce two most commonly used retractions based on polar and QR decompositions respectively. As shown in [KFT13], both these retractions admit a procedure for which the inverse retraction can be conveniently computed. In the following, let $X \in \mathrm{St}(n,k)$ and $V \in T_X\mathrm{St}(n,k)$. The *P-factor retraction* is the metric projection retraction for the Stiefel manifold [AM12, Proposition 3.4] and is given by

$$R_X^\pi(V) = \mathrm{pf}(X + V),$$

where for any $A \in \mathbb{R}^{n \times k}$ we define $\mathrm{pf}(A) = P$ such that $A = PS$ is the polar decomposition of the matrix A with $P \in \mathrm{St}(n,k)$ and $S \in \mathbb{S}_k^+$. This decomposition is uniquely defined provided the matrix A has full rank. This is always the case for $X + V$. In fact, from the condition given in Example 1.16 satisfied by any tangent vector V at X, we have that $(X+V)^\top(X+V) = I_k + V^\top V$, which is a positive definite matrix. Thus $X+V$ is always of full rank and the P-factor retraction is defined on the whole tangent space. The polar factor P can be obtained from the SVD of the matrix $A = \sum_{i=1}^k \sigma_i u_i v_i^\top$ as $\mathrm{pf}(A) = \sum_{i=1}^k u_i v_i^\top$ or with the matrix square root $R_X^\pi(V) = (X+V)(I + V^\top V)^{1/2}$. An algorithm to compute the inverse P-factor retraction is proposed in [KFT13, Algorithm 2] and requires solving a Lyapunov matrix equation.

The so-called *Q-factor retraction* is defined as [Bou23, §7.3]:

$$R_X^\mathrm{Q} = \mathrm{qf}(X + V)$$

where $\mathrm{qf}(A) = Q \in \mathbb{R}^{n \times k}$ such that $A = QR$ is the unique (thin) QR decomposition of a matrix $A \in \mathbb{R}^{n \times k}$. The QR decomposition is made unique for any full rank A by enforcing

the diagonal entries of R to be positive. Hence, analogously to the P-factor retraction, the Q-factor retraction is also defined on the whole tangent space. Contrarily to the P-factor retraction, the Q-factor retraction does not meet the second-order condition. A procedure to compute the Q-factor inverse retraction is proposed in [KFT13, Algorithm 1] and consists of solving a linear matrix equation for an upper triangular matrix which is required to have positive diagonal entries.

For completeness, we point out the orthographic retraction for the Stiefel manifold has been experimented with in [KFT13, Algorithm 3]. The procedure requires solving a quadratic Lyapunov matrix equation through a potentially lengthy iterative process, and is not considered any further.

2.4.3   Symmetric positive definite matrices

The Riemannian geometry of the set of $n \times n$ symmetric positive definite matrices $\mathbb{S}_n^+$ as presented in Section 1.5.2 is not inherited from the Euclidean geometry of the vector space of symmetric matrices $\mathbb{S}_n$. Nevertheless, the metric projection retraction can be defined on $\mathbb{S}_n^+$. Since $\mathbb{S}_n^+$ is an open subset of $\mathbb{S}_n$ it is only locally defined. If $V \in T_X\mathbb{S}_n^+ \simeq \mathbb{S}_n$ has a sufficiently small norm, then $X + V$ is symmetric positive definite, meaning the metric projection of $X + V$ onto $\mathbb{S}_n$ is the identity. Hence, we have $R_X^\pi(V) = X + V$ and is well-defined for any V such that $X + V$ is positive definite. The restriction on the domain makes this retraction inconvenient.

The analytical expression provided in Section 1.5.2 for the exponential map of $\mathbb{S}_n^+$ under its bi-invariant metric makes it a valid candidate as a retraction, although potentially expensive to compute. A cheaper alternative is found by approximating the matrix exponential in (1.12) with its truncated Taylor series up to order 2. For every $X \in \mathbb{S}_n^+$, this defines a retraction with the following expression [JVV12, §4.1.3]

$$R_X(V) = X + V + \frac{1}{2}VX^{-1}V, \quad \forall V \in T_x\mathbb{S}_n^+. \tag{2.10}$$

2.4.4   Fixed-rank matrices

A substantial number of retractions are known for the fixed-rank manifold presented in Section 1.5.3. The overview paper by Absil and Oseledets [AO15] provides an extensive list, and we refer the reader to that reference for a complete overview, including implementation details and complexity analyses. We here introduce only the two projective retractions. In the following we consider any $X = U\Sigma V^\top \in \mathcal{M}_k$ and $Z = (M, U_p, V_p) \in T_X\mathcal{M}_k$ as in (1.15) and (1.16).

Although the fixed-rank matrix manifold is not a closed subset of $\mathbb{R}^{m \times n}$, the closest-point projection is always uniquely defined in a neighborhood of $\mathcal{M}_k$ as a consequence of the general result [AM12, Lemma 3.1]. Hence, for sufficiently small Z, the rank-k truncated

SVD denoted $\Pi_{\mathcal{M}_k}$ given by (1.18) defines that metric projection retraction for $\mathcal{M}_k$

$$R_X^\pi(Z) := \underset{Y \in \mathcal{M}_k}{\arg\min} \|X + Z - Y\|_F = \Pi_{\mathcal{M}_r}(X + V). \tag{2.11}$$

More specifically, this retraction is well-defined provided $\sigma_k(X + Z) > \sigma_{k+1}(X + Z)$. By stability of singular values, we evince that $\|Z\|_{\mathrm{op}} := \sigma_1(Z) < \sigma_k(X)/2$ guarantees the well-posedness of the retraction. From the structure of $Z \in T_X \mathcal{M}_k$ seen in (1.16), the matrix $X + Z$ is of rank at most $2k$, in fact in can be written with block matrices as

$$X + Z = \begin{bmatrix} U & U_p \end{bmatrix} \begin{bmatrix} \Sigma + M & I_k \\ I_k & 0 \end{bmatrix} \begin{bmatrix} V^\top \\ V_p^\top \end{bmatrix}$$

This allows for an efficient implementation of the projective retraction, as detailed in Algorithm 2.7: the rank-k truncation can be carried out with two orthonormalizations of sizes $m \times k$ and $n \times k$ and the computation of an SVD of size $2k \times 2k$. Note that the thin QR decompositions at line 1 and 2 could be replaced by a polar decompositions without changing the output of the algorithm. As advocated in [Van13] and as formerly implemented in the library Manopt [BMAS14], an artificial machine precision value $\varepsilon_{\mathrm{mach}}$ ($= 2^{-52} \simeq 10^{-16}$ in double precision) can be added at line 4 to all newly computed singular values as a safeguard from falling onto $\mathcal{M}_{k-1}$ in the unfortunate occurrence of zero singular values.

Algorithm 2.7 Metric projection retraction on $\mathcal{M}_k$.

Input: $X = U_0 \Sigma_0 V_0^\top \in \mathcal{M}_k$, $Z = (M, U_p, V_p) \in T_X \mathcal{M}_k$

1: $[Q_u, R_u] = \mathrm{qr}(U_p, 0)$; $\triangleright$ thin QR decomposition

2: $[Q_v, R_v] = \mathrm{qr}(V_p, 0)$;

3: $[U_s, \Sigma_s, V_s] = \mathrm{SVD}\left(\begin{bmatrix} \Sigma + M & R_v^\top \\ R_u & 0 \end{bmatrix} \right)$;

4: $\Sigma_1 = \Sigma_s(1:k, 1:k)$; $\triangleright$ optional: $\Sigma_1 = \Sigma_1 + \varepsilon_{\mathrm{mach}}$

5: $U_1 = \begin{bmatrix} U_0 & Q_u \end{bmatrix} U_s(:, 1:k)$;

6: $V_1 = \begin{bmatrix} V_0 & Q_v \end{bmatrix} V_s(:, 1:k)$;

7: **return** : $U_1 \Sigma_1 V_1^\top =: R_X^\pi(V)$;

For the fixed-rank manifold, a closed-form expression for the solution of the optimization problem (2.5) defining the orthographic retraction is given in [AM12, Proposition 4.11] and its implementation detailed in [AO15, §3.2]. We report the computation of this orthographic retraction as the output of Algorithm 2.8. As for the case of the metric projection retraction, the orthonormalizations in the first two lines of Algorithm 2.8 could be performed by polar decomposition.

Algorithm 2.8 Orthographic retraction on $\mathcal{M}_k$.

Input: $X = U_0 \Sigma_0 V_0^\top \in \mathcal{M}_k$, $Z = (M, U_p, V_p) \in T_X \mathcal{M}_k$

1: $[\tilde{U}_1, S_u] = \mathrm{qr}(U_0(\Sigma_0 + M) + U_p, 0);$ $\triangleright$ thin QR decomposition
2: $[\tilde{V}_1, S_v] = \mathrm{qr}(V_0(\Sigma_0 + M^\top) + V_p, 0);$
3: $\tilde{\Sigma}_1 = S_u(\Sigma_0 + M)^{-1} S_v^\top$
4: $[U_s, \Sigma_s, V_s] = \mathrm{SVD}(\tilde{\Sigma}_1);$
5: $U_1 \leftarrow \tilde{U}_1 U_s,\ \Sigma_1 \leftarrow \Sigma_s,\ V_1 \leftarrow \tilde{V}_1 V_s;$
6: **return** : $U_1 \Sigma_1 V_1^\top =: R_X^o(V);$

3 Retractions as a curve generating device

The developments of the present thesis leverage the tool of retractions as presented in Chapter 2 by viewing it as a device capable of generating portions of smooth manifold curves. This chapter is specifically dedicated to this viewpoint on retractions. We start by discussing the well-known fact that suitable tangent space curves can be used to generate manifold curves with prescribed initial position, velocity and acceleration. Then we introduce r-endpoint retraction curves: a new class of curves built using a retraction and its inverse of which the endpoint can be prescribed. This chapter also gathers some tools used in the thesis to analyze these curves and the algorithms that use them. Notably, we introduce the novel concept of retraction-convex sets, a generalization of geodesically convex sets where retraction curves replace geodesics in the definition. We then state and analyze Lipschitz continuity properties of the retraction and, in turn, of the retraction curves previously introduced. A detailed discussion on the definition and the properties of the inverse retraction high-order differentials is included next. Finally, we prove a result quantifying how accurately suitable retractions curves with prescribed velocity and acceleration can approximate a given smooth manifold curve.

3.1 Retractions on tangent space curves

While very general, the definition of a retraction (see Definition 2.9) determines a device that can be used to generate manifold curves with prescribed initial position, velocity and, provided the retraction is second-order, also acceleration. More in general, curves defined on a tangent space of the manifold can be mapped by means of a retraction to the manifold. The resulting manifold curve retains derivative information of the tangent space curve, as specified by the following proposition.

Proposition 3.1 ([Bou23, Exercise 5.46])**.** *Let $\mathcal{M}$ denote a smooth manifold and R a retraction on $\mathcal{M}$. For any $x \in \mathcal{M}$ consider a smooth tangent space curve $w : J \subset \mathbb{R} \to T_x\mathcal{M}$ defined in a neighborhood of zero such that $w(0) = 0$. For sufficiently small $|t|$ the retraction curve $c(t) = R_x(w(t))$ is well-defined and satisfies*

$$\begin{cases} c(0) = x, \\ \dot{c}(0) = w'(0). \end{cases}$$
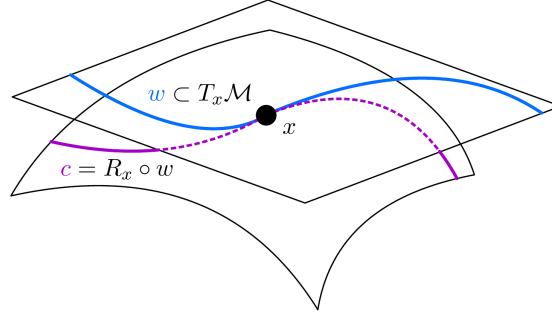
Figure 3.1: Mapping a tangent space curve w to manifold curve γ with a retraction R.

If additionally $\mathcal{M}$ is Riemannian and R is a second-order retraction then we also have

$$\ddot{c}(0) = w''(0).$$

Before dedicating the rest of this section to the proof of Proposition 3.1, let us state two useful application of this result. The first describes a simple way of generating retraction curves with prescribed initial velocity and acceleration. The second one can be seen as the converse of Proposition 3.1, in that it specifies when derivative information of a manifold curve is maintained when mapped through an inverse retraction.

Corollary 3.2. *For every $x \in \mathcal{M}$ and $v, a \in T_x\mathcal{M}$ the curve*

$$\sigma(t) = R_x\left(tv + \frac{t^2}{2}a\right)$$

is well-defined for all $|t|$ sufficiently small and satisfies
  1. $\sigma(0) = x$ and $\dot{\sigma}(0) = v$ for any retraction R,
  2. $\ddot{\sigma}(0) = a$ if R is a second-order retraction.

Corollary 3.3. *Given a smooth manifold curve $\gamma : J \subset \mathbb{R} \to \mathcal{M}$ and a retraction R, for any $t \in J$ the tangent space curve $h \mapsto \hat{\gamma}(h) := R_{\gamma(t)}^{-1}(\gamma(t+h))$ is well-defined for every $|h|$ sufficiently small and satisfies*
  1. $\hat{\gamma}(0) = 0$ and $\hat{\gamma}'(0) = \dot{\gamma}(t)$ for any retraction R,
  2. $\hat{\gamma}''(0) = \ddot{\gamma}(t)$ if R is a second-order retraction.

Proof. For any $t \in J$, the well-posedness of the curve $\hat{\gamma}$ in an open interval containing $h = 0$ follows from the smoothness of the curve γ and the well-posedness of the inverse retraction in an open neighborhood of $\gamma(t)$. The definition of retractions immediately implies $\hat{\gamma}(0) = R_{\gamma(t)}^{-1}(\gamma(t)) = 0$. The other requirements are shown by contradiction. Assume that $\hat{\gamma}'(0) \neq \dot{\gamma}(t)$. By Proposition 3.1, it holds that for $\sigma(h) := R_{\gamma(t)}(\hat{\gamma}(h))$ we have $\dot{\sigma}(0) = \hat{\gamma}'(0)$. But by construction $\sigma(h) = \gamma(t+h)$ for every sufficiently small h, hence $\dot{\sigma}(0) = \dot{\gamma}(t)$, a contradiction. The same argument holds for the acceleration condition by invoking Proposition 3.1 with a second-order retraction. $\qquad\square$

Proof of Proposition 3.1

The proof of the acceleration condition in Proposition 3.1 makes use of a smooth local frame around x, i.e. a collection of $d = \dim(\mathcal{M})$ smooth vector fields E_i, $i = 1, \ldots, d$, defined in a neighborhood $\mathcal{U}$ of x such that $\{E_i(y)\}_{i=1}^d$ forms a basis of $T_y\mathcal{M}$ for all $y \in \mathcal{U}$. A smooth local frame can be defined around any $x \in \mathcal{M}$ [Lee13, Example 8.10-(b)] and if $\mathcal{M}$ is endowed with a Riemannian metric it is always possible [Lee18, Proposition 2.8] to make it orthonormal with respect to the metric, i.e. to satisfy

$$\langle E_i(y), E_j(y) \rangle_y = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \qquad \forall\, y \in \mathcal{U},$$

by applying the Gram-Schmidt algorithm. The resulting orthonormal local frame is also smooth by the smoothness of the Riemannian metric and of the Gram-Schmidt process.

Proof of Proposition 3.1. By the definition of retractions, we know R_x is defined in an open neighborhood of $0 \in T_x\mathcal{M}$. The tangent vector $w(t)$ is contained in that open neighborhood for sufficiently small $|t|$ since w is smooth and $w(0) = 0$. This grants the well-posedness of the curve $c(t) = R_x(w(t))$ in a neighborhood of zero. The first defining property of retractions, yields $c(0) = R_x(w(0)) = R_x(0) = x$. Where defined, the tangent vector to the retraction curve at t is given by chain rule as

$$\dot{c}(t) = \mathrm{D}R_x(w(t))[w'(t)].$$

Using the local rigidity property of retractions, we find $\dot{c}(0) = \mathrm{D}R_x(0)[w'(0)] = w'(0)$. Assume now $\mathcal{M}$ has a Riemannian metric and let $\{E_i\}_{i=1}^d$ for $d = \dim(\mathcal{M})$ denote an orthonormal smooth local frame in a neighborhood $\mathcal{U}$ of x. For any $v \in T_x\mathcal{M}$ and $u \in T_v T_x\mathcal{M} \simeq T_x\mathcal{M}$, from the definition of differential we know that $\mathrm{D}R_x(v)[u]$ belongs to $T_{R_x(v)}\mathcal{M}$ and can be decomposed as

$$\mathrm{D}R_x(v)[u] = \sum_{i=1}^d \mathrm{D}R_x(v)[u_i E_i(x)] = \sum_{i=1}^d u_i \sum_{j=1}^d r_{ij}(v) E_j(R_x(v)),$$

where $u_i = \langle u, E_i(x) \rangle_x$ is the component of u along $E_i(x)$ and

$$r_{ij}(v) = \langle \mathrm{D}R_x(v)[E_i(x)], E_j(R_x(v)) \rangle_{R_x(v)}$$

is a well-defined scalar for any sufficiently small $v \in T_x\mathcal{M}$, for each $i, j = 1, \ldots, d$. More specifically, by smoothness of R_x, the functions r_{ij} are defined and smooth on an open neighborhood $V \subset T_x\mathcal{M}$ containing the origin such that $R_x(v) \in \mathcal{U}$. Hence, for any $|t|$ sufficiently small so that $w(t) \in V$, we may rewrite $\dot{c}(t)$ as

$$\dot{c}(t) = \sum_{i,j=1}^d w_i(t) r_{ij}(w(t)) E_j(c(t)),$$

where we have introduced the components functions $w_i(t) = \langle w(t), E_j(c(t)) \rangle_{c(t)}$, for all $i = 1, \ldots, d$. The smoothness of w implies the smoothness of the component functions $w_i(t)$ and the functions r_{ij} are smooth on the set V by the smoothness of the retraction, the local frame and the Riemannian metric. Then, by $\mathbb{R}$-linearity and the product-rule for induced covariant differentiation (see Theorem 1.25) we have that

$$\ddot{c}(t) = \sum_{i,j=1}^{d} \frac{\mathrm{D}}{\mathrm{d}t} \left(w_i'(t) r_{ij}(w(t)) E_j(c(t)) \right)$$

$$= \sum_{i,j=1}^{d} w_i''(t) r_{ij}(w(t)) E_j(c(t)) + w_i'(t) \left(\left. \frac{\mathrm{d}r_{ij}(w(t))}{\mathrm{d}t} \right|_{t=0} E_j(c(t)) + r_{ij}(w(t)) \frac{\mathrm{D}}{\mathrm{d}t} E_j(c(t)) \right).$$

The derivative of $r_{ij} \circ w$ can be written as the differential of r_{ij} at $w(t)$ along $w'(t)$ as

$$\left. \frac{\mathrm{d}r_{ij}(w(t))}{\mathrm{d}t} \right|_{t=0} = \mathrm{D}r_{ij}(w(t))[w'(t)].$$

Since E_j is a smooth vector field, its covariant derivative along c can be expressed by chain rule with the connection

$$\frac{\mathrm{D}}{\mathrm{d}t} E_j(c(t)) = \nabla_{\dot{c}(t)} E_j(c(t)),$$

and depends on the pointwise value of $\dot{c}(t)$. Note that by local rigidity of retractions we have $r_{ij}(0) = \langle E_i(x), E_j(x) \rangle_x = \delta_{ij}$. Hence, in $t = 0$ reordering terms gives

$$\ddot{c}(0) = \sum_{i}^{d} w_i''(0) E_i(x) + w_i'(0) \nabla_{w'(0)} E_i(x) + w_i'(0) \sum_{j=1}^{d} \mathrm{D}r_{ij}(w(0))[w'(0)] E_j(x). \quad (3.1)$$

Let us now consider the tangent space line $\tilde{w}(t) = tw'(0)$. If R is a second-order retraction, we know that $\tilde{c}(t) := R_x(t\tilde{w}'(0))$ has zero acceleration in $t = 0$. But since $\tilde{w}(0) = w(0)$, $\tilde{w}'(0) = w'(0)$ and $\tilde{w}''(0) = 0$, writing equation (3.1) for $\tilde{c}$ with this replacement tells us that

$$0 = w_i'(0) \nabla_{w'(0)} E_j(x) + w_i'(0) \sum_{j=1}^{d} \mathrm{D}r_{ij}(w(0))[w'(0)] E_j(x), \quad \forall i = 1, \ldots, d.$$

Therefore, if R is a second-order retraction then the second and third term of equation (3.1) vanish and we can conclude

$$\ddot{c}(0) = \sum_{i}^{d} w_i''(0) E_i(x) = w''(0).$$

$\square$

3.2   A family of endpoint retraction curves

The local rigidity property of retractions leads to formulating a precise definition for a local inverse of the retraction, see Section 2.2.1. Other than being a convenient theoretical tool, inverse retractions can also be efficiently computed, as in the case of the orthographic retraction presented in Section 2.2.2. This opens the way to using inverse retractions also in numerical algorithms. In the following, we introduce a new procedure that combines a retraction and its inverse to define a one-parameter family of curves that share some prescribed endpoints. Throughout this section, we assume that any use of the retraction and its inverse is well-defined and we postpone the discussion on well-posedness to Section 3.3.
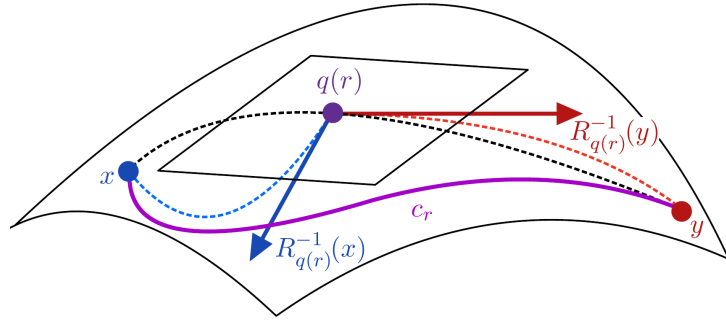
Figure 3.2: r-endpoint retraction curve.

Definition 3.4. *Given $x, y \in \mathcal{M}$ and $r \in [0,1]$, the r-endpoint retraction curve joining x and y is defined as*

$$c_r(t; x, y) = R_{q(r)}\left((1-t)R_{q(r)}^{-1}(x) + tR_{q(r)}^{-1}(y)\right), \quad \forall\, t \in [0,1],$$

where $q(r) = R_x(rR_x^{-1}(y))$.

Whenever x and y are sufficiently close, the r-endpoint retraction curve can be defined for any $r \in [0,1]$. A more precise statement on the well-posedness of r-endpoint retraction curves is provided in the next section by Proposition 3.8. The fundamental endpoint property for all r-endpoint curves in the family is stated next, together with a property concerning the initial and final velocities for the cases $r = 0$ and $r = 1$ respectively.

Proposition 3.5. *The r-endpoint retraction curve family satisfies the following properties:*
  (i) $c_r(0; x, y) = x$ and $c_r(1; x, y) = y$ for every $r \in [0,1]$;
  (ii) $\dot{c}_0(0; x, y) = R_x^{-1}(y)$ and $\dot{c}_1(1; x, y) = -R_y^{-1}(x)$.

Proof. For any $r \in [0,1]$, the definition of c_r directly implies (i) by invoking the first

property of retractions:

$$c_r(0; x, y) = R_{q(r)}\left(R_{q(r)}^{-1}(x)\right) = x,$$
$$c_r(1; x, y) = R_{q(r)}\left(R_{q(r)}^{-1}(y)\right) = y.$$

To show (ii), first note that

$$\dot{c}_r(t; x, y) = \mathrm{D}R_{q(r)}\left((1 - t)R_{q(r)}^{-1}(x) + tR_{q(r)}^{-1}(y)\right)\left[R_{q(r)}^{-1}(y) - R_{q(r)}^{-1}(x)\right].$$

Because we have $q(0) = x$ and $q(1) = y$, noting that $R_x^{-1}(x) = 0$ and $R_y^{-1}(y) = 0$ allows concluding, by local rigidity of the retraction:

$$\dot{c}_0(0; x, y) = \mathrm{D}R_x(0)\left[R_x^{-1}(y)\right] = R_x^{-1}(y),$$
$$\dot{c}_1(1; x, y) = \mathrm{D}R_y(0)\left[-R_y^{-1}(x)\right] = -R_y^{-1}(x).$$

$\square$

The choice of $r \in [0, 1]$ in the r-endpoint retraction curve determines the anchor point for the retraction that defines c_r. For values of $r \in (0, 1)$, taking the anchor point on the image of the curve $q(r) = R_x(rR_x^{-1}(y))$ is an arbitrary choice. One could have, for example, taken it on the curve $\tilde{q}(r) = R_y(rR_y^{-1}(x))$. This curve is in general different from q and leads to a different curve family. Nevertheless, this alternative curve family obtained from replacing q with $\tilde{q}$ is included in our Definition 3.4 by considering the curves $c_r(\cdot; y, x)$. In other words, the image of the curve c_r is not invariant under permutation of the endpoints, or also, $c_r(t; x, y)$ does not coincide with $c_r(1 - t; y, x)$ in general. Only extreme members of the family are related in this way, as $c_0(t; x, y) = c_1(1 - t; y, x)$.

3.3   Retraction-convex sets

To guarantee that the retraction and the inverse retractions involved in the definition of r-endpoint retraction curve c_r and, in turn, c_r itself are well-defined, we introduce the notion of a retraction-convex set. Retraction-convex sets can be seen as a generalization for geodesically convex sets. Among different possible notions of geodesic convexity, its strongest form goes by the name of strongly geodesic convexity in [Bou23, Definition 11.17] and is defined as follows. It uses the notion of length-minimizing geodesic introduced along with Proposition 1.33.

Definition 3.6 ([Lee18, §6]). *A subset $\mathcal{V} \subset \mathcal{M}$ is a geodesically convex set if for every $x, y \in \mathcal{U}$ there exists a unique length-minimizing geodesic $\gamma : [0, 1] \to \mathcal{M}$ such that $\gamma(0) = x$, $\gamma(1) = y$ and $\gamma(t) \in \mathcal{V}$ for all $t \in [0, 1]$.*

In contrast with this definition, since endpoint retraction curves are not unique, the definition of retraction-convexity replaces the unique length-minimizing geodesic by a
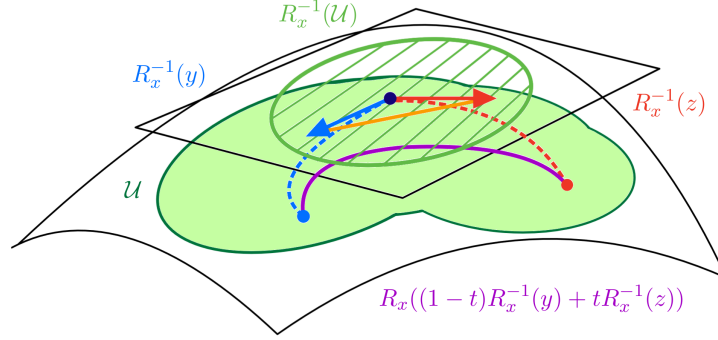
Figure 3.3: Retraction-convex set.

collection of endpoint retraction curves. The requirement then becomes that all their images remain in the set, for every pair of endpoints taken in the set.

Definition 3.7. *A set $\mathcal{U} \subset \mathcal{M}$ is said to be* retraction-convex *if for any $x, y, z \in \mathcal{U}$*

(i) *the inverse retraction $R_x^{-1}(y)$ is well-defined,*

(ii) *$R_x((1-t)R_x^{-1}(y) + tR_x^{-1}(z))$ is well-defined for every $t \in [0, 1]$ and belongs to $\mathcal{U}$.*

An equivalent characterization of retraction-convexity of $\mathcal{U}$ is that the inverse retraction R_x^{-1} is well-defined on $\mathcal{U}$ and the image of $\mathcal{U}$ through the inverse retraction is a convex subset of $T_x\mathcal{M}$ for every $x \in \mathcal{U}$.

Note that in the case where the retraction is the exponential map, retraction-convexity is not necessarily equivalent to geodesic convexity. On the one hand, for a geodesically convex set $\mathcal{V}$ and any $x \in \mathcal{V}$, the inverse retraction $\mathrm{Exp}_x^{-1} = \mathrm{Log}_x$ is indeed well-defined on $\mathcal{V}$. On the other hand, $\mathrm{Log}_x(\mathcal{V})$ is only required to be star-shaped: for every $t \in [0, 1]$ the point $\mathrm{Exp}_x(t\mathrm{Log}_x(y))$ belongs to $\mathcal{V}$, hence $t\mathrm{Log}_x(y)$ belongs to $\mathrm{Log}_x(\mathcal{V})$ for every $t \in [0, 1]$. Furthermore, even if the tangent space segment $(1-t)\mathrm{Log}_x(y) + t\mathrm{Log}_x(z)$ is contained in the domain of Exp_x for $t \in [0, 1]$, the image of this segment through the exponential map is not necessarily a length-minimizing geodesic and therefore its image may not belong entirely to $\mathcal{V}$.

Another definition for the retraction-convexity for a set $\mathcal{U} \subset \mathcal{M}$ appears in [Bou23, §11.8]. It requires that for every $x, y \in \mathcal{U}$ there exists $v \in T_x\mathcal{M}$ such that $R_x(v) = y$ and $R_x(tv) \in \mathcal{U}$ for every $t \in [0, 1]$. This definition is less restrictive than Definition 3.7 in that the retraction is not required to be invertible as specified in Section 2.2.1: the tangent vector v such that $R_x(v) = y$ need not be unique. Furthermore, only endpoint retraction curves of the form $t \mapsto R_x(tv)$ are bound to remain in the set. Hence, this definition is not sufficient to grant the well-posedness of r-endpoint retraction curves. This motivates the need for the stronger notion of retraction-convexity introduced in Definition 3.7. By comparing definitions in the other direction, it is possible to show a retraction-convex set $\mathcal{U}$ according to Definition 3.7 is retraction-convex in the sense of [Bou23, §11.8], as we now briefly argue. For any $x, y \in \mathcal{U}$, by condition (i) of Definition 3.7 we can take

$v = R_x^{-1}(y)$ to have $R_x(v) = y$. Then, by condition (ii),

$$R_x(tv) = R_x((1-t)\underbrace{R_x^{-1}(x)}_{=0} + tR_x^{-1}(y))$$

is guaranteed to belong to $\mathcal{U}$ for every $t \in [0,1]$.

Retraction-convexity for a set should not be confused with the more commonly encountered notion of retraction-convexity for smooth scalar fields [HGA15, Definition 3.1] used in the convergence analysis of Riemannian optimization algorithms.

From now on, a retraction-convex set is to be intended exclusively in the sense of Definition 3.7. The existence of retraction-convex sets is the object of the next section. As a motivation for pursuing this goal, let us show that having a retraction-convex set is sufficient to guarantee the well-posedness of any r-endpoint retraction curve between any pair of points in the set. As a consequence, we also get that the curve's image is contained in the retraction-convex set.

Proposition 3.8. *Let $\mathcal{U}$ be retraction-convex and $x, y \in \mathcal{U}$. Then for any $r \in [0,1]$ the r-endpoint retraction curve $t \mapsto c_r(t; x, y)$ is well-defined for every $t \in [0,1]$ and belongs to $\mathcal{U}$.*

Proof. By Definition 3.7 (i), $R_x^{-1}(y)$ is well-defined. Now using Definition 3.7 (ii) with $z = x$ and $t = 1 - r$, we find that $q(r) = R_x(rR_x^{-1}(y))$ is well-defined and belongs to $\mathcal{U}$. Therefore, the inverse retractions $R_{q(r)}^{-1}(x)$ and $R_{q(r)}^{-1}(y)$ are also well-defined. Finally, applying once again Definition 3.7 (ii) ensures the well-posedness of $c_r(t; x, y)$ and that it is contained in $\mathcal{U}$, for every $t \in [0,1]$. $\qquad\square$

3.3.1   Existence of retraction-convex sets

Having established that r-endpoint retraction curves exist in a retraction-convex set, let us show that retraction-convex sets actually exist. In the following section, we prove that for every point on the manifold any sufficiently small metric ball is retraction-convex.

Regarding the related notion of geodesic convexity, any manifold point is known to admit a geodesically convex neighborhood, as shown for instance by do Carmo [dC92, Proposition 4.2].[1] The proof is based on the Gauss lemma [dC92, Lemma 3.5], a result from Riemannian geometry that crucially relies on the exponential map which does not extend to general retraction curves. Hence, a different approach has to be taken to show existence of retraction-convex sets. The following proof relies mainly on the properties of the manifold distance function and shares common features with the proof of [dC92, Theorem 3.3.7] that shows the existence of totally retractive neighborhoods as defined in [HAG15, §3.3].

Theorem 3.9. *On a Riemannian manifold $\mathcal{M}$, for every $x \in \mathcal{M}$ there exists $\bar{\rho} > 0$ such that for any $\rho < \bar{\rho}$ the manifold ball $B(x, \rho) = \{y \in \mathcal{M} : d(x, y) < \rho\}$ is a retraction-convex set.*

---

[1]Note that [dC92] uses the term strongly convex set instead of geodesically convex set.

The proof of Theorem 3.9 makes use of the following lemma.

Lemma 3.10. *For every $x \in \mathcal{M}$ and any $v \in T_x\mathcal{M}$ we have*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} d(x, R_x(tv))^2\big|_{t=0} = 2 \|v\|_x^2 .$$

Proof. Setting $c(t) := R_x(tv)$, by Proposition 1.33 we know the distance between x and $c(t)$ can be expressed for sufficiently small $|t|$ as

$$d(x, c(t))^2 = \|\mathrm{Log}_x(c(t))\|_x^2 .$$

The compatibility with the metric of induced covariant differentiation stated in Theorem 1.25 allows us to explicitly compute

$$\frac{\mathrm{d}}{\mathrm{d}t} d(x, c(t))^2 = 2 \left\langle \mathrm{DLog}_x(c(t))[\dot{c}(t)], \mathrm{Log}_x(c(t)) \right\rangle_x ,$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} d(x, c(t))^2 = 2\langle \frac{\mathrm{D}}{\mathrm{d}t}(\mathrm{DLog}_x(c(t))[\dot{c}(t)]), \mathrm{Log}_x(c(t))\rangle_x \\ + 2 \|\mathrm{DLog}_x(c(t))[\dot{c}(t)]\|_x^2 . \tag{3.2}$$

Using that $c(0) = x$ and $\mathrm{Log}_x(c(0)) = \mathrm{Log}_x(x) = 0$, the first term of (3.2) vanishes at $t = 0$. By local rigidity of the exponential map, the inverse function theorem implies that $\mathrm{DLog}_x(x) = I_{T_x\mathcal{M}}$. The proof is completed by noting that by Proposition 3.1 we have $\dot{c}(0) = v$ and therefore

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} d(x, c(t))^2\big|_{t=0} = 2 \|\mathrm{DLog}_x(x)[v]\|_x^2 = 2 \|v\|_x^2 .$$

$\square$

Proof of Theorem 3.9. Given $x \in \mathcal{M}$, consider the function $f : T\mathcal{M} \to \mathbb{R}$

$$f(y, v) = d(x, R_y(v))^2.$$

The squared Riemannian distance function is smooth [Lee18, Lemma 6.8] on a neighborhood of $\{(x, x) : x \in \mathcal{M}\}$, the diagonal of $\mathcal{M} \times \mathcal{M}$, and by Proposition 2.12 the retraction R is a diffeomorphism from an open neighborhood of $T\mathcal{M}$ containing $(x, 0)$ to an open neighborhood of the diagonal of $\mathcal{M} \times \mathcal{M}$. Therefore, by composition the function f is smooth on an open neighborhood of $(x, 0)$ in $T\mathcal{M}$.

We first establish a (local) convexity property of f. Using $\nabla^2 f(x, w)$ to denote the (Euclidean) Hessian of $f(x, \cdot)$ at $w \in T_x\mathcal{M}$ for a fixed $x \in \mathcal{M}$, the result of Lemma 3.10 can be rephrased as

$$\left\langle \nabla^2 f(x, 0)[v], v \right\rangle_x = 2 \|v\|_x^2 .$$

In particular, this shows that $\nabla^2 f(x, 0)$ is positive definite. By continuity of f in both

arguments, there exists an open neighborhood $V \subset T\mathcal{M}$ containing $(x,0)$ such that $\nabla^2 f(y,v)$ remains positive definite at every $(y,v) \in V$.

For an arbitrary subset $S \subseteq T\mathcal{M}$, we let

$$\begin{aligned}\pi_{\mathcal{M}}(S) &:= \{y \in \mathcal{M} \,:\, \exists\, v \in T_y\mathcal{M} \text{ such that } (y,v) \in S\}\,, \\ \pi_{T_y\mathcal{M}}(S) &:= \{v \in T_y\mathcal{M} \,:\, (y,v) \in S\}\,. \end{aligned} \tag{3.3}$$

For every $y \in \pi_{\mathcal{M}}(V)$ the function $f(y,\cdot)$ is convex on any convex subset of $\pi_{T_y\mathcal{M}}(V)$. The set V is an open set in the atlas topology of $T\mathcal{M}$, associated with the natural smooth structure of $T\mathcal{M}$ [Lee13, Proposition 3.18]. This means that for any local chart $(\mathcal{U}, \phi)$ of $\mathcal{M}$, the set $\psi_\phi\left(\pi_{\mathcal{M}}^{-1}(\mathcal{U}) \cap V\right)$ is an open subset of $\mathbb{R}^{2d}$, with $d = \dim(\mathcal{M})$ and where

$$\pi_{\mathcal{M}}^{-1}(\mathcal{U}) = \{(y,v) \in T\mathcal{M} \,:\, y \in \mathcal{U},\, v \in T_y\mathcal{M}\} = \coprod_{y \in \mathcal{U}} T_y\mathcal{M}$$

and

$$\begin{aligned}\psi_\phi \,:\, \pi_{\mathcal{M}}^{-1}(\mathcal{U}) \cap V &\to \mathbb{R}^{2d} \\ (y,v) &\mapsto (\phi(y), \lambda) \end{aligned}$$

with $\lambda \in \mathbb{R}^d$ being the coordinates of v in the basis of $T_y\mathcal{M}$ formed by the partial derivatives of the inverse local charts, that is $v = \sum_{i=1}^d \lambda_i \mathrm{D}\phi^{-1}(\phi(y))[e_i]$.

Now consider a local chart such that $x \in \mathcal{U}$. Because $\psi_\phi\left(\pi_{\mathcal{M}}^{-1}(\mathcal{U}) \cap V\right)$ is an open subset of $\mathbb{R}^{2d}$ that contains $(\phi(x), 0)$, there exist $\varepsilon_{\mathcal{M}} > 0$ and $\varepsilon_T > 0$ such that

$$B_d(\phi(x), \varepsilon_{\mathcal{M}}) \times B_d(0, \varepsilon_T) \subseteq \psi_\phi\left(\pi_{\mathcal{M}}^{-1}(\mathcal{U}) \cap V\right)\,,$$

where $B_d(\cdot, \cdot)$ denotes an open ball (in the Euclidean norm) of $\mathbb{R}^d$. By continuity of ψ_ϕ, the preimage

$$\begin{aligned}W := &\psi_\phi^{-1}(B_d(\phi(x), \varepsilon_{\mathcal{M}}) \times B_d(0, \varepsilon_T)) \\ =& \left\{(y,v) \in T\mathcal{M} : y \in \phi^{-1}\left(B_d(\phi(x), \varepsilon_{\mathcal{M}})\right),\, v = \sum_{i=1}^d \lambda_i \partial_i \phi^{-1}(\phi(y)),\, \|\lambda\|_2 < \varepsilon_T \right\} \end{aligned}$$

is an open subset of V that contains $(x,0)$. From Proposition 2.12, we know the map $E : (x,v) \in T\mathcal{M} \to (x, R_x(v)) \in \mathcal{M} \times \mathcal{M}$ is a diffeomorphism on the open subset $\mathcal{D} \subset T\mathcal{M}$, the disjoint union of tangent space disks centered at the origin whose radius is determined by a continuous function Δ. Since E is also a diffeomorphism on the intersection $W \cap \mathcal{D}$, which is open, the set $E(W \cap \mathcal{D})$ is an open neighborhood of (x,x). Hence, the constant

$$\bar{\rho} := \sup\left\{\rho \geq 0 : B(x,\rho) \times B(x,\rho) \subseteq E(W \cap \mathcal{D})\right\}\,.$$

is strictly positive. The statement of the theorem is proved by showing that the set $B(x,\rho)$ is retraction-convex for any $\rho < \bar{\rho}$.

Requirement (i) for retraction-convexity regarding the invertibility of the retraction follows from the invertibility of E on $B(x, \rho) \times B(x, \rho)$. For any $y, z \in B(x, \rho)$ there exists a unique $v \in T_y \mathcal{M}$ such that $E(y, v) = (y, z)$ and, hence, $v = R_y^{-1}(z)$ is well-defined.

To establish requirement (ii), consider arbitrary $w, y, z \in B(x, \rho)$. We first note that the set

$$\pi_{T_y \mathcal{M}}(W \cap \mathcal{D}) = \left\{ v \in T_y \mathcal{M} : v = \sum_{i=1}^{D} \lambda_i \partial_i \phi^{-1}(\phi(y)), \|\lambda\|_2 < \varepsilon_T, \|v\|_y < \Delta(y) \right\}$$

is convex (as the intersection of two convex sets) and in the domain of R_y. Both $R_y^{-1}(w)$ and $R_y^{-1}(z)$ are contained in $\pi_{T_y \mathcal{M}}(W \cap \mathcal{D})$ and hence the same holds for their convex linear combination. The convexity of $f(y, \cdot)$ on $\pi_{T_y \mathcal{M}}(W \cap \mathcal{D})$, a convex subset of $\pi_{T_y \mathcal{M}}(V)$, implies for every $t \in [0, 1]$ that

$$\begin{aligned} d\big(x, R_y((1-t)R_y^{-1}(w) + tR_y^{-1}(z))\big)^2 &= f\big(y, (1-t)R_y^{-1}(w) + tR_y^{-1}(z)\big) \\ &\leq (1-t)f\big(y, R_y^{-1}(w)\big) + tf\big(y, R_y^{-1}(z)\big) \\ &= (1-t)d(x, w)^2 + td(x, z)^2 < \rho^2, \end{aligned}$$

which proves the retraction-convexity of $B(x, \rho)$. $\qquad \square$

3.3.2 Retraction-convexity radius function

Given a $\bar{\rho} > 0$ satisfying Theorem 3.9, then the property is also satisfied by $\eta \bar{\rho}$, for any $\eta \in (0, 1)$. Thus, to fix a single value for each $x \in \mathcal{M}$, we denote

$$\bar{\rho}(x) := \sup \left\{ \bar{\rho} > 0 : B(x, \rho) \text{ is retraction-convex } \forall \rho < \bar{\rho} \right\}. \tag{3.4}$$

The quantity $\bar{\rho}$ can be interpreted as a function defined on the whole manifold which is strictly positive at every point by Theorem 3.9. We refer to it as the retraction-convexity radius function. Throughout the thesis we occasionally invoke the following assumption.

Assumption 3.11. *For every compact set $K \in \mathcal{M}$, there exists a constant $\rho_{\min}(K) > 0$ such that $\bar{\rho}(x) \geq \rho_{\min}(K)$ for all $x \in K$.*

Since retractions are smooth, intuition suggests that the function $\bar{\rho}$ should be continuous, thereby guaranteeing the validity of this technical assumption.

3.4 Retractions and Lipschitz continuity

A map f between metric space (X, d_X) and (Y, d_Y) is Lipschitz continuous if there exists a constant $L > 0$ such that

$$d_Y(f(a), f(b)) \leq L d_X(a, b), \quad \forall a, b \in X.$$

It is only locally Lipschitz continuous if this condition is verified in a neighborhood of each point in the domain of f. One interpretation of this definition is that it provides an estimate on the behavior of the map near a point. For this reason, Lipschitz continuity is an essential tool in the error analysis of approximation algorithms such as numerical integration and interpolation.

The smoothness of retractions implies some Lipschitz continuity properties that we fully spell out in this section and that we use in the analyses of the upcoming chapters.

3.4.1 Upper bound on the norm of the differentials of the retraction

For differentiable maps between Riemannian manifolds, Lipschitz continuity can be related with the maximum magnitude of the derivatives. For example, for a scalar field f on a Riemannian manifold, Lipschitz continuity is equivalent to having $\|\mathrm{grad} f(x)\|_x \leq L$ [Bou23, Proposition 10.43]. We will establish an analogous link for the Lipschitz continuity of retractions and inverse retractions.

Let us first introduce some notation for the derivatives of the retractions and its inverse and the sets over which they are defined. Recall that Proposition 2.12 states the existence of a continuous function $\Delta : \mathcal{M} \to (0, +\infty]$ describing conveniently a domain

$$\mathcal{D} = \{(x, v) \in T\mathcal{M} : \|v\|_x < \Delta(x)\}$$

on which the map $E : (x, v) \mapsto (x, R_x(v))$ is a diffeomorphism. We denote $\mathcal{I} = E(\mathcal{D})$ and let us introduce for every $x \in \mathcal{M}$ the sets

$$\mathcal{D}_x := \pi_{T_x\mathcal{M}}(\mathcal{D}) = \{v \in T_x\mathcal{M} : (x, v) \in \mathcal{D}\},$$
$$\mathcal{I}_x := R_x(\mathcal{D}_x) = \{R_x(v) : v \in \mathcal{D}_x\}.$$

On these sets, we can define the differentials of the retraction and inverse retraction with respect to each of their arguments separately. Given $(x, u) \in \mathcal{D}$ and $(x, y) \in \mathcal{I}$, for all $v \in T_x\mathcal{M}$ and $w \in T_y\mathcal{M}$ we

$$\mathrm{D}_1 R_x(u)[v] = \frac{\mathrm{d}}{\mathrm{d}t} R_{\sigma_{x,v}(t)}(u)\big|_{t=0} \in T_{R_x(u)}\mathcal{M},$$
$$\mathrm{D}_2 R_x(u)[v] = \frac{\mathrm{d}}{\mathrm{d}t} R_x(u + tv)\big|_{t=0} \in T_{R_x(u)}\mathcal{M},$$
$$\mathrm{D}_1 R_x^{-1}(y)[v] = \frac{\mathrm{d}}{\mathrm{d}t} R_{\sigma_{x,v}(t)}^{-1}(y)\big|_{t=0} \in T_x\mathcal{M},$$
$$\mathrm{D}_2 R_x^{-1}(y)[w] = \frac{\mathrm{d}}{\mathrm{d}t} R_x^{-1}(\sigma_{y,w}(t))\big|_{t=0} \in T_x\mathcal{M}, \tag{3.5}$$

where $\sigma_{x,v}$ is any continuously differentiable manifold curve defined in a neighborhood of $t = 0$ such that $\sigma_{x,v}(0) = x \in \mathcal{M}$, $\dot{\sigma}_{x,v}(0) = v \in T_x\mathcal{M}$. The so-defined differentials are all linear maps between tangent spaces that depend continuously on their arguments. For instance $\mathrm{D}_1 R_x(u)[\cdot]$ is a linear map from $T_x\mathcal{M}$ to $T_{R_x(u)}\mathcal{M}$ that is continuous with respect to both x and u. For any such linear map $A[\cdot] : T_x\mathcal{M} \to T_y\mathcal{M}$, the Riemannian
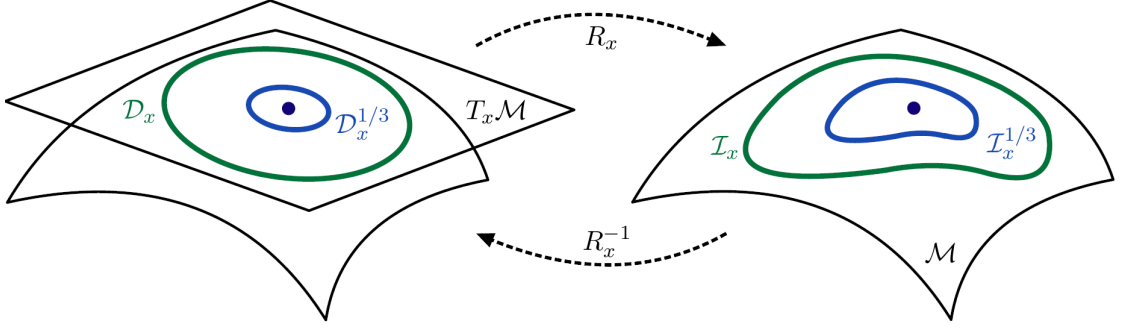
Figure 3.4: The sets $\mathcal{D}_x$, $\mathcal{I}_x$, $\mathcal{D}_x^{1/3}$, $\mathcal{I}_x^{1/3}$.

metric induces an operator norm defined by

$$\|A\| := \sup_{w \in T_x\mathcal{M}, w \neq 0_x} \left\{ \frac{\|A\,[w]\|_y}{\|w\|_x} \right\}. \tag{3.6}$$

As Proposition 3.12 below illustrates and as one would naturally expect, the operator norm of the retraction's differentials is an upper bound to the magnitude of the retraction's derivatives. In the following, we denote for any $\varepsilon \in (0,1)$.

$$\mathcal{D}^\varepsilon := \{(x,v) \in T\mathcal{M} : \|v\|_x < \varepsilon\Delta(x)\},$$

Similarly, we set

$$\mathcal{D}_x^\varepsilon := \pi_{T_x\mathcal{M}}(\mathcal{D}^\varepsilon), \tag{3.7}$$

the restriction to the tangent space at x as defined by (3.3) and

$$\mathcal{I}_x^\varepsilon = R_x(\mathcal{D}_x^\varepsilon). \tag{3.8}$$

Proposition 3.12. *Consider an arbitrary $\varepsilon \in (0,1)$. For any compact subset $K \subset \mathcal{M}$ there exist positive constants $L_1(K,\varepsilon)$, $M_1(K,\varepsilon)$, $L_2(K,\varepsilon)$ and $M_2(K,\varepsilon)$ such that for every $x \in K$ the retraction R satisfies*
   (i) $\|\mathrm{D}_1 R_x(u)[v]\|_{R_x(u)} \leq L_1(K,\varepsilon)\|v\|_x$ for every $u \in \mathcal{D}_x^\varepsilon$ and $v \in T_x\mathcal{M}$
   (ii) $\|\mathrm{D}_2 R_x(u)[v]\|_{R_x(u)} \leq L_2(K,\varepsilon)\|v\|_x$ for every $u \in \mathcal{D}_x^\varepsilon$ and $v \in T_x\mathcal{M}$,
   (iii) $\left\|\mathrm{D}_1 R_x^{-1}(y)[v]\right\|_x \leq M_1(K,\varepsilon)\|v\|_x$ for every $y \in \mathcal{I}_x^\varepsilon$ and $v \in T_x\mathcal{M}$,
   (iv) $\left\|\mathrm{D}_2 R_x^{-1}(y)[w]\right\|_x \leq M_2(K,\varepsilon)\|w\|_y$ for every $y \in \mathcal{I}_x^\varepsilon$ and $w \in T_y\mathcal{M}$.

By smoothness of the retraction's differentials on their domain of definition, their operator norms are continuous scalar functions on the same domains. Hence, on suitably chosen compact subset, they attain a maximum which becomes the constant appearing in Proposition 3.12. Then, the proof of Proposition 3.12 becomes very short once these compact subsets of $\mathcal{D}$ and $\mathcal{I}$ have been established. This is the object of the following Lemma 3.13, of which a more general form appears in [Bou23, Exercise 10.31].

Lemma 3.13. *For every compact set $K \subset \mathcal{M}$ and every $\varepsilon \in (0,1)$, the sets*

$$\overline{\mathcal{D}_K^\varepsilon} := \{(x,v) \in \mathcal{D} : x \in K, \ \|v\|_x \leq \varepsilon\Delta(x)\} \subset T\mathcal{M},$$
$$\overline{\mathcal{I}_K^\varepsilon} := E(\overline{\mathcal{D}_K^\varepsilon}) = \left\{(x, R_x(v)) : (x,v) \in \overline{\mathcal{D}_K^\varepsilon}\right\}$$

are compact sets.

Proof. We have that $E(\overline{\mathcal{D}_K^\varepsilon}) = K \times \overline{\mathcal{I}_K^\varepsilon}$, where E is the diffeomorphism of Proposition 2.12. Diffeomorphisms preserve compactness and the product of two sets is compact if and only if the factor sets are compact. Therefore, since K is compact, $\overline{\mathcal{D}_K^\varepsilon}$ is compact if and only if $\overline{\mathcal{I}_K^\varepsilon}$ is compact. Let us prove $\overline{\mathcal{D}_K^\varepsilon}$ is sequentially compact, equivalent to being compact by the assumed second countability of the manifold topology.
Consider any sequence $\{(x_n, v_n)\}_{n\in\mathbb{N}} \subset \overline{\mathcal{D}_K^\varepsilon}$. Then the sequence $\{x_n\}_{n\in\mathbb{N}}$ is contained in the compact set K and so there exists a convergent subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that $x_{n_k} \to x \in K$. Then, there exists $N > 0$ such that for every $k > N$, $d(x, x_{n_k}) < \operatorname{inj}(x)$, the injectivity radius of the exponential map at x. Hence, by Proposition 1.33, there exists a unique length-minimizing geodesic σ_k joining x and x_{n_k}. The length-minimizing property tells us that $L(\sigma_k) = d(x, x_{n_k})$. Consider the parallel transport map (see Definition 1.34) along σ_k denoted $\mathrm{P}_{0\leftarrow 1}^{\sigma_k} : T_{x_{n_k}}\mathcal{M} \to T_x\mathcal{M}$ and let us define $w_k = \mathrm{P}_{0\leftarrow 1}^{\gamma_k} v_{n_k} \in T_x\mathcal{M}$. Since $\|v_{n_k}\|_{x_{n_k}} \leq \varepsilon\Delta(x_{n_k})$ and $\mathrm{P}_{0\leftarrow 1}^{\gamma_k}$ is an isometry [Bou23, Proposition 10.36], then $\|w_k\|_x \leq \varepsilon\Delta(x_{n_k})$. By the continuity of Δ, we know $\{\varepsilon\Delta(x_{n_k})\}_{k\in\mathbb{N}}$ is a convergent, hence bounded sequence. Therefore $\{w_k\}_{k\in\mathbb{N}} \subset T_x\mathcal{M}$ is a bounded sequence, and thus admits a convergent subsequence $\{w_{k_j}\}_{j\in\mathbb{N}}$ such that $w_{k_j} \to w \in T_x\mathcal{M}$. But since $\|w_{k_j}\|_x \leq \varepsilon\Delta(x_{n_{k_j}})$, we have that $\|w\|_x \leq \varepsilon\Delta(x)$ and consequently that $(x, w) \in \overline{\mathcal{D}_K^\varepsilon}$. The standard Riemannian metric on the tangent bundle associated to any Riemannian manifold, also known as Sasaki metric [GHL04, p. 80], allows the definition of a distance function on the tangent bundle as per Definition 1.21. As discussed in [EM06, p. 240], the infimum over the lengths all paths between two points in the tangent bundle can be upper-bounded by a combination of the length of the underlying manifold curve and the length of the resulting tangent space curve once transported to one of the endpoints. From this we have that

$$d\left((x, w), \left(x_{n_{k_j}}, v_{n_{k_j}}\right)\right) \leq \sqrt{L(\sigma_{k_j})^2 + \left\|w - \mathrm{P}_{0\leftarrow 1}^{\sigma_{k_j}} v_{n_{k_j}}\right\|_x^2}$$

Since the right-hand side converges to zero as $j \to +\infty$, this shows that $\left\{\left(x_{n_{k_j}}, v_{n_{k_j}}\right)\right\}$ converges to $(x, w) \in \overline{\mathcal{D}_K^\varepsilon}$, concluding the proof. $\qquad\square$

Proof of Proposition 3.12. The result follows from the smoothness of the retraction and its inverse on their domain and of the continuity of the operator norm. The Lipschitz constants are found by taking the maxima of the operator norm for the differentials of the retraction on the compact sets $\overline{\mathcal{D}_K^\varepsilon}$ and $\overline{\mathcal{I}_K^\varepsilon}$. $\qquad\square$

3.4.2 Lipschitz continuity of the retraction on compact sets

We now turn the result of Proposition 3.12 into a Lipschitz continuity result for the retraction R_x restricted to $\mathcal{D}_x^\varepsilon$ and for the inverse retraction R_x^{-1} restricted to $\mathcal{I}_x^\varepsilon$. We recall that these subsets are defined in (3.7) and (3.8) using the function Δ introduced in Proposition 2.12 characterizing the domain over which a retraction is a diffeomorphism, see also Figure 3.4 for a pictorial representation. For simplicity of exposition, a particular choice $\varepsilon = 1/3$ is considered in Proposition 3.14. Note however that the result and the following proof hold for any $\varepsilon \in \left(0, \frac{1}{2}\right)$. A statement very similar to Proposition 3.14, can be found in [RW12, Lemma 6], under a local equicontinuity assumption of the retraction derivatives. Our proof, leverages Proposition 2.12 to avoid this assumption.

Proposition 3.14. *For every compact subset $K \subset \mathcal{M}$, there exist constants $L_R(K) > 0$ and $M_R(K) > 0$ such that for all $x \in K$ it holds that*
  (i) $d(R_x(u), R_x(v)) \leq L_R(K)\|u - v\|_x$, for any $u, v \in \mathcal{D}_x^{1/3}$,
  (ii) $\left\|R_x^{-1}(y) - R_x^{-1}(z)\right\|_x \leq M_R(K)d(y, z)$, for any $y, z \in \mathcal{I}_x^{1/3}$.

Proof. (i) Consider any $x \in K$, $u, v \in \mathcal{D}_x^{1/3}$ and define a manifold curve joining $R_x(u)$ and $R_x(v)$ as $\delta(\tau) := R_x(u + \tau(v - u))$. It is well-defined by the convexity of $\mathcal{D}_x^{1/3}$. By definition, the manifold distance between $R_x(u)$ and $R_x(v)$ is bounded by the length of the curve δ. Thus, we have

$$d(R_x(u), R_x(v)) \leq L(\delta) = \int_0^1 \|D_2 R_x(u + \tau(v - u))\,[v - u]\|_{\delta(\tau)}\,\mathrm{d}\tau$$
$$\leq L_2(K, 1/3)\|v - u\|_x,$$

where the final inequality follows from Proposition 3.12-(ii). Hence $L_R(K) = L_2(K, 1/3)$.
(ii) Consider $y, z \in \mathcal{I}_x^{1/3}$. Since $\mathcal{M}$ is connected, by definition of the distance function we know there exists a sequence of piecewise smooth manifold curves $\gamma_k : [0, 1] \to \mathcal{M}$, $k \in \mathbb{N}$, such that
$$\gamma_k(0) = y, \quad \gamma_k(1) = z, \quad \forall\, k \in \mathbb{N}$$
and
$$\lim_{k \to \infty} L(\gamma_k) = \inf_{k \in \mathbb{N}} \{L(\gamma_k)\} = d(y, z).$$

If the image of the curve γ_k, is fully contained in $\mathcal{I}_x^{2/3}$, we can define the tangent space curve $q(t) := R_x^{-1}(\gamma_k(t))$ for every $t \in [0, 1]$ and deduce that

$$\left\|R_x^{-1}(y) - R_x^{-1}(z)\right\|_x \leq \int_0^1 \left\|D_2 R_x^{-1}(\gamma_k(\tau))\,[\dot{\gamma}_k(\tau)]\right\|_x\,\mathrm{d}\tau$$

$$\leq M_2(K, 2/3) \int_0^1 \|\dot{\gamma}_k(\tau)\|_{\gamma_k(\tau)}\,\mathrm{d}\tau,$$

$$= M_2(K, 2/3)L(\gamma_k)$$
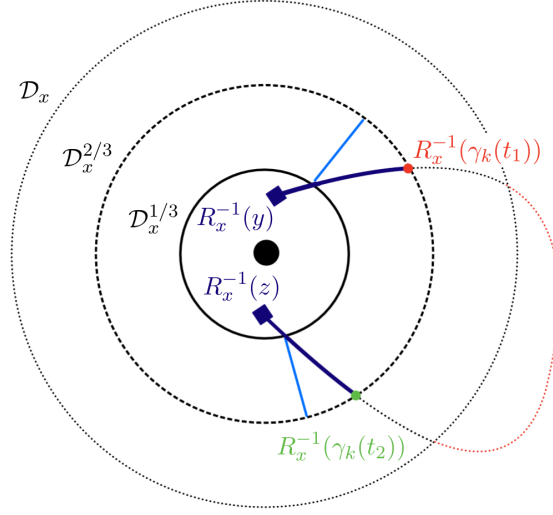
Figure 3.5: Case where γ_k is not fully contained in $\mathcal{I}_x^{2/3}$.

If the image of the curve γ_k is not fully contained in $\mathcal{I}_x^{2/3}$, then let

$$t_1 = \inf\left\{ t \in [0,1] : \gamma_k(t) \notin \mathcal{I}_x^{2/3} \right\},$$
$$t_2 = \sup\left\{ t \in [0,1] : \gamma_k(t) \notin \mathcal{I}_x^{2/3} \right\}.$$

Therefore we can define the tangent space curve $q(t) = R_x^{-1}(\gamma_k(t))$ only for $t \in [0, t_1) \cup (t_2, 1]$. Since $y, z \in \mathcal{I}_x^{1/3}$, it follows that

$$\left\| R_x^{-1}(y) - R_x^{-1}(z) \right\|_x \leq 2\Delta(x)/3$$

Furthermore, the tangent space curve q traverses the tangent space spherical annulus of width $\Delta(x)/3$ back and forth, hence its length must exceed $2\Delta(x)/3$, i.e.

$$2\Delta(x)/3 \leq \int_{[0,t_1)\cup(t_2,1]} \|\dot{q}(\tau)\|_x \, \mathrm{d}\tau.$$

Hence, combing these inequalities leads to

$$
\begin{aligned}
\left\| R_x^{-1}(y) - R_x^{-1}(z) \right\|_x &\leq \int_{[0,t_1)\cup(t_2,1]} \left\| \mathrm{D}_2 R_x^{-1}(\gamma_k(\tau)) \left[\dot{\gamma}_k(\tau) \right] \right\|_x \, \mathrm{d}\tau, \\
&\leq M_2(K, 2/3) \int_{[0,t_1)\cup(t_2,1]} \|\dot{\gamma}_k(\tau)\|_{\gamma_k(\tau)} \, \mathrm{d}\tau, \\
&\leq M_2(K, 2/3) L(\gamma_k).
\end{aligned}
$$

We have shown that for every $k \in \mathbb{N}$

$$\left\| R_x^{-1}(y) - R_x^{-1}(z) \right\|_x \leq M_2(K, 2/3) L(\gamma_k).$$
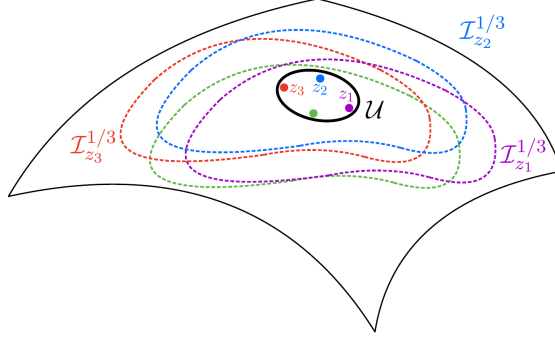
Figure 3.6: A non-empty retraction-convex set $\mathcal{U}$ such that $\mathcal{U} \subset \mathcal{I}_z^{1/3}$ for every $z \in \mathcal{U}$ as in Proposition 3.15

Therefore the result remains true upon taking the infimum over $k \in \mathbb{N}$, which yields

$$\left\| R_x^{-1}(y) - R_x^{-1}(z) \right\|_x \leq M_2(K, 2/3) d(x, y)$$

as desired and shows $M_R(K) = M_2(K, 2/3)$. $\qquad\square$

3.4.3 Lipschitz continuity of retraction curves

As a first application of the Lipschitz continuity results for retractions established in the previous section, let us develop on Lipschitz continuity properties of the r-endpoint retraction curves. We first show the Lipschitz continuity of the r-endpoint retraction curve with respect to each of its arguments and then to all its arguments jointly.

Proposition 3.15. *Let $\mathcal{U} \subset \mathcal{M}$ be non-empty retraction-convex such that*

$$\mathcal{U} \subset \mathcal{I}_z^{1/3}, \quad \forall\, z \in \mathcal{U}. \tag{3.9}$$

There exist positive constants L_t, L_r and L_{xy} depending on the retraction verifying:
(i) $d\left(c_r(t_1; x, y), c_r(t_2; x, y)\right) \leq L_t d(x, y)|t_1 - t_2|, \ \forall\, x, y \in \mathcal{U}, \ \forall\, r, t_1, t_2 \in [0, 1].$
(ii) $d\left(c_{r_1}(t; x, y), c_{r_2}(t; x, y)\right) \leq L_r d(x, y)|r_1 - r_2|, \ \forall\, x, y \in \mathcal{U}, \ \forall\, r_1, r_2, t \in [0, 1].$
(iii) $d\left(c_r(t; x_1, y_1), c_r(t; x_2, y_2)\right) \leq L_{xy}(d(x_1, x_2) + d(y_1, y_2)), \ \forall\, x_1, x_2, y_1, y_2 \in \mathcal{U}, \ \forall\, r, t \in [0, 1].$

Proof. For any fixed $x, y \in \mathcal{U}$, first observe that (3.9) implies that the evaluation of $c_r(t; x, y)$ for any $r, t \in [0, 1]$ requires only the evaluation of the retraction on $\mathcal{D}_z^{1/3}$ and of the inverse retraction on $\mathcal{I}_z^{1/3}$ for some $z = R_x(r R_x^{-1}(y)) \in \mathcal{U}$, see Figure 3.6. Furthermore, (3.9) implies the intersection

$$K = \bigcap_{z \in \mathcal{U}} \overline{\mathcal{I}_z^{1/3}}, \quad \text{with } \overline{\mathcal{I}_z^{1/3}} = \left\{ R_z(v) : \|v\|_z \leq \Delta(z)/3 \right\},$$

is non-empty. In fact, it contains any $z \in \mathcal{U}$, which is assumed to be non-empty. The set K is also compact as the intersection of sets that are compact by Lemma 3.13. We introduce the Lipschitz constants of the retraction and inverse retraction on the set K

given by Proposition 3.12 with the more concise notation

$$L_1 = L_1(K, 1/3), \quad L_2 = L_2(K, 1/3),$$
$$M_1 = M_1(K, 1/3), \quad M_2 = M_2(K, 2/3).$$

Note that we took $M_2(K, 2/3)$ instead of $M_2(K, 1/3)$ in order to use also Proposition 3.14 with these constants. Since $M_2(K, 1/3) \leq M_2(K, 2/3)$, Proposition 3.12 remains valid.
(i) Using Proposition 3.14 with the above constants we find

$$d\left(c_r(t_1; x, y), c_r(t_2; x, y)\right)$$
$$= d\Bigg(R_{q(r)} \left(R_{q(r)}^{-1}(x) + t_1 \left(R_{q(r)}^{-1}(y) - R_{q(r)}^{-1}(x) \right) \right),$$
$$R_{q(r)} \left(R_{q(r)}^{-1}(x) + t_2 \left(R_{q(r)}^{-1}(y) - R_{q(r)}^{-1}(x) \right) \right) \Bigg)$$
$$\leq L_2 \| R_{q(r)}^{-1}(y) - R_{q(r)}^{-1}(x) \|_{q(r)} |t_1 - t_2|$$
$$\leq L_2 M_2 d(x, y) |t_1 - t_2|.$$

Therefore we can choose $L_t = L_2 M_2$.
(ii) The smooth curve $\delta(\tau) := c_{r(\tau)}(t; x, y)$, where $r(\tau) = (1 - \tau) r_1 + \tau r_2$, $\tau \in [0, 1]$, joins $c_{r_1}(t; x, y)$ and $c_{r_2}(t; x, y)$. Therefore we can bound

$$d\left(c_{r_1}(t; x, y), c_{r_2}(t; x, y)\right) \leq L(\delta) \leq \max_{\tau \in [0,1]} \left\{ \left\| \dot{\delta}(\tau) \right\|_{\delta(\tau)} \right\}. \tag{3.10}$$

Denoting $\xi(\tau) := (1 - t) R_{q(r(\tau))}^{-1}(x) + t R_{q(r(\tau))}^{-1}(y) \in T_{q(r(\tau))}\mathcal{M}$, we can rewrite more concisely as

$$\delta(\tau) = R_{q(r(\tau))}(\xi(\tau)),$$
$$q(r(\tau)) = R_x(r(\tau) R_x^{-1}(y)).$$

Therefore by chain rule we find

$$\dot{\delta}(\tau) = D_1 R_{q(r(\tau))}\left(\xi(\tau)\right)\left[(r_2 - r_1)\dot{q}(r(\tau))\right] + D_2 R_{q(r(\tau))}\left(\xi(\tau)\right)\left[\dot{\xi}(\tau)\right],$$

with

$$\dot{q}(r(\tau)) = D_2 R_{q(r(\tau))}(r(\tau) R_x^{-1}(y))\left[R_x^{-1}(y)\right]$$

and

$$\dot{\xi}(\tau) = (1 - t) D_1 R_{q(r(\tau))}^{-1}(x)\left[(r_2 - r_1)\dot{q}(r(\tau))\right] + t D_1 R_{q(r(\tau))}^{-1}(y)\left[(r_2 - r_1)\dot{q}(r(\tau))\right].$$

Using Proposition 3.12 and Proposition 3.14 we can bound the norms of these tangent

vectors as follows:

$$\|\dot{q}(r(\tau))\|_{q(r(\tau))} \leq L_2 \|R_x^{-1}(y)\|_x$$
$$\leq L_2 M_2 d(x, y),$$

$$\left\|\dot{\xi}(\tau)\right\|_{q(r(\tau))} \leq (|1 - t| + |t|) M_1 \|\dot{q}(r(\tau))\|_{q(r(\tau))} |r_1 - r_2|$$
$$\leq 2 M_1 L_2 M_2 d(x, y)|r_1 - r_2|,$$

$$\left\|\dot{\delta}(\tau)\right\|_{\delta(\tau)} \leq L_1 \|\dot{q}(r(\tau))\|_{q(r(\tau))} |r_1 - r_2| + L_2 \left\|\dot{\xi}(\tau)\right\|_{q(r(\tau))}$$
$$\leq (L_1 L_2 M_2 + 2 L_2^2 M_1 M_2) d(x, y)|r_1 - r_2|.$$

Together with (3.10), this shows that $L_r = (L_1 L_2 + 2 L_2^2 M_1) M_2$.
(iii) We define a manifold curve joining $c_r(t; x_1, y_1)$ and $c_r(t; x_2, y_2)$ as

$$\delta(\tau) = c_r(t; \delta_x(\tau), \delta_y(\tau)), \tau \in [0, 1],$$

where the curves $\delta_x(\tau) := R_{x_1}(\tau R_{x_1}^{-1}(x_2))$, $\delta_y(\tau) := R_{y_1}(\tau R_{y_1}^{-1}(y_2))$ are well-defined since endpoints belong a to retraction-convex set. Then, as previously

$$d\left(c_r(t; x_1, y_1), c_r(t; x_2, y_2)\right) \leq L(\delta) \leq \max_{\tau \in [0,1]} \left\{ \left\|\dot{\delta}(\tau)\right\|_{\delta(\tau)} \right\}.$$

We rewrite δ in the more compact form

$$\delta(\tau) = R_{p(\tau)}\left(\xi(\tau)\right),$$

where

$$p(\tau) := R_{\delta_x(\tau)}(r R_{\delta_x(\tau)}^{-1}\left(\delta_y(\tau)\right)),$$
$$\xi(\tau) := (1 - t) R_{p(\tau)}^{-1}(\delta_x(\tau)) + t R_{p(\tau)}^{-1}(\delta_y(\tau)).$$

Differentiating these quantities with respect to τ yields

$$\dot{\delta}(\tau) = \mathrm{D}_1 R_{p(\tau)}(\xi(\tau))\left[\dot{p}(\tau)\right] + \mathrm{D}_2 R_{p(\tau)}(\xi(\tau))\left[\dot{\xi}(\tau)\right],$$

$$\dot{p}(\tau) = \mathrm{D}_1 R_{\delta_x(\tau)}(rR^{-1}_{\delta_x(\tau)}(\delta_y(\tau)))\left[\dot{\delta}_x(\tau)\right] +$$

$$\mathrm{D}_2 R_{\delta_x(\tau)}(rR^{-1}_{\delta_x(\tau)}(\delta_y(\tau)))\bigg[r\mathrm{D}_1 R^{-1}_{\delta_x(\tau)}(\delta_y(\tau))\left[\dot{\delta}_x(\tau)\right] +$$

$$r\mathrm{D}_2 R^{-1}_{\delta_x(\tau)}(\delta_y(\tau))\left[\dot{\delta}_y(\tau)\right]\bigg],$$

$$\dot{\xi}(\tau) = (1-t)\bigg(\mathrm{D}_1 R^{-1}_{p(\tau)}(\delta_x(\tau))\left[\dot{p}(\tau)\right] + \mathrm{D}_2 R^{-1}_{p(\tau)}(\delta_x(\tau))\left[\dot{\delta}_x(\tau)\right]\bigg)$$

$$+ t\bigg(\mathrm{D}_1 R^{-1}_{p(\tau)}(\delta_y(\tau))\left[\dot{p}(\tau)\right] + \mathrm{D}_2 R^{-1}_{p(\tau)}(\delta_y(\tau))\left[\dot{\delta}_y(\tau)\right]\bigg),$$

$$\dot{\delta}_x(\tau) = \mathrm{D}_2 R_{x_1}\left(\tau R^{-1}_{x_1}(x_2)\right)\left[R^{-1}_{x_1}(x_2)\right],$$

$$\dot{\delta}_y(\tau) = \mathrm{D}_2 R_{y_1}\left(\tau R^{-1}_{y_1}(y_2)\right)\left[R^{-1}_{y_1}(y_2)\right].$$

Using Proposition 3.12 and Proposition 3.14 we establish the following bounds:

$$\left\|\dot{\delta}_x(\tau)\right\|_{\delta(\tau)} \leq L_2\left\|R^{-1}_{x_1}(x_2)\right\| \leq L_2 M_2 d(x_1, x_2),$$

$$\left\|\dot{\delta}_y(\tau)\right\|_{\delta(\tau)} \leq L_2\left\|R^{-1}_{y_1}(y_2)\right\| \leq L_2 M_2 d(y_1, y_2),$$

$$\left\|\dot{p}(\tau)\right\|_{p(\tau)} \leq (L_1 + L_2 M_1)\left\|\dot{\delta}_x(\tau)\right\|_{\delta_x(\tau)} + L_2 M_2\left\|\dot{\delta}_y(\tau)\right\|_{\delta_y(\tau)},$$

$$\left\|\dot{\xi}(\tau)\right\|_{p(\tau)} \leq 2M_1\left\|\dot{p}(\tau)\right\|_{p(\tau)} + M_2\left\|\dot{\delta}_x(\tau)\right\|_{\delta_x(\tau)} + M_2\left\|\dot{\delta}_y(\tau)\right\|_{\delta_y(\tau)},$$

$$\left\|\dot{\delta}(\tau)\right\|_{\delta(\tau)} \leq L_1\left\|\dot{p}(\tau)\right\|_{p(\tau)} + L_2\left\|\dot{\xi}(\tau)\right\|_{p(\tau)}. \tag{3.11}$$

By suitably plugging the previous inequalities into the right-hand side of (3.11), we find

$$\left\|\dot{\delta}(\tau)\right\|_{\delta(\tau)} \leq L_x d(x_1, x_2) + L_y d(y_1, y_2)$$

where L_x and L_y are polynomials of L_1, L_2, M_1 and M_2. Then, we can take $L_{xy} = \max\{L_x, L_y\}$. $\qquad\square$

Corollary 3.16. *Let $\mathcal{U} \subset \mathcal{M}$ be any retraction-convex set as in Lemma 3.15. For any $x_1, x_2, y_1, y_2 \in \mathcal{U}$ and for every $r_1, r_2, t_1, t_2 \in [0, 1]$, we have*

$$d\left(c_{r_1}(t_1; x_1, y_1), c_{r_2}(t_2; x_2, y_2)\right) \leq (d(x_1, y_1) + d(x_2, y_2))\left(\frac{1}{2}L_t|t_1 - t_2| + \frac{1}{2}L_r|r_1 - r_2|\right)$$

$$+ (d(x_1, x_2) + d(y_1, y_2)) L_{xy}.$$

Proof. Using the triangular inequality and Lemma 3.15 we have

$$
\begin{aligned}
&d\left(c_{r_1}(t_1; x_1, y_1), c_{r_2}(t_2; x_2, y_2)\right)\\
\leq &d\left(c_{r_1}(t_1; x_1, y_1), c_{r_1}(t_2; x_1, y_1)\right) + d\left(c_{r_1}(t_2; x_1, y_1), c_{r_2}(t_2; x_2, y_2)\right)\\
\leq &L_t d(x_1, y_1)|t_1 - t_2| + d\left(c_{r_1}(t_2; x_1, y_1), c_{r_1}(t_2; x_2, y_2)\right)\\
&+ d\left(c_{r_1}(t_2; x_2, y_2), c_{r_2}(t_2; x_2, y_2)\right)\\
\leq &L_t d(x_1, y_1)|t_1 - t_2| + L_{xy}\left(d(x_1, x_2) + d(y_1, y_2)\right) + L_r d(x_2, y_2)|r_1 - r_2|. \quad (3.12)
\end{aligned}
$$

Exchanging (r_1, t_1, x_1, y_1) and (r_2, t_2, x_2, y_2) and repeating the same procedure leads to

$$
\begin{aligned}
d\left(c_{r_2}(t_2; x_2, y_2), c_{r_1}(t_1; x_1, y_1)\right) \leq &L_t d(x_2, y_2)|t_1 - t_2| + L_{xy}\left(d(x_1, x_2) + d(y_1, y_2)\right)\\
&+ L_r d(x_1, y_1)|r_1 - r_2|.
\end{aligned} \quad (3.13)
$$

Then averaging (3.12) and (3.13) proves the result. $\qquad\square$

3.5 Higher order differentials of inverse retractions

In the previous section, the smoothness of retractions unfolded into Lipschitz continuity results that proved to be useful in controlling the norm of the velocity field of retraction based curves. As a further tool of analysis, we here discuss high-order differentials of retractions, narrowing down the discussion to the inverse retraction. In fact, we aim at using the local inverse of a retraction R as a chart to locally linearize the manifold. As stated in Proposition 2.12, for every $x \in \mathcal{M}$ the local inverse of R can be used to smoothly map the neighborhood $\mathcal{I}_x = \{R_x(v) : \|v\|_x < \Delta(x)\}$ to the tangent space at x. Given a smooth manifold curve γ defined on the open interval $J \subset \mathbb{R}$ whose image is entirely contained in $\mathcal{I}_x$, for a given $x \in \mathcal{M}$, consider the tangent space curve

$$
\hat{\gamma}(\tau) := R_x^{-1}(\gamma(\tau)), \quad \forall\, \tau \in J. \quad (3.14)
$$

From the definition of inverse retraction's differential with respect to the second argument given by (3.5), we know that

$$
\hat{\gamma}'(t) = \mathrm{D}_2 R_x^{-1}(\gamma(t))\left[\dot{\gamma}(t)\right].
$$

In the following, we aim at providing a suitable theoretical foundation for the relation between high-order derivatives of $\hat{\gamma}$ and high-order covariant derivatives of the manifold curve γ. Then, we show how upper bounds on the covariant derivatives of the manifold curve translate into bounds for the derivatives of the tangent space curve.
The differential $\mathrm{D}_2 R_x^{-1}(y)$ is well-defined on the open subset of $\mathcal{M} \times \mathcal{M}$ introduced in Proposition 2.12 as

$$
E(\mathcal{D}) = \{(x, R_x(v)) : \|v\|_x < \Delta(x)\}.
$$

When defined, $\mathrm{D}_2 R_x^{-1}(y)$ is a linear map from $T_y\mathcal{M}$ to $T_{R_x^{-1}(y)}T_x\mathcal{M} \simeq T_x\mathcal{M}$. The

smoothness of the retraction implies that $D_2 R^{-1}$ is smooth as a map from a subset of $\mathcal{M} \times \mathcal{M}$ to a subset of the vector bundle

$$\mathcal{L}(T\mathcal{M}, T\mathcal{M}) = \{(x, y, L) : (x, y) \in \mathcal{M} \times \mathcal{M}, \, L : T_y\mathcal{M} \to T_x\mathcal{M} \text{ linear}\}$$

which admits a smooth manifold structure described in [Bou23, Proposition 10.60][2]. To define the covariant derivative of $D_2 R^{-1}$ we first construct a smooth tensor field on $\mathcal{M} \times \mathcal{M}$. Then, we leverage the definition of covariant differentiation of tensor fields given by Definition 1.26 to establish a definition for the covariant derivative of $D_2 R^{-1}$.

A tensor field on $\mathcal{M} \times \mathcal{M}$ involving the inverse retraction differential

First note that the tangent space of $\mathcal{M} \times \mathcal{M}$ at (x, y) can be identified with the product of tangent spaces $T_x\mathcal{M} \times T_y\mathcal{M}$ [Lee13, Proposition 3.14]. Hence, every smooth vector field $U \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$ can be decomposed for every $\forall (x, y) \in \mathcal{M} \times \mathcal{M}$ as

$$U(x, y) \simeq (U_1(x, y), U_2(x, y)),$$

with $U_1(x, y) \in T_x\mathcal{M}$ and $U_2(x, y) \in T_y\mathcal{M}$. We will denote the restriction to the first or the second component as $\pi_i(U) := U_i$, for $i = 1, 2$. Conversely, every smooth vector field $V_1, W_2 \in \mathfrak{X}(\mathcal{M})$ can be lifted to $V, W \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$ by appending the zero vector field, i.e. defining for every $(x, y) \in \mathcal{M} \times \mathcal{M}$

$$V(x, y) \simeq (V_1(x), 0), \quad W(x, y) \simeq (0, W_2(y)).$$

To indicate the lift operation we will use the notation $V = \ell_1(V_1)$ and $W = \ell_2(W_2)$. Finally, we indicate by z_i the concatenation of π_i followed by ℓ_i, that is $z_1(U) = (U_1, 0)$ and $z_2(V) = (0, V_2)$.

For any $U, V \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$, consider the scalar function on $\mathcal{M} \times \mathcal{M}$ given by

$$(x, y) \to T(U(x, y), V(x, y)) := \left\langle U_1(x, y), D_2 R_x^{-1}(y)[V_2(x, y)] \right\rangle_x. \tag{3.15}$$

This function is well-defined on the open subset $E(\Delta) \subset \mathcal{M} \times \mathcal{M}$ and by the smoothness of the retraction it is smooth where defined. Since the value of the function depends linearly on the pointwise value of U and V, the map

$$(U, V) \mapsto T(U, V)$$

is a smooth tensor field in the sense of Definition 1.10. If we endow $\mathcal{M} \times \mathcal{M}$ with the product Riemannian metric [Bou23, Example 3.57] given by

$$\left\langle (v_x, v_y), (w_x, w_y) \right\rangle_{(x, y)} := \left\langle v_x, v_y \right\rangle_x + \left\langle w_x, w_y \right\rangle_y, \tag{3.16}$$

---

[2]Although in this reference the linear map goes from $T_x\mathcal{M}$ to $T_y\mathcal{M}$.

for every $(v_x, v_y), (w_x, w_y) \in T_{(x,y)}(\mathcal{M} \times \mathcal{M})$, we may rewrite using the previously intro-
duced notations

$$
\begin{aligned}
T(U,V)\big|_{(x,y)} &= \big\langle z_1(U(x,y)), \ell_1\big(D_2 R_x^{-1}(y)[z_2(V(x,y))]\big)\big\rangle_{(x,y)}, \\
&=: \big\langle z_1(U), \ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big\rangle\big|_{(x,y)},
\end{aligned}
\tag{3.17}
$$

where the second line is to be understood as a shorthand notation for the first line.

3.5.1   Second-order covariant derivative of the inverse retraction

With the Riemannian metric (3.16), the product connection [Bou23, Equation (5.6)] on
$\mathcal{M} \times \mathcal{M}$ defined with the Riemannian connection on $\mathcal{M}$ coincides with the Riemannian
connection on $\mathcal{M} \times \mathcal{M}$ [Bou23, Exercise 5.13]. In turn, the Riemannian connection on
$\mathcal{M} \times \mathcal{M}$ determines the covariant differentiation of the tensor field T defined in (3.15)
as per Definition 1.26. The covariant derivative of T is a third-order tensor field defined
for every $U, V, W \in \mathfrak{X}(\mathcal{M})$ as

$$
\begin{aligned}
\nabla T(U,V,W) = W\big\langle z_1(U), \ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big\rangle \\
- \big\langle z_1(\nabla_W U), \ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big\rangle - \big\langle z_1(U), \ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big\rangle.
\end{aligned}
$$

Using the compatibility of covariant differentiation with the Riemannian metric we evince

$$
\begin{aligned}
\nabla T(U,V,W) &= \big\langle \nabla_W(z_1(U)), \ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big\rangle \\
&\quad + \big\langle z_1(U), \nabla_W\big(\ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big)\big\rangle \\
&\quad - \big\langle z_1(\nabla_W U), \ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big\rangle \\
&\quad - \big\langle z_1(U), \ell_1\big(D_2 R^{-1}[z_2(\nabla_W V)]\big)\big\rangle.
\end{aligned}
\tag{3.18}
$$

The definition of the product connection, see [Bou23, Equation (5.6)], prescribes that
the first component of $\nabla_W U$ depends only on the first component of U. Hence, we have
that $\nabla_W(z_1(U)) = z_1(\nabla_W U)$ and $z_2(\nabla_W V) = \nabla_W(z_2(V))$, analogously. This means
that the first and the third term of (3.18) cancel out and we find

$$
\begin{aligned}
\nabla T(U,V,W) &= \big\langle z_1(U), \nabla_W\big(\ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big) - \ell_1\big(D_2 R^{-1}[\nabla_W(z_2(V))]\big)\big\rangle \\
&=: \big\langle z_1(U), \nabla_W D_2 R^{-1}[z_2(V)]\big\rangle,
\end{aligned}
\tag{3.19}
$$

where for any $V, W \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$ we have defined a vector field on $E(\mathcal{D}) \subset \mathcal{M} \times \mathcal{M}$
defined by

$$
\nabla_W D_2 R^{-1}[z_2(V)] := \nabla_W\big(\ell_1\big(D_2 R^{-1}[z_2(V)]\big)\big) - \ell_1\big(D_2 R^{-1}[\nabla_W(z_2(V))]\big).
$$

The value of $\nabla T(U,V,W)$ at $(x,y) \in \mathcal{M} \times \mathcal{M}$ depends linearly on the pointwise value
of the input vector fields. In the same way, $\nabla_W D_2 R^{-1}[z_2(V)]$ at (x,y) depends linearly
on the pointwise value of V and W. More specifically, it depends only on the pointwise

value of V_2. To stress this fact, we maintain z_2 in the abbreviated notation. It may depend on the value of W_1 but we will be mostly interested in the case where $W_1 = 0$. Indeed, we formulate the following definition.

Definition 3.17. *For every $(x, y) \in \mathcal{M} \times \mathcal{M}$ such that $R_x^{-1}(y)$ is well-defined, the second-order covariant differential with respect to the second argument of the inverse retraction is the multilinear map*

$$\mathrm{D}_2^2 R_x^{-1}(y) : T_y\mathcal{M} \times T_y\mathcal{M} \to T_x\mathcal{M}$$

defined for every $v, w \in T_y\mathcal{M}$ as

$$\mathrm{D}_2^2 R_x^{-1}(y)[v, w] := \pi_1(\nabla_W \mathrm{D}_2 R^{-1}[z_2(V)]\big|_{(x,y)}) \tag{3.20}$$

for any $V, W \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$ such that $\pi_2(V(x, y)) = v$ and $\pi_2(W(x, y)) = w$.

Note that the multilinear operator (3.20) is not necessarily symmetric in the sense that changing the order of v and w may affect the result. Going back to the overarching goal of this section, we are now in position to express the acceleration of the tangent space curve $\hat{\gamma}$.

Proposition 3.18. *If the manifold curve γ is twice differentiable, the acceleration of the tangent space curve (3.14) is given by*

$$\hat{\gamma}''(\tau) = \mathrm{D}_2 R_x^{-1}(\gamma(\tau))\left[\ddot{\gamma}(\tau)\right] + \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau)], \quad \forall \tau \in J.$$

Proof. Fix $\tau \in J$ and consider any arbitrary $u \in T_x\mathcal{M}$. We have that

$$\left\langle u, \hat{\gamma}''(\tau)\right\rangle_x = \frac{\mathrm{d}}{\mathrm{d}s}\left\langle u, \hat{\gamma}'(s)\right\rangle_x\big|_{s=\tau} = \frac{\mathrm{d}}{\mathrm{d}s}\left\langle u, \mathrm{D}R_x^{-1}(\gamma(s))[\gamma'(s)]\right\rangle_x\big|_{s=\tau}.$$

This can be interpreted as evaluating the directional derivative of the tensor field (3.17) along the curve $c(s) := (x, \gamma(s)) \in \mathcal{M} \times \mathcal{M}$ with input vector fields U, V along the curve c defined for every $s \in J$ by $U(s) = (u, 0)$ and $V(s) = (0, \dot{\gamma}(s))$. Then

$$\left\langle u, \hat{\gamma}'(s)\right\rangle = \left\langle z_1(U(s)), \ell_1(\mathrm{D}_2 R^{-1}[z_2(V(s))])\right\rangle = T(U(s), V(s)).$$

Using the consequence of chain rule property of induced covariant differentiation for tensor field reported in (1.3) we get

$$\frac{\mathrm{d}}{\mathrm{d}s}T(U(s), V(s)) = \nabla T(U(s), V(s), \dot{c}(s)) + T\left(\frac{\mathrm{D}}{\mathrm{d}s}U(s), V(s)\right) + T\left(U(s), \frac{\mathrm{D}}{\mathrm{d}s}V(s)\right).$$

For every $s \in J$, the vector field U is constant, therefore $\frac{\mathrm{D}}{\mathrm{d}s}U(s) = 0$ and the second term vanishes. From the definition of the product connection and the expression of the

velocity vector field of c given by $\dot{c}(s) = (0, \dot{\gamma}(s))$, we evince $\frac{D}{ds}V(s) = (0, \ddot{\gamma}(s))$. Hence

$$\frac{d}{ds}T(U(s), V(s))\big|_{s=\tau} = \big\langle z_1(u, 0), \nabla_{(0, \dot{\gamma}(\tau))}D_2R^{-1}[z_2(0, \dot{\gamma}(\tau))]\big\rangle_{(x,y)} + T\big((u, 0), (0, \ddot{\gamma}(\tau))\big).$$

In the first term we recognize the second-order differential of the inverse retraction introduced by Definition 3.17 and we conclude

$$\begin{aligned}\big\langle u, \hat{\gamma}''(\tau)\big\rangle_x &= \big\langle u, \pi_1(\nabla_{(0, \dot{\gamma}(\tau))}D_2R^{-1}[(0, \dot{\gamma}(\tau))])\big\rangle_x + \big\langle u, D_2R^{-1}[\ddot{\gamma}(\tau)]\big\rangle_x \\ &= \big\langle u, D_2^2R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau)] + D_2R_x^{-1}(\gamma(\tau))\,[\ddot{\gamma}(\tau)]\big\rangle_x.\end{aligned}$$

$\square$

3.5.2 Third and fourth-order covariant derivative of the inverse retraction

The procedure which leads to the second-covariant derivative of the inverse retraction can be repeated to define the third and fourth-order derivatives, as we now briefly outline. We apply Definition 1.26 to compute the covariant derivative of the third-order tensor field ∇T given in (3.19) and define the resulting fourth-order tensor field as

$$\begin{aligned}\nabla^2 T(U, V, W, Z) = {}&Z\big\langle z_1(U), \nabla_W D_2R^{-1}[z_2(V)]\big\rangle - \big\langle z_1(\nabla_Z U), \nabla_W D_2R^{-1}[z_2(V)]\big\rangle \\ &- \big\langle z_1(U), \nabla_{\nabla_Z W}D_2R^{-1}[z_2(V)]\big\rangle - \big\langle z_1(U), \nabla_W D_2R^{-1}[z_2(\nabla_Z V)]\big\rangle,\end{aligned}$$

for every $U, V, W, Z \in \mathfrak{X}(\mathcal{M})$. As for the case of the second-order differential, the second term vanishes when developing the first term using the compatibility with the Riemannian metric of the connection. The resulting expression is used to define

$$\begin{aligned}\nabla_Z \nabla_W D_2R^{-1}[z_2(V)] := {}&\nabla_Z\big(\nabla_W D_2R^{-1}[z_2(V)]\big) - \big(\nabla_{\nabla_Z W}D_2R^{-1}[z_2(V)]\big) \\ &- \big(\nabla_W D_2R^{-1}[\nabla_W(z_2(V))]\big).\end{aligned}$$

Analogously differentiating $\nabla^2 T$ leads to defining

$$\begin{aligned}\nabla_Q \nabla_Z \nabla_W D_2R^{-1}[z_2(V)] := {}&\nabla_Q\big(\nabla_Z \nabla_W D_2R^{-1}[z_2(V)]\big) - \nabla_{\nabla_Q Z}\nabla_W D_2R^{-1}[z_2(V)] \\ &- \nabla_Z \nabla_{\nabla_Q W}D_2R^{-1}[z_2(V)] - \nabla_Z \nabla_W D_2R^{-1}[\nabla_Q(z_2(V))].\end{aligned}$$

With these multilinear map, we give the analogous of Definition 3.17 for the third and fourth-order covariant differential of the retraction.

Definition 3.19. *For every $(x, y) \in \mathcal{M} \times \mathcal{M}$ such that $R_x^{-1}(y)$ is well-defined, the third and fourth-order covariant differential with respect to the second argument of the inverse*

retraction are the multilinear maps

$$\mathrm{D}_2^3 R_x^{-1}(y) : (T_y\mathcal{M})^3 \to T_x\mathcal{M}$$
$$\mathrm{D}_2^4 R_x^{-1}(y) : (T_y\mathcal{M})^4 \to T_x\mathcal{M}$$

defined for every $v, w, z, q \in T_y\mathcal{M}$ *as*

$$\mathrm{D}_2^3 R_x^{-1}(y)[v, w, z] := \pi_1(\nabla_Z \nabla_W \mathrm{D}_2 R^{-1}[z_2(V)]\big|_{(x,y)})$$
$$\mathrm{D}_2^4 R_x^{-1}(y)[v, w, z, q] := \pi_1(\nabla_Q \nabla_Z \nabla_W \mathrm{D}_2 R^{-1}[z_2(V)]\big|_{(x,y)})$$

for any $V, W, Z, Q \in \mathfrak{X}(\mathcal{M}\times\mathcal{M})$ *such that* $\pi_2(V(x,y)) = u$, $\pi_2(W(x,y)) = w$, $\pi_2(Z(x,y)) = z$, $\pi_2(Q(x,y)) = q$.

Finally, analogously to Proposition 3.18, these multilinear operators are shown to intervene in the expression of the third and fourth derivative of the curve $\hat{\gamma}$.

Proposition 3.20. *If the manifold curve* γ *is three times continuously differentiable, the third derivative of the tangent space curve* (3.14) *is given by*

$$\hat{\gamma}'''(\tau) = \mathrm{D}_2 R_x^{-1}(\gamma(\tau))\,[\,\dddot{\gamma}(\tau)] + 2\mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\ddot{\gamma}(\tau), \dot{\gamma}(\tau)]$$
$$+ \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \ddot{\gamma}(\tau)] + \mathrm{D}_2^3 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau), \dot{\gamma}(\tau)], \quad \forall\,\tau \in J.$$

If γ *is four-times continuously differentiable, then* $\forall\,\tau \in J$ *it holds that*

$$\hat{\gamma}^{(4)}(\tau) = \mathrm{D}_2 R_x^{-1}(\gamma(\tau))\,[\,\ddddot{\gamma}(\tau)] + 3\mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\dddot{\gamma}(\tau), \ddot{\gamma}(\tau)]$$
$$+ 3\mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\,\dddot{\gamma}(\tau), \dot{\gamma}(\tau)] + 3\mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dddot{\gamma}(\tau)]$$
$$+ 3\mathrm{D}_2^3 R_x^{-1}(\gamma(\tau))[\ddot{\gamma}(\tau), \dot{\gamma}(\tau), \dot{\gamma}(\tau)] + 2\mathrm{D}_2^3 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \ddot{\gamma}(\tau), \dot{\gamma}(\tau)]$$
$$+ \mathrm{D}_2^3 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau), \ddot{\gamma}(\tau)] + \mathrm{D}_2^4 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau), \dot{\gamma}(\tau), \dot{\gamma}(\tau)].$$

Proof. We detail only the computations for the third derivative since the case of the fourth derivative is analogous.

In the proof of Proposition 3.18 we have established that for an arbitrary $u \in T_x\mathcal{M}$ and $\tau \in J$,

$$\langle u, \hat{\gamma}''(\tau)\rangle_x = \nabla T(U(\tau), V(\tau), W(\tau)) + T\Big(\frac{\mathrm{D}}{\mathrm{d}\tau}U(\tau), V(\tau)\Big) + T\Big(U(\tau), \frac{\mathrm{D}}{\mathrm{d}\tau}V(\tau)\Big),$$

where U, V and W are vector field along the curve $c(s) = (x, \gamma(s))$ defined for every $s \in J$ by $U(s) = (u, 0)$ and $V(s) = W(s) = (0, \dot{\gamma}(s)) = \dot{c}(s)$. We differentiate once again

by using (1.3) and we find

$$
\begin{aligned}
\left\langle u, \hat{\gamma}'''(\tau) \right\rangle_x = {}& \nabla^2 T(U(\tau), V(\tau), W(\tau), \dot{c}(\tau)) + \nabla T\Big(\frac{\mathrm{D}}{\mathrm{d}\tau} U(\tau), V(\tau), W(\tau) \Big) \\
& + \nabla T\Big(U(\tau), \frac{\mathrm{D}}{\mathrm{d}\tau} V(\tau), W(\tau) \Big) + \nabla T\Big(U(\tau), V(\tau), \frac{\mathrm{D}}{\mathrm{d}\tau} W(\tau) \Big) \\
& + \nabla T\Big(\frac{\mathrm{D}}{\mathrm{d}\tau} U(\tau), V(\tau), \dot{c}(\tau) \Big) + T\Big(\frac{\mathrm{D}^2}{\mathrm{d}\tau^2} U(\tau), V(\tau) \Big) + T\Big(\frac{\mathrm{D}}{\mathrm{d}\tau} U(\tau), \frac{\mathrm{D}}{\mathrm{d}\tau} V(\tau) \Big) \\[2mm]
& + \nabla T\Big(U(\tau), \frac{\mathrm{D}}{\mathrm{d}\tau} V(\tau), \dot{c}(\tau) \Big) + T\Big(\frac{\mathrm{D}}{\mathrm{d}\tau} U(\tau), \frac{\mathrm{D}}{\mathrm{d}\tau} V(\tau) \Big) + T\Big(U(\tau), \frac{\mathrm{D}^2}{\mathrm{d}\tau^2} V(\tau) \Big)
\end{aligned}
$$

From the fact U is constant along c we have that $\frac{\mathrm{D}}{\mathrm{d}\tau} U(\tau) = \frac{\mathrm{D}^2}{\mathrm{d}\tau^2} U(\tau) = 0$. Since $V(s) = W(s) = \dot{c}(s) = (0, \dot{\gamma}(s))$ we evince $\frac{\mathrm{D}}{\mathrm{d}\tau} V(\tau) = \frac{\mathrm{D}}{\mathrm{d}\tau} W(\tau) = (0, \ddot{\gamma}(\tau))$ and $\frac{\mathrm{D}^2}{\mathrm{d}\tau^2} V(\tau) = (0, \dddot{\gamma}(\tau))$. By linearity of tensors, this implies all terms involving these quantities vanish. Finally, replacing the tensor fields with their definition, we recognize the expressions of the second and third-order covariant differentials of the retraction given by Definition 3.17 and 3.19

$$
\begin{aligned}
\left\langle u, \hat{\gamma}'''(\tau) \right\rangle_x = {}& \left\langle u, \mathrm{D}_2^3 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau), \dot{\gamma}(\tau)] \right\rangle_x + \left\langle u, \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\ddot{\gamma}(\tau), \dot{\gamma}(\tau)] \right\rangle_x \\
& + \left\langle u, \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \ddot{\gamma}(\tau)] \right\rangle_x + \left\langle u, \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\ddot{\gamma}(\tau), \dot{\gamma}(\tau)] \right\rangle_x \\
& + \left\langle u, \mathrm{D}_2 R_x^{-1}(\gamma(\tau))[\dddot{\gamma}(\tau), \dot{\gamma}(\tau)] \right\rangle_x .
\end{aligned}
$$

Reordering and collecting terms proves the result. $\qquad\square$

3.5.3 Operator norms for the high-order differentials of the retraction

The second, third and fourth-order differentials of an inverse retraction introduced in Definition 3.17 and Definition 3.19 are multilinear operators on a product of vector spaces. For any linear operator $A : (T_y\mathcal{M})^k \to T_x\mathcal{M}$ we may define its operator norm as follows

$$
\|A\| := \sup_{\substack{w_1,\dots,w_k \in T_y\mathcal{M} \\ w_1,\dots,w_k \neq 0_x}} \left\{ \frac{\|A[w_1,\dots,w_k]\|_x}{\|w_1\|_y \cdots \|w_k\|_y} \right\}. \tag{3.21}
$$

Since $T_y\mathcal{M}$ and $T_x\mathcal{M}$ have finite dimension, the norm for any such A is always finite [Lim21, p. 610]. Hence, the multilinear operators $\mathrm{D}_2^2 R_x^{-1}(y)$, $\mathrm{D}_2^3 R_x^{-1}(y)$ and $\mathrm{D}_2^4 R_x^{-1}(y)$ have a finite operator norm for every $(x, y) \in \mathcal{M} \times \mathcal{M}$ for which they are defined. By the smoothness of retractions, these operator norms as a function of (x, y) are smooth. Therefore, in analogy with Proposition 3.12 concerning the first-order differentials of the retraction, in the following we show that the operator norms can be uniformly upper-bounded upon varying x in a compact set and y on each $\mathcal{I}_x^\varepsilon$.

Proposition 3.21. *Consider an arbitrary $\varepsilon \in (0, 1)$ and a retraction R. For any compact*

subset $K \subset \mathcal{M}$ *there exist positive constants* $M_{2,2}(K, \varepsilon)$, $M_{2,3}(K, \varepsilon)$ *and* $M_{2,4}(K, \varepsilon)$ *such that for every* $x \in K$ *and* $y \in \mathcal{I}_x^\varepsilon$ *and for any* $t, u, v, w \in T_y \mathcal{M}$, *the high-order covariant differentials of the inverse retraction satisfy*

(i) $\left\| \mathrm{D}_2^2 R_x^{-1}(y) [t, u] \right\|_x \leq M_{2,2}(K, \varepsilon) \left\| t \right\|_y \left\| u \right\|_y$,

(ii) $\left\| \mathrm{D}_2^3 R_x^{-1}(y) [t, u, v] \right\|_x \leq M_{2,3}(K, \varepsilon) \left\| t \right\|_y \left\| u \right\|_y \left\| v \right\|_y$,

(iii) $\left\| \mathrm{D}_2^4 R_x^{-1}(y) [t, u, v, t] \right\|_x \leq M_{2,4}(K, \varepsilon) \left\| t \right\|_y \left\| u \right\|_y \left\| v \right\|_y \left\| w \right\|_y$.

Proof. The result is reached analogously to the result of Proposition 3.12. Take any $k \in \{2, 3, 4\}$ and denote $\|\mathrm{D}_2^k R_x^{-1}(y)\|$ the operator norm (3.21) of the multilinear operator. On the one hand, we know the function $(x, y) \mapsto \|\mathrm{D}_2^k R_x^{-1}(y)\|$ is smooth where it is defined. On the other hand, by Lemma 3.13, the set $\overline{\mathcal{I}_K^\varepsilon} = \{(x, R_x(v)) : x \in K, \|v\|_x \leq \varepsilon \Delta(x)\}$ is compact in $\mathcal{M} \times \mathcal{M}$. Hence the constant $M_{2,k}(K, \varepsilon)$ is found by taking the maximum of $\|\mathrm{D}_2^k R_x^{-1}(y)\|$ for every $(x, y) \in \overline{\mathcal{I}_K^\varepsilon}$. $\qquad \square$

3.6 Approximating power of retraction curves

In Corollary 3.2, it is shown that retractions can be used to generate manifold curves with prescribed initial position, velocity and, provided the retraction is second-order, also the initial acceleration. If the prescribed data is sampled from a smooth manifold curve γ at a parameter t, the resulting retraction curve approximates the curve γ in a neighborhood of t. Proposition 3.22 below quantifies the local discrepancy between the original curve and the retraction curve in terms of the Riemannian distance function.

Proposition 3.22. *Suppose Assumption 3.11 holds and let R denote a retraction on a smooth manifold $\mathcal{M}$. For any manifold curve γ of class C^2 defined on an open interval $J \supset [0, 1]$ there exists $C_2 > 0$ and $\bar{h} > 0$ depending only on γ and R such that for all $h < \bar{h}_2$*

$$d\big(\gamma(t + h), R_{\gamma(t)}(h\dot{\gamma}(t))\big) \leq C_2 h^2, \quad \forall\, t \in [0, 1].$$

If furthermore R is a second-order retraction and γ is of class C^3, there exists $C_3 > 0$ and $\bar{h}_3 > 0$ depending only on γ and R such that for all $h < \bar{h}_3$

$$d\Big(\gamma(t + h), R_{\gamma(t)}\big(h\dot{\gamma}(t) + \frac{h^2}{2}\ddot{\gamma}(t)\big)\Big) \leq C_3 h^3, \quad \forall\, t \in [0, 1].$$

We first state a technical result necessary for the proof. It is meant to quantify in term of Riemannian distance function the size of the region $\mathcal{I}_x^{1/3}$ where the inverse retraction at x is a diffeomorphism and admits a Lipschitz constant that was made explicit in Proposition 3.12. For every $x \in \mathcal{M}$, we introduce the constant

$$\bar{\nu}(x) = \sup \left\{ \nu > 0 : B(x, \nu) \subset \mathcal{I}_x^{1/3} \right\}.$$

The set $\mathcal{I}_x^{1/3}$ is open in $\mathcal{M}$, hence $\bar{\nu}(x) > 0$ at every x. Provided Assumption 3.11 holds, i.e. that the retraction-convexity radius $\bar{\rho}$ is lower bounded by a strictly positive constants on any compact set, the same can be said for the quantity $\bar{\nu}$.

Lemma 3.23. *For any compact set $K \subset \mathcal{M}$ such that $\rho_{\min}(K) = \inf_{x \in K} \bar{\rho}(x) > 0$ it holds that $\nu_{\min}(K) := \inf_{x \in K} \bar{\nu}(x) > 0$.*

Proof. Let us denote $\Delta_{\min}(K) := \inf_{x \in K} \Delta(x)$, where Δ is the strictly positive and continuous function used in Proposition 2.12 to describe the domain over which a retraction is a diffeomorphism. Since Δ is known to be continuous and strictly positive, we know $\Delta_{\min}(K) > 0$ for every compact subset K. We indicate by $L_R(K)$ and $M_R(K)$ the Lipschitz constants of the retraction and its inverse on K, as given by Lemma 3.14.

We need to show that there exists $\nu_0 > 0$ such that for every $x \in K$ and every $y \in B(x, \nu_0)$ we have $y \in \mathcal{I}_x^{1/3}$, which by definition means $R_x^{-1}(y)$ is well-defined and that its norm is bounded by $\Delta(x)/3$. If $\nu_0 < \rho_{\min}(K)$ then $R_x^{-1}(y)$ is always well-defined by retraction-convexity. Hence, we proceed by contradiction assuming the second condition fails: for every integer n large enough so that $\frac{1}{n} < \rho_{\min}(K)$, there exists $x_n \in K$ and $y_n \in B(x_n, \frac{1}{n})$ such that $\|R_{x_n}^{-1}(y_n)\|_{x_n} > \frac{\Delta(x_n)}{3}$. By retraction-convexity of $B(x_n, \frac{1}{n})$, we know that

$$z_n = R_{x_n}\left(\frac{\Delta_{\min}(K)}{6} \frac{R_{x_n}^{-1}(y_n)}{\|R_{x_n}^{-1}(y_n)\|_{x_n}}\right)$$

is well-defined and belongs to $B(x_n, \frac{1}{n})$. By construction $z_n \in \mathcal{I}_{x_n}^{1/3}$ and $x_n \in K$, hence by Lipschitz continuity of the retraction

$$\frac{\Delta_{\min}(K)}{6} = \|R_{x_n}^{-1}(z_n)\|_{x_n} \leq M_R(K)d(x_n, z_n) < \frac{M_R(K)}{n}.$$

Therefore $n\frac{\Delta_{\min}(K)}{6} < M_R(K)$ for every $n > \frac{1}{\rho_{\min}(K)}$, contradicting the finiteness of $M_R(K)$. $\square$

Proof of Proposition 3.22. Denote $\rho_{\min}(\gamma)$ the infimum of the retraction-convexity radius function on the image of the curve γ on the interval $[0, 1]$. By Assumption 3.11, $\rho_{\min}(\gamma)$ is strictly positive and by Lemma 3.23 also the infimum of $\bar{\nu}$ on the curve γ, denoted $\nu_{\min}(\gamma)$. For any fixed $t \in [0, 1]$, we know that for any $\rho < \min\{\rho_{\min}, \nu_{\min}\} =: \rho^*$, the metric ball $B(\gamma(t), \rho)$ is retraction-convex and is contained in $\mathcal{I}_{\gamma(t)}^{1/3}$. This enables using for any choice of $t \in [0, 1]$ the same Lipschitz constants of the retraction $L_2(\gamma, 1/3)$ and $M_2(\gamma, 1/3)$ as defined in Proposition 3.12 and the operator norms of the inverse retraction differentials $M_{2,2}(\gamma, 1/3)$ and $M_{2,3}(\gamma, 1/3)$ as defined in Proposition 3.21, where the compact set in consideration is the image of the curve γ on the interval $[0, 1]$.

By smoothness of the curve and the Lipschitz continuity of the retraction, there exists $\bar{h} > 0$ so that $\gamma(t + h)$, $\sigma_1(h) := R_{\gamma(t)}(h\dot{\gamma}(t))$ and $\sigma_2(h) := R_{\gamma(t)}(h\dot{\gamma}(t) + \frac{1}{2}h^2\ddot{\gamma}(t))$ all belong to $B(\gamma(t), \rho^*)$ for all $h \in (-\bar{h}, \bar{h})$ and for all $t \in [0, 1]$, see Figure 3.7. By retraction-convexity of $B(\gamma(t), \rho^*)$, we can define the retraction curves

$$e_i(\tau) = R_{\gamma(t)}(\hat{e}_i(\tau)), \quad \tau \in [0, 1],$$

with $\hat{e}_i(\tau) = (1 - \tau)R_{\gamma(t)}^{-1}(\gamma(t + h)) + \tau R_{\gamma(t)}^{-1}(\sigma_i(h))$. By construction, both curves join $\gamma(t + h)$ and $\sigma_i(h)$, for $i = 1, 2$ for every $|h| < \bar{h}$. The Riemannian distance between the
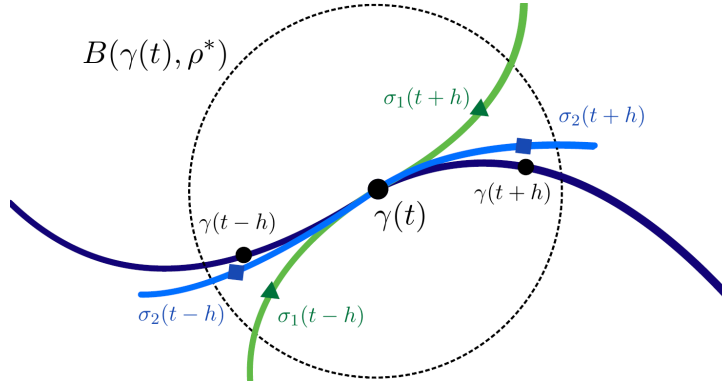
Figure 3.7: Consider h small enough so that $\gamma(t+h)$, $\sigma_1(h) := R_{\gamma(t)}(h\dot{\gamma}(t))$ and $\sigma_2(h) := R_{\gamma(t)}(h\dot{\gamma}(t) + \frac{1}{2}h^2\ddot{\gamma}(t))$ all belong to $B(\gamma(t), \rho^*)$ for every $h \in (-\bar{h}, \bar{h})$ and for all $t \in [0, 1]$.

endpoints is upper-bounded by the length of the curve e_i, which can itself by bounded by the maximum derivative norm:

$$d(\gamma(t+h), \sigma_i(h)) \leq L(e_i) = \int_0^1 \|\dot{e}_i(\tau)\|_{e_i(\tau)} \, \mathrm{d}\tau \leq \max_{\tau \in [0,1]} \|\dot{e}_i(\tau)\|_{e_i(\tau)}.$$

But since $\hat{e}_i(\tau) \in \mathcal{D}_{\gamma(t)}^{1/3} = R_{\gamma(t)}^{-1}\big(\mathcal{I}_{\gamma(t)}^{1/3}\big)$ for every $\tau \in [0, 1]$ and $|h| < \bar{h}$, we can bound

$$\|\dot{e}_i(\tau)\|_{e_i(\tau)} = \left\|\mathrm{D}R_{\gamma(t)}(\hat{e}_i(\tau))[R_{\gamma(t)}^{-1}(\sigma_i(h)) - R_{\gamma(t)}^{-1}(\gamma(t+h))]\right\|_{\sigma_i(\tau)}$$

$$\leq L_2(\gamma, 1/3)\left\|R_{\gamma(t)}^{-1}(\sigma_i(h)) - R_{\gamma(t)}^{-1}(\gamma(t+h))\right\|_{\gamma(t)},$$

where $L_2(\gamma, 1/3)$ is the Lipschitz constant introduced in Proposition 3.12. By Corollary 3.3 and smoothness of the inverse retraction, the tangent space curve $\hat{\gamma}(h) := R_{\gamma(t)}^{-1}(\gamma(t+h))$ admits, depending on the smoothness of γ, the following Taylor expansions around $h = 0$

$$\hat{\gamma}(h) = \begin{cases} h\dot{\gamma}(t) + \frac{h^2}{2}\hat{\gamma}''(s_1) & \text{if } \gamma \in C^2, \\ h\dot{\gamma}(t) + \frac{h^2}{2}\ddot{\gamma}(t) + \frac{h^3}{6}\hat{\gamma}'''(s_2) & \text{if } \gamma \in C^3, \end{cases}$$

for some s_1, s_2 between 0 and h. Note that the second expansion holds only if R is a second-order retraction. Then, by definition of σ_1 and σ_2 we evince

$$\left\|R_{\gamma(t)}^{-1}(\sigma_1(h)) - R_{\gamma(t)}^{-1}(\gamma(t+h))\right\|_{\gamma(t)} = \frac{h^2}{2}\left\|\hat{\gamma}''(s_1)\right\|_{\gamma(t)}$$

$$\left\|R_{\gamma(t)}^{-1}(\sigma_2(h)) - R_{\gamma(t)}^{-1}(\gamma(t+h))\right\|_{\gamma(t)} = \frac{h^3}{6}\left\|\hat{\gamma}'''(s_2)\right\|_{\gamma(t)}$$

It remains to bound the second and third derivatives of $\hat{\gamma}$ uniformly in both $s \in (-\bar{h}, \bar{h})$ and $t \in [0, 1]$. From Proposition 3.18, we know that

$$\hat{\gamma}''(\tau) = \mathrm{D}_2 R_x^{-1}(\gamma(\tau))\left[\ddot{\gamma}(\tau)\right] + \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau)]$$

Using the triangular inequality and introducing the operator norm upper bounds $M_2(\gamma, 1/3)$ and $M_{2,2}(\gamma, 1/3)$ defined in Proposition 3.12 and Proposition 3.21 respectively, we find

$$\left\| \gamma''(s) \right\|_{\gamma(t)} \leq M_2(\gamma, 1/3) \left\| \ddot{\gamma} \right\|_\infty + M_{2,2}(\gamma, 1/3) \left\| \dot{\gamma} \right\|_\infty^2 , \qquad (3.22)$$

where $\dot{\gamma}(t) := \max_{\tau \in [0,1]} \left\| \dot{\gamma}(\tau) \right\|_{\gamma(\tau)}$ and $\ddot{\gamma}(t) := \max_{\tau \in [0,1]} \left\| \ddot{\gamma}(\tau) \right\|_{\gamma(\tau)}$. The bound (3.22) holds uniformly in s and t and concludes the first part of the proof with $\bar{h}_2 = \bar{h}$ and

$$C_2 = \frac{L_2(\gamma, 1/3)}{2} \Big(M_2(\gamma, 1/3) \left\| \ddot{\gamma} \right\|_\infty + M_{2,2}(\gamma, 1/3) \left\| \dot{\gamma} \right\|_\infty^2 \Big).$$

Consider now γ three times differentiable. From Proposition 3.18 we know that

$$\begin{aligned} \hat{\gamma}'''(\tau) = {}& \mathrm{D}_2 R_x^{-1}(\gamma(\tau)) \left[\dddot{\gamma}(\tau) \right] + \mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\ddot{\gamma}(\tau), \dot{\gamma}(\tau)] \\ & + 2\mathrm{D}_2^2 R_x^{-1}(\gamma(\tau))[\ddot{\gamma}(\tau), \dot{\gamma}(\tau)] + \mathrm{D}_2^3 R_x^{-1}(\gamma(\tau))[\dot{\gamma}(\tau), \dot{\gamma}(\tau), \dot{\gamma}(\tau)]. \end{aligned}$$

Once again, using triangular inequality and introducing $M_{2,3}(\gamma, \varepsilon)$ defined in Proposition 3.21 yields

$$\left\| \hat{\gamma}'''(s) \right\|_{\gamma(t)} \leq M_2(\gamma, 1/3) \left\| \dddot{\gamma} \right\|_\infty + 3M_{2,2}(\gamma, 1/3) \left\| \dot{\gamma} \right\|_\infty \left\| \ddot{\gamma} \right\|_\infty + M_{2,3}(\gamma, 1/3) \left\| \dot{\gamma} \right\|_\infty^3 ,$$

where $\dddot{\gamma}(t) := \max_{\tau \in [0,1]} \left\| \dddot{\gamma}(\tau) \right\|_{\gamma(\tau)}$. This concludes the proof as we have shown that $\bar{h}_3 = \bar{h}$ and

$$C_3 = \frac{L_2(\gamma, 1/3)}{6} \Big(M_2(\gamma, 1/3) \left\| \dddot{\gamma} \right\|_\infty + 3M_{2,2}(\gamma, 1/3) \left\| \dot{\gamma} \right\|_\infty \left\| \ddot{\gamma} \right\|_\infty + M_{2,3}(\gamma, 1/3) \left\| \dot{\gamma} \right\|_\infty^3 \Big).$$

$\square$

3.7 Conclusions

The results of this chapter support the idea that retractions can be used as a general tool to build portions of manifold curves of which the properties can be conveniently analyzed. In addition to previously known retraction curves with prescribed velocity and acceleration in one point, we present the novel class of r-endpoint retraction curves for which the endpoints can be prescribed. Their well-posedness is granted by the existence of retraction-convex sets. This new theoretical tool together with the Lipschitz continuity properties of the retraction and the higher-order differentials of the inverse retraction already demonstrated to be effective in analyzing retraction-based procedures, as in the proof of Proposition 3.22 concerning the local approximating power of retraction curves.

4 Riemannian Continuation

Homotopy theory is the study of continuous deformations between smooth maps and geometric structures in general. This field blossomed into various branches one of which is numerical continuation, homotopy continuation, or simply continuation. It is a well-established set of algorithmic techniques to approximate the solutions to a set of nonlinear equations. The paradigm of numerical continuation is based on tracking a continuous deformation between the known solution set of a given equation and the sought solution set of the equation of interest. This leading idea diversified into numerous algorithms, of which an overview can be found in the reference textbooks [AG90, Deu11]. Numerical continuation has also been extended to unconstrained and constrained optimization in Euclidean spaces. This chapter introduces a generalization of numerical continuation methods for its application to Riemannian optimization. The exposition is based on the content of [SK22a].

4.1 Continuation and optimization

Let us first recall the setting of numerical continuation for nonlinear equations. Given a nonlinear equation

$$F(x) = 0, \tag{4.1}$$

for a smooth function $F : \mathbb{R}^d \to \mathbb{R}^d$, numerical continuation [AG90, Deu11] is used to track solutions of (4.1) when the problem is smoothly perturbed. This can be useful for, e.g., ensuring global convergence of the Newton method by progressively transforming a simple problem with a known solution into (4.1). More specifically, one considers a parametrized family of equations,

$$G(x, \lambda) = 0, \quad \forall \lambda \in [0, 1], \tag{4.2}$$

with $G : \mathbb{R}^d \times [0, 1] \to \mathbb{R}^d$ such that $G(x, 1) = F(x)$ holds and a solution $x_0 \in \mathbb{R}^d$ of $G(x_0, 0) = 0$ can be easily determined. The function G is also known as a *homotopy*. Under suitable assumptions, the solution set

$$G^{-1}(0) = \left\{ (x, \lambda) \in \mathbb{R}^d \times [0, 1] : G(x, \lambda) = 0 \right\}$$
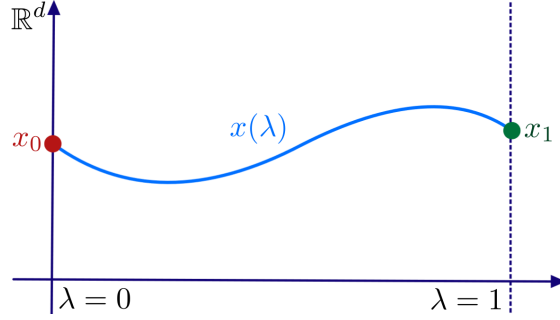
Figure 4.1: Idealized setting of Euclidean continuation: $G^{-1}(0)$ contains only one well-behaved curve.

to the parametric problem (4.2) contains a smooth $x(\lambda)$, $\lambda \in [0,1]$, connecting $x_1 = x(1)$, the solution to the original problem, to $x_0 = x(0)$, see Figure 4.1. Homotopy methods are also relevant in optimization. Optimization methods for a given, smooth target function $f : \mathbb{R}^n \to \mathbb{R}$ often aim at retrieving critical points, that is, solutions to

$$F(x) = \nabla f(x) = 0.$$

Homotopy methods can be useful for, e.g., ensuring global convergence (to a critical point) by tracking critical points of a parametrized optimization problem, which amounts to considering

$$G(x, \lambda) = \nabla f(x, \lambda) = 0, \quad \forall \lambda \in [0,1]. \tag{4.3}$$

This approach to optimization problems has been widely studied in the literature, both for unconstrained and constrained optimization problems [KH84, GWZ84]. Among others, this has led to almost always globally convergent methods for non convex optimization [DO05] and nonlinear programming [GWZ84, Wat01]. Another use of homotopy methods is to improve the convergence behavior of a method by, e.g. defining a homotopy in which a regularization term is reduced progressively [LX15].

The Riemannian optimization counterpart of the homotopy (4.3) is

$$\mathrm{grad} f(x, \lambda) = 0, \quad \forall \lambda \in [0,1], \tag{4.4}$$

where $f : \mathcal{M} \times [0,1] \to \mathbb{R}$ is a parameter dependent smooth vector field on a Riemannian manifold $\mathcal{M}$ and $\mathrm{grad} f(x, \lambda)$ denotes the Riemannian gradient of $f(\cdot, \lambda)$ at x. Continuation methods for (4.4) need to ensure that x stays on $\mathcal{M}$. In this work, we use tools from Riemannian optimization to design path-following algorithms achieving this demand. A related question has been explored in the more restricted setting of time-varying convex optimization on Hadamard manifolds [MMN^{+}20], making use of the exponential map. In [Man12], a theoretical study of parameter-dependent Riemannian optimization is performed; the resulting homotopy-based algorithm involves local charts in order to utilize standard continuation algorithms on Euclidean spaces. In this work, we develop

continuation methods within the framework of Riemannian optimization as presented in [AMS08], which allows for the convenient design of efficient numerical methods in a general setting.

Contributions and outline of the chapter. After recalling in Section 4.2 the general structure of a path-following predictor-corrector continuation algorithm for nonlinear equations on Euclidean spaces, we introduce in Section 4.3 the setting of parametric Riemannian optimization and provide sufficient conditions for the numerical continuation problem to be well-posed. We then translate to the Riemannian setting the predictor-corrector algorithm to address them. We analyze the prediction phase, a key step of the algorithm and also propose a step size adaptation strategy. Finally, Sections 4.4 and 4.5 are dedicated to the application of the algorithm to two classical Riemannian optimization problems, respectively the computation of the Karcher mean and the low-rank matrix completion problem.

4.2 Euclidean predictor-corrector continuation

To motivate our Riemannian continuation algorithm, let us first recall the standard predictor-corrector continuation approach; see, e.g. [AG90, chapter 2].

Considering the parametric nonlinear equation (4.2), let us assume that 0 is a regular value of G, that is, the differential

$$\mathrm{D}G(x,\lambda) = [G_x(x,\lambda)|G_\lambda(x,\lambda)] \in \mathbb{R}^{d\times d+1},$$

where $G_x(x,\lambda) \in \mathbb{R}^{d\times d}$ and $G_\lambda(x,\lambda) \in \mathbb{R}^d$ denote respectively the differentials with respect to x and λ, has full rank for each $(x,\lambda) \in G^{-1}(0)$. Then the constant-rank level set theorem [Lee13, Theorem 5.12] asserts the set $G^{-1}(0)$ is an embedded submanifold of $\mathbb{R}^{d+1}$ of dimension 1 or, in other words, the union of disjoint curves. Under the stronger assumption that $G_x(x,\lambda) \in \mathbb{R}^{d\times d}$ has full rank, the implicit function theorem [KP13, Theorem 3.3.1] implies that it is possible to parametrize each solution curve as a function $x(\lambda)$. Moreover, its derivative is given by

$$x'(\lambda) = -G_x(x(\lambda),\lambda)^{-1}\left[G_\lambda(x(\lambda),\lambda)\right]. \tag{4.5}$$

In turn, the solution curve in (4.2) can be obtained from solving the following implicit ODE:

$$\begin{cases} G_x(x,\lambda)\left[x'\right] + G_\lambda(x,\lambda) = 0, & \forall\,\lambda \in [0,1] \\ x(0) = x_0. \end{cases} \tag{4.6}$$

This equation is sometimes called Davidenko equation [Dav53]. The path-following approach consists of numerically integrating (4.6) from time $\lambda = 0$ to $\lambda = 1$. The existence of the solution to (4.6) is discussed in [KP13, Theorem 4.2.1]; see also Theorem 4.1 below. Given an approximation $x_k \simeq x(\lambda_k)$ of the solution curve at point λ_k, a predictor-corrector continuation algorithm first performs a prediction step, which obtains a possibly
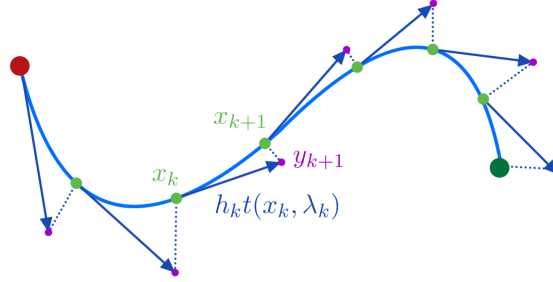
Figure 4.2: Euclidean Predictor-Corrector algorithm.

very rough estimate y_{k+1} of the solution curve at the next point λ_{k+1}. This is followed by a correction phase which aims at projecting this estimate back to the solution curve, see Figure 4.2.

The most common choices for the *prediction step* are:

$$classical\ prediction\ :\ y_{k+1} = x_k \qquad (4.7)$$

$$tangential\ prediction\ :\ y_{k+1} = x_k + (\lambda_{k+1} - \lambda_k)t(x_k, \lambda_k), \qquad (4.8)$$

where the tangent vector $t(x_k, \lambda_k) := x'(\lambda_k)$ is obtained from (4.5). This requires the solution of a linear system, a cost that is offset by increased prediction accuracy, see [Deu11, p.238-239] and Section 4.3.3. Note that (4.8) coincides with one step of the Euler method applied to (4.6).

In the *correction phase*, the refinement of the estimate y_{k+1} is performed by applying a nonlinear equation solver, typically a Newton-type method, on the equation $G(x, \lambda_{k+1}) = 0$ with initial guess y_{k+1}. A sufficiently small step size $\lambda_{k+1} - \lambda_k$ leads to a prediction that is accurate enough to yield (very) fast convergence. Various step size selection strategies have been developed in the literature, see [AG90, Deu11] and Section 4.3.4.

4.3   Continuation for Riemannian optimization

In this section, we consider a Riemannian optimization problem depending on a scalar parameter. The parameter can be intrinsic to the problem (e.g., time) or has been artificially added to form a homotopy. Examples of homotopies for Riemannian optimization problems are given in Sections 4.4 and 4.5.

4.3.1   Riemannian Davidenko equation

Let $\mathcal{M}$ denote a d-dimensional Riemannian manifold endowed with a Riemannian metric and its associated Riemannian connection. The parameter-dependent objective function

$$
\begin{aligned}
f\ :\ \mathcal{M} \times [0,1] &\to \mathbb{R} \\
(x, \lambda) &\mapsto f(x, \lambda)
\end{aligned}
\qquad (4.9)
$$

is assumed to be smooth in both arguments (at least of class C^2). For fixed $\lambda \in [0, 1]$, the Riemannian gradient and the Riemannian Hessian at x of the scalar field $f(\cdot, \lambda)$ are respectively denoted $\mathrm{grad} f(x, \lambda)$ and $\mathrm{Hess} f(x, \lambda)$ and their definitions are reported in Sections 2.1.1 and 2.1.2.

Consider the numerical continuation problem (4.4) of tracking critical points of the objective function as the parameter λ varies. Theorem 4.1 below is inspired by [KP13, Theorem 4.2.1] and gives sufficient conditions for the existence and parametrizability with respect to λ of a differentiable manifold curve joining a critical point $x_0 \in \mathcal{M}$ at $\lambda = 0$ and a critical point at $\lambda = 1$. For the purpose of the analysis, we assume that $\mathcal{M}$ is *complete*.

Theorem 4.1. *Let $\mathcal{M}$ be a complete Riemannian manifold, $\mathcal{U}$ be an open subset of $\mathcal{M}$ and V an open subset of $\mathcal{M} \times \mathbb{R}$ such that $\mathcal{U} \times [0, 1] \subset V$. Consider a scalar field $f \in C^2(V, \mathbb{R})$. Assume that there exist $x_0 \in \mathcal{U}$ such that $\mathrm{grad} f(x_0, 0) = 0$ and a constant $L > 0$ such that $B(x_0, L) \subseteq \mathcal{U}$. Moreover, suppose that for every $(z, \lambda) \in \mathcal{U} \times [0, 1]$ it holds that*

(i) $\mathrm{rank}\,(\mathrm{Hess} f(z, \lambda)) = d$,

(ii) $\left\| \mathrm{Hess} f(z, \lambda)^{-1} \left[\frac{\partial}{\partial \lambda} \mathrm{grad} f(z, \lambda) \right] \right\|_z < L$.

Then there exist an open interval $J \supset [0, 1]$ and a curve $x \in C^1(J, \mathcal{M})$ verifying

$$x(0) = x_0, \quad \mathrm{grad} f(x(\lambda), \lambda) = 0, \quad \forall \lambda \in [0, 1].$$

This curve satisfies the initial value problem

$$\begin{cases} \mathrm{Hess} f(x(\lambda), \lambda)\,[\dot{x}(\lambda)] + \frac{\partial\,\mathrm{grad} f(x(\lambda), \lambda)}{\partial \lambda} = 0, \quad \forall \lambda \in [0, 1], \\ x(0) = x_0. \end{cases} \tag{4.10}$$

As illustrated in Figure 4.3, Hypothesis (i) guarantees the parametrizability with respect to λ by ensuring the implicit ODE (4.10) is well-defined. For fixed λ, an analogous assumption is required for local quadratic convergence of the Riemannian Newton method [AMS08, Theorem 6.3.2]. Hypothesis (ii) ensures that the manifold curve can be parametrized up to $\lambda = 1$ as the limit point of the curve for $\lambda \to \lambda^*$, for any $0 < \lambda^* < 1$, is guaranteed to stay in the region $\mathcal{U}$ where the Riemannian Hessian is still of full rank. These hypotheses are global a priori assumptions that are difficult to verify in practice. Yet, for a large class of problems it is reasonable to assume the Riemannian Hessian is of full rank at the starting point $(x_0, 0)$, and therefore the solution curve is at least parametrizable on a possibly smaller interval $[0, \tau] \subseteq [0, 1]$. In the following, we call the initial value problem (4.10) the *Riemannian Davidenko equation*. Note that by Hypothesis (i), if x_0 is a local minimum, then the solution curve to the Riemannian Davidenko equation is a manifold curve of local minima. If we further assume the objective function to be geodesically convex [Bou23, Chapter 11] for each $\lambda \in [0, 1]$, this implies that the solution curve consists of global minima.

The proof of Theorem 4.1 subsumes the proof for the Euclidean case [KP13, Theorem
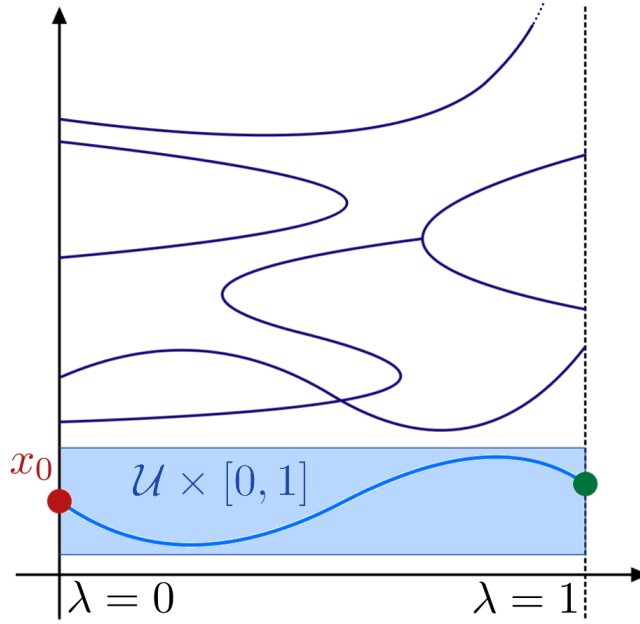
Figure 4.3: The shaded region is where the assumptions of Theorem 4.1 guarantee there exists a curve of critical points starting at x_0 which reaches $\lambda = 1$ without crossing other curves nor folding back.

4.2.1] by considering a pullback of the objective function to $\mathbb{R}^d$ through a local chart. The proof appearing in [SK22a] tacitly uses a property of the local chart representation at a critical point, as pointed out in [ZB22, Theorem 3.2]. Therefore, we first recall the definition and detail some technical results about the local coordinate representation of the objective function.

Preliminary results: local chart representation of the objective function

In the setup of Theorem 4.1, consider an open subset $\mathcal{N} \subset \mathcal{M}$ associated with a local chart $\varphi : \mathcal{N} \to \mathbb{R}^d$ such that $x_0 \in \mathcal{N}$ and $\mathcal{N} \times [0, 1] \subseteq V$. The inverse of the local chart is defined on $\hat{\mathcal{N}} := \varphi(\mathcal{N})$ and is denoted for convenience $\psi : \hat{\mathcal{N}} \to \mathcal{N}$. Every $x \in \mathcal{N}$ has local coordinates $\hat{x} = \varphi(x)$ associated with the chart φ. The other way around, we denote $x = \psi(\hat{x})$ for every $\hat{x} \in \hat{\mathcal{N}}$. The local coordinate representation of the parameter dependent objective function (4.9) through the local chart φ is defined as

$$\hat{f}(\hat{x}, \lambda) := f(\psi(\hat{x}), \lambda), \quad \forall\, \hat{x} \in \hat{\mathcal{N}},\, \lambda \in [0, 1]\,.$$

The following result specifies the link between the (Euclidean) gradient of $\hat{f}$ with respect to $\hat{x}$ and the Riemannian gradient of f with respect to x. This is a well-known result appearing for example in [Bou23, Exercise 10.73], where the local chart is a retraction.

Lemma 4.2. *The Euclidean gradient of $\hat{f}$ with respect to $\hat{x}$ is given by*

$$\nabla\hat{f}(\hat{x},\lambda) = \mathrm{D}\psi^*(\hat{x})\left[\mathrm{grad}f(\psi(\hat{x}),\lambda)\right], \quad \forall\,\hat{x}\in\hat{\mathcal{N}},\,\lambda\in[0,1]\,,$$

where $\mathrm{D}\psi^*(\hat{x}) : T_x\mathcal{M} \to T_{\hat{x}}\mathbb{R}^d \simeq \mathbb{R}^d$ denotes the adjoint of the inverse local chart's differential. The adjoint is taken with respect to the Riemannian metric on $T_x\mathcal{M}$ and the Euclidean metric on $\mathbb{R}^d$

Proof. First note that by chain rule we have $\mathrm{D}_{\hat{x}}\hat{f} = \mathrm{D}_x f \circ \mathrm{D}\psi$. Then, applying the definition of gradient first for $\hat{f}$ and then for f we find that for every $\hat{v}\in T_{\hat{x}}\mathbb{R}^d \simeq \mathbb{R}^d$

$$\begin{aligned}
\langle\nabla\hat{f}(\hat{x},\lambda),\hat{v}\rangle = \mathrm{D}_{\hat{x}}\hat{f}(\hat{x},\lambda)\left[\hat{v}\right] &= \mathrm{D}_x f(\psi(\hat{x}),\lambda)\left[\mathrm{D}\psi(\hat{x})[\hat{v}]\right] \\
&= \langle\mathrm{grad}f(\psi(\hat{x}),\lambda),\mathrm{D}\psi(\hat{x})[\hat{v}]\rangle_x \\
&= \langle\mathrm{D}\psi(\hat{x})^*[\mathrm{grad}f(\psi(\hat{x}),\lambda)],\hat{v}\rangle,
\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product on $\mathbb{R}^d$. $\qquad\square$

For convenience, let us denote $\nabla\hat{f}(\cdot,\lambda) =: \hat{F}(\cdot,\lambda)$ for all $\lambda\in[0,1]$. As we now show, the differential with respect to $\hat{x}$ of $\hat{F}$, in other words the Hessian of $\hat{f}$, has a simple relationship with the Riemannian Hessian of f when $\hat{F}(\hat{x},\lambda) = 0$, that is when $\mathrm{grad}f(x,\lambda) = 0$. When the Riemannian gradient is not zero, this is no longer the case and the differential of $\hat{F}$ with respect to $\hat{x}$ contains an extra term, see [Bou23, Exercise 10.73].

Lemma 4.3. *For every $(\hat{x},\lambda)\in\hat{\mathcal{N}}\times[0,1]$ such that $\mathrm{grad}f(\psi(\hat{x}),\lambda) = 0$, the differential of $\hat{F}$ with respect to $\hat{x}$ is given by*

$$\mathrm{D}_{\hat{x}}\hat{F}(\hat{x},\lambda) = \mathrm{D}\psi(\hat{x})^* \circ \mathrm{Hess}f(\psi(\hat{x}),\lambda) \circ \mathrm{D}\psi(\hat{x}).$$

Proof. From the definition of differential, see Definition 1.8, for any $\hat{v}\in\mathbb{R}^d$ we have that

$$\mathrm{D}_{\hat{x}}\hat{F}(\hat{x},\lambda)[\hat{v}] = \frac{\mathrm{d}}{\mathrm{d}t}\hat{F}(\hat{\gamma}_{\hat{x},\hat{v}}(t),\lambda)\big|_{t=0},$$

for any smooth curve $\hat{\gamma}_{\hat{x},\hat{v}}$ in $\mathbb{R}^d$ such that $\hat{\gamma}_{\hat{x},\hat{v}}(0) = \hat{x}$ and $\hat{\gamma}'_{\hat{x},\hat{v}}(0) = \hat{v}$. By computing the inner-product with any $\hat{w}\in\mathbb{R}^d$, we evince

$$\langle\mathrm{D}_{\hat{x}}\hat{F}(\hat{x},\lambda)[\hat{v}],\hat{w}\rangle = \langle\frac{\mathrm{d}}{\mathrm{d}t}\hat{F}(\hat{\gamma}_{\hat{x},\hat{v}}(t),\lambda)\big|_{t=0},\hat{w}\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{F}(\hat{\gamma}_{\hat{x},\hat{v}}(t),\lambda),\hat{w}\rangle\big|_{t=0}.$$

In the proof of Lemma 4.2 we have established

$$\langle\hat{F}(\hat{\gamma}_{\hat{x},\hat{v}}(t),\lambda),\hat{w}\rangle = \langle\mathrm{grad}f(\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t)),\lambda),\mathrm{D}\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t))[\hat{w}]\rangle_{\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t))}\,.$$

Hence, from the compatibility of induced covariant differentiation with the Riemannian

metric, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\langle \mathrm{grad} f(\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t))),\lambda), \mathrm{D}\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t))[\hat{w}]\rangle_{\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t))} \Big)\big|_{t=0} =$$
$$\langle \frac{\mathrm{D}}{\mathrm{d}t}\, \mathrm{grad} f(\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t)),\lambda)\big|_{t=0}, \mathrm{D}\psi(\hat{x})[\hat{w}]\rangle_x + \langle \mathrm{grad} f(\psi(\hat{x}),\lambda), \frac{\mathrm{D}}{\mathrm{d}t}\Big(\mathrm{D}\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t))[\hat{w}]\Big)\big|_{t=0}\rangle_x.$$

The vanishing Riemannian gradient assumption in $(\psi(x),\lambda)$ crucially guarantees the second term is zero. Finally, using the properties of the induced covariant differentiation concludes the proof:

$$\langle \mathrm{D}_{\hat{x}}\hat{F}(\hat{x},\lambda)[\hat{v}], \hat{w}\rangle = \langle \frac{\mathrm{D}}{\mathrm{d}t}\, \mathrm{grad} f(\psi(\hat{\gamma}_{\hat{x},\hat{v}}(t)),\lambda)\big|_{t=0}, \mathrm{D}\psi(\hat{x})[\hat{w}]\rangle_x$$
$$= \big\langle \nabla_{\mathrm{D}\psi(\hat{x})[\hat{w}]}\, \mathrm{grad} f(\psi(\hat{x}),\lambda), \mathrm{D}\psi(\hat{x})[\hat{w}]\big\rangle_x$$
$$= \langle \mathrm{Hess} f(\psi(\hat{x}),\lambda)[\mathrm{D}\psi(\hat{x})[\hat{v}]], \mathrm{D}\psi(\hat{x})[\hat{w}]\rangle_x$$
$$= \langle \mathrm{D}\psi(\hat{x})^* \circ \mathrm{Hess} f(\psi(\hat{x}),\lambda) \circ \mathrm{D}\psi(\hat{x})[\hat{v}], \hat{w}\rangle_x.$$

$\square$

Proof of Theorem 4.1

The following proof of Theorem 4.1 has been adapted from the one appearing in [SK22a] to explicitly use the results of Lemmas 4.2 and 4.3, which were not included in the published article.

Proof of Theorem 4.1. Since $\mathrm{grad} f(x_0, 0) = 0$, by Lemma 4.2 it holds that

$$\nabla \hat{f}(\hat{x}_0, 0) = \hat{F}(\hat{x}_0, 0) = 0.$$

for $\hat{x}_0 = \varphi(x_0)$. Then, by Lemma 4.3 we have

$$\mathrm{D}_{\hat{x}}\hat{F}(\hat{x}_0, 0) = \mathrm{D}\psi(\hat{x}_0)^* \circ \mathrm{Hess} f(x_0, 0) \circ \mathrm{D}\psi(\hat{x}_0).$$

Since local charts are diffeomorphisms, Hypothesis (i) on the invertibility of the Riemannian Hessian guarantees that $\mathrm{D}_{\hat{x}}\hat{F}(\hat{x}_0, 0)$ is of full rank. By the implicit function theorem [KP13, Theorem 3.3.1] there exists an open interval I containing 0 and a continuously differentiable curve $\lambda \mapsto \hat{x}(\lambda) \in \mathbb{R}^d$ such that $\hat{x}(0) = \hat{x}_0$, $\hat{x}(\lambda) \in \hat{\mathcal{N}}$ for every $\lambda \in I$ and

$$\hat{F}(\hat{x}(\lambda), \lambda) = 0, \quad \forall \lambda \in I. \tag{4.11}$$

By chain rule, differentiating (4.11) with respect to λ yields

$$\mathrm{D}_{\hat{x}}\hat{F}(\hat{x}(\lambda), \lambda)[\hat{x}'(\lambda)] + \mathrm{D}_\lambda \hat{F}(\hat{x}(\lambda), \lambda) = 0, \quad \forall \lambda \in I.$$

By the smoothness of $\lambda \mapsto \hat{x}(\lambda)$ and the assumed continuity of $\mathrm{Hess} f$ on $\mathcal{U} \times [0, 1]$, since

$D_{\hat{x}}\hat{F}(\hat{x}(0),0)$ is invertible, there exists $I' \subset I$ containing 0 such that $D_{\hat{x}}\hat{F}(x(\lambda),\lambda)$ is invertible for all $\lambda \in I'$. Hence, it holds that

$$\hat{x}'(\lambda) = -D_{\hat{x}}\hat{F}(\hat{x}(\lambda),\lambda)^{-1}\big[D_\lambda\hat{F}(\hat{x}(\lambda),\lambda)\big], \quad \forall\,\lambda \in I'. \tag{4.12}$$

Define $x(\lambda) := \psi(\hat{x}(\lambda))$ for all $\lambda \in I'$. By Lemma 4.2, we know

$$\mathrm{grad}f(x(\lambda),\lambda)) = 0, \quad \forall\,\lambda \in I'.$$

Furthermore, differentiation of $\hat{F}$ with respect to λ transfers to the Riemannian gradient as

$$D_\lambda\hat{F}(\hat{x}(\lambda),\lambda) = D\psi^*(\hat{x}(\lambda))\left[\frac{\partial\,\mathrm{grad}f(x(\lambda),\lambda)}{\partial\lambda}\right], \quad \forall\,\lambda \in I.$$

Therefore, using Lemma 4.3 we evince from (4.12) that the derivative of the manifold curve for every $\lambda \in I'$ is

$$\begin{aligned}
\dot{x}(\lambda) &= D\psi(\hat{x}(\lambda))[\hat{x}'(\lambda)] \\
&= -D\psi(\hat{x}(\lambda))\Big[D_{\hat{x}}\hat{F}(\hat{x}(\lambda),\lambda)^{-1}\big[D_\lambda\hat{F}(\hat{x}(\lambda),\lambda)\big]\Big] \\
&= -D\psi(\hat{x}(\lambda))\circ D\psi(\hat{x}(\lambda))^{-1}\circ\mathrm{Hess}f(x(\lambda),\lambda)^{-1} \\
&\quad \circ (D\psi(\hat{x}(\lambda))^*)^{-1}\circ D\psi^*(\hat{x}(\lambda))\left[\frac{\partial\,\mathrm{grad}f(x(\lambda),\lambda)}{\partial\lambda}\right] \\
&= -\mathrm{Hess}f(x(\lambda),\lambda)^{-1}\left[\frac{\partial\,\mathrm{grad}f(x(\lambda),\lambda)}{\partial\lambda}\right]. \tag{4.13}
\end{aligned}$$

Summarizing the above, there exists $\lambda_{\mathrm{L}} < 0$ and $\lambda_{\mathrm{U}} > 0$ such that
  (1) x is defined on $J := (\lambda_{\mathrm{L}},\lambda_{\mathrm{U}})$,
  (2) $\mathrm{grad}f(x(\lambda),\lambda) = 0 \quad \forall\,\lambda \in J$,
  (3) x is continuously differentiable on J with $\dot{x}$ given by (4.13),
  (4) $x(\lambda) \in \mathcal{U}, \quad \forall\,\lambda \in J$.
Define the following

$$\lambda^* = \sup\{\lambda_{\mathrm{U}} : \text{there exists } x \text{ such that (1), (2), (3) and (4) are verified}\}.$$

By the discussion above $I' \subset J$, hence $\lambda^* > 0$. If $\lambda^* > 1$, the result is proved. Therefore assume that $0 < \lambda^* \le 1$. Due to condition (4) and Hypothesis (ii) we obtain from (4.13) that $\|\dot{x}(\lambda)\|_{x(\lambda)} < L$ holds for every $\lambda \in (\lambda_{\mathrm{L}},\lambda^*)$. This implies

$$\tilde{L} := \lim_{\lambda\uparrow\lambda^*} d(x_0,x(\lambda)) \le \lim_{\lambda\uparrow\lambda^*}\int_0^\lambda\|x'(\tau)\|_{x(\tau)}\mathrm{d}\tau < \lim_{\lambda\uparrow\lambda^*}\int_0^\lambda L\mathrm{d}\tau \le L.$$

Given a sequence $\{\lambda_k\} \subset J$ with $\lambda_k \to \lambda^*$, it follows in an analogous fashion that $\{x(\lambda_k)\}$ is a Cauchy sequence. By completeness of $\mathcal{M}$, the sequence admits a limit point x^* such that $d(x_0,x^*) \le \tilde{L} < L$. By assumption $B(x_0,L) \subseteq \mathcal{U}$, hence $x^* \in \mathcal{U}$.

Now, since by continuity $\mathrm{grad} f(x^*, \lambda^*) = 0$, using another local chart $\phi : \mathcal{N}' \to \mathbb{R}^d$ such that $x^* \in \mathcal{N}'$, we can repeat the above procedure and apply the implicit function theorem to the (Euclidean) gradient of

$$\tilde{f}(\hat{z}, \lambda) := f(\phi^{-1}(\hat{z}), \lambda)$$

at $(\phi(x^*), \lambda^*)$ and thus extend $x(\lambda)$ to a larger interval. This contradicts the definition of λ^*. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

4.3.2 Riemannian predictor-corrector continuation

The Riemannian predictor-corrector continuation algorithm mimics the Euclidean version from Section 4.2 by numerically integrating the Riemannian Davidenko equation (4.10). For the moment, we consider N steps with fixed step size $h_k = 1/N$, for $k = 1, \ldots, N$. A possible adaptive step size strategy is discussed in Section 4.3.4.

Algorithm 4.1 Riemannian Newton Continuation (RNC)

Input: $f(\cdot, \lambda) \in \mathfrak{F}(\mathcal{M})$ for $\lambda \in [0, 1]$, $x_0 \in \mathcal{M}$ such that $\mathrm{grad} f(x_0, 0) = 0$, $N_{\mathrm{steps}} \in \mathbb{N}$,
$\qquad \varepsilon_{\mathrm{tol}} > 0$, $N_{\mathrm{inner}} \in \mathbb{N}$.

1: $h_0 = \frac{1}{N_{\mathrm{steps}}}$, $\lambda_0 = 0$, $k = 0$;
2: **while** $\lambda_k < 1$ **do**
3: $\quad$ **if** tangentialPrediction **then**
4: $\quad\quad$ Solve $\mathrm{Hess} f(x_k, \lambda_k)[t_k] = -\frac{\partial \,\mathrm{grad} f(x_k, \lambda_k)}{\partial \lambda}$ for $t_k \in T_{x_k}\mathcal{M}$;
5: $\quad\quad$ **if** adaptStepSize **then**
6: $\quad\quad\quad$ Determine the new step size h_k with Algorithm 4.2.
7: $\quad\quad$ **end if**
8: $\quad\quad$ $y_{k+1} = R_{x_k}(h_k t_k)$;
9: $\quad$ **else**
10: $\quad\quad$ $y_{k+1} = x_k$;
11: $\quad$ **end if**
12: $\quad$ $\lambda_{k+1} = \min\{\lambda_k + h_k, 1\}$;
13: $\quad$ $x_{k+1} = \mathrm{RiemannianNewton}(y_{k+1}, f(\cdot, \lambda_{k+1}), \varepsilon_{\mathrm{tol}}, N_{\mathrm{inner}})$;
14: $\quad$ **if** $\| \mathrm{grad} f(x_{k+1}, \lambda_{k+1}) \| > \mathrm{tol}$ **then**
15: $\quad\quad$ Error("Correction failed at $\lambda = \lambda_{k+1}$.");
16: $\quad$ **else**
17: $\quad\quad$ $k = k + 1$;
18: $\quad$ **end if**
19: **end while**
20: **return** x_k;

Prediction $\quad$ The classical continuation scheme (4.7) can be trivially extended to the Riemannian case without any adjustment. The initial guess for the subsequent correction phase is simply

$$y_{k+1} = x_k, \tag{4.14}$$

the iterate at the previous step of the algorithm.

The Riemannian extension of the tangential prediction strategy (4.8) is slightly more involved. It consists of performing a step in the direction of the tangent vector of the solution curve. This tangent vector can be computed from the Davidenko equation as

$$t(x_k, \lambda_k) := -\mathrm{Hess}f(x_k, \lambda_k)^{-1} \left[\frac{\partial \, \mathrm{grad}f(x_k, \lambda_k)}{\partial \lambda} \right] \in T_{x_k}\mathcal{M}. \qquad (4.15)$$

Note that this entails the solution of a linear system with the Riemannian Hessian. If its solution by a direct solver (e.g., via the Cholesky decomposition) is too expensive, especially for manifolds of higher dimension, matrix-free Krylov type methods [vdV09, §5] can be used instead.

In the Euclidean case, a tangent vector was simply added to the current iterate. In the Riemannian case, in analogy the optimization methods, this can be replaced by a retraction step. Letting R denote any retraction on $\mathcal{M}$, the Riemannian tangential prediction step is defined as

$$y_{k+1} = R_{x_k}(h_k t(x_k, \lambda_k)), \qquad (4.16)$$

where we recall that h_k denotes the step size. This type of update strategy which guarantees to remain on the manifold can be used in the context of numerically integrating differential equations on manifolds [Hai01, KV19], see also Chapter 6.
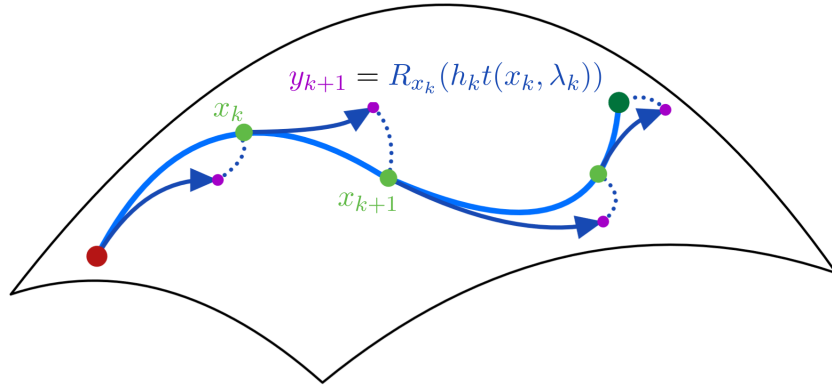
Figure 4.4: The RNC algorithm.

Correction   In analogy to the Euclidean case from Section 4.2, we rely on a second-order method for refining the estimate y_{k+1} such that it becomes a (nearly) critical point of $f(\cdot, \lambda_{k+1})$. The tolerance on the Riemannian gradient norm is chosen small enough to closely track the solution curve, typically 10^{-6}. Among the Riemannian optimization methods introduced in Section 2.3, the Riemannian Newton (RN) method can be used for this purpose; its basic form is described in Algorithm 2.3. Note that the Riemannian Newton method can be replaced by any locally superlinearly convergent method, e.g., the Riemannian Trust Region (RTR) method (Algorithm 2.4) or the Riemannian BFGS

method (Algorithm 2.6). These methods can take full advantage of sufficiently accurate initial guess provided by the prediction, yielding a fast correction phase. Although a first-order method such as Riemannian gradient descent (Algorithm 2.1) or Riemannian conjugate gradient (Algorithm 2.5) could in principle be used, they would not benefit the warm-starting fully as they do not exhibit accelerated convergence near a critical point.

Riemannian-Newton Continuation (RNC)   The whole predictor-corrector scheme for Riemannian manifolds is sketched in Algorithm 4.1 and illustrated in Figure 4.4. The optional adaptive step size strategy in line 6 is explained in Section 4.3.4 below.

4.3.3   Prediction order analysis

An accurate prediction step leads to fast convergence in the correction step. The concept of order is used in the Euclidean case [Deu11, p.238-239] to qualitatively capture this accuracy. The following definition extends this concept to the Riemannian case by considering the prediction path $y(h) \in \mathcal{M}$, $h > 0$, obtained from the prediction step by varying the step size h.

Definition 4.4 (Prediction order). *Let $x(\lambda)$ be the solution curve defined by* (4.10) *for $\lambda \in [0, 1]$. A prediction path $y(h)$ such that $y(0) = x(\lambda)$ is said to be of* order p *if there exists a constant $\eta_p > 0$ independent of h, such that*

$$d(x(\lambda + h), y(h)) \leq \eta_p h^p, \quad \forall \lambda \in [0, 1),$$

holds for all sufficiently small $h > 0$.

In the following we prove that the prediction orders for the Riemannian classical and tangential prediction schemes match the ones in the Euclidean case. More specifically, the following propositions show that classical prediction (4.14) has order 1 while tangential prediction (4.16) has order 2.

Proposition 4.5. *The classical prediction path $y_c(h) = x(\lambda)$ has order 1.*

Proof. Applying the definition of distance function, we obtain for sufficiently small $h > 0$ that

$$d(x(\lambda + h), y_c(h)) = d\left(x(\lambda + h), x(\lambda)\right) \leq \int_{\lambda}^{\lambda+h} \|\dot{x}(\tau)\| \mathrm{d}\tau \leq h \max_{\tau \in [\lambda, \lambda+h]} \{\|\dot{x}(\tau)\|\}$$
$$\leq h \max_{\tau \in [0,1]} \{\|\dot{x}(\tau)\|\}$$

$\square$

Proposition 4.6. *If $\lambda \mapsto x(\lambda)$ is twice continuously differentiable on an open interval $J \supset [0, 1]$, the tangential prediction path*

$$y_t(h) = R_{x(\lambda)}(ht(x(\lambda), \lambda))$$

has order 2.

Proof. By Theorem 4.1, the derivative of x at λ is equal to $t(x(\lambda), \lambda)$. Then, the result follows by applying Proposition 3.22 to the solution curve x. $\qquad\square$

4.3.4 Step size adaptation via asymptotic expansion

The selection of the step size h_k in Algorithm 4.1 has a great impact on its efficiency. A good step size selection should find a balance between the two conflicting goals of attaining fast convergence in each correction step and maintaining a low number of correction steps.

An overview of existing strategies for the Euclidean case can be found in [AG90, Deu11]. In the following, we focus on the case of tangential prediction. We propose to generalize to the Riemannian setting a step size selection scheme which is summarized in [AG90, §6.1]. It aims at guaranteeing the three following conditions: (i) the distance between the prediction point y_{k+1} and the corresponding solution point x_{k+1} is below a prescribed tolerance, (ii) the RN method on $f(\cdot, \lambda_k + h_k)$ started at the prediction point y_{k+1} is sufficiently contractive and (iii) the curvature of the solution curve between x_k and x_{k+1} is below a prescribed tolerance. For the Euclidean case, an analogous approach intended to fulfill condition (ii) is used in the numerical continuation software package HOMPACK [Wat79], while the strategy we now describe targets the three above conditions simultaneously.

Given any $(w, \lambda) \in \mathcal{M} \times [0, 1]$ such that $\mathrm{Hess} f(w, \lambda)$ is full rank, we denote the prediction vector and the RN update vector respectively as

$$t(w, \lambda) = -\mathrm{Hess} f(w, \lambda)^{-1} \Big[\frac{\partial}{\partial \lambda} \, \mathrm{grad} f(w, \lambda) \Big],$$

and

$$n(w, \lambda) = -\mathrm{Hess} f(w, \lambda)^{-1} \big[\, \mathrm{grad} f(w, \lambda) \big].$$

Given $(x(\lambda), \lambda)$ on the solution curve, recall the tangential prediction point as a function of step size $h > 0$ is

$$y(h) = R_{x(\lambda)}(ht(x(\lambda), \lambda)). \tag{4.17}$$

An approximation of the distance between $y(h)$ and $x(\lambda + h)$ can be obtained from the norm of the first RN update vector, see Figure 4.5. We shall denote it

$$\delta(x(\lambda), \lambda, h) := \|n(y(h), \lambda + h)\|_{y(h)}. \tag{4.18}$$

Let us indicate the first iterate of the RN method started at $y(h)$ as

$$z(h) = R_{y(h)} \left(n(y(h), \lambda + h) \right). \tag{4.19}$$
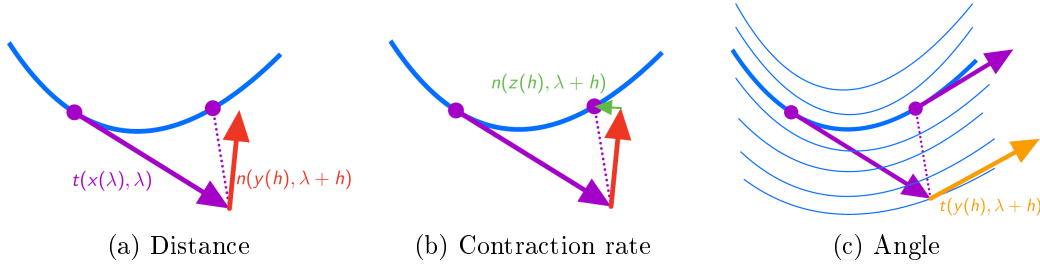
(a) Distance (b) Contraction rate (c) Angle

Figure 4.5: Step size adaptation performance indicators.

Then, the first contraction rate of the RN is defined as

$$\kappa(x(\lambda), \lambda, h) := \frac{\|n(z(h), \lambda + h)\|_{z(h)}}{\|n(y(h), \lambda + h)\|_{y(h)}}. \tag{4.20}$$

Upon convergence of the RN method for $f(\cdot, \lambda + h)$, this ratio is smaller than 1. Finally, the curvature of the solution curve between two points $x(\lambda)$ and $x(\lambda + h)$ can be approximated with

$$\alpha(x(\lambda), \lambda, h) := \mathrm{acos}\left(\frac{\left\langle \mathcal{T}_{ht(x(\lambda), \lambda)}(t(x(\lambda), \lambda)), t(y(h), \lambda + h) \right\rangle_{y(h)}}{\|\mathcal{T}_{ht(x(\lambda), \lambda)}(t(x(\lambda), \lambda))\|_{y(h)} \|t(y(h), \lambda + h)\|_{y(h)}} \right) \tag{4.21}$$

the angle between the prediction vector at the solution curve point $(x(\lambda), \lambda)$ and the prediction vector at the prediction point $(y(h), \lambda + h)$. Note that in order to measure their relative angle we transport $t(x(\lambda), \lambda) \in T_{x(\lambda)}\mathcal{M}$ to $T_{y(h)}\mathcal{M}$ using the linear map

$$\mathcal{T}_{ht(x(\lambda), \lambda)} : T_{x(\lambda)}\mathcal{M} \to T_{y(h)}\mathcal{M}$$

given by a vector transport associated with the retraction R, see Definition 2.13. Note that (4.21) is well-defined only if $t(x(\lambda), \lambda) \neq 0$ and $t(y(h), \lambda + h) \neq 0$. However, the second follows from the first by continuity of $h \to t(y(h), \lambda + h)$ provided h is small enough.

The following lemma inspired by [AG90, Lemmas 6.1.2, 6.1.8] is the cornerstone of the step selection strategy. It provides a Taylor expansion with respect to h around $h = 0$ of the indicators (4.18), (4.20), (4.21). For convenience of notation, in the following we drop the explicit dependence of x on λ.

Lemma 4.7. *Assume f is seven times continuously differentiable in both arguments and that for each (x, λ) of the solution curve we have*

$$\frac{\mathrm{D}^2}{\mathrm{d}h^2}\mathrm{grad}f(y(h), \lambda + h)\Big|_{h=0} \neq 0, \tag{4.22}$$

where $\frac{\mathrm{D}^2}{\mathrm{d}h^2}$ denote the second covariant derivative along the prediction path (4.17). Then, there exist functions $\delta_2(x, \lambda)$, $\kappa_2(x, \lambda)$, $\alpha_1(x, \lambda)$ only depending on x and λ such that the

following holds:

(i) *The norm of the first Newton update vector* $\delta(x, \lambda, h) = \|n(y(h), \lambda + h)\|_{y(h)}$ *verifies*

$$\delta(x, \lambda, h) = \delta_2(x, \lambda)h^2 + O\left(h^3\right).$$

(ii) *If R is second-order retraction, then the contraction rate of the Newton method, namely,*

$$\kappa(x, h) = \frac{\|n(z(h), \lambda + h)\|_{z(h)}}{\|n(y(h), \lambda + h)\|_{y(h)}},$$

verifies

$$\kappa(x, \lambda, h) \leq \kappa_2(x, \lambda)h^2 + o\left(h^2\right).$$

(iii) *If $t(x, \lambda) \neq 0$, the prediction angle*

$$\alpha(x, \lambda, h) = \mathrm{acos}\left(\frac{\left\langle \mathcal{T}_{ht(x,\lambda)}(t(x, \lambda)), t(y(h), \lambda + h) \right\rangle_{y(h)}}{\|\mathcal{T}_{ht(x,\lambda)}(t(x, \lambda))\|_{y(h)} \|t(y(h), \lambda + h)\|_{y(h)}}\right) \tag{4.23}$$

is well-defined and, provided

$$\frac{\mathrm{D}}{\mathrm{d}h}\mathcal{T}_{ht(x,\lambda)}(t(x, \lambda))\big|_{h=0} - \frac{\mathrm{D}}{\mathrm{d}h}t(y(h), \lambda + h)\big|_{h=0} \neq ct(x, \lambda), \quad \forall c \in \mathbb{R}, \tag{4.24}$$

it verifies

$$\alpha(x, \lambda, h) = \alpha_1(x, \lambda)h + O\left(h^2\right).$$

The proof of Lemma 4.7 is quite lengthy and is reported in the next section. Let us first describe the step size selection strategy inspired by this result. Given positive constants $\delta_{\max}$, $\kappa_{\max}$ and $\alpha_{\max}$, we aim at finding the largest $h_k > 0$ such that

$$\delta(x_k, \lambda_k, h_k) \leq \delta_{\max}, \quad \kappa(x_k, \lambda_k, h_k) \leq \kappa_{\max}, \quad \alpha(x_k, \lambda_k, h_k) \leq \alpha_{\max}.$$

Since κ is a ratio that should be smaller than 1 when the Newton method enters its superlinear convergence regime, only values $\kappa_{\max} \in (0, 1)$ make sense but in practice one would choose a value that is away from 1, say $1/4$. Similarly, α is an angle that should remain small and so $\alpha_{\max} \in (0, \pi/2)$. However, once again, a typical value is away from $\pi/2$, for instance $\alpha_{\max} = 10°$. The last indicator δ is an estimate for the distance to the next minimizer and so the choice for $\delta_{\max}$ is application dependent.

Given a trial step size $\tilde{h}_k$ (obtained, e.g., from the previous step), Lemma 4.7 allows us

to estimate

$$\delta_2(x_k, \lambda_k) \simeq \tilde{\delta}_2(x_k, \lambda_k) := \frac{\delta(x_k, \lambda_k, \tilde{h}_k)}{\tilde{h}_k^2}, \tag{4.25}$$

$$\kappa_2(x_k, \lambda_k) \simeq \tilde{\kappa}_2(x_k, \lambda_k) = \frac{\kappa(x_k, \lambda_k, \tilde{h}_k)}{\tilde{h}_k^2}. \tag{4.26}$$

$$\alpha_1(x_k, \lambda_k) \simeq \tilde{\alpha}_1(x_k, \lambda_k) = \frac{\alpha(x_k, \lambda_k, \tilde{h}_k)}{\tilde{h}_k}, \tag{4.27}$$

Then, imposing

$$\tilde{\delta}_2(x_k, \lambda_k) h_k^2 \le \delta_{\max}, \quad \tilde{\kappa}_2(x_k, \lambda_k) h_k^2 \le \kappa_{\max}, \quad \tilde{\alpha}_1(x_k, \lambda_k) h_k \le \alpha_{\max},$$

yields

$$h_k \le \tilde{h}_k \min\left\{ \sqrt{\frac{\delta_{\max}}{\tilde{\delta}(x_k, \lambda_k)}}, \sqrt{\frac{\kappa_{\max}}{\tilde{\kappa}(x_k, \lambda_k)}}, \frac{\alpha_{\max}}{\tilde{\alpha}(x_k, \lambda_k)} \right\}.$$

This is the criterion to adjust step size, but not to make too drastic changes, the relative change of the step size is clamped to the interval the interval $[1/2, 2]$, i.e. the step size is at most halved or doubled. The resulting procedure is summarized in Algorithm 4.2. Note that this comes at the non-negligible cost of (approximately) solving 3 extra linear systems involving the Riemannian Hessian.

Algorithm 4.2 Adaptive step size for RNC Algorithm 4.1

Input: $f(\cdot, \lambda) \in \mathfrak{F}(\mathcal{M})$ for $\lambda \in [0, 1]$, $(x_k, \lambda_k) \in \mathcal{M} \times (0, 1)$ such that $\mathrm{grad} f(x_k, \lambda_k) = 0$, $t(x_k, \lambda_k)$ as in (4.15), $\alpha_{\max} \in (0, \pi)$, $\delta_{\max} > 0$, $\kappa_{\max} \in (0, 1)$, $\tilde{h}_k > 0$.

1: $y_k = R_{x_k}(\tilde{h}_k t(x_k, \lambda_k))$;
2: Solve $\mathrm{Hess} f(y_k, \lambda_k + \tilde{h}_k)[t(y_k, \lambda + \tilde{h}_k)] = -\frac{\partial \mathrm{grad} f(y_k, \lambda_k + \tilde{h}_k)}{\partial \lambda}$;
3: Solve $\mathrm{Hess} f(y_k, \lambda_k + \tilde{h}_k)[n(y_k, \lambda + \tilde{h}_k)] = -\mathrm{grad} f(y_k, \lambda_k + \tilde{h}_k)$;
4: $z_k = R_{y_k}(n(y_k, \lambda + \tilde{h}_k))$;
5: Solve $\mathrm{Hess} f(z_k, \lambda_k + \tilde{h}_k)[n(z_k, \lambda + \tilde{h}_k)] = -\mathrm{grad} f(z_k, \lambda_k + \tilde{h}_k)$;
6: Compute $\tilde{\delta}_2$, $\tilde{\kappa}_2$ and $\tilde{\alpha}_1$ using (4.25), (4.26) and (4.27).
7: $h_k = \tilde{h}_k \max\left\{ \frac{1}{2}, \min\left\{ \sqrt{\frac{\delta_{\max}}{\tilde{\delta}_2}}, \sqrt{\frac{\kappa_{\max}}{\tilde{\kappa}_2}}, \frac{\alpha_{\max}}{\tilde{\alpha}_1}, 2 \right\} \right\}$;
8: **return** h_k

4.3.5 Proofs of Lemma 4.7

A slightly different version of Lemma 4.7 appears in [SK22a, Lemma 3.5] and is proved using a local representation of the problem via a local chart. In this proof, the assumption [SK22a, Equation 3.12], which prescribes

$$\left. \frac{\mathrm{D}^2}{\mathrm{d}h^2} n(y(h), \lambda + h) \right|_{h=0} \neq 0, \tag{4.28}$$

is invoked to guarantee the leading coefficient of the Taylor expansion in Lemma 4.7-(i) is not zero. However, the link between (4.28) and the coefficient is quite technical and is not worked out in full detail in [SK22a]. Therefore, we here propose an alternative coordinate-free proof based on covariant differentiation for which the use of that assumption is made more explicit. As a positive side effect, this new technique allows replacing assumption (4.28) with the more interpretable requirement (4.22), to which it is equivalent as we show in the new proof. At the same time, the new proof techniques comes at a certain cost. On the one hand, the smoothness requirement of the objective function is made more stringent: from four to seven times continuously differentiable. On the other hand, the second result is established provided the retraction that is used is a second-order retraction. Finally, assumption (4.24) was added to ensure the leading coefficient of the Taylor expansion of result (iii) does not vanish. While the first two modifications are due to the alternative proof technique and do not seem necessary in the proof of [SK22a, Lemma 3.5], assumption (4.24) is required also in the original proof but was not explicitly stated [1]. We report a detailed version of the original proof from [SK22a, Lemma 3.5] at the end of this section.

We also point out that the angle indicator used in [SK22a, Lemma 3.5], which is the object of result (iii), is defined transporting the prediction tangent vector $t(y(h), \lambda + h)$ to the tangent space at x using the notion of vector transport, see Definition 2.13. However, what we use as a vector transport in [SK22a, Lemma 3.5] is actually an instance of a more general concept, going by the name of transporter [Bou23, §10.5]. The main difference between a vector transport and a transporter is that the latter is associated with a retraction. Given a vector transport $\mathcal{T}$ associated to a retraction R and a given $v \in T_x\mathcal{M}$ in the domain of the retraction, the map $\mathcal{T}_v$ is a linear map between $T_x\mathcal{M}$ and $T_{R_x(v)}\mathcal{M}$. In contrast, a transporter can be used to define linear maps between arbitrary pairs of tangent spaces for which their base points are sufficiently close. Rather than reformulating [SK22a, Lemma 3.5] to explicitly state the use of transporters, we propose in Lemma 4.7 a modification where the vector transport is used. Accordingly, the angle indicator (4.23) is now measured in the tangent space at $y(h) = R_x(ht(x, \lambda))$ by transporting $t(x, \lambda)$ with the use of vector transport along the vector $ht(x, \lambda)$.

Proof based on covariant differentiation

For a given (x, λ) the indicators δ, κ and α are scalar functions of h of which we aim at performing the Taylor expansion around $h = 0$. These functions involve the Riemannian metric and smooth vectors fields restricted to the manifold curves $y(h)$ and $z(h)$, defined by (4.17) and (4.19). Hence, the derivatives of the scalar functions involve induced covariant differentiation along such curves. In order to explicitly manipulate the covariant derivative of vector fields involving the Riemannian Hessian, e.g. the Newton update vector, we introduce the covariant derivative of the Riemannian Hessian based on [Bou23, Example 10.78].

---

[1] The Arxiv version of [SK22a] has been updated with assumption (4.24) stated explicitly.

The second-order covariant derivative of a smooth scalar field $g \in \mathfrak{F}(\mathcal{M})$ is a second order tensor field that can be represented with the Riemannian Hessian:

$$\nabla^2 g(U, V) = \langle \mathrm{Hess}g(x)\,[U]\,, V \rangle, \quad \forall U, V \in \mathfrak{X}(\mathcal{M}).$$

Then, applying Definition 1.26, the third-order covariant derivative of g is given by

$$\nabla^3 g(U, V, W) = W \langle \mathrm{Hess}g[U], V \rangle - \langle \mathrm{Hess}g[\nabla_W U], V \rangle - \langle \mathrm{Hess}g[U], \nabla_W V \rangle.$$

Expanding the first term using the compatibility of the Riemannian connection with the metric we get, for every $U, V, W \in \mathfrak{X}(\mathcal{M})$,

$$
\begin{aligned}
\nabla^3 g(U, V, W) &= \langle \nabla_W\left(\mathrm{Hess}g[U]\right), V \rangle + \langle \mathrm{Hess}g[U], \nabla_W V \rangle \\
&\quad - \langle \mathrm{Hess}g[\nabla_W U], V \rangle - \langle \mathrm{Hess}g[U], \nabla_W V \rangle \\
&= \langle \nabla_W\left(\mathrm{Hess}g[U]\right), V \rangle - \langle \mathrm{Hess}g[\nabla_W U], V \rangle \\
&=: \langle \nabla_W \mathrm{Hess}g[U], V \rangle,
\end{aligned}
$$

where $\nabla_W \mathrm{Hess}g$ is defined from the above expression by

$$\nabla_W \mathrm{Hess}g[U] = \nabla_W\left(\mathrm{Hess}g[U]\right) - \mathrm{Hess}g[\nabla_W U].$$

We refer to this terms as the covariant derivative of the Riemannian Hessian. It appears whenever we compute the covariant derivative of a vector field through the Riemannian Hessian:

$$\nabla_W\left(\mathrm{Hess}g[U]\right) = \nabla_W \mathrm{Hess}g[U] + \mathrm{Hess}g[\nabla_W U].$$

In the same way as the value of $\nabla_W U$ at any $p \in \mathcal{M}$ depends on W only through its value at p, the covariant derivative of the Riemannian Hessian at p along W only depends on $W(p) =: w$. Hence, for every $w \in T_p\mathcal{M}$ it is legitimate to define $\nabla_w \mathrm{Hess}g(p)$ as a linear map from $T_p\mathcal{M}$ to itself.

Proof of Lemma 4.7. Result (i)
Let us denote $g(h) = \|n(y(h), \lambda + h)\|_{y(h)}^2$. Since the Riemannian metric and the retraction are assumed to be infinitely smooth, the smoothness of the function g is determined by the smoothness of n with respect to its arguments. The expression of n features the Riemannian Hessian of f with respect to its first argument, hence if $f \in C^7$, then $g \in C^5$ and we can compute its Taylor expansion around zero up to a residual of order 5:

$$g(h) = g(0) + h g'(0) + \frac{h^2}{2} g''(0) + \frac{h^3}{6} g'''(0) + \frac{h^4}{24} g^{(4)}(0) + O\left(h^5\right).$$

By the defining property of retractions $y(0) = R_x(0) = x$. Then,

$$g(0) = \|n(x, \lambda)\|_x^2 = \left\| \mathrm{Hess}f(x, \lambda)^{-1} \Big[\underbrace{\mathrm{grad}f(x, \lambda)}_{=0} \Big] \right\|_x^2 = 0.$$

By compatibility of covariant differentiation with the Riemannian metric,

$$g'(0) = \frac{\mathrm{d}}{\mathrm{d}h} \langle n(y(h), \lambda + h), n(y(h), \lambda + h) \rangle \big|_{h=0} = 2 \Big\langle \frac{\mathrm{D}}{\mathrm{d}h} n(y(h), \lambda + h) \big|_{h=0}, n(x, \lambda) \Big\rangle_x = 0.$$

Using $n(x, \lambda) = 0$, the second derivative in zero amounts to

$$\begin{aligned} g''(0) = \frac{\mathrm{d}^2}{\mathrm{d}h} \langle n(y(h), \lambda + h), n(y(h), \lambda + h) \rangle \big|_{h=0} &= 2 \Big\| \frac{\mathrm{D}}{\mathrm{d}h} n(y(h), \lambda + h) \big|_{h=0} \Big\|_x^2 \\ &\quad + 2 \Big\langle \frac{\mathrm{D}^2}{\mathrm{d}h} n(y(h), \lambda + h) \big|_{h=0}, n(x, \lambda) \Big\rangle_x \\ &= 2 \Big\| \frac{\mathrm{D}}{\mathrm{d}h} n(y(h), \lambda + h) \big|_{h=0} \Big\|_x^2. \end{aligned}$$

The map $n : \mathcal{M} \times [0, 1] \mapsto T\mathcal{M}$ is restricted to $h \to (y(h), \lambda + h)$ implying that its covariant derivative is the sum of two terms

$$\frac{\mathrm{D}n(y(h), \lambda + h)}{\mathrm{d}h} = \nabla_{y'(h)} n(y(h), \lambda + h) + \frac{\partial}{\partial \lambda} n(y(h), \lambda + h).$$

We can now use to the notion of covariant derivative of tensor fields following from Definition 1.26 and explicitly write the above as

$$\begin{aligned} \frac{\mathrm{D}n(y(h), \lambda + h)}{\mathrm{d}h} = &- \nabla_{y'(h)} \left(\mathrm{Hess}f(y(h), \lambda + h)^{-1} \right) \left[\mathrm{grad}f(y(h), \lambda + h) \right] \\ &- \mathrm{Hess}f(y(h), \lambda + h)^{-1} \left[\nabla_{y'(h)} \mathrm{grad}f(y(h), \lambda + h) \right] \\ &- \frac{\partial}{\partial \lambda} \left(\mathrm{Hess}f(y(h), \lambda + h)^{-1} \right) \left[\mathrm{grad}f(y(h), \lambda + h) \right] \\ &- \mathrm{Hess}f(y(h), \lambda + h)^{-1} \left[\frac{\partial}{\partial \lambda} \mathrm{grad}f(y(h), \lambda + h) \right] \end{aligned} \tag{4.29}$$

The second term simplifies thanks to the definition of Riemannian Hessian:

$$\begin{aligned} \mathrm{Hess}f(y(h), \lambda + h)^{-1} \left[\nabla_{y'(h)} \mathrm{grad}f(y(h), \lambda + h) \right] &= \mathrm{Hess}f(y(h), \lambda + h)^{-1} \left[\mathrm{Hess}f(y(h), \lambda + h) \left[y'(h) \right] \right] \\ &= y'(h) = \mathrm{D}_2 R_x(ht(x, \lambda)) \left[t(x, \lambda) \right] \end{aligned}$$

Furthermore, evaluating (4.29) in $h = 0$, the first and third terms vanish while in the last term we recognize the expression of $t(x, \lambda)$. This reveals that the first covariant

derivative of n vanishes in $h = 0$

$$\frac{\mathrm{D}}{\mathrm{d}h}n(y(h), \lambda + h)\big|_{h=0} = - \nabla_{y'(0)}\left(\mathrm{Hess}f(x, \lambda)^{-1}\right)\left[\,\mathrm{grad}f(x, \lambda)\right] - y'(0)$$
$$- \frac{\partial}{\partial\lambda}\left(\mathrm{Hess}f(x, \lambda)^{-1}\right)\left[\,\mathrm{grad}f(x, \lambda)\right]$$
$$- \mathrm{Hess}f(x, \lambda)^{-1}\left[\frac{\partial}{\partial\lambda}\,\mathrm{grad}f(x, \lambda)\right]$$
$$= - \mathrm{D}R_x(0)[t(x, \lambda)] + t(x, \lambda)$$
$$= - t(x, \lambda) + t(x, \lambda)$$
$$= 0.$$

Therefore $g''(0) = 0$. Differentiating once again, we get the covariant derivative of n of order three.

$$g'''(0) = \frac{\mathrm{d}}{\mathrm{d}h}\left(2\left\|\frac{\mathrm{D}}{\mathrm{d}h}n(y(h), \lambda + h)\right\|^2_{y(h)} + 2\left\langle\frac{\mathrm{D}^2}{\mathrm{d}h}n(y(h), \lambda + h), n(x, \lambda)\right\rangle_{y(h)}\right)\bigg|_{h=0}$$
$$= 2\left\langle\frac{\mathrm{D}^3}{\mathrm{d}h^3}n(y(h), \lambda + h)\big|_{h=0}, n(x, \lambda)\right\rangle_x + 6\left\langle\frac{\mathrm{D}^2}{\mathrm{d}h^2}n(y(h), \lambda + h)\big|_{h=0}, \frac{\mathrm{D}}{\mathrm{d}h}n(y(h), \lambda + h)\big|_{h=0}\right\rangle_x.$$

But since $n(x, \lambda) = 0$, $\frac{\mathrm{D}}{\mathrm{d}h}n(y(h), \lambda + h)\big|_{h=0} = 0$, we find $g'''(0) = 0$. Finally, our assumption indicates that the fourth-order term is the leading term of the Taylor expansion. Indeed:

$$g^{(4)}(0) = 2\left\langle\frac{\mathrm{D}^4}{\mathrm{d}h^4}n(y(h), \lambda + h)\Big|_{h=0}, n(x, \lambda)\right\rangle_x$$
$$+ 8\left\langle\frac{\mathrm{D}^3}{\mathrm{d}h^3}n(y(h), \lambda + h)\Big|_{h=0}, \frac{\mathrm{D}}{\mathrm{d}h}n(y(h), \lambda + h)\big|_{h=0}\right\rangle_x$$
$$+ 6\left\|\frac{\mathrm{D}^2}{\mathrm{d}h^2}n(y(h), \lambda + h)\Big|_{h=0}\right\|^2_x$$
$$= 6\left\|\frac{\mathrm{D}^2}{\mathrm{d}h^2}n(y(h), \lambda + h)\Big|_{h=0}\right\|^2_x.$$

By assumption (4.22), this is non-zero. In fact, using that $n(x, \lambda) = 0$ and $\frac{\mathrm{D}}{\mathrm{d}h}n(y(h), \lambda +$

$\left. h) \right|_{h=0} = 0$ we get

$$
\begin{aligned}
0 \neq \frac{D^2}{dh^2} \mathrm{grad} f(y(h), \lambda + h) \Big|_{h=0} &= \frac{D^2}{dh^2} \Big(- \mathrm{Hess} f(y(h), \lambda + h) \left[n(y(h), \lambda + h) \right] \Big) \Big|_{h=0} \\
&= - \frac{D}{dh} \Big(\frac{D}{dh} \mathrm{Hess}(y(h), \lambda + h) \left[n(y(h), \lambda + h) \right] \\
&\quad + \mathrm{Hess}(y(h), \lambda + h) \left[\frac{D}{dh} n(y(h), \lambda + h) \right] \Big) \Big|_{h=0} \\
&= - \frac{D^2}{dh^2} \left(\mathrm{Hess}(y(h), \lambda + h) \right) \Big|_{h=0} \left[n(x, \lambda) \right] \\
&\quad - 2 \frac{D}{dh} \Big(\mathrm{Hess}(y(h), \lambda + h) \Big) \Big|_{h=0} \left[\frac{D}{dh} n(y(h), \lambda + h) \Big|_{h=0} \right] \\
&\quad - \mathrm{Hess} f(x, \lambda) \left[\frac{D^2}{dh^2} n(y(h), \lambda + h) \Big|_{h=0} \right] \\
&= - \mathrm{Hess} f(x, \lambda) \left[\frac{D^2}{dh^2} n(y(h), \lambda + h) \Big|_{h=0} \right].
\end{aligned}
$$

This implies $\left\| \frac{D^2}{dh^2} n(y(h), \lambda + h) \big|_{h=0} \right\|_x^2 \neq 0$ by invertibility of the Riemannian Hessian. Therefore we have shown

$$
\delta(x, \lambda, h)^2 = g(h) = \frac{1}{4} \left\| \frac{D^2}{dh^2} n(y(h), \lambda + h) \Big|_{h=0} \right\|_x^2 h^4 + O\left(h^5\right).
$$

We now apply the square root and use its Taylor expansion in $\frac{1}{4} \left\| \frac{D^2}{dh^2} n(y(h), \lambda + h) \big|_{h=0} \right\|_x^2 h^4 \neq 0$ on the right hand side and we recover

$$
\delta(x, \lambda, h) = \delta_2(x, \lambda) h^2 + O\left(h^3\right)
$$

with $\delta_2(x, \lambda) = \frac{1}{2} \left\| \frac{D^2}{dh^2} n(y(h), \lambda + h) \big|_{h=0} \right\|_x$.

Result (ii)
The main effort in proving this result consists in showing that there exists a constant $z_4(x, \lambda) > 0$ such that

$$
\| n(z(h), \lambda + h) \|_{z(h)} \leq z_4(x, \lambda) h^4 + O(h^5). \tag{4.30}
$$

Then using result (i) and the first-order Taylor expansion of $\frac{1}{t}$ around $t = \delta_2(x, \lambda)$ leads to the following bound

$$
\kappa(x, \lambda, h) \leq \frac{z_4(x, \lambda) h^4 + O(h^5)}{\delta_2(x, \lambda) h^2 + O(h^3)} = h^2 \left(\frac{z_4(x, \lambda) + O(h)}{\delta_2(x, \lambda) + O(h)} \right) = h^2 \frac{z_4(x, \lambda)}{\delta_2(x, \lambda)} + O(h^3).
$$

This would then conclude the proof with $\kappa_2(x, \lambda) = \frac{z_4(x, \lambda)}{\delta_2(x, \lambda)}$. In order to prove (4.30),

first note that

$$\|n(z(h), \lambda + h)\|_{z(h)} \leq \|\mathrm{Hess}f(z(h), \lambda + h)\|_{z(h)} \|\mathrm{grad}f(z(h), \lambda + h)\|_{z(h)},$$

where the first factor is the Hessian's operator norm, as defined by (3.6). Since the operator norm is Lipschitz continuous and f and z are smooth, we find

$$\|\mathrm{Hess}f(z(h), \lambda + h)\|_{z(h)} = \|\mathrm{Hess}f(x, \lambda)\|_x + O(h).$$

For the second factor, observe that for a fixed h we can interpreted it as evaluating a smooth scalar field defined on the tangent space at $y(h)$. Indeed, we have

$$\|\mathrm{grad}f(z(h), \lambda + h)\|_{z(h)} = q \circ R_{y(h)}(n(y(h), \lambda + h)),$$

where q is the smooth scalar field $q(p) = \|\mathrm{grad}f(p, \lambda + h)\|_p$. Since the retraction is assumed to be second-order, the Taylor expansion at the origin of $q \circ R_{y(h)}$ with respect to the tangent vector argument can be written in terms of the Riemannian gradient and Hessian of q as stated in [Bou23, Proposition 5.44]:

$$\begin{aligned}
\|\mathrm{grad}f(z(h), \lambda + h)\|_{z(h)} = {} & q(y(h)) + \langle \mathrm{grad}q(y(h)), n(y(h), \lambda + h)\rangle_{y(h)} \\
& + \frac{1}{2} \langle \mathrm{Hess}q(y(h))[n(y(h), \lambda + h)], n(y(h), \lambda + h)\rangle_{y(h)} \\
& + O(\|n(y(h), \lambda + h)\|_{y(h)}^3).
\end{aligned}$$

The first three terms of the Taylor expansion above coincide with the first three terms of the Taylor expansion of $l(\tau) := q(R_{y(h)}(\tau n(y(h), \lambda)))$ around $\tau = 0$ evaluated at $\tau = 1$, i.e.

$$\|\mathrm{grad}f(z(h), \lambda + h)\|_{z(h)} = l(0) + l'(0) + \frac{1}{2}l''(0) + O(\|n(y(h), \lambda + h)\|_{y(h)}^3). \qquad (4.31)$$

Let us denote $\tilde{z}(\tau) = R_{y(h)}(\tau n(y(h), \lambda + h))$, satisfying $\tilde{z}(0) = y(h)$, $\tilde{z}(1) = z(h)$, and $\dot{\tilde{z}}(0) = n(y(h), \lambda + h)$. Then, we get

$$\begin{aligned}
l'(0) &= \frac{\mathrm{d}}{\mathrm{d}\tau}\sqrt{\langle \mathrm{grad}f(\tilde{z}(\tau), \lambda + h), \mathrm{grad}f(\tilde{z}(\tau), \lambda + h)\rangle_{\tilde{z}(\tau)}}\Bigg|_{\tau=0} \\
&= \frac{2\left\langle \mathrm{Hess}f(\tilde{z}(0), \lambda + h)\left[\dot{\tilde{z}}(0)\right], \mathrm{grad}f(\tilde{z}(0), \lambda + h)\right\rangle_{\tilde{z}(0)}}{2\|\mathrm{grad}f(\tilde{z}(0), \lambda + h)\|_{\tilde{z}(0)}} \\
&= \frac{-\langle \mathrm{grad}f(y(h), \lambda + h), \mathrm{grad}f(y(h), \lambda + h)\rangle_{y(h)}}{\|\mathrm{grad}f(y(h), \lambda + h)\|_{y(h)}} \\
&= -\|\mathrm{grad}f(y(h), \lambda + h)\|_{y(h)}.
\end{aligned}$$

This implies that

$$l(0) + l'(0) = \|\mathrm{grad}f(y(h), \lambda + h)\|_{y(h)} - \|\mathrm{grad}f(y(h), \lambda + h)\|_{y(h)} = 0.$$

Finally, we compute

$$
\begin{aligned}
l''(\tau) &= \frac{\mathrm{d}}{\mathrm{d}\tau} \frac{\left\langle \mathrm{Hess}f(\tilde{z}(\tau), \lambda+h)\left[\dot{\tilde{z}}(\tau)\right], \mathrm{grad}f(\tilde{z}(\tau), \lambda+h)\right\rangle_{\tilde{z}(\tau)}}{\|\mathrm{grad}f(\tilde{z}(\tau), \lambda+h)\|_{\tilde{z}(\tau)}} \\
&= \frac{l(\tau)}{l(\tau)^2}\left(\left\langle \nabla_{\dot{\tilde{z}}(\tau)}\mathrm{Hess}f(\tilde{z}(\tau), \lambda+h)\left[\dot{\tilde{z}}(\tau)\right], \mathrm{grad}f(\tilde{z}(\tau), \lambda+h)\right\rangle_{\tilde{z}(\tau)} \right. \\
&\qquad\qquad + \left\langle \mathrm{Hess}f(\tilde{z}(\tau), \lambda+h)\left[\ddot{\tilde{z}}(\tau)\right], \mathrm{grad}f(\tilde{z}(\tau), \lambda+h)\right\rangle_{\tilde{z}(\tau)} \\
&\qquad\qquad \left. + \left\langle \mathrm{Hess}f(\tilde{z}(\tau), \lambda+h)\left[\dot{\tilde{z}}(\tau)\right], \mathrm{Hess}f(\tilde{z}(\tau), \lambda+h)\left[\dot{\tilde{z}}(\tau)\right]\right\rangle_{\tilde{z}(\tau)} \right) \\
&\quad - \frac{l'(\tau)}{l(\tau)^2}\left\langle \mathrm{Hess}f(\tilde{z}(\tau), \lambda+h)\left[\dot{\tilde{z}}(\tau)\right], \mathrm{grad}f(\tilde{z}(\tau), \lambda+h)\right\rangle_{\tilde{z}(\tau)}.
\end{aligned}
$$

Now we evaluate the above in $\tau = 0$. Note that $\ddot{\tilde{z}}(0) = 0$ by the second-order property of the retraction and we have shown $l(0) = -l'(0)$. Therefore, replacing $\tilde{z}(0)$ by $y(h)$ and $\dot{\tilde{z}}(0)$ by $n(y(h), \lambda+h)$ we find

$$
\begin{aligned}
l''(0) = \frac{1}{l(0)}\Bigg(&\left\langle \nabla_{n(y(h),\lambda+h)}\mathrm{Hess}f(y(h), \lambda+h)\left[n(y(h), \lambda+h)\right], \mathrm{grad}f(y(h), \lambda+h)\right\rangle_{y(h)} \\
&+ \left\langle \mathrm{Hess}f(y(h), \lambda+h)\left[n(y(h), \lambda+h)\right], \mathrm{Hess}f(y(h), \lambda+h)\left[n(y(h), \lambda+h)\right]\right\rangle_{y(h)} \\
&+ \left\langle \mathrm{Hess}f(y(h), \lambda+h)\left[n(y(h), \lambda+h)\right], \mathrm{grad}f(y(h), \lambda+h)\right\rangle_{y(h)} \Bigg)
\end{aligned}
$$

and from the expression of $n(y(h), \lambda+h)$, the two last terms cancel out, showing

$$
l''(0) = \frac{1}{l(0)}\left\langle \nabla_{n(y(h),\lambda+h)}\mathrm{Hess}f(y(h), \lambda+h)\left[n(y(h), \lambda+h)\right], \mathrm{grad}f(y(h), \lambda+h)\right\rangle_{y(h)}.
$$

$$
l''(0) = \frac{\left\langle \nabla_{n(y(h),\lambda+h)}\mathrm{Hess}f(y(h), \lambda+h)\left[n(y(h), \lambda+h)\right], \mathrm{grad}f(y(h), \lambda+h)\right\rangle_{y(h)}}{\|\mathrm{grad}f(y(h), \lambda+h)\|}.
$$

By smoothness of f, the tensor field $(U, V, W) \mapsto \left\langle \nabla_U \mathrm{Hess}f[V], W\right\rangle$ is Lipschitz continuous in the sense of [Bou23, Proposition 10.83] with constant $L_3(f)$. This allows to bound

$$
l''(0) \leq L_3(f)\frac{\|n(y(h), \lambda+h)\|_{y(h)}^2\,\|\mathrm{grad}f(y(h), \lambda+h)\|}{\|\mathrm{grad}f(y(h), \lambda+h)\|_{y(h)}} = L_3(f)\,\|n(y(h), \lambda+h)\|_{y(h)}^2.
$$

From (4.31), we can now evince

$$
\|\mathrm{grad}f(z(h), \lambda+h)\|_{z(h)} \leq \frac{L_3(f)}{2}\,\|n(y(h), \lambda+h)\|_{y(h)}^2 + O(\|n(y(h), \lambda+h)\|_{y(h)}^3).
$$

Finally, using result (i) we have

$$\|\mathrm{grad}f(z(h),\lambda+h)\|_{z(h)} \leq \frac{L_3(f)\delta_2(x,\lambda)^2}{2}h^4 + O(h^5),$$

which shows (4.30) and concludes the proof of result (ii).

Result (iii)
Let us introduce for a fixed $(x,\lambda) \in \mathcal{M} \times [0,1]$ the scalar function

$$m(h) = \frac{\left\langle \mathcal{T}_{ht(x,\lambda)}(t(x,\lambda)), t(y(h),\lambda+h) \right\rangle_{y(h)}}{\|\mathcal{T}_{ht(x,\lambda)}(t(x,\lambda))\|_{y(h)}\|t(y(h),\lambda+h)\|_{y(h)}},$$

so that $\alpha(x,\lambda,h) = \mathrm{acos}(m(h))$. This is well-defined since $t(x,\lambda) \neq 0$ by assumption, which by continuity also implies that $t(y(h),\lambda+h) \neq 0$ provided h is sufficiently small. We will show that

$$m(h) = 1 - \frac{m''(0)}{2}h^2 + O(h^3). \tag{4.32}$$

Then, invoking the Puiseux series

$$\mathrm{acos}(1-y) = \sqrt{2y} + \frac{y^{3/2}}{6\sqrt{2}} + O\left(y^{5/2}\right),$$

and the Taylor expansion of the square root in $m''(0)h^2 \neq 0$ we may conclude that for h sufficiently small

$$\begin{aligned}
\alpha(x,\lambda,h) &= \mathrm{acos}\left(1 - \frac{m''(0)}{2}h^2 + O(h^3)\right)\\
&= \sqrt{m''(0)h^2 + O(h^3)} + \frac{(m''(0)\frac{h^2}{2} + O(h^3))^{3/2}}{6\sqrt{2}} + O\left((m''(0)\frac{h^2}{2} + O(h^3))^{5/2}\right)\\[2mm]
&= \sqrt{m''(0)}h + \frac{1}{2\sqrt{m''(0)}h}O(h^3) + O(h^3) + O(h^5)\\
&= \sqrt{m''(0)}h + O(h^2),
\end{aligned}$$

and therefore $\alpha_1(x,\lambda) = \sqrt{m''(0)}$. In order to prove (4.32), we introduce the notation

$$m(h) = \frac{\left\langle T(h), P(h) \right\rangle_{y(h)}}{\|T(h)\|_{y(h)}\|P(h)\|_{y(h)}} = \frac{A(h)}{B(h)},$$

where $T(h) := \mathcal{T}_{ht(x,\lambda)}(t(x,\lambda))$ is the transported prediction vector and $P(h) := t(y(h),\lambda+h)$ is the new prediction vector. These are both smooth vector fields along the curve $y(h)$. We indicate with dots their covariant differentiation along the curve, for instance

$\dot{P}(h) = \frac{\mathrm{D}}{\mathrm{d}h} t(y(h), \lambda + h)$. Note that

$$T(0) = P(0) = t(x, \lambda) =: t \tag{4.33}$$

and it is non-zero by assumption. Therefore $A(0) = B(0) = \|t\|_x^2$ implying that $m(0) = 1$. The first derivative of m therefore simplifies in $h = 0$ to

$$m'(0) = \frac{A'(0) - B'(0)}{\|t\|_x^2}.$$

We have

$$A'(0) = \left\langle \dot{T}(0), P(0) \right\rangle_x + \left\langle T(0), \dot{P}(0) \right\rangle_x,$$

and

$$B'(0) = \frac{1}{2\|t\|} \left(2\left\langle \dot{T}(0), T(0) \right\rangle_x \|t\|_x + 2\left\langle \dot{P}(0), P(0) \right\rangle_x \|t\|_x \right)$$
$$= \left\langle \dot{T}(0), P(0) \right\rangle_x + \left\langle T(0), \dot{P}(0) \right\rangle_x = A'(0).$$

Therefore we have $m'(0) = 0$. The second derivative of m reads as

$$m''(h) = \frac{\left(A''(h)B(h) - A(h)B''(h) \right)B(h) - 2B(h)B'(h)\left(A'(h)B(h) - A(h)B'(h) \right)}{B(h)^4}.$$

Thanks to the above, in $h = 0$ it simplifies to

$$m''(0) = \frac{A''(0) - B''(0)}{\|t\|^2}. \tag{4.34}$$

The second derivative of A is readily obtained in $h = 0$ as

$$A''(0) = \left\langle \ddot{T}(0) + \ddot{P}(0), t \right\rangle_x + 2\left\langle \dot{T}(0), \dot{P}(0) \right\rangle_x.$$

For B, we have

$$B''(h) = \frac{\mathrm{d}}{\mathrm{d}h} \left(\frac{\left\langle \dot{T}(h), T(h) \right\rangle \|P(h)\|_{y(h)}^2 + \left\langle \dot{P}(h), P(h) \right\rangle \|T(h)\|_{y(h)}^2}{B(h)} \right)$$
$$= \frac{B(h)}{B(h)^2} \left(\left(\left\langle \ddot{T}(h), T(h) \right\rangle_{y(h)} + \left\| \dot{T}(h) \right\|_{y(h)}^2 \right) \|P(h)\|_{y(h)}^2 + 2\left\langle \dot{T}(h), T(h) \right\rangle_{y(h)} \left\langle \dot{P}(h), P(h) \right\rangle_{y(h)} \right.$$

$$\left(\left\langle \ddot{P}(h), P(h) \right\rangle_{y(h)} + \left\| \dot{P}(h) \right\|_{y(h)}^2 \right) \|T(h)\|_{y(h)}^2 + 2\left\langle \dot{P}(h), P(h) \right\rangle_{y(h)} \left\langle \dot{T}(h), T(h) \right\rangle_{y(h)} \right)$$

$$- \frac{B'(h)}{B(h)^2} \left(\left\langle \dot{T}(h), T(h) \right\rangle \|P(h)\|_{y(h)}^2 + \left\langle \dot{P}(h), P(h) \right\rangle \|T(h)\|_{y(h)}^2 \right).$$

Evaluating in $h = 0$ leads to a simpler expression using (4.33) and the expression for $B'(0)$:

$$
\begin{aligned}
B''(0) = \frac{1}{\|t\|^2} \Bigg(&\left(\langle \ddot{T}(0), t \rangle_x + \|\dot{T}(0)\|_x^2 \right) \|t\|_x^2 + 2 \langle \dot{T}(0), t \rangle_x \langle \dot{P}(0), t \rangle_x \\
&\left(\langle \ddot{P}(0), t \rangle_x + \|P'(0)\|_x^2 \right) \|t\|_x^2 + 2 \langle \dot{P}(0), t \rangle_x \langle \dot{T}(0), t \rangle_x \Bigg) \\
&- \frac{B'(0)}{\|t\|^4} \left(\langle \dot{T}(0), t \rangle_x \|t\|_x^2 + \langle \dot{P}(0), t \rangle_x \|t\|_x^2 \right) \\
= &\langle \ddot{T}(0), t \rangle_x + \|\dot{T}(0)\|_x^2 + \langle \ddot{P}(0), t \rangle_x + \|\dot{P}(0)\|_x^2 \\
&\frac{1}{\|t\|_x^2} \left(4 \langle \dot{T}(0), t \rangle_x \langle \dot{P}(0), t \rangle_x - \left(\langle \dot{T}(0), t \rangle_x + \langle \dot{P}(0), t \rangle_x \right)^2 \right) \\
= &\langle \ddot{T}(0) + \ddot{P}(0), t \rangle_x + \|\dot{T}(0)\|_x^2 + \|\dot{P}(0)\|_x^2 - \frac{1}{\|t\|_x^2} \left(\langle \dot{T}(0), t \rangle_x - \langle \dot{P}(0), t \rangle_x \right)^2.
\end{aligned}
$$

Therefore, plugging the expression for $A''(0)$ and $B''(0)$ in (4.34) we find

$$
\begin{aligned}
m''(0) = \frac{1}{\|t\|_x^2} \Bigg(&\langle \ddot{T}(0) + \ddot{P}(0), t \rangle_x + 2 \langle \dot{T}(0), \dot{P}(0) \rangle_x - \langle \ddot{T}(0) + \ddot{P}(0), t \rangle_x \\
&- \|\dot{T}(0)\|_x^2 - \|\dot{P}(0)\|_x^2 + \frac{1}{\|t\|_x^2} \left(\langle \dot{T}(0), t \rangle_x - \langle \dot{P}(0), t \rangle_x \right)^2 \Bigg) \\
= &\frac{1}{\|t\|_x^4} \left(\left(\langle \dot{T}(0) - \dot{P}(0), t \rangle_x \right)^2 - \|\dot{T}(0) - \dot{P}(0)\|_x^2 \|t\|_x^2 \right) \\
= &\frac{1}{\|t\|_x^2} \left(\cos \left(\theta_x(\dot{T}(0) - \dot{P}(0), t) \right)^2 - 1 \right) \|\dot{T}(0) - \dot{P}(0)\|_x^2 \|t\|_x^2 \\
= &- \sin \left(\theta_x(\dot{T}(0) - \dot{P}(0), t) \right)^2 \frac{\|\dot{T}(0) - \dot{P}(0)\|_x^2}{\|t\|_x^2}.
\end{aligned}
$$

where we have used the notation $\cos(\theta_p(u, v)) = \frac{\langle u, v \rangle_p}{\|u\|_p \|v\|_p}$, for any $p \in \mathcal{M}$ and $u, v \in T_p \mathcal{M}$. By assumption (4.24), $m''(0)$ is not zero as $\dot{T}(0) - \dot{P}(0)$ is not zero and never colinear with t. This shows (4.32) and concludes the proof. $\square$

Proof based on local charts representation

The following proof of Lemma 4.7 is taken from the one in [SK22a, Lemma 3.5] and mimics the proof of the Euclidean case from [AG90, §6.1]. As anticipated, the problem is mapped to $\mathbb{R}^d$ making use of a local chart. However, in contrast with the proof technique used for Theorem 4.1, instead of defining the pullback of the objective function, the involved entities such as tangent vectors and linear maps are individually given a coordinate representation, as we now explain. A similar strategy is used in the local convergence proof of the Riemannian Newton in [AMS08, Theorem 6.3.2].

Consider a local chart $(\mathcal{U}, \varphi)$ on $\mathcal{M}$. To every point $p \in \mathcal{U}$ we associate $\hat{p} := \varphi(p) \in \mathbb{R}^d$

and for any tangent vector $w \in T_p\mathbb{R}^d$, the differential of the local chart at p allows mapping it to $\hat{w} := D\varphi(w)[w] \in \mathbb{R}^d$. Conversely, for every $\hat{p} \in \hat{\mathcal{U}} := \varphi(\mathcal{U})$, we write $p := \varphi^{-1}(\hat{p})$ and any vector $\hat{w} \in T_{\hat{p}}\mathbb{R}^d \simeq \mathbb{R}^d$ becomes $w = D\varphi^{-1}(\hat{p})[\hat{w}]$. As in the proof of Theorem 4.1, the addition or removal of a hat decoration on a symbol implicitly indicates mapping it to local coordinates and vice-versa. The Riemannian metric can be applied to tangent vectors in local coordinates using the Gramian matrix $\hat{G}(\hat{p})$ defined for every $\hat{p} \in \hat{\mathcal{U}}$ by

$$(\hat{G}(\hat{p}))_{ij} = \left\langle D\varphi^{-1}(\hat{p})[e_i], D\varphi^{-1}(\hat{p})[e_j] \right\rangle_p, \quad \forall i, j = 1, \ldots, d$$

where e_i are the canonical vectors of $\mathbb{R}^d$. Then, for every $u, v \in T_p\mathcal{M}$, it holds that $\langle u, v \rangle_p = \hat{u}^\top \hat{G}(\hat{p}) \hat{v}$. The same holds for the norm in local coordinates: it is denoted $\|\hat{v}\|_{\hat{p}} := (\hat{v}^\top \hat{G}(\hat{p})\hat{v})^{1/2}$ and is equal to $\|v\|_p$. By smoothness of the Riemannian metric and of the local charts, the function $\hat{p} \to \hat{G}(\hat{p})$ is also smooth. Any linear map $A : T_p\mathcal{M} \to T_q\mathcal{M}$ between tangent spaces of some $p, q \in \mathcal{U}$ translates to local coordinates as $\hat{A} := D\varphi(q) \circ A \circ D\varphi^{-1}(\hat{p})$. Finally, the retraction can also be read in local coordinates. Given a sufficiently small $v \in T_p\mathcal{M}$ for some $p \in \mathcal{U}$, the retracted point $R_p(v)$ is well-defined and belongs to $\mathcal{U}$. In this scenario, the retraction in local coordinates becomes

$$\hat{R}_{\hat{p}}(\hat{v}) := \varphi(R_p(D\varphi^{-1}(\hat{p})[\hat{v}])).$$

As also mentioned in [AMS08, Theorem 6.3.2], we point out that the local rigidity property of the retraction transfers to its local chart version, i.e.

$$D\hat{R}_{\hat{p}}(0)[\hat{w}] = \hat{w}, \quad \forall \hat{p} \in \hat{\mathcal{U}}, \hat{w} \in \mathbb{R}^d.$$

Proof of Lemma 4.7 based on local coordinates. Consider a sufficient small h so that the solution point x, the predicted point $y(h)$ and the first RN iterate $z(h)$ are contained in the domain of the same local chart $(\mathcal{U}, \varphi)$. Their coordinates representations are denoted $\hat{x}, \hat{y}(h)$ and $\hat{z}(h)$, respectively. For the tangential prediction vector t and the RN update vector n we stick to the convention and write their coordinate representation as $\hat{t}$ and $\hat{n}$ while for the Riemannian gradient and its derivatives with respect to λ the notation is simplified to

$$\hat{F}(\hat{p}, \lambda) := D\varphi(p)[\text{grad}f(p, \lambda)], \quad \hat{F}_\lambda(\hat{p}, \lambda) := D\varphi(p)\left[\frac{\partial}{\partial \lambda} \text{grad}f(p, \lambda)\right],$$

for all $\hat{p} \in \hat{\mathcal{U}}$ and $\lambda \in [0, 1]$. The coordinate representation of the Riemannian Hessian of $f(\cdot, \lambda)$ at $p \in \mathcal{U}$ is denoted $\hat{H}(\hat{p}, \lambda) : \hat{\mathcal{U}} \times [0, 1] \to \mathbb{R}^{d \times d}$ and defined as

$$\hat{H}(\hat{p}, \lambda) := D\varphi(p) \circ \text{Hess}f(p, \lambda) \circ D\varphi(\hat{p})^{-1},$$

Note that this representation is different from the one considered in Lemma 4.3. For the transport map along the prediction vector at (p, λ) with step size h we adopt the

notation $\mathcal{T}(p, \lambda, h) := \mathcal{T}_{ht(p,\lambda)}$ and consider its coordinates representation

$$\hat{\mathcal{T}}(\hat{p}, \lambda, h) := D\varphi(R_p(ht(p, \lambda))) \circ \mathcal{T}(p, \lambda, h) \circ D\varphi(\hat{p})^{-1}.$$

Note that the function $\hat{F}_\lambda$ defined above coincides with the derivative of $\hat{F}$ with respect to λ. However, the differential of $\hat{F}$ with respect to its first argument, denoted $\hat{F}_{\hat{x}}$, does *not* necessarily coincide with $\hat{H}$, the coordinate representation of the Hessian. Indeed, for any $\hat{p} \in \hat{\mathcal{U}}$ and $\hat{w} \in \mathbb{R}^d$ and any curve $\hat{\gamma}_{\hat{p},\hat{w}}$ such that $\hat{\gamma}_{\hat{p},\hat{w}}(0) = \hat{p}$ and $\hat{\gamma}'_{\hat{p},\hat{w}}(0) = \hat{w}$ one obtains

$$
\begin{aligned}
\hat{F}_{\hat{x}}(\hat{p}, \lambda)[\hat{w}] &= \hat{H}(\hat{p}, \lambda)[\hat{w}] + \frac{d}{dt} D\varphi(\varphi^{-1}(\hat{\gamma}_{\hat{p},\hat{w}}(t)))\big|_{t=0} \left[\text{grad} f(p, \lambda)\right] \\
&= \hat{H}(\hat{p}, \lambda)[\hat{w}] + \hat{A}(\hat{p})\big[\hat{F}(\hat{p}, \lambda), \hat{w}\big],
\end{aligned}
\tag{4.35}
$$

where $\hat{A}(\hat{p})$ is the bilinear form

$$\hat{A}(\hat{p})\left[\hat{v}, \hat{w}\right] = \frac{d}{dt} D\varphi(\varphi^{-1}(\hat{\gamma}_{\hat{p},\hat{w}}(t)))\big|_{t=0} \left[(D\varphi(\varphi^{-1}(\hat{p})))^{-1}\left[\hat{v}\right]\right].$$

On the solution curve, we have $\hat{F}(\hat{x}, \lambda) = 0$, so the second term in (4.35) vanishes for $\hat{p} = \hat{x}$ and we find $\hat{F}_{\hat{x}}(\hat{x}, \lambda) = \hat{H}(\hat{x}, \lambda)$. This fact echoes the link between the Hessian of the pullback and the Riemannian Hessian of the objective function at a critical point discussed in Lemma 4.3. By the definitions above, the coordinate representations of $y(h)$, $z(h)$, $t(w, \lambda)$ and $n(w, \lambda)$ have the following convenient expressions

$$\hat{y}(h) = \hat{R}_{\hat{x}}(h\hat{t}(\hat{x}, \lambda)), \tag{4.36}$$

$$\hat{z}(h) = \hat{R}_{\hat{y}(h)}(\hat{n}(\hat{y}(h), \lambda + h)), \tag{4.37}$$

$$\hat{t}(\hat{w}, \lambda) = -\hat{H}(\hat{w}, \lambda)^{-1}\big[\hat{F}_\lambda(\hat{w}, \lambda)\big], \tag{4.38}$$

$$\hat{n}(\hat{w}, \lambda) = -\hat{H}(\hat{w}, \lambda)^{-1}\big[\hat{F}(\hat{w}, \lambda)\big]. \tag{4.39}$$

Using the previous definitions and omitting the explicit dependence of $\hat{\mathcal{T}}$ on $\hat{x}$ and λ we evince the following expressions:

$$
\begin{aligned}
\delta(x, \lambda, h) &= \sqrt{\hat{n}(\hat{y}(h), \lambda + h)^\top \hat{G}(\hat{y}(h))\hat{n}(\hat{y}(h), \lambda + h)}, \\
\kappa(x, \lambda, h) &= \frac{\sqrt{\hat{n}(\hat{z}(h), \lambda + h)^\top \hat{G}(\hat{z}(h))\hat{n}(\hat{z}(h), \lambda + h)}}{\delta(x, \lambda, h)}, \\
\alpha(x, \lambda, h) &= \text{acos}\left(\frac{(\hat{\mathcal{T}}(h)\hat{t}(\hat{x}, \lambda))^\top \hat{G}(\hat{y}(h))\hat{t}(\hat{y}(h), \lambda + h)}{\|\hat{\mathcal{T}}(h)\hat{t}(\hat{x}, \lambda)\|_{\hat{y}(h)}\|\hat{t}(\hat{y}(h), \lambda + h)\|_{\hat{y}(h)}}\right)
\end{aligned}
$$

We are now in the position to perform Taylor expansion with respect to h of these functions. The dependence on h is smooth by smoothness of the operations involved in

their definition.

Result (i)
Combining the characterizations of $\hat{y}$ and $\hat{n}$ in (4.36) and (4.39) we have

$$\hat{n}(\hat{y}(h), \lambda + h) = -\hat{H}(\hat{R}_{\hat{x}}(h\hat{t}(\hat{x}, \lambda)), \lambda + h)^{-1}\big[\hat{F}(\hat{R}_{\hat{x}}(h\hat{t}(\hat{x}, \lambda)), \lambda + h)\big].$$

Let us expand both terms separately.

$$\begin{aligned}
\hat{F}(\hat{R}_{\hat{x}}(h\hat{t}(\hat{x}, \lambda)), \lambda + h) &= \hat{F}(\hat{x}, \lambda) + h\Big(\hat{F}_{\hat{x}}(\hat{x}, \lambda)\big[\hat{t}(\hat{x}, \lambda)\big] + \hat{F}_\lambda(\hat{x}, \lambda)\Big) + h^2 c_1(\hat{x}, \lambda) + O\left(h^3\right) \\
&= h^2 c_1(\hat{x}, \lambda) + O\left(h^3\right),
\end{aligned}$$
$$(4.40)$$

where the second equality follows from $\hat{F}(\hat{x}, \lambda) = 0$, $\hat{F}_{\hat{x}}(\hat{x}, \lambda) = \hat{H}(\hat{x}, \lambda)$ and (4.38).
Now note that
$$\hat{H}(\hat{R}_{\hat{x}}(h\hat{t}(\hat{x}, \lambda)), \lambda + h) = \hat{H}(\hat{x}, \lambda) + O\left(h\right).$$

Then by smoothness of matrix inversion

$$\hat{H}(\hat{R}_{\hat{x}}(h\hat{t}(\hat{x}, \lambda)), \lambda + h)^{-1} = \hat{H}(\hat{x}, \lambda)^{-1} + O\left(h\right).$$

Combined with (4.40) one has

$$\hat{n}(\hat{y}(h), \lambda + h) = h^2 c_2(\hat{x}, \lambda) + O\left(h^3\right). \tag{4.41}$$

with $c_2(\hat{x}, \lambda) = -\hat{H}(\hat{x}, \lambda)^{-1}[c_1(\hat{x}, \lambda)]$. Noting that $\hat{G}(\hat{y}(h)) = \hat{G}(\hat{x}) + O\left(h\right)$ we obtain

$$\begin{aligned}
\delta(x, \lambda, h) &= \sqrt{\hat{n}(\hat{y}(h), \lambda + h)^\top \hat{G}(\hat{y}(h))\hat{n}(\hat{y}(h), \lambda + h)} \\
&= (h^4 c_3(\hat{x}, \lambda)^2 + O\left(h^5\right))^{1/2} \\
&= h^2 c_3(\hat{x}, \lambda) + O\left(h^3\right),
\end{aligned}$$

where $c_3(\hat{x}, \lambda) := (c_2(\hat{x}, \lambda)^\top \hat{G}_{\hat{x}} c_2(\hat{x}, \lambda))^{1/2}$. The last equality follows from the Taylor expansion of the square root in $c_3(\hat{x}, \lambda)^2$. This is possible provided $c_3(\hat{x}, \lambda)$ does not vanish. From the proof of Lemma 4.7 based on covariant differentiation, and by uniqueness of the Taylor expansion of

$$\|n(y(h), \lambda + h)\|_{y(h)}^2 = \|\hat{n}(\hat{y}(h), \lambda + h)\|_{\hat{y}(h)}^2$$

we have that $c_3(x, \lambda)^2 = \frac{1}{4}\big\|\frac{\mathrm{D}^2}{\mathrm{d}h^2} n(y(h), \lambda + h)\big|_{h=0}\big\|^2$. Thanks to assumption (4.22), this was shown to be non-zero and concludes the proof of (i) with $\delta_2(x, \lambda) := c_3(\varphi(x), \lambda)$.

Result (ii)
To obtain the expansion for κ, we combine result (i) with the expansion of the Newton

direction evaluated in $\hat{z}(h)$. For this purpose, note that by combining (4.37) and (4.39)

$$\hat{n}(\hat{z}(h), \lambda + h) = -\hat{H}(\hat{R}_{\hat{y}(h)}(\hat{n}(\hat{y}(h), \lambda + h), \lambda + h)^{-1} \left[\hat{F}(\hat{R}_{\hat{y}(h)}(\hat{n}(\hat{y}(h), \lambda + h), \lambda + h)) \right].$$
(4.42)

The Taylor expansion with respect to $\hat{n}(\hat{y}(h), \lambda + h)$ of the right-hand side term gives

$$\hat{F}\left(\hat{R}_{\hat{y}(h)}(\hat{n}(\hat{y}(h), \lambda + h)), \lambda + h \right) = \hat{F}(\hat{y}(h), \lambda + h) + \hat{H}(\hat{y}(h), \lambda + h)\left[\hat{n}(\hat{y}(h), \lambda + h)\right]$$

$$+ \hat{A}(\hat{y}(h)) \left[\hat{F}(\hat{y}(h), \lambda + h), \hat{n}(\hat{y}(h), \lambda + h) \right]$$

$$+ \frac{1}{2}\hat{F}_{\hat{x}\hat{x}}(\hat{y}(h), \lambda + h)\left[\hat{n}(\hat{y}(h), \lambda + h), \hat{n}(\hat{y}(h), \lambda + h)\right]$$

$$+ \frac{1}{2}\hat{F}_{\hat{x}}(\hat{y}(h), \lambda + h)\left[\mathrm{D}^2 \hat{R}_{\hat{x}}(0)\left[\hat{n}(\hat{y}(h), \lambda + h), \hat{n}(\hat{y}(h), \lambda + h)\right]\right] + O\left(\|\hat{n}(\hat{y}(h), \lambda + h)\|^3\right).$$

The first two summands cancel out owing to (4.39). Furthermore, by smoothness of the retraction and of the local charts, we have

$$\hat{A}(\hat{y}(h)) = \hat{A}(\hat{x}) + O\left(h\right),$$
$$\hat{F}_{\hat{x}\hat{x}}(\hat{y}(h), \lambda + h) = \hat{F}_{\hat{x}\hat{x}}(\hat{x}, \lambda) + O\left(h\right),$$
$$\hat{F}_{\hat{x}}(\hat{y}(h), \lambda + h) \circ \mathrm{D}^2 \hat{R}_{\hat{y}(h)}(0) = \hat{F}_{\hat{x}}(\hat{x}, \lambda) \circ \mathrm{D}^2 \hat{R}_{\hat{x}}(0) + O\left(h\right).$$

By plugging in the Taylor expansions of $\hat{n}(\hat{y}(h), \lambda + h)$ and $\hat{F}(\hat{y}(h), \lambda + h)$ given by (4.40) and (4.41) respectively we obtain

$$\hat{F}(\hat{z}(h), \lambda + h) = h^4 c_4(\hat{x}, \lambda) + O\left(h^5\right),$$

for some $c_4(\hat{x}, \lambda)$ not depending on h.

Now, for the left-hand side term in (4.42), the Taylor expansion with respect to $\hat{n}(\hat{y}(h), \lambda + h)$ gives

$$\hat{H}(\hat{z}(h), \lambda + h) = \hat{H}(\hat{y}(h), \lambda + h) + O\left(\|\hat{n}(\hat{y}(h), \lambda + h)\|\right) = \hat{H}(\hat{x}, \lambda) + O\left(h\right),$$

and thus

$$\hat{H}(\hat{z}(h), \lambda + h)^{-1} = \hat{H}(\hat{x}, \lambda)^{-1} + O\left(h\right).$$

Therefore

$$\hat{n}(\hat{z}(h), \lambda + h) = h^4 c_5(\hat{x}, \lambda) + O\left(h^5\right),$$

where $c_5(\hat{x}, \lambda) = -\hat{H}(\hat{x}, \lambda)^{-1}\left[c_4(\hat{x}, \lambda)\right]$. Finally, noticing that $\hat{G}(\hat{z}(h)) = \hat{G}(\hat{x}) + O(h)$, we can approximate the numerator of κ as

$$\|\hat{n}(\hat{z}(h), \lambda + h)\|_{\hat{z}(h)} = \left(h^8 c_6(\hat{x}, \lambda)^2 + O\left(h^9\right)\right)^{1/2} = h^4 c_6(\hat{x}, \lambda) + O\left(h^4\right),$$

with $c_6(\hat{x}, \lambda) = \|c_5(\hat{x}, \lambda)\|_{\hat{x}}$. This allows us to conclude that

$$\kappa(x, \lambda, h) = \frac{h^4 c_6(\hat{x}, \lambda) + O(h^4)}{h^2 c_3(\hat{x}, \lambda) + O(h^3)} = h^2 c_7(\hat{x}, \lambda) + O(h^2),$$

with $c_7(\hat{x}, \lambda) = \frac{c_6(\hat{x}, \lambda)}{c_3(\hat{x}, \lambda)}$, where we used the Taylor expansion of the inverse function in $c_3(\hat{x}, \lambda)$, which is non-zero as noted for result (i). This proves the expansion (ii) with $\kappa_2(x, \lambda) = c_7(\varphi(x), \lambda)$.

Result (iii)

The proof is essentially identical to the one based on covariant differentiation up to small variations due to the presence of the matrix $\hat{G}(\hat{p})$ in the inner products. With the coordinate representation of the involved quantities, for a fixed (x, λ) we can express α as

$$\alpha(x, \lambda, h) = \mathrm{acos}(\hat{m}(h))$$

where

$$\hat{m}(h) := \frac{\hat{T}(h)^\top \hat{G}_{\hat{y}}(h) \hat{P}(h)}{\sqrt{\hat{T}(h)^\top \hat{G}_{\hat{y}}(h) \hat{T}(h) \hat{P}(h)^\top \hat{G}_{\hat{y}}(h) \hat{P}(h)}} =: \frac{\hat{A}(h)}{\hat{B}(h)}$$

and

$$\hat{T}(h) := \hat{\mathcal{T}}(h)\hat{t}(\hat{x}, \lambda), \quad \hat{P}(h) := \hat{t}(\hat{y}(h), \lambda + h), \quad \hat{G}_{\hat{y}}(h) = \hat{G}(\hat{y}(h)).$$

The well-posedness of $\hat{m}$ is guaranteed by the assumption $t(x, \lambda) \neq 0$: in fact this implies $\hat{t}(\hat{x}, \lambda) \neq 0$ and in turn $\hat{T}(h) \neq 0$ as well as $\hat{P}(h) \neq 0$, by continuity provided h is small enough. Let us show that for sufficiently small h it holds that

$$\hat{m}(h) = 1 - \frac{1}{2}\hat{m}''(0) + O(h^3).$$

Then, the result follows using the Puiseux series for $\mathrm{acos}(1 - y)$ in the same way as in the proof based in covariant differentiation. Retracing the steps of the coordinate-free proof, we have that

$$\hat{T}(0) = \hat{P}(0) = \hat{t}(\hat{x}, \lambda) =: \hat{t}.$$

Also $\hat{G}_{\hat{y}}(0) = \hat{G}(\hat{x})$. This implies that $\hat{A}(0) = \hat{B}(0) = \|\hat{t}\|_{\hat{x}}$ and so $\hat{m}(0) = 1$. From this, we deduce that in $h = 0$, the derivative of $\hat{m}$ simplifies to

$$\hat{m}'(0) = \frac{\hat{A}'(0) - \hat{B}'(0)}{\|\hat{t}\|_x^2}.$$

Applying the product rule and the fact $\hat{T}(0) = \hat{P}(0) = \hat{t}$ we get, on the one hand,

$$\hat{A}'(0) = \hat{T}'(0)^\top \hat{G}_{\hat{y}}(\hat{x})\hat{t} + \hat{t}^\top \hat{G}'_{\hat{y}}(0)\hat{t} + \hat{t}^\top \hat{G}_{\hat{y}}(\hat{x})\hat{P}'(0)$$

On the other hand we have

$$
\begin{aligned}
\hat{B}'(h) = \frac{1}{2\hat{B}(h)} \Bigg(& \left(2\hat{T}'(h)^\top \hat{G}_{\hat{y}}(h)\hat{T}(h) + \hat{T}(h)^\top \hat{G}'_{\hat{y}}(h)\hat{T}(h) \right) \left\| \hat{P}(h) \right\|^2_{\hat{y}(h)} \\
& + \left(2\hat{P}'(h)^\top \hat{G}_{\hat{y}}(h)\hat{P}(h) + \hat{P}(h)^\top \hat{G}'_{\hat{y}}(h)\hat{P}(h) \right) \left\| \hat{T}(h) \right\|^2_{\hat{y}(h)} \Bigg), \quad (4.43)
\end{aligned}
$$

which in $h = 0$ simplifies to

$$
\hat{B}'(0) = \hat{T}'(0)^\top \hat{G}(\hat{x})\hat{t} + \hat{t}^\top \hat{G}'_{\hat{y}}(0)\hat{t} + \hat{t}^\top \hat{G}(\hat{x})\hat{P}'(0) = \hat{A}'(0).
$$

Therefore, it holds that $\hat{m}'(0) = 0$. In the same way as in the coordinate free proof, since $\hat{A}(0) = \hat{B}(0) = \left\| \hat{t} \right\|^2_{\hat{x}}$ and $\hat{A}'(0) = \hat{B}'(0)$ we have

$$
\hat{m}''(0) = \frac{\hat{A}''(0) - \hat{B}''(0)}{\left\| \hat{t} \right\|^2_{\hat{x}}}.
$$

Differentiating $\hat{A}'$ once again in $h = 0$ yields

$$
\begin{aligned}
\hat{A}''(0) = {} & \hat{T}''(0)^\top \hat{G}(\hat{x})\hat{t} + \hat{t}^\top \hat{G}(\hat{x})\hat{P}''(0) + \hat{t}^\top \hat{G}''_{\hat{y}}(0)\hat{t} \\
& 2\left(\hat{T}'(0)^\top \hat{G}'_{\hat{y}}(0)\hat{t} + \hat{t}^\top \hat{G}'_{\hat{y}}(0)\hat{P}'(0) + \hat{T}'(0)^\top \hat{G}(\hat{x})\hat{P}'(0) \right).
\end{aligned}
$$

The derivative of (4.43) is obtained from a lengthy calculations that we therefore omit. The expression simplifies when evaluated in $h = 0$ and most terms of $\hat{A}''(0)$ cancel out with terms in $\hat{B}''(0)$, in particular all terms involving a second derivative. Ultimately, we find

$$
\begin{aligned}
\hat{m}''(0) = {} & \frac{1}{\left\| \hat{t} \right\|^2_{\hat{x}}} \left(2\hat{T}'(0)^\top \hat{G}(\hat{x})\hat{P}'(0) - \left\| \hat{T}'(0) \right\|^2_{\hat{x}} - \left\| \hat{P}'(0) \right\|^2_{\hat{x}} \right) \\
& + \frac{1}{\left\| \hat{t} \right\|^4_x} \left(\left(\hat{T}'(0) - \hat{P}'(0) \right)^\top \hat{G}(\hat{x})\hat{t} \right)^2 \\
= {} & \frac{\left(\left(\hat{T}'(0) - \hat{P}'(0) \right)^\top \hat{G}(\hat{x})\hat{t} \right)^2 - \left\| \hat{T}'(0) - \hat{P}'(0) \right\|^2_{\hat{x}} \left\| \hat{t} \right\|^2_{\hat{x}}}{\left\| \hat{t} \right\|^4_x} \\
= {} & -\sin\left(\hat{\theta}_{\hat{x}}(\hat{T}'(0) - \hat{P}'(0), \hat{t}) \right)^2 \frac{\left\| \hat{T}'(0) - \hat{P}'(0) \right\|^2_{\hat{x}}}{\left\| \hat{t} \right\|^2_{\hat{x}}} \\
= {} & -\sin\left(\theta_x(\dot{T}(0) - \dot{P}(0), t) \right)^2 \frac{\left\| \dot{T}(0) - \dot{P}(0) \right\|^2_x}{\left\| t \right\|^2_x}.
\end{aligned}
$$

where we have introduced $\hat{\theta}_{\hat{p}}(\hat{u}, \hat{v}) = \frac{\hat{u}^\top \hat{G}(\hat{p})\hat{v}}{\|\hat{u}\|_{\hat{p}} \|\hat{v}\|_{\hat{p}}}$ and the last step follows from definition of the coordinates representations. Assumption (4.24) implies this is non-zero and this concludes the proof with $\alpha_1(x, \lambda) = \sqrt{\hat{m}''(0)}$. $\qquad \square$

4.4   Application to the Karcher mean of symmetric positive definite matrices

In this section, we apply the RNC method, Algorithm 4.1, to a classical problem of Riemannian optimization: the computation of the Karcher mean, also referred to as Riemannian center of mass. This concept first appeared in [GK73] and was studied thoroughly in [Kar77]. The Karcher mean was introduced as a generalization of the concept of center of mass for mass distributions defined on a manifold. Here we focus on the case where the mass in concentrated in discrete points with uniform weights. In other words, given K points $p_1, \ldots, p_K$ on a Riemannian manifold $\mathcal{M}$, their Karcher mean is defined as

$$\underset{q \in \mathcal{M}}{\arg\min} \left\{ \sum_{i=1}^{K} d(q, p_i)^2 \right\}, \tag{4.44}$$

where $d(q, p_i)$ is the Riemannian distance function on $\mathcal{M}$ (see Definition 1.21). As a consequence of a result applicable for arbitrary mass distributions [Kar77, Theorem 1.2], if the manifold is complete , simply connected and of non-positive sectional curvature[2], a class of manifolds sometimes called Cartan-Hadamard manifolds, the Karcher mean exists and is unique for every set of points on the manifold. On a general Riemannian manifold, problem (4.44) admits a unique solution provided that all p_i are sufficiently close to each other. A possible justification comes from the study of weighted geodesic averages of K points [AGSW16], defined as the solution of

$$\underset{q \in \mathcal{M}}{\arg\min} \left\{ \sum_{i=1}^{K} w_i d(q, p_i)^2 \right\},$$

for some weights w_i such that $\sum_{i=1}^{K} w_i = 1$. Around every point on the manifold there exists a sufficiently small neighborhood where all weighted geodesic averages of any set of points in that neighborhood are uniquely defined [AGSW16, Proposition 12]; see also [San16, Theorem 3.1] for the case of arbitrary mass distributions. This includes the Karcher mean of the points, which is equivalently obtained as the weighted geodesic average with equal weights $w_i = 1/K$.

We focus on the Karcher mean of $n \times n$ real symmetric positive definite matrices, viewed as points of the $\mathbb{S}_n^+$ manifold endowed with the bi-invariant Riemannian metric presented in Section 1.5.2. Under this geometry, $\mathbb{S}_n^+$ is a Cartan-Hadamard manifold, hence the Karcher mean is always uniquely defined. The Karcher mean is one of many possibilities for generalizing the concept of average to a collection of symmetric positive definite matrices. A survey and comparison of existing methods can be found in [JVV12].

Computing the Karcher mean of K matrices of $\mathbb{S}_n^+$ can be addressed using Riemannian optimization. The objective function can be given an explicit expression which in turn allows computing its Riemannian gradient and Riemannian Hessian. With the chosen

---

[2]The definition of sectional curvature is reported in Section 6.4.1

Riemannian metric, the Riemannian distance function admits the analytical expression given in (1.13) and the Karcher mean problem (4.44) for $A_1, \ldots, A_K \in \mathbb{S}_n^+$ becomes

$$\operatorname*{arg\,min}_{X \in \mathbb{S}_n^+} \left\{ f(X) := \sum_{i=1}^{K} \| \log(A_i^{-\frac{1}{2}} X A_i^{-\frac{1}{2}}) \|_F^2 \right\}. \tag{4.45}$$

The Riemannian gradient can be derived by rewriting the objective function. Recalling that similarity transformations commute with the matrix logarithm, we have that $\log(A_i^{-1/2} X A_i^{-1/2}) = A_i^{1/2} \log(A_i^{-1} X) A_i^{-1/2}$, and cyclic invariance of the trace yields

$$\| \log(A_i^{-\frac{1}{2}} X A_i^{-\frac{1}{2}}) \|_F^2 = \operatorname{Tr}\left(\log\left(A_i^{-1} X\right)^2 \right).$$

From the identity [Moa05, Proposition 2.1]

$$\frac{\mathrm{d}\operatorname{Tr}\left(\log\left(X(t)\right)^2 \right)}{\mathrm{d}t} = 2\operatorname{Tr}\left(\log(X(t)) X(t)^{-1} X'(t) \right)$$

and the expression of the bi-invariant metric (1.9), the Riemannian gradient of f under this Riemannian geometry of $\mathbb{S}_n^+$ reads as [JVV12, §4.1.2]

$$\operatorname{grad}f(X) = \sum_{i=1}^{K} 2X \log(A_i^{-1} X). \tag{4.46}$$

The Riemannian connection associated with the bi-invariant metric of $\mathbb{S}_n^+$ admits an explicit expression [JVV12, §4.1.4]. This makes it possible to express the Riemannian Hessian of f at along any $V \in T_X \mathbb{S}_n^+$ from its Definition 2.5 as

$$\operatorname{Hess}f(X)[V] = \sum_{i=1}^{N} 2X \mathrm{D}\log(A_i^{-1} X) \left[A_i^{-1} V \right] + V \log(A_i^{-1} X) - \log(A_i^{-1} X) V,$$

where $\mathrm{D}\log(X)[\cdot]$ is the Fréchet derivative of the matrix logarithm, the linear operator defined by

$$\lim_{h \to 0} \frac{\log(X + hY) - \log(X) - \mathrm{D}\log(X)[Y]}{h} = 0.$$

It can be estimated numerically as shown in [AMHR13].

4.4.1 Homotopy for the Karcher mean problem

The bi-invariant Riemannian manifold structure for $\mathbb{S}_n^+$ guarantees the Karcher mean of positive definite matrices is always uniquely defined. The objective function (4.45) is in fact strictly geodesically convex on $\mathbb{S}_n^+$ [Bha07, Exercise 6.1.13], the generalization of convexity for manifold scalar field [Bou23, Definition 11.5]. This implies that any standard Riemannian optimization algorithm that is globally convergent can successfully solve the problem since the objective function admits a unique critical point which coincides with

the only global minimum [Bou23, Theorem 11.7]. Nevertheless, we use this application as a model problem for the purpose of testing the RNC algorithm and illustrating its behavior.
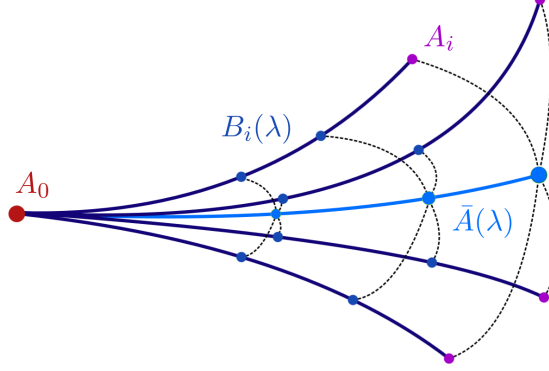

Figure 4.6: Homotopy for the Karcher mean problem. The parameter dependent Karcher mean is denoted $\bar{A}(\lambda)$.

We propose the following homotopy for the Karcher mean of $A_1, \ldots, A_K \in \mathbb{S}_n^+$. We define K smooth curves $B_i : [0,1] \to \mathbb{S}_n^+$ such that

$$B_i(1) = A_i, \quad \forall\, i = 1, \ldots, K,$$

and such that the Karcher mean of $B_1(0), \ldots, B_K(0)$ can be solved easily. In particular, this is the case when all starting points are equal, $B_1(0) = \cdots = B_K(0) = A_0$, their Karcher mean is simply A_0. We report results for experiments where $A_0 = I_{n \times n}$. Other choices for A_0 are possible, e.g. the planar approximations of the Karcher mean discussed in [JVV12]. The curves B_i are endpoint geodesic curve joining A_0 with A_i. From their analytical expression (1.11), we can write

$$B_i(\lambda) = A_0 \exp(\lambda \log(A_0^{-1} A_i)).$$

Then, the parametric Karcher mean optimization problem is given by

$$\operatorname*{arg\,min}_{X \in \mathbb{S}_n^+} \left\{ f(X, \lambda) := \sum_{i=1}^{K} \| \log(B_i(\lambda)^{-\frac{1}{2}} X B_i(\lambda)^{-\frac{1}{2}}) \|_F^2 \right\}, \quad \forall\, \lambda \in [0,1]. \qquad (4.47)$$

The Riemannian gradient of the parameter dependent objective (4.47) is obtained from the original Riemannian gradient (4.46) by simply replacing the matrices A_i with the curves $B_i(\lambda)$. Then, its derivative with respect to the parameter λ, needed for performing tangential continuation, is given by

$$\frac{\partial \operatorname{grad} f(X, \lambda)}{\partial \lambda} = -2 \sum_{i=1}^{K} X \mathrm{D}\log(X^{-1} B_i(\lambda)) \left[X^{-1} B_i'(\lambda) \right],$$

where $B_i'(\lambda) = A_0 \exp(\lambda \log(A_0^{-1} A_i)) \log(A_0^{-1} A_i)$.

4.4.2   Numerical results

All numerical experiments presented in this chapter have been performed in Matlab
2019b, using the Matlab Riemannian optimization library Manopt [BMAS14]. It contains
a factory class for the bi-invariant Riemannian structure of $\mathbb{S}_n^+$ presented in section 1.5.2
of which we use the retraction (2.10) and the vector transport given by parallel transport
along geodesics (1.14) for the following experiments.

All considered instances of the Karcher mean problem consist of a set of $K = 75$ sym-
metric positive definite matrices of size $n = 10$ that are built from their eigenvalue
decomposition

$$A_i = V_i D_i V_i^\top, \quad \forall\, i = 1, \dots, K,$$

where V_i is a random orthogonal matrix and D_i a diagonal matrix. For the diagonal
entries, 9 are chosen at random in the interval $[1, 2]$ and the last one is chosen such that
the matrices have a large but still moderate condition number (approximately 10^3). Fig-
ure 4.7 compares the direct optimization with the standard Riemannian Newton (RN)
method and the continuation approach on the homotopy (4.47) using Riemannian New-
ton Continuation (RNC) with a fixed number of steps $N_{\text{steps}} = 3$ (see Algorithm 4.1).
The correction step, here the RN method, is terminated when the gradient norm is below
$\varepsilon_{\text{tol}} = 10^{-6}$ or the number of Newton iterations exceeds $N_{\text{inner}} = 5000$. For this example,
it turns out that the RN method enters a superlinear convergence regime from the be-
ginning (as seen from the concavity of the black convergence curve) and thus solves the
problem in very few iterations. For such a simple instance, the continuation approach
does not offer advantages.

In order to better highlight the advantage of the RNC algorithm, we choose a somewhat
pathological instance: the diagonal matrices D_i are chosen such that their condition
number is 10^8. Half of the diagonal entries are exponentially distributed in $[0.1, 1]$ and
the other half exponentially distributed in $\left[10^6, 10^7\right]$. In turn, the optimization problem
is highly ill-conditioned, leading to stagnation in the initial phase of the RN method; see
Figure 4.8. In contrast, the RNC algorithm with fixed number of steps $N_{\text{steps}} = 2$ does
not suffer from such stagnation during the correction phase. In turn, the total number
of RN iterations is reduced. Tangential prediction leads to slightly better convergence
compared to classical prediction, but it also comes at the cost of solving an extra linear
system, which leads to a less favorable computational time; see Table 4.1. The number
of fixed steps in Figure 4.8 is chosen to best highlight the slight improvement of RNC
over direct RN optimization. However, for this particular application the advantage
disappears when an automatic step sizing strategy is used. Nevertheless, the step size
adaptation results for different set of hyperparameters $(\kappa_{\max}, \alpha_{\max}, \delta_{\max})$ in Table 4.1
illustrate the need for a compromise to be found between the number of corrections and
the length of each correction. This is further demonstrated by Figure 4.9 where the
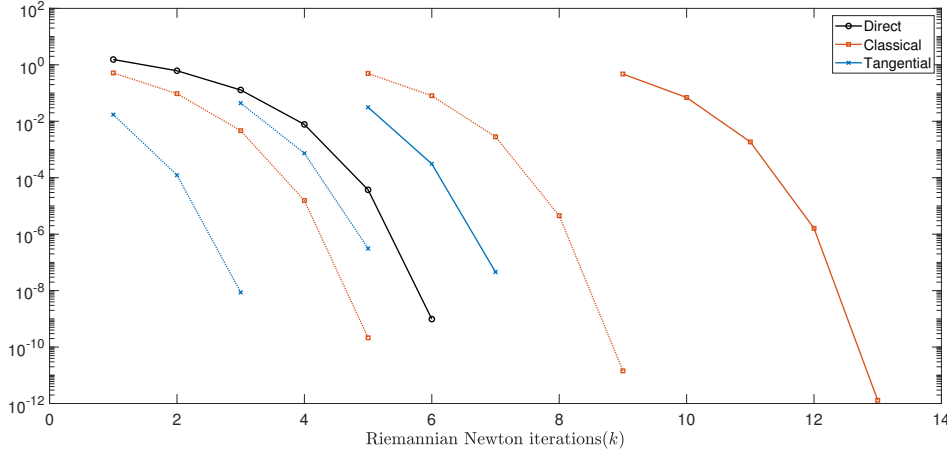computational effort for fixed step size RNC is reported for different values of N_{steps}.

Figure 4.7: Convergence of the Riemannian gradient norm versus RN iterations for a non-pathological instance of the Karcher mean problem. The iterations needed by the (plain) RN method is compared to the total number of RN correction steps needed by fixed step size classical and tangential prediction RNC ($N_{steps} = 3$). The Riemannian gradient norm for $\lambda = 1$ is plotted with solid lines, whereas we use dashed lines for intermediate values of λ.
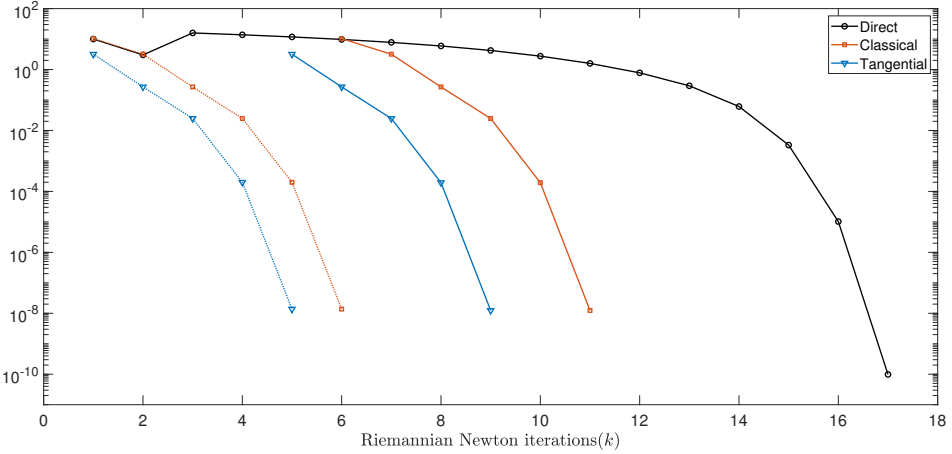

Figure 4.8: Convergence of the Riemannian gradient norm (of the original problem in solid lines and of each intermediate problem in dashed lines) versus RN iterations for the pathological instance of the Karcher mean problem. RN method is compared with fixed step size classical and tangential prediction RNC algorithm ($N_{steps} = 2$).
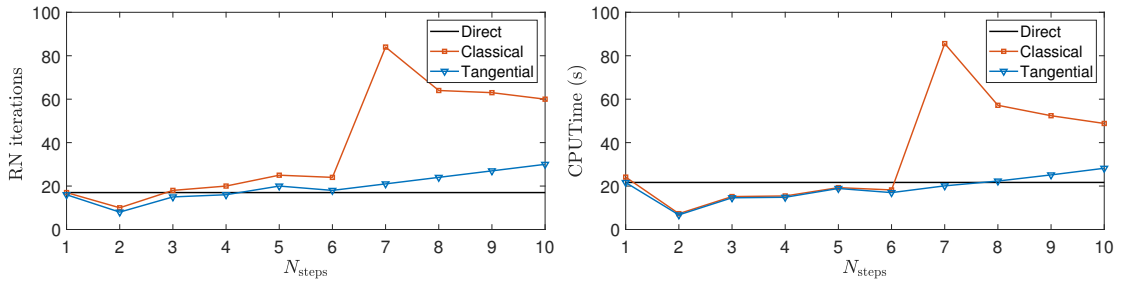

Figure 4.9: RN iterations (left) and computation time (right) versus the number of continuation steps for the fixed step size RNC on the pathological instance of the Karcher mean problem.

Table 4.1: Summary of the number of iterations and computation times for the numerical experiments on the Karcher mean pathological instance. The hyperparameters $(\kappa_{\max}, \alpha_{\max}, \delta_{\max})$ for the step size adaptive experiments (1), (2) and (3) are respectively $(0.6, 3°, 10)$, $(0.3, 1.5°, 5)$ and $(0.15, 0.75°, 2.5)$.

	Karcher mean		
Direct Optimization (RN)	1	17	20.04
Fixed step size classical RNC	2	11	6.65
Fixed step size tangential RNC	2	9	6.32
Step size adaptive RNC (1)	3	22	45.66
Step size adaptive RNC (2)	3	17	32.36
Step size adaptive RNC (3)	6	25	56.96
	Corrections	Correction iterations	Time (s)

4.5 Application to low-rank matrix completion

The problem of matrix completion has received a lot of attention in recent years due to its countless applications. Collaborative filtering, system identification, classification and global positioning are just a few of its success stories. The problem consists of recovering a matrix based on a small fraction of observed entries. One of the most famous instances of matrix completion is the so-called *Netflix problem*: the observed entries are ratings that users gave to some movies and the goal is to predict the missing ratings. In turn, this allows proposing to each user personalized recommendations based on the predicted ratings. Realistic predictions are possible based on the observation that preferences are often summarized with few recurrent preference patterns, much fewer than the number of users and movies. In other words, the matrix of ratings exhibits a low-rank structure. In all the above mentioned applications, the data matrix is at least numerically low rank. The consistency of this hypothesis with reality is one of the reason for the success of matrix completion.

Let us outline the classical model for matrix completion. The goal is to recover a matrix $A \in \mathbb{R}^{m \times n}$ based on a set of entries with indices (i, j) in the observation set $\Omega \subset \{1, \ldots, m\} \times \{1, \ldots, n\}$. If we define the observation operator $P_\Omega : \mathbb{R}^{m \times n} \mapsto \mathbb{R}^{m \times n}$ with

$$P_\Omega(A) = \begin{cases} A_{ij} & \text{if } (i,j) \in \Omega \\ 0 & \text{if } (i,j) \notin \Omega, \end{cases}$$

we can write the original matrix completion problem as

$$\underset{X \in \mathbb{R}^{m \times n}}{\arg \min} \{\text{rank}(X)\},$$

$$\text{s.t. } P_\Omega(X) = P_\Omega(A).$$

This problem being NP hard, convex relaxations have been proposed. For instance nuclear norm minimization instead of rank minimization was proved under mild hypothesis to effectively recover the matrix with high probability [CT10]. A recent review on existing methods based on this idea, among others, can be found in [NKS19].

In the following, we consider the approach based on a Riemannian optimization formulation of the matrix completion problem introduced by [Van13]. Rather than minimizing the rank of the matrix while matching exactly the observed entries, the rank of the matrix is fixed and the discrepancy on the observed entries is minimized. This is an optimization problem constrained to the manifold of fixed-rank matrices that can be written, for a given choice of rank k, as

$$\underset{X \in \mathcal{M}_k}{\arg\min} \left\{ \frac{1}{2} \|P_\Omega(X) - A_\Omega\|_F^2 \right\}, \tag{4.48}$$

with $A_\Omega = P_\Omega(A)$ and where $\mathcal{M}_k$ is the manifold of rank k matrices, as presented in Section 1.5.3. Using Riemannian optimization under the embedded submanifold geometry of $\mathcal{M}_k \subset \mathbb{R}^{m \times n}$ to solve (4.48), showed to outperform state of the art techniques in terms of accuracy, speed and scalability. The low-rank assumption, prescribes taking $k \ll \min\{m, n\}$, but in general its value must be tuned by trial and error. Alternatively, the problem of choosing the rank can be addressed with a rank-adaptive strategy exploiting this fixed-rank formulation complemented with a greedy rank update [UV15]. Noting the gradient of the objective function seen as a scalar field on $\mathbb{R}^{m \times n}$ is $P_\Omega(X - A)$, we readily obtain an expression for the Riemannian gradient by projecting it onto $T_x \mathcal{M}_k$ with the orthogonal projector (1.17):

$$\mathrm{grad} f(X) = \Pi(X) P_\Omega(X - A).$$

By the chosen Riemannian submanifold structure for $\mathcal{M}_k$, the Riemannian Hessian of (4.48) can be obtained from the Euclidean Hessian as stated in Proposition 2.8 by using the differential of the tangent space projection. Alternatively, one can consider the pullback $\hat{f} = f \circ R_X$ for a given retraction R and a point $X \in \mathcal{M}$. Provided the retraction is second-order, the Riemannian Hessian of f at X coincides with the (Euclidean) Hessian of $\hat{f}$ at $0 \in T_X \mathcal{M}_k$ [AMS08, Proposition 5.5.5]. Therefore, the Riemannian Hessian can be isolated from the second-order term in the Taylor expansion of $\hat{f}$ around the zero vector. This strategy is pursued in [Van13] and leads to

$$\begin{aligned} \mathrm{Hess} f(X)[Z] =& P_U P_\Omega(Z) P_V \\ &+ P_U^\perp \left[P_\Omega(Z) + P_\Omega(X - A) V_p \Sigma^{-1} V^\top \right] P_V \\ &+ P_U \left[P_\Omega(Z) + U \Sigma^{-1} U_p^\top P_\Omega(X - A) \right] P_V^\perp. \end{aligned}$$

For the numerical experiments, we opt for the truncated SVD metric projection retraction (2.11) and use the vector transport associated with this retraction obtained by orthogonal projection to the destination tangent space (2.8).

4.5.1 Homotopy for the matrix completion

The homotopy we propose for the matrix completion problem consists of replacing A_Ω in (4.48) with a smooth curve $B_\Omega(\lambda) \in \mathbb{R}^{m \times n}$, $\lambda \in [0,1]$, such that $B_\Omega(1) = A_\Omega$. If we take $B_\Omega(0) = P_\Omega(A_0)$, for some known matrix A_0 of rank k, then the first point of the continuation solution curve is A_0 itself. Recall that $\Pi_{\mathcal{M}_k} : \mathbb{R}^{m \times n} \to \mathcal{M}_k$ denotes the rank-k truncated singular value decomposition. We use $A_0 = \Pi_{\mathcal{M}_k}(\mathcal{F}(A_\Omega))$ where $\mathcal{F}$ is a map that does not alter the known entries of A_Ω and imputes the unknown entries via a heuristic procedure. For example, it is common to use zeros for the unknown entries when initializing Riemannian optimization applied to (4.48) [NKS19, BA15]. In our experiments, we found it more effective to replace missing entries by averaging neighboring known values.

Then, the parametric matrix completion problem is given by

$$\underset{X \in \mathcal{M}_k}{\arg\min} \left\{ f(X, \lambda) = \frac{1}{2} \| P_\Omega(X) - B_\Omega(\lambda) \|_F^2 \right\}, \quad \forall \, \lambda \in [0,1], \tag{4.49}$$

with

$$B_\Omega(\lambda) = (1 - \lambda) P_\Omega(\Pi_{\mathcal{M}_k}(\mathcal{F}(A_\Omega))) + \lambda A_\Omega. \tag{4.50}$$

From the parameter dependent expression of the Riemannian gradient of (4.49), the linearity of P_Ω and of the tangent space projection $\Pi(X) : \mathbb{R}^{m \times n} \to T_X \mathcal{M}_k$, we obtain

$$\frac{\partial \operatorname{grad} f(X, \lambda)}{\partial \lambda} = \Pi(X) \left(A_\Omega - P_\Omega(\Pi_{\mathcal{M}_k}(\mathcal{F}(A_\Omega))) \right).$$

4.5.2 Numerical results

We apply the RNC Algorithm to an instance of the matrix completion problem where the matrix A is obtained by sampling a bi-variate smooth function g on a regular grid of $[a,b] \times [c,d] \subset \mathbb{R}^2$,

$$A_{i,j} = g \left(a + i \frac{(b-a)}{m-1}, c + j \frac{(d-c)}{n-1} \right), \quad \forall \, i = 0, \ldots, m-1, \forall \, j = 1, \ldots, n-1.$$

We then set $A_\Omega = P_\Omega(A)$, with a randomly generated observation operator P_Ω. We choose the number of known entries accordingly with the rank chosen for $\mathcal{M}_k$ using the oversampling rate defined as

$$\mathrm{OS} = \frac{|\Omega|}{\dim(\mathcal{M}_k)} = \frac{|\Omega|}{k(m + n - k)},$$

where $|\Omega|$ is the cardinality of Ω. The matrix A is known to exhibit exponentially decaying singular values, which, as we will see, deteriorates the convergence of direct
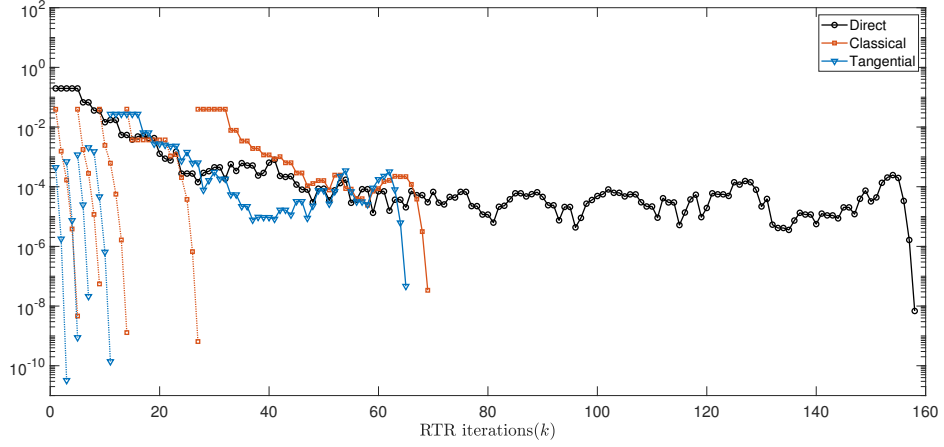
Figure 4.10: Convergence of the Riemannian gradient norm (of the original problem in solid lines and of each intermediate problem in dashed lines) versus RTR iterations on the matrix completion problem. We compare (plain) RTR optimization initialized at A_0 with fixed step size classical and tangential prediction RNC algorithm ($N_{\text{steps}} = 5$) on the matrix completion problem.

Riemannian optimization methods for (4.48). In particular, we consider the function

$$g(x, y) = e^{-\frac{(x-y)^2}{\sigma}}$$

with $\sigma = 0.1$. This function is sampled on $[-1, 1]^2$ with a regular grid of $m = n = 300$ points in each direction. We choose the rank $k = 15$ and set OS $= 3$, implying that 29.25% of the entries are observed.

As the standard RN method tends to fail for this kind of problems, we substituted it with the Riemannian Trust Region algorithm (RTR) (see Algorithm 2.4), both as a corrector at line 11 of Algorithm 4.1 and as a direct optimization scheme.

The results of the direct optimization with RTR initialized at A_0 compared with fixed step size continuation $N_{\text{steps}} = 5$ on the homotopy using the instance curve (4.50) can be seen in Figure 4.10. For all experiments we set tol $= 10^{-7}$ and $N_{\text{inner}} = 5000$. The direct method suffers a long stagnation before entering the superlinear convergence regime. The same stagnation occurs in the last corrections of the continuation procedures, yet less severely and thus the continuation scheme showed to be globally faster both in number of RTR iterations and computation time as summarized in Table 4.2. The table also reports experiments conducted with two other widely used methods for low-rank matrix completion, namely the Riemannian Conjugate Gradient (see Algorithm 2.5), referred to as LRGeomCG [Van13], and the alternating least-squares approach LMAFit [WYZ12]. To make a fair comparison, both use the same initial condition A_0 and the stopping criterion is based on the final relative residual on the known entries that the direct RTR method achieves. In Figure 4.11, the best compromise in terms of computation time of fixed step size RNC between the number of continuation steps and the number of steps of each correction is found to be for $N_{\text{steps}} = 3$. If we increase the number of continuation
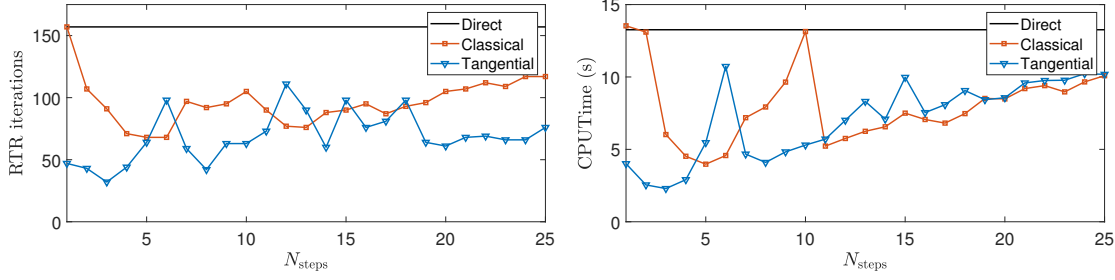
Figure 4.11: RTR iterations (left) and computation time (right) versus the number of continuation steps for the fixed step size RNC on the matrix completion problem.
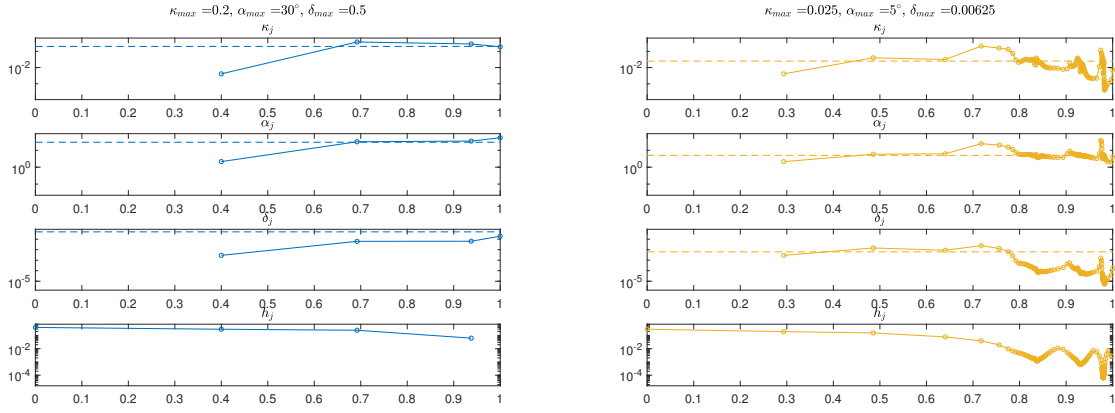

Figure 4.12: Step size selection on the matrix completion problem. Indicators (4.18), (4.20), (4.21) measured after running algorithm 4.2 for selecting the step size (bottom plot), are plotted against the corresponding continuation parameter λ. The dashed lines are the hyperparameters $\kappa_{\max}$, $\alpha_{\max}$, $\delta_{\max}$ used in the step size adaptation procedure for each experiment.

steps, convergence on each correction requires fewer steps so the total number of RTR does not increase significantly, however the computation time increases due to the fixed costs of each prediction and correction. As also confirmed by the step size adaptation experiments[3] (Figures 4.12 and 4.13), the solution curve to the homotopy generated by the instance curve (4.50) is initially trivial to trace. Indeed, in the first part of the homotopy very few RTR iterations per correction are necessary for the classical prediction and even less for the tangential prediction. We clearly get a sense of the increasing difficulty by observing the results of Figure 4.12. Shorter and shorter step sizes are chosen in order to satisfy the step size selection criteria. Finally, as seen from the last plot in Figure 4.13, completely removing the stagnation from the correction phase requires to enforce very strict step size selection criteria causing very small step sizes to be taken and numerous intermediate corrections to be performed. All in all, the most effective setting is the step size adaptive configuration with a permissive step size selection criteria (first plot in Figure 4.13), which still exhibited transient stagnations. We therefore conclude

---

[3]These result are taken from [SK22a] and so the considered angle indicator uses the notion of transporter, see the discussion in Section 4.3.5. Rerunning the experiments for the angle indicator considered in Lemma 4.7 produces barely distinguishable results.
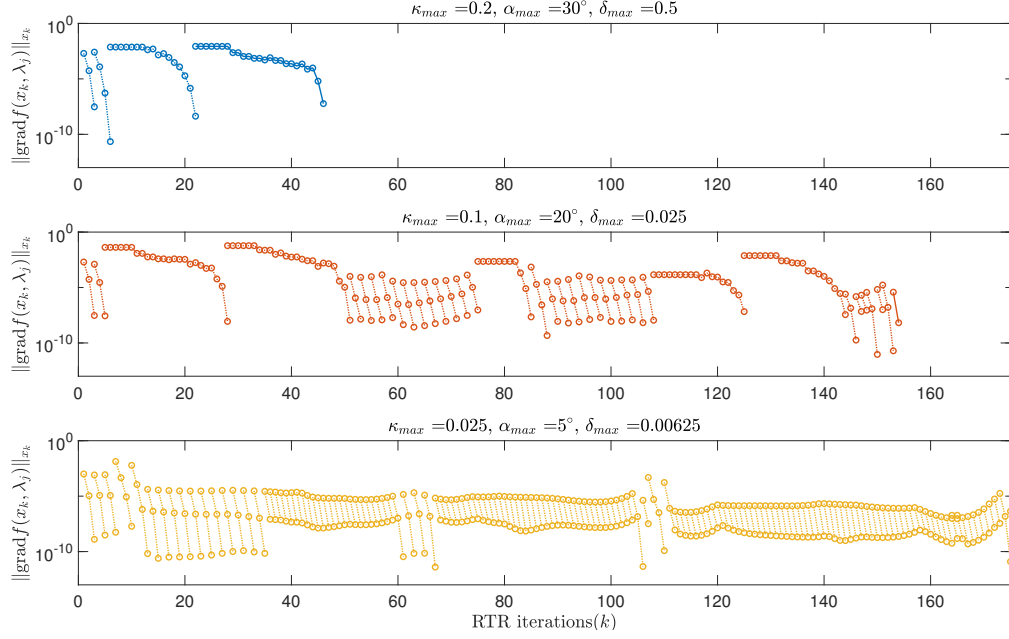
Figure 4.13: Convergence of the Riemannian gradient norm of each intermediate problem versus RTR iterations on the matrix completion problem. The step size adaptive RNC algorithm is compared for different step size adaptation hyperparameters.

that continuation is effective when the stagnation in the correction is mitigated, while removing completely this behavior requires an effort that is not worthwhile.

4.6   Conclusions

This chapter proposes a generalization of numerical continuation to the setting of Riemannian optimization and sufficient conditions for the existence of a solution curve are given. The central contribution is the RNC Algorithm 4.1, a path-following predictor-corrector algorithm relying on the concept of retraction for the prediction combined with superlinearly converging Riemannian optimization routines such as Riemannian Newton method or the Riemannian Trust Region algorithm for the correction. We have generalized to the Riemannian case a step size adaptation strategy relying on the asymptotic expansion of some performance indicators of the correction. Furthermore, we have provided the analysis of the prediction phase motivating the choice of tangential prediction over classical prediction.

The behavior of our algorithm has been illustrated for the problem of computing the Karcher mean of positive definite matrices and for low-rank matrix completion. Particular homotopies have been proposed for both problems, thereby suggesting a more general approach for achieving this task: defining smooth curves of problem instances starting from an easily solvable one and ending at the instance of interest. This proved to be successful in particular for the matrix completion problem, where a fast decay of

Table 4.2: Summary of the number of iterations and computation time for the numerical experiments on the matrix completion problem. The parameters for the step size adaptive experiments (1), (2) and (3) are the same as in Figure 4.13, from top to bottom.

	Matrix completion		
Direct (RTR)	1	159	10.67
Direct (LRGeomCG)	1	1117	5.78
Direct (LMAFit)	1	17309	15.65
Fixed step size classical RNC	5	68	3.39
Fixed step size tangential RNC	3	32	1.96
Step size adaptive RNC (1)	4	46	4.96
Step size adaptive RNC (2)	32	154	70.01
Step size adaptive RNC (3)	143	175	259.30
	Corrections	Correction iterations	Time (s)

singular values leads to a challenging optimization task. The step size adaptation proved to effectively control the Newton update vector norm, the Newton contraction rate and the prediction vectors angle allowing for the correction algorithms to directly exhibit superlinear convergence. However this came at a relatively high computational cost due to the small step sizes required making the fixed step size continuation or permissive step size selection more competitive.

5 │ Hermite interpolation

This chapter introduces a new retraction-based method to solve the Hermite interpolation problem on a manifold: find a smooth curve which interpolates sample points of a given manifold curve while also matching the derivative of the curve at the interpolation points. The method is based on the class of endpoint retraction curves introduced in Section 3.2 and therefore is applicable when a retraction with efficiently computable local inverse is available. This relieves the need of the Riemannian exponential and logarithmic maps to solve the Hermite interpolation problem, as previously existing methods required. This is particularly interesting for manifolds such as the fixed-rank matrix manifold for which computing these maps is a challenging task. The content of this chapter is an excerpt of the submitted preprint [SK22b], whose material is also presented in Chapter 3.

5.1 Overview on manifold curve interpolation

Data processing on non-Euclidean spaces has become a well-established tool in many fields of science and engineering. In particular, there has been a rising interest to interpolate data on a manifold with a curve contained in the manifold. This is motivated by numerous applications in robotics [PR95] [BCC21], computer vision [BKSL17], medical imaging [GSA14] [KDLS21], statistics [MMH$^+$22] and model-order reduction [Ams10], just to mention a few. For example, motion-planning of a robotic manipulator can be carried out by interpolating points on the Lie group of rigid motions SE(3) [PR95]. In statistical modeling, estimating covariance matrices between discrete samples of a random field can be viewed as interpolation on the manifold of symmetric positive definite matrices [MMH$^+$22]. Reduced-order modeling in engineering typically involves projecting high-dimensional dynamics onto low-dimensional subspaces and interpolating such subspaces on the Grassman manifold is an important task for parameter-dependent systems [Ams10].

Over the last two decades, several ways of performing manifold interpolation computationally have been proposed. These methods are tailored to meet different requirements of the application, concerning the regularity of the interpolating curve and the nature of the interpolation constraints. In the following, we focus on continuously differentiable interpolation curves that match prescribed data *and* velocities at each point; this is
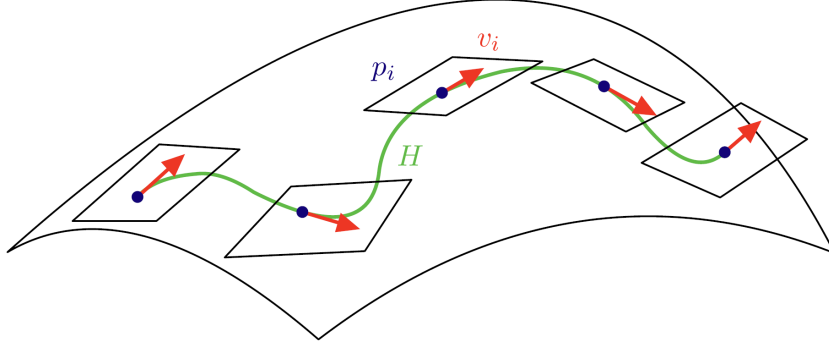
Figure 5.1: The Hermite interpolation problem on manifolds.

commonly known as *Hermite interpolation*.

The mathematical formulation of the manifold Hermite interpolation problem is the following. In this chapter, $\mathcal{M}$ denotes a D-dimensional smooth manifold.

Problem 5.1. *Given $N + 1$ tangent bundle data points $\{(p_i, v_i)\}_{i=0}^{N} \in T\mathcal{M}$ and scalar parameters $t_0 < t_1 < \cdots < t_N$, find a continuously differentiable curve $H : [t_0, t_N] \to \mathcal{M}$ such that*

$$\begin{cases} H(t_i) = p_i, \\ \dot{H}(t_i) = v_i, \end{cases} \quad \forall\, i = 0, \ldots, N.$$

5.2 Related work

In a Euclidean space, the classical solution of the Hermite interpolation problem utilizes piecewise cubic polynomials [Far02]. This solution can be characterized in (at least) three different ways: (i) it minimizes the integral of the squared acceleration over the set of admissible interpolating curves, (ii) it is the unique piecewise cubic polynomial interpolating the points and the derivatives, (iii) it can be constructed with a geometric algorithm introduced by de Casteljau [dC59] involving iterated linear interpolation between suitably chosen control points. As we explain in the following, each of these characterizations can be generalized to manifolds. However, unlike for the Euclidean case, each characterization and extension produces a different curve.

(i) Extending the variational characterization is straightforward: the second derivative (acceleration) is replaced by the covariant derivative of the velocity vector and the search space is constrained to curves on the manifold that satisfy the interpolation constraints [CSLC95, ZN19]. The solution to this variational problem, also known as Riemannian cubic, can be computed by numerical methods for boundary value differential problems, such as shooting methods [BCC21].

(ii) Exploiting the polynomial characterization of the solution requires one to recast manifold interpolation into an Euclidean setting. Several strategies have been explored for this purpose, e.g., a local linearization can be obtained from a (local) bijection between the manifold and a tangent space. For Hermite interpolation, Zimmermann [Zim20]

proposes to use the Riemannian logarithmic map and its differential to map points and derivatives to the tangent space at one of the data points. Standard techniques, including polynomial interpolation, can then be used to construct the interpolation curve on the tangent space and map it back to the manifold with the Riemannian exponential map. We refer to [ZB22] for a recent extension to the multivariate setting.

(iii) The geometric nature of the de Casteljau algorithm lends itself to an intrinsic definition on manifolds. In fact, it is sufficient to replace the linear segments that define the algorithm with geodesic segments, based on the observation that geodesics are the manifold generalization of straight lines. This idea was pioneered by Park and Ravani in [PR95]; it has been adapted to solve the Hermite interpolation problem on spheres and compact Lie groups by Crouch and Leite [CKSL99] and on general Riemannian manifolds by Popiel and Noakes [PN07]. Rodriguez et al. [RSLJ05] proposed a similar approach, where the classical de Casteljau algorithm is interpreted and generalized as the weighted average of two curve branches satisfying the interpolating conditions. This blending technique is adapted to the manifold setting using geodesics branches and geodesic averaging, which – at least in principle – allows one to perform Hermite interpolation on a large class of manifolds even if the work [RSLJ05] itself focuses on compact Lie groups and the sphere.

The non-exhaustive list of algorithms above aims at illustrating that most approaches proposed so far focus on relatively simple manifolds: compact Lie groups, symmetric spaces like the sphere or complete Riemannian manifolds. Most of these techniques require to have closed-form expressions or at least a numerically tractable method for computing endpoint geodesics or the Riemannian exponential and logarithmic maps. For instance, for the case of the Stiefel manifold, the exponential map under the canonical metric has a closed-form expression [EAS99, §2.4.2] and an algorithm to approximate the corresponding logarithmic map has been proposed in [Zim17]. In contrast, for the manifold of fixed-rank matrices we are not aware of a computationally efficient way to realize the logarithmic map; a closed-form expression for the exponential map (under a suitable quotient geometry) is given in [AAM14, §6].

A first step towards relaxing computational requirements has been put forth by Polthier and Nava-Yazdani [NYP13] by generalizing the de Casteljau algorithm to work with generic endpoint curves instead of geodesics as building blocks. This enlarges the applicability of the algorithm to, e.g., polygonal surfaces. On the other hand, velocity constraints cannot be taken into account and the concatenation of two such generalized de Casteljau curves may result in a non-differentiable junction. These drawbacks of [NYP13] were addressed by Krakowski et al. [KMSLB17] for the case of the Stiefel manifold with a generalized de Casteljau algorithm that uses a novel class of endpoint curves, termed quasi-geodesics. This technique allows one to interpolate points on the Stiefel manifold with a globally continuously differentiable curve but the velocity can only be prescribed at the starting point.
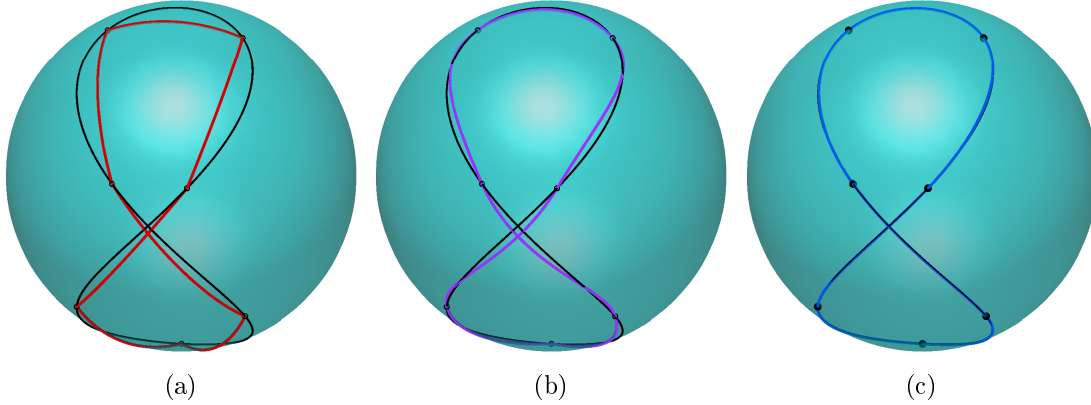
(a)   (b)   (c)

Figure 5.2: Simple interpolation problem on S^2 for different manifold interpolation techniques illustrating the high accuracy of the proposed Hermite interpolation method. The interpolation data is sampled exactly from the figure-eight curve on the sphere (black) given by $\gamma(t) = (\sin(t)\cos(t); \sin(t)\sin(t); \cos(t))$ at seven equally spaced values of $t \in [-\pi, \pi]$. The piecewise geodesic scheme (a) is defined with endpoint geodesics between the sample points. The spline interpolant (b) is a generalization of polynomial spline interpolation proposed in [GSA14]. The resulting curve is continuously differentiable by using the Riemannian de Casteljau algorithm with geodesics [PN07]. The proposed retraction-based Hermite interpolant (c) described in Section 5.3 is computed here using the projective retraction (2.9).

Contributions and outline of the chapter. Extending upon [KMSLB17, NYP13] in being able to handle velocity constraints, the proposed approach is only based on retractions through the usage of the novel class of endpoint curves introduced in Section 3.2. For a large class of manifolds and retractions, the inverse retraction is available in closed-form. Whenever this is the case, our method makes it possible to solve the Hermite interpolation problem in a numerically efficient way on a larger class of manifolds than those available until now. In Section 5.3, we develop the retraction-based Hermite (RH) interpolation scheme on manifolds. An analysis of the method is carried out in Section 5.4. In particular, the method is shown to be well-posed whenever consecutive data points are in a retraction-convex set. Furthermore, a convergence analysis of the interpolation error of the RH scheme is proposed, extending a well-known result on the error convergence of Hermite interpolation on Euclidean spaces for sufficiently smooth data. Finally, in Section 5.5, we demonstrate several applications of our novel interpolation scheme for both the manifold of fixed-rank matrices and the Stiefel manifold.

5.3   Generalized de Casteljau algorithm with retractions

The classical de Casteljau algorithm [dC59] is a geometric procedure to construct polynomial curves in $\mathbb{R}^D$. To describe the algorithm, let $\sigma_1(t; x, y) := (1-t)x + ty$ denote the linear interpolation between two points $x, y \in \mathbb{R}^D$. Given $N+1$ so called *control points*
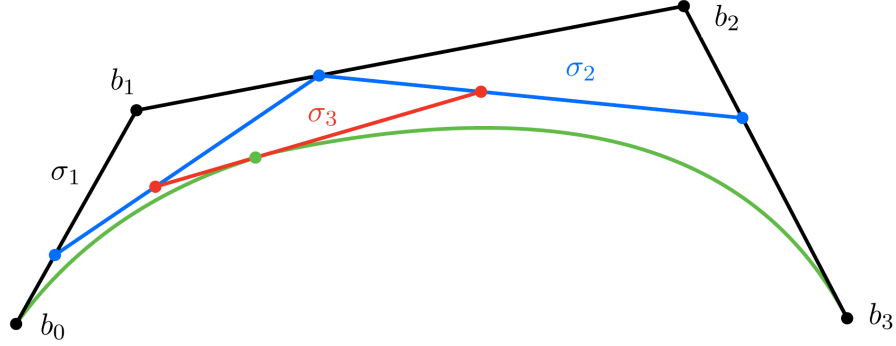
Figure 5.3: The classical de Casteljau algorithm with four control points in $\mathbb{R}^2$. The resulting curve in green is a polynomial curve of degree three.

$b_0, \ldots, b_N \in \mathbb{R}^D$, the relation

$$\sigma_k(t; b_i, \ldots, b_{i+k}) := \sigma_1(t; \sigma_{k-1}(t; b_i, \ldots, b_{i+k-1}), \sigma_{k-1}(t; b_{i+1}, \ldots, b_{i+k})), \quad i = 0, \ldots, N-k,$$

is applied recursively for $k = N, N-1, \ldots, 2$ to define a polynomial curve σ_N of degree N such that

$$\begin{aligned}
\sigma_N(0; b_0, \ldots, b_N) &= b_0, \\
\sigma_N(1; b_0, \ldots, b_N) &= b_N, \\
\dot\sigma_N(0; b_0, \ldots, b_N) &= N\dot\sigma_1(0; b_0, b_1) = N(b_1 - b_0), \\
\dot\sigma_N(1; b_0, \ldots, b_N) &= N\dot\sigma_1(1; b_{N-1}, b_N) = N(b_N - b_{N-1}),
\end{aligned}$$

where $\dot\sigma_N$ denotes the derivative of σ_N with respect to t. Only the first and last control points are interpolated while the other points influence the shape of the curve. See Figure 5.3 for an illustration of the de Casteljau algorithm with four control points in $\mathbb{R}^2$. In [PR95] an extension to control points on a manifold $\mathcal{M}$ was proposed that replaces σ_1 by the endpoint geodesic joining x and y. The same recursive relation then yields a manifold curve verifying analogous properties: (a) it is smooth, (b) it interpolates the first and last points, and (c) the derivatives at the first and last points only depend on the first and last two control points, respectively. Property (c) allows one to conveniently control the endpoint derivatives via the choice of b_1, b_{N-1}, a property that makes the de Casteljau algorithm useful for Hermite interpolation; see, e.g., [PN07].

Along the line of work by Krakowski et al. [KMSLB17], we consider a generalization of the de Casteljau algorithm that allows for arbitrary smooth manifold curves at each step of the recursion (instead of constructing everything on top of geodesics). If each curves chosen to define the algorithm join the prescribed endpoints, then properties (a) and (b) are trivially satisfied. In [KMSLB17, Proposition 13], sufficient conditions for the chosen curves to produce property (c) are given for $N = 2$. The following proposition extends these results to $N = 3$.
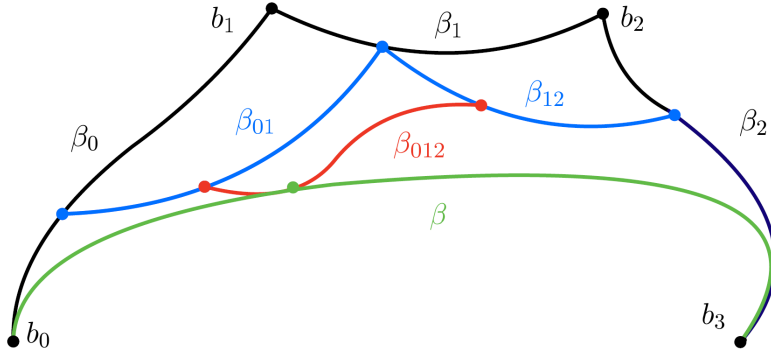
Figure 5.4: Generalized de Casteljau with 4 control points described in Proposition 5.2.

Proposition 5.2. *For $b_0, b_1, b_2, b_3 \in \mathcal{M}$ consider, as in Figure 5.4:*

- *smooth $\beta_i : [0,1] \to \mathcal{M}$ joining b_i and b_{i+1} for each $i = 0, 1, 2$,*
- *smooth $\beta_{01} : [0,1]^2 \to \mathcal{M}$ such that $\beta_{01}(\cdot, t)$ joins $\beta_0(t)$ and $\beta_1(t)$ for every $t \in [0,1]$,*
- *smooth $\beta_{12} : [0,1]^2 \to \mathcal{M}$ such that $\beta_{12}(\cdot, t)$ joins $\beta_1(t)$ and $\beta_2(t)$ for every $t \in [0,1]$,*
- *smooth $\beta_{012} : [0,1]^3 \to \mathcal{M}$ such that $\beta_{012}(\cdot, s, t)$ joins $\beta_{01}(s,t)$ and $\beta_{12}(s,t)$ for every $s, t \in [0,1]$.*

If, additionally,

(i) $\beta_{01}(s, 0) = \beta_0(s)$ and $\beta_{12}(s, 1) = \beta_2(s)$,

(ii) $\beta_{012}(s, 0, 0) = \beta_0(s)$ and $\beta_{012}(s, 1, 1) = \beta_2(s)$,

then the generalized de Casteljau manifold curve

$$\beta(t) = \beta_{012}(t, t, t)$$

satisfies

$$\beta(0) = b_0, \quad \beta(1) = b_3, \tag{5.1}$$

and

$$\dot{\beta}(0) = 3\dot{\beta}_0(0), \quad \dot{\beta}(1) = 3\dot{\beta}_2(1). \tag{5.2}$$

Proof. The interpolation condition (5.1) follows directly from the definitions:

$$\beta(0) = \beta_{012}(0, 0, 0) = \beta_{01}(0, 0) = \beta_0(0) = b_0$$

and, analogously, $\beta(1) = b_3$. To prove (5.2), we first note that

$$\dot{\beta}(t) = \frac{\mathrm{d}}{\mathrm{d}s}\beta_{012}(s, t, t)\big|_{s=t} + \frac{\mathrm{d}}{\mathrm{d}s}\beta_{012}(t, s, t)\big|_{s=t} + \frac{\mathrm{d}}{\mathrm{d}s}\beta_{012}(t, t, s)\big|_{s=t}.$$

At $t = 0$, inserting the definitions of the curves as well as conditions (i) and (ii) we thus

obtain

$$\dot{\beta}(0) = \frac{\mathrm{d}}{\mathrm{d}s}\beta_{012}(s,0,0)\big|_{s=0} + \frac{\mathrm{d}}{\mathrm{d}s}\beta_{012}(0,s,0)\big|_{s=0} + \frac{\mathrm{d}}{\mathrm{d}s}\beta_{012}(0,0,s)\big|_{s=0}$$

$$= \dot{\beta}_0(0) + \frac{\mathrm{d}}{\mathrm{d}s}\beta_{01}(s,0)\big|_{s=0} + \frac{\mathrm{d}}{\mathrm{d}s}\beta_{01}(0,s)\big|_{s=0} = 3\dot{\beta}_0(0).$$

Analogously, one establishes $\dot{\beta}(1) = 3\dot{\beta}_2(1)$, which completes the proof.

$\square$

When one employs the same type of endpoint curve to define each $\beta_0, \ldots, \beta_{012}$ in Proposition 5.2, such as endpoint geodesics or endpoint quasi-geodesics as proposed by [KMSLB17], then conditions (i) and (ii) are trivially satisfied. With this simplification, a result of the form (5.2) for a generalized de Casteljau algorithm of arbitrary order N can be found in [NYP13, Theorem 8]. However, to be useful for Hermite interpolation, we need to have explicit relationships between the control points and $\dot{\beta}_0(0)$ as well as $\dot{\beta}_2(1)$. While both are available for geodesics, only the first relationship can be made explicit for quasi-geodesics. Consequently, the interpolation problem considered in [KMSLB17] incorporates a velocity constraint at the first interpolation point only. In the next section, we show how to choose the defining curves $\beta_0, \ldots, \beta_{012}$ in order to satisfy Proposition 5.2 and have an explicit relationships between the control points and $\dot{\beta}_0(0)$, $\dot{\beta}_2(1)$. This is made possible by choosing suitable members of the class of endpoint retraction curves introduced in Section 3.2.

5.3.1 The de Casteljau algorithm with endpoint retraction curves

In Section 3.2 we introduced a class of retraction-based curves which allow connecting pairs of manifold points. Recall from Definition 3.4, for any $x, y \in \mathcal{M}$ and $r \in [0,1]$, we have defined the retraction curve

$$c_r(t; x, y) = R_{q(r)}\left((1-t)R_{q(r)}^{-1}(x) + tR_{q(r)}^{-1}(y) \right), \quad \forall\, t \in [0,1],$$

with $q(r) = R_x(rR_x^{-1}(y))$ which joins the points x and y for every $r \in [0,1]$. As noted in Proposition 3.5 (ii), when $r = 0$ or $r = 1$ the derivative of the retraction curve at $t = 0$ and $t = 1$ respectively is explicitly related to the endpoint. This is the crucial condition allowing to use the curves c_r with suitable values of r to define curves that satisfy the conditions of Proposition 5.2 and, in turn, to define a suitable generalization of the de Casteljau algorithm. From Proposition 3.5 (ii) and (5.2), it follows that $\beta_0(\cdot) = c_0(\cdot; b_0, b_1)$ and $\beta_2(\cdot) = c_1(\cdot; b_2, b_3)$ are canonical choices for joining b_0 with b_1 and b_2 with b_3, respectively. The other curves must be suitably chosen from the r-endpoint retraction curve family in order to satisfy Proposition 5.2.

Proposition 5.3. *The following choices of β_0, β_1, β_2, β_{01}, β_{12} and β_{012} satisfy the conditions of Proposition 5.2:*
  - $\beta_0(t) = c_0(t; b_0, b_1)$, $\beta_1(t) = c_{r_1(t)}(t; b_1, b_2)$, $\beta_2(t) = c_1(t; b_2, b_3)$,

- $\beta_{01}(s,t) = c_{r_{01}(s,t)}(s; \beta_0(t), \beta_1(t))$, $\beta_{12}(s,t) = c_{r_{12}(s,t)}(s; \beta_1(t), \beta_2(t))$,
- $\beta_{012}(u,s,t) = c_{r_{012}(u,s,t)}(u; \beta_{01}(s,t), \beta_{12}(s,t))$,

for any smooth functions $r_1 : [0,1] \to [0,1]$, $r_{01}, r_{12} : [0,1]^2 \to [0,1]$ *and* $r_{012} : [0,1]^3 \to [0,1]$ *such that*

$$r_{01}(s,0) = 0, \quad r_{12}(s,1) = 1, \quad r_{012}(s,0,0) = 0, \ r_{012}(s,1,1) = 1.$$

Moreover, the resulting manifold curve $\beta(t) = \beta_{012}(t,t,t)$ *satisfies*

$$\dot{\beta}(0) = 3\dot{c}_0(0; b_0, b_1) = 3R_{b_0}^{-1}(b_1), \quad \dot{\beta}(1) = 3\dot{c}_1(1; b_2, b_3) = -3R_{b_3}^{-1}(b_2). \tag{5.3}$$

Proof. Proposition 3.5 (i) implies that the curves β_0, β_1, β_2, $\beta_{01}(\cdot, t)$, $\beta_{12}(\cdot, t)$ and $\beta_{012}(\cdot, s, t)$ have the correct endpoints for every $s, t \in [0,1]$ and any choice of r_1, r_{01}, r_{12} and r_{012}. Direct computation shows the remaining requirements (i) and (ii) of Proposition 5.2:

(i) $\beta_{01}(s,0) = c_{r_{01}(s,0)}(s; \beta_0(0), \beta_1(0)) = c_0(s; b_0, b_1) = \beta_0(s)$,
   $\beta_{12}(s,1) = c_{r_{12}(s,1)}(s; \beta_1(1), \beta_2(1)) = c_1(s; b_2, b_3) = \beta_2(s)$,

(ii) $\beta_{012}(s,0,0) = c_{r_{012}(s,0,0)}(s; \beta_{01}(0,0), \beta_{12}(0,0)) = c_0(s; b_0, b_1) = \beta_0(s)$,
   $\beta_{012}(s,1,1) = c_{r_{012}(s,1,1)}(s; \beta_{01}(1,1), \beta_{12}(1,1)) = c_1(s; b_2, b_3) = \beta_2(s)$.

Finally, the relation (5.3) follows from combining (5.2) with Proposition 3.5 (ii). □

Proposition 5.3 offers a great degree of flexibility in choosing r_1, r_{01}, r_{12}, and r_{012}. For practical purposes, a simple choice that leads to a computationally less expensive evaluation of the curve is preferable. We propose to choose

$$r_1(s) = 1/2, \ r_{01}(s,t) = 0, \ r_{12}(s,t) = 1, \ r_{012}(u,s,t) = t. \tag{5.4}$$

As detailed below in Algorithm 5.2, this choice essentially requires 7 retractions and 5 inverse retractions per evaluation of the generalized de Casteljau manifold curve, ignoring the cost for preprocessing.

At this point, the choice of $r_1(s) = 1/2$ appears to be ad hoc, especially because Proposition 5.3 imposes no constraint on r_1. In Section 5.5.1, we argue that this choice of r_1 is crucial for the scheme to attain favorable convergence properties. From now on, we restrict ourselves to (5.4).

Definition 5.4. *Given control points* $b_0, b_1, b_2, b_3 \in \mathcal{M}$ *we use* $\beta(\cdot; b_0, b_1, b_2, b_3)$ *to denote the* generalized de Casteljau curve *constructed in Proposition 5.3 with the choice* (5.4) *for* $r_1, r_{01}, r_{12}, r_{012}$.

5.3.2   The retraction-based Hermite (RH) interpolation scheme

The generalized de Casteljau curve of Definition 5.4 can now be used to solve the Hermite interpolation problem, Problem 5.1, by choosing suitable control points.

Proposition 5.5. *Given* $(p_0, v_0), (p_1, v_1) \in T\mathcal{M}$, *define*

$$p_0^+ = R_{p_0}\left(\frac{1}{3}v_0\right), \quad p_1^- = R_{p_1}\left(-\frac{1}{3}v_1\right)$$

and let $\alpha(t) \equiv \alpha(t; p_0, v_0, p_1, v_1) := \beta(t; p_0, p_0^+, p_1^-, p_1)$ *denote the generalized de Casteljau curve according to Definition 5.4. Then*

$$\alpha(0) = p_0, \ \alpha(1) = p_1, \quad \dot{\alpha}(0) = v_0, \ \dot{\alpha}(1) = v_1.$$

Proof. The result follows from combining Propositions 5.2 and 5.3:

$$\alpha(0) = \beta(0; p_0, p_0^+, p_1^-, p_1) = p_0, \ \alpha(1) = \beta(1; p_0, p_0^+, p_1^-, p_1) = p_1,$$
$$\dot{\alpha}(0) = \dot{\beta}(0; p_0, p_0^+, p_1^-, p_1) = 3R_{p_0}^{-1}(p_0^+) = 3R_{p_0}^{-1}\left(R_{p_0}\left(\frac{1}{3}v_0\right)\right) = v_0,$$
$$\dot{\alpha}(1) = \dot{\beta}(1; p_0, p_0^+, p_1^-, p_1) = -3R_{p_1}^{-1}(p_1^-) = -3R_{p_1}^{-1}\left(R_{p_1}\left(-\frac{1}{3}v_1\right)\right) = v_1.$$

$$\square$$

As an immediate consequence of Proposition 5.5, the following corollary shows how α is used piecewise to define the *retraction-based Hermite (RH) interpolant* H that addresses Problem 5.1.

Corollary 5.6. *Letting* $h_i := t_{i+1} - t_i$ *for* $i = 0, \ldots, N-1$, *the manifold curve* $H : [t_0, t_N] \to \mathcal{M}$ *defined piecewise by*

$$H(t)\big|_{[t_i, t_{i+1}]} = \alpha\left(\frac{t - t_i}{h_i}; p_i, h_i v_i, p_{i+1}, h_i v_{i+1}\right), \quad i = 0, \ldots, N-1, \qquad (5.5)$$

is a solution to Problem 5.1.

Algorithms 5.1 and 5.2 summarize the construction of H and its evaluation, respectively. We separate the computations needed for evaluating the RH interpolant (online phase) from those that can be precomputed, stored and used in every evaluation (offline phase).

Algorithm 5.1 Offline Phase (precompute quantities defining the RH interpolant)

Input: Tangent bundle data points $\{(p_i, v_i)\}_{i=0}^{N} \in T\mathcal{M}$, $t_0 < t_1 < \cdots < t_N$.

1: **for** $i = 0, \ldots, N-1$ **do**
2:      $h_i = t_{i+1} - t_i$;
3:      $p_i^+ = R_{p_i} \left(\frac{1}{3} h_i v_i \right)$;
4:      $p_{i+1}^- = R_{p_{i+1}} \left(-\frac{1}{3} h_i v_{i+1} \right)$;
5:      $q_i = R_{p_i^+} \left(\frac{1}{2} R_{p_i^+}^{-1} \left(p_{i+1}^- \right) \right)$;   ▷ Anchor for the middle segment β_1.
6:      $w_i^+ = R_{q_i}^{-1} \left(p_i^+ \right)$;   ▷ Tangent vector from q_i to p_i^+
7:      $w_{i+1}^- = R_{q_i}^{-1} \left(p_{i+1}^- \right)$;   ▷ Tangent vector from q_i to p_{i+1}^-
8: **end for**
9: **return** : $\left\{ q_i, w_i^+, w_{i+1}^- \right\}_{i=0}^{N-1}$;

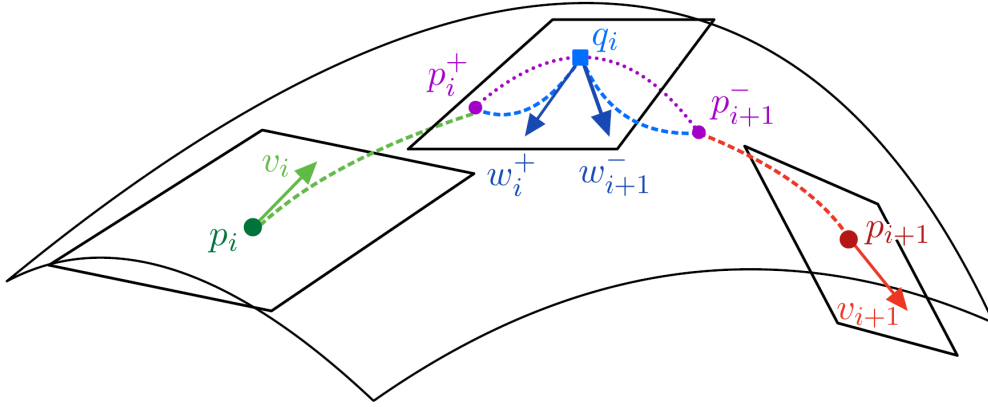
Figure 5.5: Illustration of offline computations performed by Algorithm 5.1.

Algorithm 5.2 Online Phase (evaluation of RH interpolant at t)

Input: $t \in [t_0, t_N]$, $\{p_i, v_i, t_i\}_{i=0}^{N}$, $\left\{ q_i, w_i^+, w_{i+1}^- \right\}_{i=0}^{N-1}$;

1: Find largest $i \in \{0, \ldots, N-1\}$ such that $t_i \leq t$;
2: $h_i = t_{i+1} - t_i$;
3: $\tau = \frac{t - t_i}{h_i}$;
4: $\beta_0 = R_{p_i} \left(\frac{\tau}{3} h_i v_i \right)$;   ▷ $1 \times R$
5: $\beta_1 = R_{q_i} \left((1 - \tau) w_i^+ + \tau w_{i+1}^- \right)$;   ▷ $1 \times R$
6: $\beta_2 = R_{p_{i+1}} \left(-\frac{\tau}{3} h_i v_{i+1} \right)$;   ▷ $1 \times R$
7: $\beta_{01} = c_0 \left(\tau; \beta_0, \beta_1 \right)$;   ▷ $1 \times R + 1 \times R^{-1}$
8: $\beta_{12} = c_1 \left(\tau; \beta_1, \beta_2 \right)$;   ▷ $1 \times R + 1 \times R^{-1}$
9: **return** $\beta = c_\tau \left(\tau; \beta_{01}, \beta_{12} \right)$;   ▷ $2 \times R + 3 \times R^{-1}$

5.4 Analysis of RH interpolation

For the purpose of deriving qualitative and asymptotic properties of RH interpolation, we suppose that the interpolation data are sampled from a continuously differentiable manifold curve $\gamma : [0,1] \to \mathcal{M}$, that is

$$p_i = \gamma(t_i), \quad v_i = \dot{\gamma}(t_i), \quad \forall \, i = 0, \dots, N, \tag{5.6}$$

for some $0 = t_0 < t_1 < \cdots < t_N = 1$. Because of its piecewise definition (see Corollary 5.6), it is sufficient to consider the RH interpolant on a single subinterval, i.e. the manifold curve $H_h : [t, t+h] \to \mathcal{M}$ defined by

$$H_h(\tau) = \alpha \left(\frac{\tau - t}{h} ; \gamma(t), h\dot{\gamma}(t), \gamma(t+h), h\dot{\gamma}(t+h) \right), \tag{5.7}$$

for sufficiently small $h > 0$ and $t \in [0, 1-h]$, satisfying $H_h(t) = \gamma(t)$, $H_h(t+h) = \gamma(t+h)$, $\dot{H}_h(t) = \dot{\gamma}(t)$, and $\dot{H}_h(t+h) = \dot{\gamma}(t+h)$. Our results for a single interval apply to the piecewise solution of (5.6) by letting $h = \max_{i=0,\dots,N-1}\{t_{i+1} - t_i\}$.

In the following analysis we make use of Assumption 3.11 to guarantee the retraction-convexity radius $\bar{\rho}$ defined by (3.4) is lower-bounded on compact sets $K \subset \mathcal{M}$. In particular, for $K = \gamma([0,1])$, knowing that

$$\rho_{\min}(\gamma) := \min_{t \in [0,1]} \bar{\rho}(\gamma(t)) > 0$$

ensures there exists a radius below which any manifold ball centered on the curve is retraction-convex. Hence, throughout the analysis section we suppose that Assumption 3.11 holds.

5.4.1 Well-posedness of RH interpolation

As stated in Theorem 3.9, any sufficiently small neighborhood on a manifold is retraction-convex. This ensures that whenever some points are sufficiently close, they all belong to a common retraction-convex set. This makes it possible to grant the well-posedness of the r-endpoint retraction curve family when the endpoint are sufficiently close, see Proposition 3.8. In turn, this implies the well-posedness of the RH interpolation scheme when the control points are sufficiently close.

Proposition 5.7. *For a retraction-convex set $\mathcal{U}$ consider control points $b_0, b_1, b_2, b_3 \in \mathcal{U}$. Then the generalized de Casteljau curve $t \mapsto \beta(t; b_0, b_1, b_2, b_3)$ constructed in Definition 5.4 is well-defined for every $t \in [0,1]$.*

Proof. We apply recursively the result of Proposition 3.8. Since $b_0, b_1, b_2, b_3 \in \mathcal{U}$, the curves $\beta_0, \beta_1, \beta_2$ are well-defined and their image is entirely contained in $\mathcal{U}$. In turn, this implies that for every $t \in [0,1]$ the curves $\beta_{01}(\cdot, t)$ and $\beta_{12}(\cdot, t)$ are well-defined and their image is entirely contained in $\mathcal{U}$. Finally, it follows that for every $s, t \in [0,1]^2$ the curve

$\beta_{012}(\cdot, s, t)$ is well-defined. $\hfill\square$

Note that this proof is valid for any generalized de Casteljau curve constructed as in Proposition 5.3.

The control points defining the RH interpolant 5.7 are directly related with the interpolation data of the equispaced Hermite interpolation problem (5.6). Hence, the sufficient condition of well-posedness expressed in terms of control points in Proposition 5.7 is translated into a sufficient condition on the step size of the equispaced sampling of the interpolation curve. If h is small enough, as Proposition 5.8 below quantifies, the control points associated with the RH interpolant (5.7) are all contained in a retraction-convex set. Then, the piecewise RH interpolant of the full curve γ is guaranteed to be globally defined.

In the following, we make use of the function Δ appearing in Proposition 2.12 to describe the domain over which a retraction is a diffeomorphism. For every compact set $K \subset \mathcal{M}$ we indicate by $\Delta_{\min}(K) := \min_{x \in K} \Delta(x)$. Since Δ is known to be continuous and strictly positive, we know $\Delta_{\min}(K) > 0$ for every compact subset K.

Proposition 5.8. *There exists a constant $h_1 > 0$ depending on the curve γ and on the retraction such that for any $0 < h < h_1$ the RH interpolant H_h defined in (5.7) is well-posed for every $t \in [0, 1 - h]$.*

Proof. By Proposition 5.5 and Corollary 5.6,

$$H_h(\tau) = \beta\left(\frac{\tau - t}{h}; p_0(t), p_0^+(t, h), p_1^-(t, h), p_1(t, h)\right),$$

with the control points

$$
\begin{aligned}
p_0(t) &:= \gamma(t), & p_1(t, h), &:= \gamma(t + h), \\
p_0^+(t, h) &:= R_{\gamma(t)}\left(h\dot{\gamma}(t)/3\right), & p_1^-(t, h) &:= R_{\gamma(t+h)}\left(-h\dot{\gamma}(t + h)/3\right).
\end{aligned}
$$

Denote by L_γ a Lipschitz constant of the curve γ on $[0, 1]$. Since γ is assumed continuously differentiable, we can take $L_\gamma = \max_{t \in [0,1]} \|\dot{\gamma}(t)\|_{\gamma(t)}$. Then, it holds that

$$d(p_0(t), p_1(t, h)) = d(\gamma(t), \gamma(t + h)) \le L_\gamma h,$$

Let us denote $\Delta_{\min}(\gamma) > 0$ the minimum of the function Δ on the compact set $\gamma([0, 1])$. If $h < \frac{\Delta_{\min}(\gamma)}{L_\gamma}$, then

$$\|h\dot{\gamma}(t)/3\|_{\gamma(t)} < \Delta(\gamma(t))/3, \quad \|h\dot{\gamma}(t + h)/3\|_{\gamma(t+h)} < \Delta(\gamma(t + h))/3.$$

This allows us to invoke the Lipschitz continuity of the retraction stated in Proposi-

tion 3.14-(i) with the constant $L_R(\gamma)$ associated with the compact set $\gamma([0,1])$ and find

$$d(p_0(t), p_0^+(t,h)) \le hL_R(\gamma) \|\dot{\gamma}(t)/3\|_{\gamma(t)} \le hL_R(\gamma)L\gamma/3,$$
$$d(p_0(t), p_1^-(t,h)) \le d(p_0(t), p_1(t,h)) + d(p_1(t,h), p_1^-(t,h))$$
$$\le L_\gamma h + hL_R(\gamma)\|\dot{\gamma}(t+h)/3\|_{\gamma(t+h)} \le L_\gamma(1 + L_R(\gamma)/3)h.$$

Hence, if $h < \frac{\Delta_{\min}(\gamma)}{L_\gamma}$, all control point of H_h are contained in $B(\gamma(t), Qh)$ with

$$Q := \max\{L_\gamma,\, L_R(\gamma)L_\gamma/3,\, L_\gamma(1 + L_R(\gamma)/3)\} = L_\gamma(1 + L_R(\gamma)/3).$$

By Assumption 3.11, the minimum of the retraction-convexity radius function $\bar{\rho}$ on the image of the curve γ denoted $\rho_{\min}(\gamma)$ is strictly positive. Therefore if we further restrict the step size by taking $h < h_1 := \min\left\{\frac{\Delta_{\min}(\gamma)}{L_\gamma}, \frac{\rho_{\min}(\gamma)}{Q}\right\}$, all the control points of H_h are contained in the retraction-convex set $B(\gamma(t), \rho_{\min}(\gamma))$. This implies the curve H_h is well-defined. $\qquad\square$

5.4.2 Interpolation error

In the Euclidean setting, asymptotic convergence rates of the maximum interpolation error are classic results of numerical analysis. For the case of piecewise cubic Hermite interpolation the error can be shown to converge as $O(h^4)$ [QSS07, §8.4], where h is the largest sampling step size. In this section, we generalize this result to the RH interpolation scheme.

Theorem 5.9. *Let $\gamma \in C^4([0,1])$ and consider H_h, the RH interpolant of γ on a subinterval $[t, t+h]$ as defined in (5.7). Under Assumption 3.11 and assuming that for $k = 2, 3, 4$ there exist constants $L_{RH}^{(k)} > 0$ and $h_2 > 0$ such that for every $0 < h < h_2$ and any $t \in [0, 1-h]$ it holds that*

$$\sup_{\tau \in [t, t+h]} \left\|\frac{\mathrm{D}^k H_h(\tau)}{\mathrm{d}\tau^k}\right\|_{H_h(\tau)} < L_{RH}^{(k)}, \tag{5.8}$$

where $\frac{\mathrm{D}^k}{\mathrm{d}\tau^k}$ denotes the order-k covariant derivative along a curve. Then, there exist constants $\kappa > 0$ and $\bar{h} > 0$ depending on the curve γ, the manifold and the retraction such that for any $0 < h < \bar{h}$ and any $t \in [0, 1-h]$, we have

$$\max_{\tau \in [t, t+h]} d(\gamma(\tau), H_h(\tau)) < \sqrt{D}\frac{\kappa}{4!}h^4,$$

where $\dim(\mathcal{M}) = D$.

We stress that the above result holds for any choice of $t \in [0, 1-h]$. As a consequence, we can bound the maximum interpolation error of the piecewise RH interpolant of the full curve γ.

Corollary 5.10. *Under the assumptions of Theorem 5.9, the piecewise RH interpolant*

H of γ defined by (5.5) with data (5.6) verifies

$$\max_{\tau \in [0,1]} d(\gamma(\tau), H(\tau)) < \sqrt{D} \frac{\kappa}{4!} h^4,$$

where $h = \max_{i=0,\dots,N-1} t_{i+1} - t_i$.

The proof of Theorem 5.9 relies on a local linearization of the manifold: both the interpolation curve H_h and the corresponding portion of the curve γ are pulled back to the tangent space at $\gamma(s)$, for some $s \in [t, t+h]$, with the Riemannian logarithmic map, the local inverse of the Riemannian exponential, see Definition 1.32. The so-obtained local coordinate system is also known as the system of normal coordinates at $\gamma(s)$ [Lee18, p. 132]. Since the Riemannian logarithm is only locally defined, we first need to ensure the image of γ, the control points, and intermediate quantities involved in the procedure remain confined to a domain of invertibility of the exponential map as $h \to 0$. We ensure that these requirements are met when considering sufficiently small h, see Lemma 5.12 below. This theoretical step size restriction is appropriate given the asymptotic nature of Theorem 5.9. In practice however, our experiments suggest that the fourth-order convergence established by the theorem can be observed as soon as the interpolation curve is well-defined; see the numerical experiments Section 5.5. We point out once again that for this to be the case, we also need the scheme to verify assumption (5.8); see Section 5.5.1 for the importance of this assumption.

Lipschitz continuity of the RH interpolation curve

As the step size h between interpolation samples converges to zero, we need to ensure the image of the H_h, the RH interpolant of γ on $[t, t+h]$, is confined to a region whose diameter also converges to zero. One way to ensure this is to show that H_h admits a Lipschitz constant that is independent of h.

Lemma 5.11. *There exist a constant $L_{RH} > 0$ and $h_3 > 0$ depending on γ and on the retraction such that for every $0 < h < h_3$ and any $t \in [0, 1-h]$*

$$d(H_h(\tau_1), H_h(\tau_2)) \leq L_{RH} |\tau_1 - \tau_2|, \quad \forall \tau_1, \tau_2 \in [t, t+h].$$

With this result, it is possible to state that, for instance, $H_h(\tau) \in B(H_h(t), r)$ for all $\tau \in [t, t+h]$ provided $h < r/L_{RH}$.

The following proof of Lemma 5.11 considers the RH interpolant given in Definition 5.4, that is with the particular choice (5.4) for the functions r_1, r_{01}, r_{12}, and r_{012}. Nevertheless, the result remains valid for any choice of these functions, as long as they are Lipschitz continuous.

Proof of Lemma 5.11. We aim at invoking the Lipschitz continuity results for the retraction and retraction curves established in Section 3.4. In particular, Corollary 3.16 states a convenient Lipschitz continuity condition for the endpoint retraction curves

$t \mapsto c_r(t; x, y)$, jointly with respect to $r, t \in [0, 1]$ and $x, y \in \mathcal{M}$. The result is valid under the assumption that the point x, y belong to a non-empty retraction-convex $\mathcal{U}$ such that $\mathcal{U} \subset \mathcal{I}_z^{1/3}$ for every $z \in \mathcal{U}$, see Figure 3.6. Let us argue there exists a $\mathcal{U}$ satisfying this property that contains all the control points of H_h, provided h is sufficiently small. Consider the set $K = \overline{\mathcal{I}_\gamma^{1/3}} = \left\{ R_{\gamma(t)}(v) : t \in [0, 1], \|v\|_{\gamma(t)} \leq \frac{\Delta(\gamma(t))}{3} \right\}$, which is compact by Lemma 3.13. Since we assumed Assumption 3.11 to hold, we know that the retraction-convexity radius $\bar{\rho}$ is lower bounded on K by $\rho_{\min}(K) > 0$. Hence by Lemma 3.23, the quantity defined for every $x \in \mathcal{M}$ as

$$\bar{\nu}(x) = \sup \left\{ \nu > 0 : B(x, \nu) \subset \mathcal{I}_x^{1/3} \right\}$$

is also lower bounded on K by a strictly positive constant that we denote $\nu_{\min}(K)$. Hence, if we require $h < h_3 := \min \left\{ h_1, \frac{\nu_{\min}(K)}{2Q} \right\}$, then for any $x, y \in B(\gamma(t), Qh)$ we have

$$d(x, y) \leq d(x, \gamma(t)) + d(y, \gamma(t)) < \frac{\nu_{\min}(K)}{2} + \frac{\nu_{\min}(K)}{2} = \nu_{\min}(K).$$

Therefore, if $h < h_3$, all control points defining H_h are contained into a retraction-convex set verifying the assumption of Corollary 3.16 and we are enabled to use the Lipschitz constants L_t, L_r and L_{xy} introduced in Corollary 3.16 for all the endpoint retraction curves appearing in the definition of H_h. To simplify notation, let us introduce

$$p_0 = \gamma(t), \quad v_0 = \dot{\gamma}(t), \quad p_1 = \gamma(t + h), \quad v_1 = \dot{\gamma}(t + h).$$

From Corollary 5.6, we have that

$$H_h(\tau) = \alpha \left(\frac{\tau - t}{h}; p_0, hv_0, p_1, hv_1 \right).$$

We shall prove that there exists L_{RH} (depending explicitly on L_t, L_r and L_t) such that for any $z_1, z_2 \in [0, 1]$

$$d(\alpha(z_1; p_0, hv_0, p_1, hv_1), \alpha(z_2; p_0, hv_0, p_1, hv_1)) \leq L_{RH}h|z_1 - z_2|. \qquad (5.9)$$

Then using (5.9), we can conclude that for any $\tau_1, \tau_2 \in [t, t + h]$

$$d(H_h(\tau_1), H_h(\tau_2)) = d\left(\alpha\left(\frac{\tau_1 - t}{h}; p_0, hv_0, p_1, hv_1 \right), \alpha\left(\frac{\tau_2 - t}{h}; p_0, hv_0, p_1, hv_1 \right) \right)$$
$$\leq L_{RH}h \frac{|\tau_1 - \tau_2|}{h} = L_{RH}|\tau_1 - \tau_2|.$$

Let us now prove (5.9) by unfolding the recursive definition of α.

$$\alpha(z; p_0, hv_0, p_1, hv_1) = \beta_{012}(z, z, z) = c_z(z; \beta_{01}(z, z), \beta_{12}(z, z)).$$

Applying several times Corollary 3.16 we find

$$d(c_{z_1}(z_1; \beta_{01}(z_1, z_1), \beta_{12}(z_1, z_1)), c_{z_2}(z_2; \beta_{01}(z_2, z_2), \beta_{12}(z_2, z_2)))$$

$$\leq |z_1 - z_2|\frac{L_t + L_r}{2} \left(d(\beta_{01}(z_1, z_1), \beta_{12}(z_1, z_1)) + d(\beta_{01}(z_2, z_2), \beta_{12}(z_2, z_2)) \right) \tag{5.10}$$

$$+ L_{xy} \left(d(\beta_{01}(z_1, z_1), \beta_{01}(z_2, z_2)) + d(\beta_{12}(z_1, z_1), \beta_{12}(z_2, z_2)) \right),$$

$$d(\beta_{01}(z_1, z_1), \beta_{12}(z_1, z_1)) = d(c_0(z_1, \beta_0(z_1), \beta_1(z_1)), c_1(z_1, \beta_1(z_1), \beta_2(z_1)))$$

$$\leq \left(\frac{1}{2}L_r + L_{xy} \right) \left(d(\beta_0(z_1), \beta_1(z_1)) + d(\beta_1(z_1), \beta_2(z_1)) \right),$$
$$\tag{5.11}$$

$$d(\beta_{01}(z_2, z_2), \beta_{12}(z_2, z_2)) \leq \left(\frac{1}{2}L_r + L_{xy} \right) \left(d(\beta_0(z_2), \beta_1(z_2)) + d(\beta_1(z_2), \beta_2(z_2)) \right),$$

$$d(\beta_{01}(z_1, z_1), \beta_{01}(z_2, z_2)) = d(c_0(z_1, \beta_0(z_1), \beta_1(z_1)), c_0(z_2, \beta_0(z_2), \beta_1(z_2)))$$

$$\leq \frac{L_t}{2}|z_1 - z_2| \left(d(\beta_0(z_1), \beta_1(z_1) + d(\beta_0(z_2), \beta_1(z_2)) \right)$$

$$+ L_{xy} \left(d(\beta_0(z_1), \beta_0(z_2)) + d(\beta_1(z_1), \beta_1(z_2)) \right),$$

$$d(\beta_{12}(z_1, z_1), \beta_{12}(z_2, z_2)) \leq \frac{L_t}{2}|z_1 - z_2| \left(d(\beta_1(z_1), \beta_2(z_1) + d(\beta_1(z_2), \beta_2(z_2)) \right)$$
$$\tag{5.12}$$

$$+ L_{xy} \left(d(\beta_1(z_1), \beta_1(z_2)) + d(\beta_2(z_1), \beta_2(z_2)) \right).$$

Plugging (5.11)-(5.12) in (5.10) and rearranging terms yields

$$d(c_{z_1}(z_1; \beta_{01}(z_1, z_1), \beta_{12}(z_1, z_1)), c_{z_2}(z_2; \beta_{01}(z_2, z_2), \beta_{12}(z_2, z_2)))$$

$$\leq |z_1 - z_2|\frac{(L_t + L_r)(L_r + 2L_{xy}) + 2L_{xy}L_t}{4} \tag{5.13}$$

$$\cdot \left(d(\beta_0(z_1), \beta_1(z_1)) + d(\beta_1(z_1), \beta_2(z_1)) + d(\beta_0(z_2), \beta_1(z_2)) + d(\beta_1(z_2), \beta_2(z_2)) \right)$$

$$+ L_{xy}^2 \left(d(\beta_0(z_1), \beta_0(z_2)) + 2d(\beta_1(z_1), \beta_1(z_2)) + d(\beta_2(z_1), \beta_2(z_2)) \right).$$

We now bound the seven distance function evaluations using once again Corollary 3.16.

$$d(\beta_0(z_1), \beta_1(z_1)) \leq \left(\frac{L_r}{4} + L_{xy} \right) \left(d(p_0, p_0^+) + d(p_0^+, p_1^-) \right),$$

$$d(\beta_0(z_2), \beta_1(z_2)) \leq \left(\frac{L_r}{4} + L_{xy} \right) \left(d(p_0, p_0^+) + d(p_0^+, p_1^-) \right),$$

$$d(\beta_1(z_1), \beta_2(z_1)) \leq \left(\frac{L_r}{4} + L_{xy} \right) \left(d(p_0^+, p_1^-) + d(p_1^-, p_1) \right),$$

$$d(\beta_1(z_2), \beta_2(z_2)) \leq \left(\frac{L_r}{4} + L_{xy} \right) \left(d(p_0^+, p_1^-) + d(p_1^-, p_1) \right),$$

$$d(\beta_0(z_1), \beta_0(z_2)) \leq L_t d(p_0, p_0^+)|z_1 - z_2|,$$

$$d(\beta_1(z_1), \beta_1(z_2)) \leq L_t d(p_0^+, p_1^-)|z_1 - z_2|,$$

$$d(\beta_2(z_1), \beta_2(z_2)) \leq L_t d(p_1^-, p_1)|z_1 - z_2|.$$

Inserting these bounds in (5.13) provides a constant $\tilde{L} > 0$ depending only on L_t, L_r and

L_{xy} such that

$$
\begin{aligned}
&d(\alpha\left(z_1; p_0, hv_0, p_1, hv_1\right), \alpha\left(z_2; p_0, hv_0, p_1, hv_1\right)) \\
&\leq (d(p_0, p_0^+) + d(p_0^+, p_1^-) + d(p_1^-, p_1))\tilde{L}|z_1 - z_2|.
\end{aligned}
\tag{5.14}
$$

Using Corollary 3.16-(i) and denoting L_γ the Lipschitz constant of the curve γ we find

$$
\begin{aligned}
d(p_0, p_0^+) &= d\left(\gamma\left(t\right), R_{\gamma(t)}\left(\frac{h}{3}\dot\gamma(t)\right)\right) \\
&\leq L_2\left\|\frac{h}{3}\dot\gamma(t)\right\| \leq \frac{h}{3}L_2 L_\gamma, \\
d(p_1^-, p_1) &= d\left(R_{\gamma(t+h)}\left(-\frac{h}{3}\dot\gamma(t+h)\right), \gamma(t+h)\right) \\
&\leq L_2\left\|\frac{h}{3}\dot\gamma(t+h)\right\| \leq \frac{h}{3}L_2 L_\gamma, \\
d(p_0^+, p_1^-) &\leq d(p_0^+, p_0) + d(p_0, p_1) + d(p_1, p_1^-) \\
&\leq \frac{2h}{3}L_2 L_\gamma + hL_\gamma.
\end{aligned}
$$

Finally, plugging these bounds into (5.14) proves (5.9) with $L_{RH} = \tilde{L}L_\gamma\left(\frac{4}{3}L_2 + 1\right)$ and concludes the proof. $\qquad\square$

Representability in normal coordinates and proof of Theorem 5.9

The Lipschitz continuity of H_h and of γ provides sufficient conditions to be able to represent the images of H_h and γ on the interval $[t, t + h]$ in normal coordinates around $\gamma(s)$, for any $s \in [t, t + h]$.

Lemma 5.12. *Denote* $r_{\min}(\gamma) := \min_{\tau \in [0,1]} \operatorname{inj}(\gamma(\tau))$, *the minimum of the injectivity radius of the Riemannian exponential map along the curve. There exists* $h_4 > 0$ *such that for every* $0 < h < h_4$ *and any* $t \in [0, 1 - h]$ *we have*

$$
\begin{aligned}
d(\gamma(s), \gamma(\tau)) &< \operatorname{inj}(\gamma), \\
d(\gamma(s), H_h(\tau)) &< \operatorname{inj}(\gamma),
\end{aligned}
\quad \forall\, s, \tau \in [t, t + h].
$$

Proof. The constant $\operatorname{inj}(\gamma)$ is strictly positive by the continuity of the injectivity radius function [Bou23, Proposition 10.18] and compactness of the image of the curve. We take any $0 < h < h_4 := \min\left\{\frac{\operatorname{inj}(\gamma)}{2L_\gamma}, h_3, \frac{\operatorname{inj}(\gamma)}{2L_{RH}}\right\}$, where h_3 is the constant introduced in Lemma 5.11, and consider an arbitrary $t \in [0, 1 - h]$. First note that for any $s, \tau \in [t, t + h]$ we have

$$
d(\gamma(s), \gamma(\tau)) \leq L_\gamma h < \frac{\operatorname{inj}(\gamma)}{2} < \operatorname{inj}(\gamma),
$$

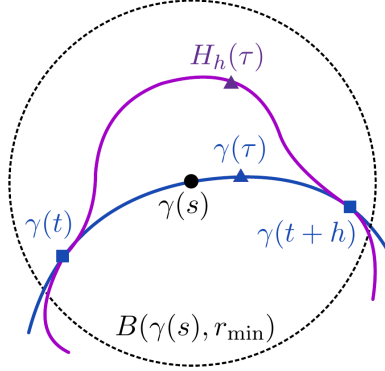Furthermore, the requirement $h < h_3$ guarantees that H_h is well-defined and, by Lemma 5.11,

Figure 5.6: Representability in normal coordinates as granted by Lemma 5.12.

that H_h is L_{RH}-Lipschitz continuous. Thus

$$d(\gamma(s), H_h(\tau)) \leq d(\gamma(s), \gamma(t)) + d(\gamma(t), H_h(\tau)) < \frac{\mathrm{inj}(\gamma)}{2} + d(H_h(t), H_h(\tau))$$
$$< \frac{\mathrm{inj}(\gamma)}{2} + L_{RH} h$$
$$< \frac{\mathrm{inj}(\gamma)}{2} + \frac{\mathrm{inj}(\gamma)}{2} = \mathrm{inj}(\gamma).$$

$\square$

Proof of Theorem 5.9. Take any $0 < h < h_4$, $t \in [0, 1-h]$ and $s \in [t, t+h]$. By Lemma 5.12, we can express γ and H_h in normal coordinates, see Figure 5.6. Define for all $\tau \in [t, t+h]$

$$\hat{\gamma}_s(\tau) = \mathrm{Exp}_{\gamma(s)}^{-1}(\gamma(\tau)),$$
$$\hat{H}_s(\tau) = \mathrm{Exp}_{\gamma(s)}^{-1}(H_h(\tau)).$$

The interpolation error in normal coordinates at $\gamma(s)$ is defined as $\hat{E}_s(\tau) = \hat{\gamma}_s(\tau) - \hat{H}_s(\tau)$ for every $\tau \in [t, t+h]$ and by construction satisfies

$$\hat{E}_s(t) = 0, \qquad\qquad \hat{E}_s(t+h) = 0,$$
$$\frac{\mathrm{d}}{\mathrm{d}\tau}\hat{E}_s(\tau)\big|_{\tau=t} = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}\tau}\hat{E}_s(\tau)\big|_{\tau=t+h} = 0. \tag{5.15}$$

Consider any orthonormal basis $\{b_i\}_{i=1}^{D}$ for $T_{\gamma}(s)\mathcal{M}$, denote $\hat{E}_{s,i}(\tau) = \langle \hat{E}_s(\tau), b_i \rangle_{\gamma(s)}$ the components functions of the error. For any fixed $\tau \in [t, t+h]$, let us defined the function

$$G_i(x) = \hat{E}_{s,i}(x) - \omega(x)\frac{\hat{E}_{s,i}(\tau)}{\omega(\tau)}, \quad \forall\, x \in [t, t+h],$$

where $\omega(x) = (x - t)^2(x - (t+h))^2$. From (5.15) and the definition of G_i we deduce

$$G_i(t) = G_i(t+h) = G_i(\tau) = 0. \tag{5.16}$$

By the smoothness of the curve γ, the interpolant H_h and of the inverse retraction we have that $G \in C^4([t, t+h])$. By Rolle's theorem, (5.16) implies G_i' admits at least two zeros in $(t, t+h)$ different from τ. Furthermore, by the interpolation condition we have $G_i'(t) = G_i'(t+h) = 0$. Hence, we have identified at least four zeros of G_i' in $[t, t+h]$. By Rolle's theorem applied to G_i', G_i'' admits at least three zeros on $[t, t+h]$. Sequentially applying Rolle's theorem to G_i'' and then to G_i''', we find there exists $\xi_i \in (t, t+h)$ such that

$$G_i^{(4)}(\xi_i) = 0$$

$$\Leftrightarrow \quad \hat{E}_{s,i}^{(4)}(\xi_i) - \underbrace{\omega^{(4)}(\xi_i)}_{=4!} \frac{\hat{E}_{s,i}(\tau)}{\omega(\tau)} = 0$$

$$\Leftrightarrow \quad \hat{E}_{s,i}(\tau) = \frac{\hat{E}_{s,i}^{(4)}(\xi_i)}{4!} \omega(\tau)$$

For any $\tau \in [t, t+h]$ we have $|\omega(\tau)| < h^4$, from which we evince

$$|\hat{E}_{s,i}(\tau)| \leq \frac{|\hat{E}_{s,i}^{(4)}(\xi_i)|}{4!} h^4.$$

Therefore, for any $s \in [t, t+h]$ we can bound the norm of interpolation error at τ given as

$$
\begin{aligned}
\left\| \hat{E}_s(\tau) \right\|_{\gamma(s)} = \left\| \sum_{i=1}^{D} \hat{E}_{s,i}(\tau) b_i \right\|_{\gamma(s)} &\leq \frac{h^4}{4!} \left(\sum_{i=1}^{D} |\hat{E}_{s,i}^{(4)}(\xi_i)|^2 \right)^{1/2} \\
&\leq \frac{h^4}{4!} \sqrt{D} \max_{i=1,\dots,D} |\hat{E}_{s,i}^{(4)}(\xi_i)| \\
&\leq \frac{h^4}{4!} \sqrt{D} \max_{i=1,\dots,D} \left\| \hat{E}_s^{(4)}(\xi_i) \right\|_{\gamma(s)} \\
&\leq \frac{h^4}{4!} \sqrt{D} \max_{\xi \in [t,t+h]} \left\| \hat{E}_s^{(4)}(\xi) \right\|_{\gamma(s)}.
\end{aligned}
$$

The appearance of the $\sqrt{D}$ factor is a consequence of the fact the ξ_i given by Rolle's theorem may different from one component to the other.

An important property of normal coordinates following from Proposition 1.33 is that radial directions map to length-minimizing geodesics, thus $d(\gamma(\tau), H(\tau)) = \left\| \hat{E}_\tau(\tau) \right\|_{\gamma(\tau)}$. So we can say

$$
\begin{aligned}
d(\gamma(\tau), H(\tau)) &\leq \frac{h^4}{4!} \sqrt{D} \max_{s,\xi \in [t,t+h]} \left\| \hat{E}_s^4(\xi) \right\|_2 \\
&\leq \frac{h^4}{4!} \sqrt{D} \max_{s,\xi \in [t,t+h]} \left\{ \left\| \hat{\gamma}_s^{(4)}(\xi) \right\|_2 + \left\| \hat{H}_s^{(4)}(\xi) \right\|_2 \right\}
\end{aligned}
$$

Since the exponential map is a particular retraction, from Proposition 3.20 we deduce

the expressions of $\hat{\gamma}_s^{(4)}$ and $\hat{H}_s^{(4)}$ are respectively a sum of contributions of the form

$$\mathrm{D}^k\mathrm{Exp}_{\gamma(s)}^{-1}(\gamma(\xi))\left[\frac{\mathrm{D}^{i_1}\gamma(\xi)}{\mathrm{d}\xi^{i_1}},\cdots,\frac{\mathrm{D}^{i_j}\gamma(\xi)}{\mathrm{d}\xi^{i_j}}\right]$$

and

$$\mathrm{D}^k\mathrm{Exp}_{\gamma(s)}^{-1}(H_h(\xi))\left[\frac{\mathrm{D}^{i_1}H_h(\xi)}{\mathrm{d}\xi^{i_1}},\cdots,\frac{\mathrm{D}^{i_k}H_h(\xi)}{\mathrm{d}\xi^{i_k}}\right]$$

with $k \in \{1, 2, 3, 4\}$ and $i_j \geq 0$ such that $i_1 + \cdots + i_k = 4$ indicating the order of manifold curve derivative and where $\mathrm{D}^k\mathrm{Exp}_{\gamma(s)}^{-1}(\cdot)$ is a multilinear operator: the linear operator corresponding to the case $k = 1$ is defined in (3.5) whereas the multilinear operators associated to the cases $k = 2$ and $k = 3, 4$ are introduced respectively in Definition 3.17 and Definition 3.19. In order to invoke for the particular case of the exponential map the Lipschitz constant of the inverse retraction given Proposition 3.12 and the upper bound for the operator norms of the inverse retraction differentials given in Propositions 3.18 and 3.20, we restrict $h \leq h_4/3$ so that the $\gamma(\xi)$ and $H_h(\xi)$ are at most $\mathrm{inj}(\gamma)/3$ away from $\gamma(s)$, for every $s \in [t, t + h]$ and $t \in [0, 1 - h]$. Then, the norm of each term can be bounded as

$$\left\|\mathrm{D}^k\mathrm{Exp}_{\gamma(s)}^{-1}(\gamma(\xi))\left[\frac{\mathrm{D}^{i_1}\gamma(\xi)}{\mathrm{d}\xi^{i_1}},\ldots,\frac{\mathrm{D}^{i_j}\gamma(\xi)}{\mathrm{d}\xi^{i_j}}\right]\right\|_{\gamma(s)} \leq M_{2,k}(\gamma, 1/3)\left\|\frac{\mathrm{D}^{i_1}\gamma}{\mathrm{d}\xi^{i_1}}\right\|_\infty \cdots \left\|\frac{\mathrm{D}^{i_k}\gamma}{\mathrm{d}\xi^{i_k}}\right\|_\infty,$$

where

$$\left\|\frac{\mathrm{D}^{i_j}\gamma}{\mathrm{d}\xi^{i_j}}\right\|_\infty = \max_{t\in[0,1]}\left\|\frac{\mathrm{D}^{i_j}\gamma(t)}{\mathrm{d}\xi^{i_j}}\right\|_{\gamma(t)}$$

and, provided $h < h_2$, using the constants given by assumption (5.8),

$$\left\|\mathrm{D}^k\mathrm{Exp}_{\gamma(s)}^{-1}(H_h(\xi))\left[\frac{\mathrm{D}^{i_1}H_h(\xi)}{\mathrm{d}\xi^{i_1}},\ldots,\frac{\mathrm{D}^{i_k}H_h(\xi)}{\mathrm{d}\xi^{i_k}}\right]\right\|_{\gamma(s)} \leq M_{2,k}(\gamma, 1/3)L_{RH}^{(i_1)}\cdot\ldots\cdot L_{RH}^{(i_k)}$$

This produces a constant $\kappa > 0$ such that for any $h < \bar{h} := \min\{h_2, h_4/3\}$ and any $t \in [0, 1 - h]$ we have

$$\max_{\tau\in[t,t+h]}d(\gamma(\tau), H_h(\tau)) < \sqrt{D}\frac{\kappa}{4!}h^4.$$

$\square$

5.5 Numerical experiments

The following section is dedicated to numerical experiments illustrating the RH interpolation method, summarized in Algorithms 5.1 and 5.2. All experiments have been carried out in Matlab 2019b leveraging the differential geometry tools of the Manopt library [BMAS14] on a laptop computer with Intel i7 CPU (1.8GHz with single-thread mode) with 8GB of RAM, 1MB of L2 cache and 8MB of L3 cache.

5.5.1   Academic examples

Our first goal is to measure experimentally the accuracy of the RH interpolation method in order to highlight the trend of convergence of the interpolation error predicted by Theorem 5.9. For this, we apply RH interpolation on two problems arising from numerical linear algebra: the computation of a smooth QR decomposition and a smooth singular value decomposition for a given smooth matrix curve $t \in [a, b] \rightarrow A(t) \in \mathbb{R}^{m \times n}$. Assuming for the moment that these smooth decompositions exist and can be computed, the experimental setup is the following. We sample the decomposition and its first-order derivative at uniformly spaced location and interpolate this data with different manifold interpolation schemes. We then vary the sampling step size h and measure the maximum interpolation error with respect to the original smooth decomposition.

Comparing with other retraction-based schemes

The RH method is compared with two interpolation schemes that only use retractions. First, the analog of the piecewise linear interpolant is defined as $L(t)\big|_{[t_i, t_{i+1}]} = c_0\left(\frac{t-t_i}{h}, p_i, p_{i+1}\right)$, where c_0 is the endpoint retraction curve defined in Definition 3.4. We then consider a naive piecewise Hermite interpolant which uses the same control points as the RH scheme but where only the endpoint curve c_0 is used as the building block for the generalized de Casteljau procedure. The resulting curve is then not expected to be continuously differentiable at the junctions.

Implementation details

To measure the error of an approximation $t \rightarrow \tilde{A}(t)$ to a matrix manifold curve $A(t)$, we consider the pointwise errors

$$\varepsilon_P(t) = \left\|A(t) - \tilde{A}(t)\right\|_F \quad \text{and} \quad \varepsilon_D(t) = \left\|\dot{A}(t) - \dot{\tilde{A}}(t)\right\|_F.$$

For simplicity and the purpose of these experiments, we compute all required derivatives via centered finite differences:

$$\dot{A}(t) \simeq \Pi_{A(t)}\left(\frac{A(t + \Delta t) - A(t - \Delta t)}{2\Delta t}\right), \quad \Delta t = 10^{-5}. \tag{5.17}$$

We point out that for the particular case of smooth parameter dependent QR and SVD factorizations, closed-form expressions for the derivatives of the factors are known [DE99] and could have been used as an alternative. Note that Theorem 5.9 uses the Riemannian distance to measure the interpolation error while we measure the error with the ambient space distance. It can be shown that in cases like ours where the manifold is embedded into an Euclidean space and endowed with the induced metric, the Euclidean distance is locally equivalent to the Riemannian distance, see e.g. [AEM07, Appendix A]. In our

context, this precisely means

$$\lim_{x \to y} \frac{d(X,Y)}{\|X - Y\|_F} = 1, \quad \forall\, X, Y \in \mathcal{M} \subseteq \mathbb{R}^{m \times n}.$$

Q-factor interpolation

For convenience, the example matrix curve we consider is the same as in [Zim20, §5.2]. It consists of the matrix polynomial

$$Y(t) = Y_0 + tY_1 + t^2 Y_2 + t^3 Y_3, \quad Y_i \in \mathbb{R}^{n \times k}, n = 500, k = 10, t \in [-1.1, 1.1], \quad (5.18)$$

where the entries of the matrices Y_i are pseudo-randomly generated from uniform distributions on $[0, 1]$, $[0, 0.5]$, $[0, 0.5]$, $[0, 0.2]$ respectively. The matrix $Y(t)$ is generically full-rank for every $t \in [-1.1, 1.1]$ and is smooth, thus owing to [DE99, Proposition 2.3] there exist unique smooth curves $t \to Q(t)$, with $Q(t)$ belonging to the Stiefel manifold $\mathrm{St}(n, k)$, and $t \to R(t) \in \mathbb{R}^{k \times k}$ with positive diagonal entries, such that $Y(t) = Q(t)R(t)$. The positivity of the diagonal entries is explicitly enforced in the experiment. We focus on interpolating the curve $Q(t)$ on $\mathrm{St}(n, k)$. At each sample location t_i we store $p_i = Q(t_i)$ and $v_i = \dot{Q}(t_i)$ obtained with (5.17).
In Figure 5.7a, we plot the pointwise and derivative error as a function of the curve parameter t when interpolating (5.18) with different schemes. While all schemes interpolate correctly the data points (left panel), only the RH scheme manages to match the derivative at sample points (right panel). Figure 5.8a illustrates the result of Theorem 5.9. Plotting the maximum pointwise interpolation error against the sampling step sizes h reveals the expected $O(h^4)$ trend for the RH scheme. Interestingly, as in the Euclidean case, the derivative error converges one order slower than the pointwise error. For these experiments we used the P-factor retraction, but analogous result are found with the Q-factor retraction. The difference between the two retractions is also negligible in terms of evaluation time as it can be seen from Table 5.1. These results also show that with the offline/online procedure proposed in Algorithms 5.1 and 5.2, the evaluation cost of the RH scheme is comparable with the one of other schemes. Note that, for a fair comparison, the other schemes have also been implemented in an offline/online fashion to minimize evaluation cost.

SVD interpolation

We interpolate the singular value decomposition of a matrix curve of constant and low rank. We consider $m = 10^4$, $n = 300$ and rank $r = 10$. We take

$$Y(t) = Y_0 + tY_1 + t^2 Y_2 + t^3 Y_3, \quad Y_i \in \mathbb{R}^{m \times r}$$
$$Z(t) = Z_0 + tZ_1 + t^2 Z_2, \quad Z_i \in \mathbb{R}^{r \times n},$$

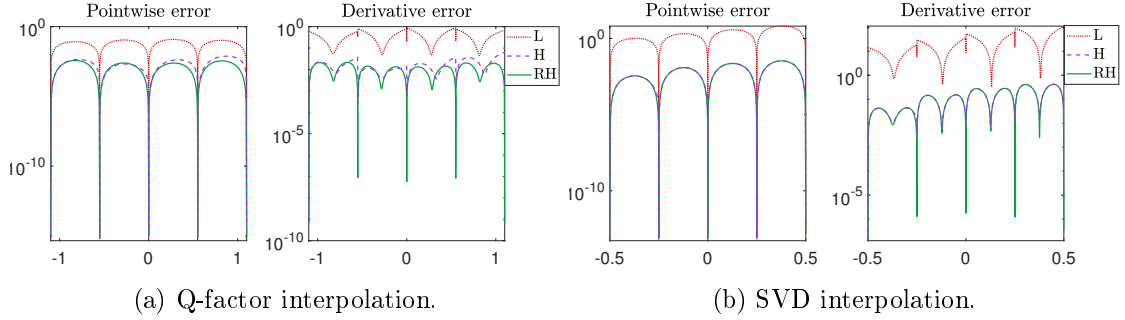(a) Q-factor interpolation. (b) SVD interpolation.

Figure 5.7: Interpolation error vs t for different retraction-based interpolants: linear interpolant (L), naive Hermite interpolant (H) and the RH interpolant.

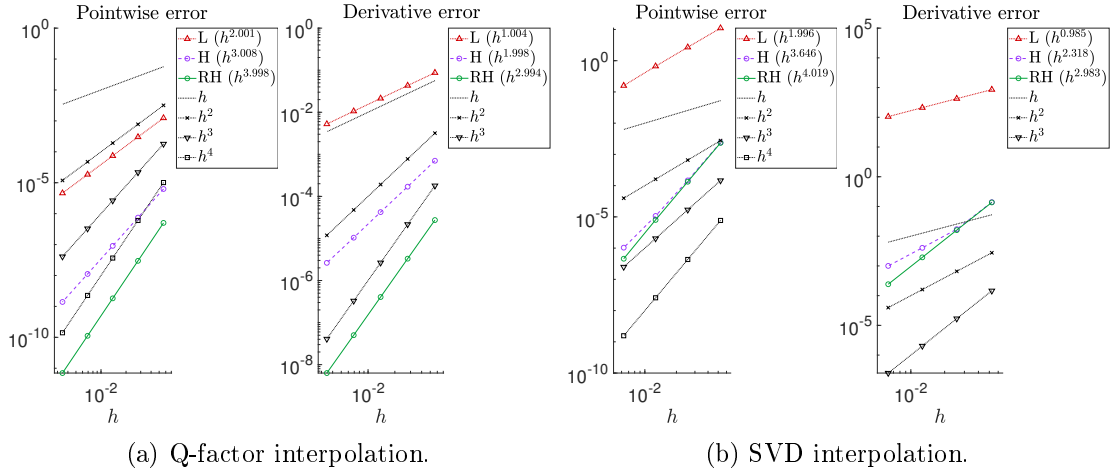(a) Q-factor interpolation. (b) SVD interpolation.

Figure 5.8: Interpolation error as a function of the sampling step size h for different retraction-based interpolants: linear interpolant (L), naive Hermite interpolant (H) and the RH interpolant.

Table 5.1: Average time per evaluation for the Q-factor and SVD interpolation experiments of Figure 5.7a and Figure 5.7b. In the last two lines we distinguish the simple and the optimized offline/online implementations of the RH scheme (see Algorithms 5.1 and 5.2).

	Q-factor interpolation		SVD interpolation
Retraction used	Q-factor	P-factor	Orthographic
Linear (L)	$1.524 \cdot 10^{-4}$	$2.049 \cdot 10^{-4}$	$1.710 \cdot 10^{-3}$
Hermite (H)	$1.540 \cdot 10^{-3}$	$1.596 \cdot 10^{-3}$	$1.147 \cdot 10^{-2}$
Retraction Hermite (RH)	$4.335 \cdot 10^{-3}$	$3.531 \cdot 10^{-3}$	$1.992 \cdot 10^{-2}$
Retraction Hermite (RH, optimized)	$2.226 \cdot 10^{-3}$	$2.188 \cdot 10^{-3}$	$1.558 \cdot 10^{-2}$

where the entries of Y_0, Z_0 and Y_1, Y_2, Y_3, Z_1, Z_2 are drawn for uniform random distributions on $[0, 1]$ and $[0, 0.5]$ respectively. Since the factors are generically full rank, the curve

$$W(t) = Y(t)Z(t)^\top, \quad t \in [-0.5, 0.5] \tag{5.19}$$

is of rank r for every t. This example has also been taken from [Zim20, §5.3] together with the suggestion of Remark 6 to ensure the smoothness of the computed SVD decomposition path $U(t)\Sigma(t)V(t)^\top = W(t)$. Note that this may cause negative values in the diagonal term.

The experimental results for the SVD path of (5.19) are reported in Figures 5.7b and 5.8b. The comments are analogous to the one we made for the Q-factor interpolation experiments: only the RH scheme manages to match the prescribed derivatives at the nodes thereby producing the expected $O(h^4)$ error convergence trend. For these experiments, the naive retraction-based Hermite interpolation scheme H produces a good approximation in terms of error, practically as good as the RH scheme. However, the curve H is not globally continuously differentiable.

The computation time per evaluation for the different schemes is reported in Table 5.1. Interestingly, while for the Stiefel manifold, a naive implementation of the RH scheme is approximately two times slower than the offline/online approach, for the low-rank manifold, the non-optimized code is only 30% more expensive. We attribute this to the fact that the inverse orthographic retraction is relatively cheap compared to the retraction and so the few inverse retractions spared by the offline/online implementation do not pay off as much.

The need for bounded derivatives

The fourth-order convergence achieved by the RH interpolation scheme was proved in Theorem 5.9 under the assumption that all derivatives up to order four of the interpolation curve remain bounded as $h \to 0$, see (5.8). As we now illustrate, this assumption can in fact not be removed. Recall that the RH interpolation scheme was built to satisfy Proposition 5.3 by making the choice (5.4). However, this choice was not unique and was motivated by the need to alleviate evaluation cost. It turns out that choice (5.4) is also important because it satisfies the bounded derivatives assumption (5.8). In Figure 5.9a we plot the maximum norm of the second, third and fourth-order derivatives as a function of the sampling size h for the RH scheme and for an alternative scheme denoted RH* where choice (5.4) is modified with $r_1 = 0$ instead of $r_1 = 1/2$. It appears that unlike the RH scheme, the RH* scheme features a fourth derivative diverging as $O(h^{-1})$. Despite producing a continuously differentiable curve which interpolates the derivatives, the alternative scheme loses one order of accuracy as can be seen from Figure 5.9b. These experiments were conducted on the Q-factor interpolation of (5.18) on the Stiefel manifold (see Figure 5.8a) but analogous results were found on the SVD interpolation instance.

The choice (5.4) is not the only one satisfying assumption (5.8). For instance, for any

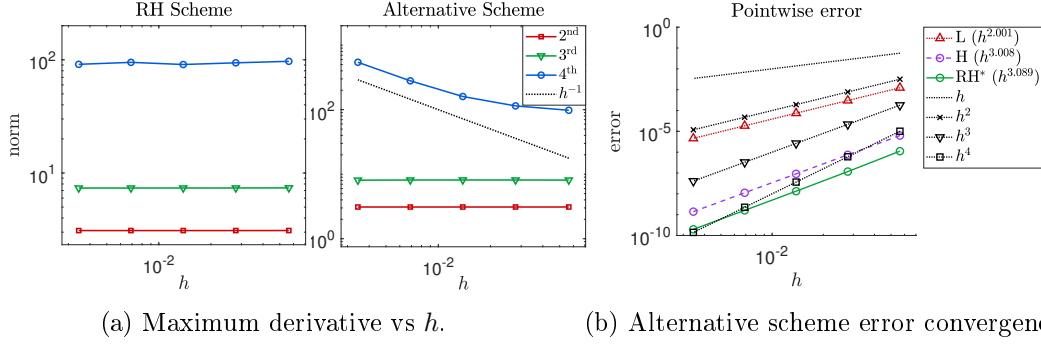(a) Maximum derivative vs h. (b) Alternative scheme error convergence.

Figure 5.9: Experiments with an alternative interpolation scheme still satisfying Proposition 5.3 but failing to satisfy assumption (5.8).

choice of $r_1 \in [0,1]$ if we take $r_{01}(s,t) = (1-r_1)t$, $r_{12}(s,t) = (1-r_1) + r_1 t$ and $r_{012}(u,s,t) = t$ we still achieve $O(h^4)$ error convergence. However, the evaluation cost of such schemes is higher and among those that we could find, choice (5.4) was the cheapest. The relationship between the choice of these functions and the fourth derivative of the scheme is intricate and we could not establish an a priori criterion to discriminate between schemes satisfying (5.8) and those violating it.

5.5.2 Applications

In this section, we illustrate two possible applications of the RH interpolation scheme. We focus on applications involving the fixed-rank manifold as we believe they are the most relevant.

Riemannian continuation

As a first application, we propose an extension of the prediction-correction numerical continuation algorithm for parameter dependent Riemannian optimization developed in Chapter 4. The accuracy of the prediction step of the continuation algorithm in approximating the solution to the next problem can be characterized by the prediction order, see Definition 4.4. Assuming the well-posedness of the RH interpolant and the result of Theorem 5.9 extend to an interval larger than the interpolation interval, using the RH interpolation scheme in the prediction step would produce a prediction order $p_{\mathrm{RH}} = 4$ compared to $p_{\mathrm{C}} = 1$ and $p_{\mathrm{T}} = 2$ for the classical and tangential prediction schemes. In practice, in the notation of Algorithm 4.1 with constant step size $h_k = h$ for every $k \geq 0$, we define the RH prediction step as

$$y_{k+1} = \begin{cases} R_{x_k}(ht_k) & k = 0, \\ H(2h; \{x_{k-1}, t_{k-1}, 0\}\{x_k, t_k, h\}), & k \geq 1. \end{cases}$$
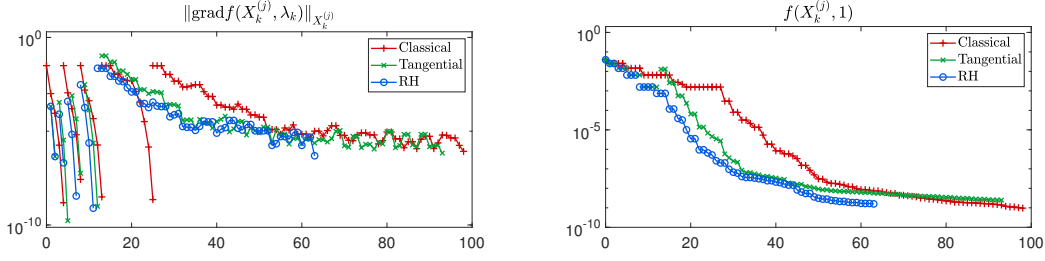
Figure 5.10: RH continuation algorithm on a favorable instance of matrix completion. Left: Riemannian gradient of the sequence of problems versus total iteration count. Right: objective function value of the final optimization problem versus total iteration count.

Table 5.2: Performance statistics of the Riemannian continuation algorithm for three different prediction schemes over 100 randomly generated instances of the matrix completion problem.

	Time (s)		Correction steps (RTR)	
	Mean	Median	Mean	Median
Classical	3.683	3.625	59.56	60
Tangential	8.013	4.944	69.95	53
RH	3.888	2.783	48.72	45

In the first step, we do standard tangential prediction and from the second prediction on, we evaluate in $\tau = 2h$ the RH interpolant of the point and derivative data $(x_{k-1}, t_{k-1}), (x_k, t_k) \in T\mathcal{M}$ located at $\tau = 0$ and $\tau = h$ respectively.

The RH prediction-correction continuation algorithm with fixed step size is applied to the same low-rank matrix completion problem considered in [SK22a, §5]. We fix the number of steps to $N_{\text{steps}} = 5$, use the Riemannian Trust Regions (RTR) algorithm as a corrector and vary the prediction scheme. We report in Table 5.2 a comparison of the computational effort required to solve the problem with each scheme. There are two factors that determine the performance of the algorithm. First, the more ill-conditioned the final optimization problem is, the more the last RTR correction encounters stagnation. Second, the more the underlying solution curve to the family of optimization problem is smooth, the more tangential and RH prediction pays off. In fact, high prediction order is achieved only when the underlying solution curve to the family of optimization problems is sufficiently smooth. Partly due to the choice of the homotopy, the solution curve often exhibited discontinuities thereby undermining the efficiency of tangential and RH prediction. Because of this, these two schemes are on average slower than classical prediction. However, when the underlying solution curve happens to be smooth, e.g. for the instance of Figure 5.10, the RH prediction benefits from the increased prediction accuracy. As can be seen from the computed medians for computation time and total RTR iteration count, the RH can significantly reduce the computational effort compared to classical and tangential prediction.

Dynamical-low-rank approximation interpolation

Dynamical low-rank approximation (DLRA) techniques are used to address the time integration of space discretized partial differential equations to mitigate computational costs and memory requirements. For 2D problems, the discretization of the PDE yields a matrix ordinary differential equation that DLRA techniques evolve on low-rank manifolds [KL07]. See also Chapter 6 for a more detailed introduction. We consider DLRA integrators for which the rank remains fixed so that their output consists of a sequence $\{\tilde{Y}_i\}_{i=0}^{N_t} \subset \mathcal{M}_k$, where k is the chosen rank, such that $\tilde{Y}_i \simeq Y(t_i)$, the solution at time t_i of the matrix differential equation $\dot{Y} = F(Y, t)$, $Y \in \mathbb{R}^{m \times n}$, $t \in [0, T]$. During the numerical integration, for each $\tilde{Y}_i$ we can store the best approximation of the vector $F(Y_i, t_i)$ on the tangent space $T_{\tilde{Y}_i} \mathcal{M}_k$, obtained as $\tilde{V}_i := \Pi_{\tilde{Y}_i}(F(\tilde{Y}_i, t_i))$. Then, the collection of triplets $\{(\tilde{Y}_i, \tilde{V}_i, t_i)\}_{i=0}^{N_t}$ can be fed to our RH interpolation scheme to obtain a continuously differentiable curve on $\mathcal{M}_k$ that approximates the best rank k solution for every time t. Given the high accuracy of the interpolation scheme, one can expect that it is sufficient to interpolate a small fraction of the time samples to obtain a satisfactory approximation of the full solution curve.
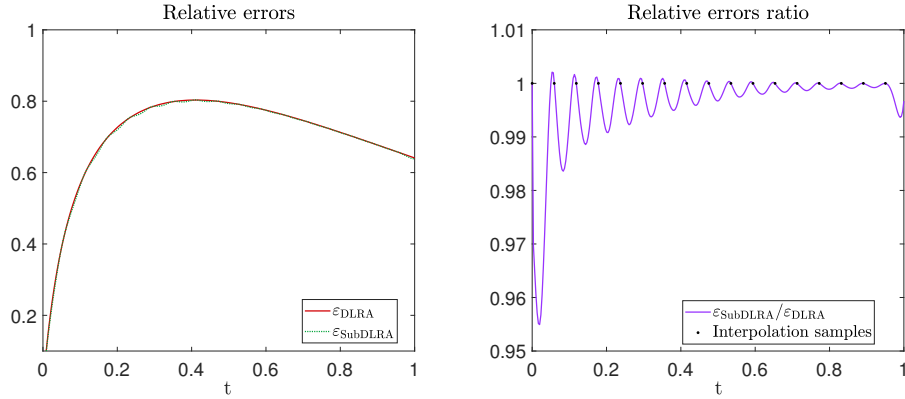

Figure 5.11: Relative error with reference solution of the interpolation of DLRA samples. The error $\varepsilon_{\mathrm{DLRA}}$ corresponds to interpolating all samples whereas $\varepsilon_{\mathrm{SubDLRA}}$ only 1 out of 20 samples.

We tested this hypothesis on the test case of [KEC23, §7.1] for which the DLRA integrator is the so-called unconventional time integrator [CL22]. For conciseness, we refer the reader to these two references for details on the radiation transport equation at hand. The time stepping scheme is also known as the KLS projector splitting integrator and is discussed in more detail in Chapter 6 (see Section 6.2). The parameters for the experiments are the same as in [KEC23, §7.1]: the integration is done on $\mathcal{M}_k \subset \mathbb{R}^{m \times n}$ with $m = 800$, $n = 100$ and $k = 15$. The only difference is that we turn off the rank adaptivity option of the integrator since we want the initial condition and all iterates to remain on $\mathcal{M}_k$. In order to obtain a rank 15 initial condition, we run the rank-adaptive version starting from the rank 1 initial condition used in [KEC23, §7.1] and store the

first sample of rank 15. From this initial condition, we run the unconventional DLRA integrator with a step size chosen to have a CFL constant of 1. This yields $\{\tilde{Y}_i\}_{i=0}^{N_t} \subset \mathcal{M}_k$. We then obtain a reference solution by performing a forward Euler integration of the ODE in $\mathbb{R}^{m \times n}$ with a step size 20 times smaller than the unconventional integrator. These samples are then projected to $\mathcal{M}_k$ with the k-truncated SVD, and we denote them $\{Y_j\}_{j=0}^{20N_t}$. Finally, we assemble three RH interpolants:

$$\tilde{Y}(t) = H\left(t; \left\{\left(\tilde{Y}_i, \tilde{V}_i, t_i\right)\right\}_{i=0}^{N_t}\right),$$

$$\hat{Y}(t) = H\left(t; \left\{\left(\tilde{Y}_{20l}, \tilde{V}_{20l}, t_{20l}\right)\right\}_{l=0}^{\lfloor N_t/20 \rfloor}\right),$$

$$Y(t) = H\left(t; \{(Y_j, V_j, t_j)\}_{j=0}^{20N_t}\right)$$

where tangent vectors $\tilde{V}_i$ and V_i are obtained as explained previously. The curve $\hat{Y}$ interpolates one every 20 samples of the DLRA solution, so roughly 5% of the integrator's output. Yet, as can be seen from Figure 5.11, the relative errors

$$\varepsilon_{\text{DLRA}}(t) = \frac{\left\|\tilde{Y}(t) - Y(t)\right\|_F}{\|Y(t)\|_F}, \quad \varepsilon_{\text{SubDLRA}}(t) = \frac{\left\|\hat{Y}(t) - Y(t)\right\|_F}{\|Y(t)\|_F},$$

are almost identical. Surprisingly, the sub-sampled interpolation curve $\hat{Y}$ can be more precise, though from a negligible amount. The real advantage comes in terms of storage requirements: the information needed to evaluate the sub-sampled interpolation curve $\hat{Y}$ occupies 20 times less memory than the information for the curve $\tilde{Y}$ and approximately 4 times less storage than the full collection of samples $\{\tilde{Y}_i\}_{i=0}^{N_t}$. This application of RH interpolation can be thought as a compression post-processing that enhances the portability of DLRA solutions.

5.6   Conclusion

In this chapter we have proposed a manifold interpolation technique to address Hermite interpolation of manifold curves. The method is general enough to be applicable to every manifold for which a retraction/inverse retraction pair is available, thereby avoiding the need for Riemannian exponential and logarithmic maps used by other interpolation schemes that take into account derivative information.

The novel notion of retraction-convex sets ensures the well-posedness of the method, provided that consecutive interpolation data points are sufficiently close. Theorem 5.9 generalizes to our scheme the classical interpolation error convergence result for polynomial Hermite interpolation in Euclidean spaces. The predicted $O(h^4)$ convergence trend has also been experimentally observed for academic interpolation problems on the Stiefel manifold and the fixed-rank matrix manifold. The high-order accuracy of the method allowed us to propose an improvement to the prediction-correction Riemannian continua-

tion algorithm introduced in Chapter 4 and to suggest a strategy to compress the output of dynamical low-rank matrix integrators.

Just like curve interpolation is a basic tool in many context of numerical analysis, we believe the RH interpolation scheme could serve as a building block for other numerical methods on manifolds. As an illustration of this, in the next chapter the interpolation scheme is used to define a numerical integration method for manifold-constrained ODEs.

6 Manifold integration

This chapter introduces two novel numerical integration schemes for manifold-constrained ordinary differential equations: the accelerated forward Euler (AFE) method and the Ralston-Hermite (RH) method. Both the proposed methods are defined using retractions and building upon the perspective developed in Chapter 3 and Chapter 5 of seeing retractions as a tool to construct portions of manifold curves approximating an underlying curve. By leveraging the results developed in Chapters 3 and 5, the two methods are argued to exhibit local truncation error of order three and numerical experiments suggests that both methods can achieve global error convergence of order two.

Despite the broad applicability of the methods, the present work focuses on the class of differential equations evolving on low-rank manifolds encountered in dynamical low-rank approximation (DLRA) [KL07]. This choice is motivated by the rising interest of these techniques in applications. Therefore, the accuracy and performance of the methods are demonstrated on classical problem instances from the DLRA literature and compared with state-of-the-art DLRA techniques. As customary for this type of problems, we also investigate the stability to small singular values of the new schemes.

This chapter takes the viewpoint of retractions on existing DLRA techniques to underline, as already noticed by several authors, that retractions are a convenient language to build numerical integration schemes on manifolds [KV19, CL23]. It helps also to shed new light on some features of DLRA. In particular, in this chapter gives a geometric interpretation of the KLS unconventional integrator [CL22] by showing that it is a second-order retraction on the fixed-rank matrix manifold which coincides with the orthographic retraction up to high-order terms.

Contributions and outline of the chapter. We first introduce and state the DLRA problem of interest in Section 6.1 together with a discussion on the error analysis results of existing methods. The link between retractions and existing DLRA numerical integration techniques, in particular the KLS scheme, is the object of Section 6.2. The two main contributions are presented in Section 6.3: the AFE method in Section 6.4 and the RH integration scheme in Section 6.5. For completeness and to complement the derivation of the AFE scheme, we provide in Section 6.4.1 a overview on the definition of the Weingarten map for embedded submanifolds of Euclidean spaces, illustrated by the
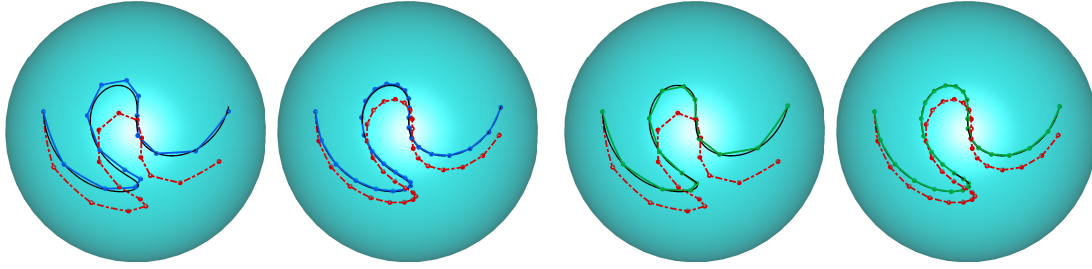
(a) Accelerated forward Euler (AFE) method.   (b) Ralston-Hermite (RH) method.

Figure 6.1: Numerical integration of a simple ordinary differential equation on S^2 with different retraction-based methods illustrating the accuracy of the methods proposed in this chapter. The differential equation is of the form $\dot{y} = \dot{\gamma}(t)$ with $y_0 = \gamma(0)$ for a given curve $\gamma \subset S^2$ (black). For $N = 15$ and $N = 30$ numerical integration steps of equal size, the retraction-based forward Euler method (6.4) is compared with the accelerated forward Euler (blue) presented in Section 6.4 and the Ralston-Hermite method (green) presented in Section 6.5.

example of the fixed-rank matrix manifold. Finally, numerical experiments are reported in Section 6.6. A preprint on the contents of this chapter is in preparation [SCK23].

6.1 Background on DLRA

The starting point of DLRA is to acknowledge that the space-discretization of a large class of partial differential equations results in ordinary differential equations defined on a high-dimensional matrix or tensor vector space $\mathcal{E}$ whose solutions exhibit a low-rank structure, for an appropriate notion of rank depending on $\mathcal{E}$. This includes equations with diffusive or advective terms but also certain types of nonlinearities. In this chapter, we consider the case where $\mathcal{E} = \mathbb{R}^{m \times n}$, with $m, n \gg 1$ and so the notion of rank is simply the matrix rank. Let us consider an initial value problem governed by a smooth vector field $F \in \mathfrak{X}(\mathbb{R}^{m \times n})$ which, for ease of exposition, does not explicitly depend on time:

$$\begin{cases} A' = F(A), & t \in [0, T], \\ A(0) = A_0 \in \mathbb{R}^{m \times n}. \end{cases} \tag{6.1}$$

In this setting, DLRA can be summarized as trying to find an approximation of the so-called *ambient solution* $A(t)$ onto the manifold of rank-k matrices, with $k \ll \min\{m, n\}$. The goal is to gain in computational efficiency without compromising the accuracy too much. Indeed, representing the approximation in factored form drastically reduces storage complexity while the low-rank approximability of the ambient solution makes low approximation error possible. Then, the central challenge of DLRA consists of computing efficiently a factored low-rank approximation of the ambient solution without having to first estimate the ambient solution and then to truncate it to an accurate rank-k approximation. In the following, we restrict to the case where the rank is fixed a priori

and does not change throughout the integration interval. Then, the DLRA problem in consideration associated to the ambient equation (6.1) can be formalized as follows.

Problem 6.1. *Given a smooth vector field $F \in \mathfrak{X}(\mathbb{R}^{m \times n})$, an initial matrix $A_0 \in \mathbb{R}^{m \times n}$ and a target rank k such that $\sigma_k(A_0) > \sigma_{k+1}(A_0)$, the DLRA problem consists of determining $t \mapsto Y(t) \in \mathcal{M}_k$ solving the following initial value problem*

$$\begin{cases} \dot{Y} = \Pi(Y)F(Y), \quad t \in [0, T], \\ Y(0) = Y_0, \end{cases} \tag{6.2}$$

where $Y_0 = \Pi_{\mathcal{M}_k}(A_0) \in \mathcal{M}_k$, the rank-$k$ truncated singular value decomposition of A_0.

The origins of the above problem are rooted in the Dirac-Frenkel variational principle, by which the dynamics of (6.1) are optimally projected onto the tangent space of the manifold:

$$\|F(Y) - \Pi(Y)F(Y)\|_F = \min_{v \in T_Y \mathcal{M}_k} \|F(Y) - V\|_F. \tag{6.3}$$

This optimality criterion together with the optimal choice of Y_0 are the local first-order approximation of the computationally demanding optimality $Y_{\text{opt}}(t) = \Pi_{\mathcal{M}_k}(A(t))$. The gap (6.3) between the original dynamics and the projected dynamics of Problem 6.1 is known as the *modeling error* [KV19]. Given appropriate smoothness requirements on F and assuming the modeling error can be uniformly bounded in a neighborhood $\mathcal{U}$ of the exact solution of (6.2) as

$$\max_{Y \in \mathcal{U} \cap \mathcal{M}_k} \|F(Y) - \Pi(Y)F(Y)\|_F \leq \varepsilon,$$

then it can be shown, see e.g. [KV19, Theorem 2], there exists a constant depending on the final time T such that

$$\|A(T) - Y(T)\|_F \leq C(T)(\delta_0 + \varepsilon).$$

In recent years, several computationally efficient numerical integration schemes to approximate the solution to Problem 6.1 have been proposed [LO14, CL22, KV19]. Their output is a time discretization of the solution, where $Y_i \in \mathcal{M}_k$ approximates $Y(i\Delta t)$, for every $i = 0, \ldots, N$, assuming a fixed step size $\Delta t = T/N$ is used. Existing error analysis results [KLW16, KV19] state that the error at final time can be bound as

$$\|Y_N - A(T)\|_F \leq \tilde{C}(T)(\delta_0 + \varepsilon + \Delta t^q)$$

for some integer $q \geq 1$ and a constant $\tilde{C}(T) > 0$ associated with each integration scheme that depends on the problem at hand. The constant q is called the *convergence order* of the time stepping method.

The most direct strategy to numerically integrate Problem 6.1 in the factored representation of Y is to derive individual evolution equations for the factors [KL07]. However, the computational gain of such approach is undermined by the high stiffness of the re-

sulting equations when the k-th singular value of the approximation becomes small. In fact, this forces unreasonable step size restrictions to guarantee stability of the integrator. Projector-splitting schemes proposed by Lubich and Oseledets [LO14] were the first remedy to this issue. Since then, a collection of methods have been designed to be stable in presence of small singular values [KLW16].

6.2 Fixed-rank retractions and DLRA numerical integrators

In this section, we introduce a selection of well-known DLRA integration techniques placing the emphasis on their connection with a particular retraction on $\mathcal{M}_k$. Indeed, if R denotes a retraction on the rank-k matrix manifold then

$$Y_{i+1} = R_{Y_i}\left(\Delta t \Pi(Y_i)F(Y_i)\right) \tag{6.4}$$

is a natural extension of the forward-Euler method [Shu86]. Several DLRA integration schemes are realized in this way for different choices of retractions.

Metric-projection retraction. The use of metric projection to numerically integrate more general manifold-constrained ODEs dates back to so-called *projection methods*, as presented for instance in the classic reference [HLW10]. Each integration step is carried out in the embedding space and is followed by the closest-point projection onto the constraint manifold.

For the fixed-rank matrix manifold embedded into $\mathbb{R}^{m \times n}$, the metric projection coincides with the rank-k truncated SVD, see Section 2.4.4. Hence, for the DLRA differential equation of Problem 6.1, one step of the *projected forward Euler* method (PFE) can be written as

$$Y_{i+1}^{\mathrm{PFE}} = R_{Y_i}^{\mathrm{SVD}}\left(\Delta t \Pi(Y_i)F(Y_i)\right).$$

Under smoothness assumption for F, this method is of order $q = 1$ [KV19, Theorem 4]. For higher accuracy, in principle, one could replace the forward Euler update given as argument to the retraction with the update of any high-order time-stepping technique such as explicit Runge-Kutta methods. This strategy is pursued in [KV19] leading to the so-called *projected Runge-Kutta* (PRK) methods. Note that PRK with $s = 1$ coincides with the PFE method above. The general form of PRK methods with $s \geq 1$ stages is

$$Y_{i+1}^{\mathrm{PRK}} = R_{Y_i}^{\mathrm{SVD}}\left(\Delta t \sum_{j=1}^{s} b_j \kappa_j\right),$$

for some weights $b_j \in \mathbb{R}$ and tangent vectors κ_j, $j = 1, \ldots, s$ obtained by suitably projecting the original intermediate updates of the Runge-Kutta method. Using the SVD retraction is particularly important because tangent vectors κ_j may not belong to the tangent space at Y_i. Yet, the definition of the metric projection retraction (2.11) remains well-posed for any $Z \in \mathbb{R}^{m \times n}$ sufficiently small. For this reason, the SVD retraction is

called an *extended retraction*, see [AO15, §2.3]. Any such extended retraction can be used for PRK methods with more than one stage.

Projector-splitting KSL retraction. The expression for the projection onto the tangent space appearing in the differential equation of Problem 6.1 is the sum of several terms, see (1.17). Projector-splitting techniques leverage this decomposition to build numerical integration schemes in which each term is sequentially integrated. This leads to the KSL projector splitting scheme introduced in [LO14]. The scheme is proved to be first-order accurate independently of the presence of small singular value. As shown in [AO15, Theorem 3.3], the evolution after one time step using the KSL scheme actually defines a second-order retraction for the fixed-rank manifold that coincides with orthographic retraction up to high-order terms. We denote R^{KSL} and its computation summarized in Algorithm 6.1. The KSL integration scheme can then simply be written as

$$Y_{i+1}^{\mathrm{KSL}} = R_{Y_i}^{\mathrm{KSL}} \left(\Delta t \Pi(Y_i) F(Y_i)\right).$$

Algorithm 6.1 KSL retraction

Input: $X = U_0 \Sigma_0 V_0^\top \in \mathcal{M}_k$, $Z = (M, U_p, V_p) \in T_X \mathcal{M}_k$
  1: (K-step) $U_1 \hat{\Sigma}_1 = U_0(\Sigma_0 + M) + U_p$ with U_1 orthonormal;
  2: (S-step) $\tilde{\Sigma}_0 = \hat{\Sigma}_1 - (U_1^\top U_p + (U_1^\top U_0)M)$;
  3: (L-step) $V_1 \Sigma_1^\top = V_0 \tilde{\Sigma}_0^\top + Z^\top U_1$ with V_1 orthonormal;
  4: Optional: $[\tilde{U}_1, \tilde{\Sigma}_1, \tilde{V}_1] = \mathrm{SVD}(\Sigma_1)$;
  5: Optional: $U_1 \leftarrow U_1 \tilde{U}_1$, $\Sigma_1 \leftarrow \tilde{\Sigma}_1$, $V_1 \leftarrow V_1 \tilde{V}_1$;
  6: **return** : $U_1 \Sigma_1 V_1^\top := R_X^{\mathrm{KSL}}(V)$;

Modified projector-splitting KLS retraction. Recently, a modification to KSL projector-splitting method was proposed [CL22] to improve its performance, while maintaining the stability and accuracy properties of KSL. The scheme goes by the name of KLS integrator and it is a modification of the KSL scheme where the L-step is performed before the S-step. This comes with the computational advantage of being able to perform the K-step and L-step in parallel without compromising first-order accuracy and stability with respect to small singular values. As we prove in the next section, one step of the KLS scheme also defines a retraction for the fixed-rank manifold. We denote it R^{KLS} and its computation is detailed in Algorithm 6.2. Then, as for other scheme so far, the KLS integration scheme for DLRA takes the simple form

$$Y_{i+1}^{\mathrm{KLS}} = R_{Y_i}^{\mathrm{KLS}} \left(\Delta t \Pi(Y_i) F(Y_i)\right).$$

6.2.1 The KLS retraction

A careful inspection of the formulas defining the orthographic retraction for the fixed-rank matrix manifold reveals a structure very similar to one step of the unconventional

integrator of [CL22], also known as the KLS scheme which we report in Algorithm 6.2. In order to facilitate the identification, we rewrote in Algorithm 6.3 the pseudo-code to compute the orthographic with a notation closer to the one used for the KLS scheme and where the term Σ_1 is computed without explicitly forming the factor S_U and S_V (see Algorithm 2.8 for the original notation). This immediately highlights that the KLS scheme and the orthographic retraction differ only in the computation of Σ_1 at line 5. Denoting Σ_1^{KLS} and Σ_1^{ORTH} the quantities computed respectively by the each algorithm, we have

$$\Sigma_1^{ORTH} - \Sigma_1^{KLS} = U_1^\top U_p (\Sigma_0 + M)^{-1} V_p^\top V_1.$$

Algorithm 6.2 KLS retraction

Input: $X = U_0 \Sigma_0 V_0^\top \in \mathcal{M}_k$, $Z = (M, U_p, V_p) \in T_X \mathcal{M}_k$

1: (K-step) $U_1 S_U = U_0 (\Sigma_0 + M) + U_p$ with U_1 orthonormal;
2: (L-step) $V_1 S_V = V_0 (\Sigma_0 + M^\top) + V_p$ with V_1 orthonormal;
3: $L = U_1^\top U_0$;
4: $R = V_1^\top V_0$;
5: (S-step) $\Sigma_1 = L \left[(\Sigma_0 + M) R^\top + V_p^\top V_1 \right] + U_1^\top U_p R^\top$; $\triangleright$ equivalent to $U_1^\top (X + Z) V_1$
6: Optional: $[\tilde{U}_1, \tilde{\Sigma}_1, \tilde{V}_1] = \mathrm{SVD}(\Sigma_1)$;
7: Optional: $U_1 \leftarrow U_1 \tilde{U}_1$, $\Sigma_1 \leftarrow \tilde{\Sigma}_1$, $V_1 \leftarrow V_1 \tilde{V}_1$;
8: **return** : $U_1 \Sigma_1 V_1^\top := R_X^{\mathrm{KLS}}(Z)$;

Algorithm 6.3 Orthographic retraction with alternative computation of Σ_1

Input: $X = U_0 \Sigma_0 V_0^\top \in \mathcal{M}_k$, $Z = (M, U_p, V_p) \in T_X \mathcal{M}_k$

1: $U_1 S_U = U_0 (\Sigma_0 + M) + U_p$ with U_1 orthonormal;
2: $V_1 S_V = V_0 (\Sigma_0 + M^\top) + V_p$ with V_1 orthonormal;
3: $L = U_1^\top U_0$;
4: $R = V_1^\top V_0$;
5: $\Sigma_1 = L \left[(\Sigma_0 + M) R^\top + V_p^\top V_1 \right] + U_1^\top U_p R^\top + U_1^\top U_p (\Sigma_0 + M)^{-1} V_p^\top V_1$,
6: Optional: $[\tilde{U}_1, \tilde{\Sigma}_1, \tilde{V}_1] = \mathrm{SVD}(\Sigma_1)$;
7: Optional: $U_1 \leftarrow U_1 \tilde{U}_1$, $\Sigma_1 \leftarrow \tilde{\Sigma}_1$, $V_1 \leftarrow V_1 \tilde{V}_1$;
8: **return** : $U_1 \Sigma_1 V_1^\top := R_X^{\mathrm{ORTH}}(Z)$;

Using this observation together with the second-order property of the orthographic retraction allows us to show the KLS procedure actually defines a second-order retraction. This link with the orthographic retraction mirrors the same observation made for the closely related KSL scheme [AO15, Theorem 3.3].

Proposition 6.2. *The procedure of Algorithm 6.2 defines a second-order retraction called the KLS retraction.*

Proof. For any $X \in \mathcal{M}_k$, if $Z \in T_X \mathcal{M}_k$ has sufficiently small norm, the orthonormalizations of the first two lines of Algorithm 6.2 are uniquely defined since the matrices to orthonormalize then have full rank. By smoothness of the orthonormalization process,

if Z is sufficiently small the matrices L and R have full rank, and analogously $L\Sigma_0 R^\top$. Hence, Σ_1 has full rank and $U_1 \Sigma_1 V_1^\top = \mathrm{R}_X^{\mathrm{KLS}}(Z)$ is uniquely and smoothly defined for any Z in a neighborhood of the origin of $T_X \mathcal{M}_k$ and belongs to $\mathcal{M}_k$. Consider the curves $t \mapsto R_X^{\mathrm{ORTH}}(tZ)$ and $t \mapsto R_X^{\mathrm{KLS}}(tZ)$, well-defined for sufficiently small t. These curves share the left and right singular vectors $U_1(t)$ and $V_1(t)$, hence their difference is given by

$$R_X^{\mathrm{ORTH}}(tZ) - R_X^{\mathrm{KLS}}(tZ) = t^2 U_1(t)^\top U_p (\Sigma_0 + tM)^{-1} V_p^\top V_1(t) = t^2 C(t),$$

where $C(t) := U_1(t)^\top U_p (\Sigma_0 + tM)^{-1} V_p^\top V_1(t)$. Let us show that $C(t) = o(t)$, i.e. $\lim_{t\to 0} C(t)/t = 0$. Since $U_1(0) = U_0$ and $V_1(0) = V_0$, by definition of tangent space of $\mathcal{M}_k$, we know that $U_1(0)^\top U_p = 0$ and $V_p^\top V_1(0) = 0$. Hence $C(0) = 0$ and therefore

$$\lim_{t\to 0} \frac{C(t)}{t} = \lim_{t\to 0} \frac{C(t) - C(0)}{t}.$$

This coincides with $C'(0)$ since C is smooth for small t. But since

$$C'(0) = U_1'(0)^\top U_p \Sigma_0^{-1} \underbrace{V_p^\top V_1(0)}_{=0} + \underbrace{U_1(0)^\top U_p}_{=0} \frac{\mathrm{d}}{\mathrm{d}t}(\Sigma_0 + tM)^{-1}\big|_{t=0} \underbrace{V_p^\top V_1(0)}_{=0}$$

$$+ \underbrace{U_1(0)^\top U_p}_{=0} \Sigma_0^{-1} V_p^\top V_1'(0) = 0$$

we can infer that $C(t) = o(t)$ and therefore that $R_X^{\mathrm{KLS}}(tZ) = R_X^{\mathrm{ORTH}}(tZ) + o(t^3)$. Introducing the geodesic $\gamma_{X,Z}(t) = \mathrm{Exp}_X(tZ)$, by the second-order property of the orthographic retraction, the second statement of Proposition 2.11 yields $R_X^{\mathrm{ORTH}}(tZ) = \gamma_{X,Z}(t) + O(t^3)$. Combining the two result leads to

$$R_X^{\mathrm{KLS}}(tZ) = R_X^{\mathrm{ORTH}}(tZ) + o(t^3) = \gamma_{X,Z}(t) + O(t^3) + o(t^3) = \gamma_{X,Z}(t) + O(t^3).$$

This implies R_X^{KLS} is a second-order retraction by using in the other direction the second statement of Proposition 2.11. □

6.3 Curve approximation time-stepping paradigm

Before proceeding to the derivation of the new integration schemes, we first briefly describe the rationale behind these methods. In order to achieve a time-stepping scheme on manifolds with high-order approximation power, we seek for high-order integration methods for ordinary differential evolving in an Euclidean spaces where the approximation at each step can be interpreted as the evaluation of a curve which locally approximates the exact solution. Provided the approximating curve can be described in geometric terms, such as interpolatory conditions, retraction curves can be used to construct the manifold counterpart of the scheme.

The accuracy of a given numerical integration techniques can be assessed by estimating how one step of the numerical scheme deviates from the exact solution of the differential

equation. For the ambient initial value problem (6.1), the error introduced by performing one step of size Δt starting from the exact solution $A(t)$ with a certain numerical integration scheme is the so-called *local truncation error*. For a fixed t it is defined as a function of the step size by

$$\tau_{\text{local}}(\Delta t) = \left\| A(t + \Delta t) - \tilde{A}(t + \Delta t) \right\|,$$

where $\tilde{A}(t + \Delta t)$ is the approximation returned by the scheme. For general linear methods such as Runge-Kutta (RK) and linear multistep methods (LMS), if the local truncation error can be bounded uniformly in t by $O(\Delta t^{q+1})$ and the scheme is stable, then performing $N \geq 1$ steps of size T/N results in a maximum error on $[0, T]$ of order $O(\Delta t^q)$ [HNW93, Theorem 8.13], provided sufficient smoothness of the exact solution. Therefore, high-order numerical integration schemes can be obtained by iterating stable procedures with small local truncation error, i.e. that approximate well the curve $\Delta t \to A(t + \Delta t)$. While numerical integration schemes of the LMS family are generally interpreted as approximations of the exact update

$$A(t + \Delta t) - A(t) = \int_t^{t+\Delta t} A'(\tau) \mathrm{d}\tau$$

obtained by interpolating the integrand $A'(\tau)$ and computing the integral exactly [HNW93, §III.1], the schemes proposed below are explicitly based on the interpolation of the exact solution curve. The accelerated forward Euler scheme comes from interpolating the exact solution curve's velocity and acceleration at each time step while the Ralston-Hermite scheme uses a Hermite interpolant to interpolate the velocity in two points.

6.4 The accelerated forward Euler method

The classic forward-Euler numerical integration schemes achieves order $q = 1$ [HNW93, Theorem 7.5]. This can be intuitively understood by observing the scheme interpolates position and velocity of the exact solution curve at initial time. In fact, if the solution to (6.1) is sufficiently smooth we have

$$
\begin{aligned}
A(t + \Delta t) &= A(t) + \Delta t A'(t) + O(\Delta t^2) \\
&= A(t) + \Delta t F(A(t)) + O(\Delta t^2).
\end{aligned}
$$

Therefore, the forward-Euler scheme is defined by the local update

$$A^{FE}(t + \Delta t) = A(t) + \Delta t F(A(t)),$$

which by construction realizes a local truncation error of order $O(\Delta t^2)$. One more step along this line of reasoning, indicates we could reach a local truncation error of order $O(\Delta t^3)$ by including the next term of the Taylor expansion in the integration scheme.

Runge-Kutta methods attain this goal by performing intermediate steps that are linearly combined. Alternatively, we can build a curve branch which interpolates the acceleration of the solution curve in t. We have

$$A(t + \Delta t) = A(t) + \Delta t A'(t) + \frac{\Delta t^2}{2} A''(t) + O(\Delta t^3)$$

$$= A(t) + \Delta t F(A(t)) + \frac{\Delta t^2}{2} DF(A(t))[F(A(t))] + O(\Delta t^3).$$

This motivates the definition of the accelerated forward Euler scheme (AFE) update as

$$A^{AFE}(t + \Delta t) = A(t) + \Delta t F(A(t)) + \frac{\Delta t^2}{2} DF(A(t))[F(A(t))]. \tag{6.5}$$

By construction, this scheme is consistent and admits local truncation error of order $O(\Delta t^3)$. The scheme is zero-stable [HNW93, Definition3.2] as it is a one step method and it is conditionally absolutely stable with the same absolute stability region of the 2-stages explicit Runge-Kutta method known as Heun's method [QSS07, §11.3.3]. Indeed, the stability polynomial of Heun's method and AFE method is in both cases given by $1 + \lambda + \frac{\lambda^2}{2}$, for $\lambda \in \mathbb{C}$. This suggest the AFE exhibits global error convergence order $q = 2$.

We now generalize the AFE scheme for the DLRA differential equation of Problem 6.1. As we later show in Proposition 6.6, the acceleration of the exact solution of (6.2) is the sum of two terms: the tangent component of the ambient acceleration and the acceleration due to the curvature of the manifold. The latter is expressed using the *Weingarten map*, a classic concept of Riemannian geometry that we briefly introduce in the following section.

6.4.1   The Weingarten map

We focus on the Weingarten map for the case of a Riemannian submanifold $\mathcal{M}$, embedded into a finite dimensional Euclidean space $\mathcal{E}$. A treatment for the general case where $\mathcal{M}$ is embedded into another Riemannian manifold $\overline{\mathcal{M}}$, can be found in [Lee18, §8] from which definitions below were taken and adapted to the present situation. We point out that the definition of the Weingarten map requires that both the embedding space and the manifold are endowed with their Levi-Civita connection. Hence, we assume that $\mathcal{E}$ is endowed with the Euclidean connection and $\mathcal{M}$ with the tangential connection induced from it, see Sections 1.3.1 and 1.4.2.

The Weingarten map is introduced jointly with the closely related concept of second fundamental form. It is a measure of the discrepancy between the Riemannian connection of a Riemannian submanifold and its embedding space. Recall that for two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$, the tangential connection is the orthogonal projection of the Euclidean connection. For a given $p \in \mathcal{M}$, this means the vector

$$D_X Y(p) - \Pi(p) D_X Y(p) = D_X Y(p) - \nabla_X Y(p)$$

belongs to the normal space at p. The function $(I-\Pi)\mathrm{D}_X Y$ is smooth on $\mathcal{M}$ and therefore defines a smooth normal vector field on $\mathcal{M}$, the set of which is denoted $\mathfrak{X}_N(\mathcal{M})$.

Definition 6.3. *The second fundamental form* $\mathbb{I} : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}_N(\mathcal{M})$ *is a symmetric bilinear map defined by* $\mathbb{I}(X,Y) = (I - \Pi)\mathrm{D}_X Y$.

Computing the inner-product between this normal field and another normal field defines an auxiliary scalar field from which the definition of the Weingarten map is extracted.

Definition 6.4. *The Weingarten map is the bilinear map*

$$\mathcal{W} : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}_N(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$$
$$(X, N) \mapsto \mathcal{W}(X, N)$$

defined by the multilinear form

$$\langle N, \mathbb{I}(X,Y)\rangle = \langle \mathcal{W}(X,N), Y\rangle, \quad \forall Y \in \mathfrak{X}(\mathcal{M}).$$

When the manifold has codimension 1, i.e. when $\mathcal{M}$ is a hypersurface, the Weingarten map is often called the *shape operator*. Recall that the value of $\nabla_X Y$ at p depends only on the value of the vector field X at p [Lee18, Proposition 4.5]. By its symmetry, the second fundamental form also depends only on values of the vector fields at p [Lee18, Proposition 8.1]. As a consequence, the same holds for the Weingarten map, which can be given a pointwise definition [dC92, p. 128]. Given $x, y \in T_p\mathcal{M}$ and $n \in N_p\mathcal{M}$, the Weingarten map at p is the bilinear map $\mathcal{W}_p : T_p\mathcal{M} \times N_p\mathcal{M} \to T_p\mathcal{M}$ defined by

$$\langle \mathcal{W}_p(x,n), y\rangle_p = \langle \mathcal{W}(X,N), Y\rangle\big|_p$$

for any $X, Y \in \mathfrak{X}(\mathcal{M})$ and $N \in \mathfrak{X}_N(\mathcal{M})$ satisfying $X(p) = x$, $Y(p) = y$ and $N(p) = n$. Thanks to the explicit link between the ambient connection and the connection on $\mathcal{M}$ expressed with the tangent space projection, the Weingarten map at p can be related with the differential of the tangent space projection.

Proposition 6.5 ([AMT13, Theorem 1]). *For any $x \in T_p\mathcal{M}$ and $n \in N_p\mathcal{M}$, the Weingarten map satisfies*

$$\mathcal{W}_p(x,n) = \mathrm{D}\Pi(p)[x]\,n = \Pi(p)\mathrm{D}\Pi(p)[x]\,z = \Pi(p)\mathrm{D}\Pi(p)[x]\,\Pi(p)^{\perp}z$$

for any $z \in T_p\mathcal{E} \simeq \mathcal{E}$ such that $\Pi(p)^{\perp}z = n$.

The differential of the projection is to be intended as $\mathrm{D}\Pi(p)[x]\,n := \frac{\mathrm{d}}{\mathrm{d}t}\left[\Pi \circ \bar{N} \circ \gamma_{p,x}(t)\right]\big|_{t=0}$, for any smooth extensions $\bar{N}$ of any normal field $N \in \mathfrak{X}_N(\mathcal{M})$ satisfying $N(p) = n$ and for any curve such that $\gamma_{p,x}(0) = p$, $\dot{\gamma}_{p,x}(0) = x$.

One practical use of the Weingarten map comes is for the computation of the Riemannian Hessian of a scalar field on an embedded submanifold. As stated in Proposition 2.8, the Riemannian Hessian of a scalar field f at $p \in \mathcal{M}$ along $x \in T_p\mathcal{M}$ can be computed as

the sum of two terms, the second of which may be expressed using the Weingarten map. For any smooth extension $\bar{f}$ of f to $\mathcal{E}$, it holds that [AMT13, Equation (10)]

$$\mathrm{Hess}\, f(p)[x] = \Pi(p)\nabla^2 \bar{f}(p)[v] + \mathcal{W}_p(x, \Pi(p)^\perp \nabla \bar{f}(p)).$$

where $\nabla \bar{f}$ and $\nabla^2 \bar{f}$ are respectively the Euclidean gradient and Euclidean Hessian of $\bar{f}$. From a theoretical viewpoint, the eigenvalues of the Weingarten map at p for a given $n \in N_p\mathcal{M}$ are real and can interpreted as principal curvatures of the manifold [Lee13, p. 238], describing the extrinsic curvature of the embedding of $\mathcal{M}$ in $\mathcal{E}$ along the normal direction n. For the rank-k matrix manifold embedded into $\mathbb{R}^{m \times n}$, depending on the conventions used to represent points and tangent vectors, many equivalent expressions are known for the Weingarten map of the fixed-rank manifold, see for example [AMT13, FL18]. In our conventions, see Section 1.5.3, for any $Y = U\Sigma V^\top \in \mathcal{M}_k$, $T = (M, U_p, V_p) \in T_Y \mathcal{M}_k$ and $N \in N_Y \mathcal{M}_k$ the Weingarten map can be computed as

$$\mathcal{W}_Y(T, N) = \left(M :\ 0_{k \times k},\ U_p :\ NV_p\Sigma^{-1},\ V_p :\ N^\top U_p \Sigma^{-1} \right).$$

This expression nicely highlights two notable features that are know for the embedding of the fixed-rank manifold into $\mathbb{R}^{m \times n}$: it is a ruled surface with unbounded curvature. Indeed, along the subspace associated to the $UMV^\top$ term, $\mathcal{M}_k$ is flat, while the curvature along the other directions grows unbounded as $\sigma_k(Y) \to 0$.

6.4.2 The integration scheme

With the definition of the Weingarten map introduced in Section 6.4.1, we have the correct vocabulary to express the acceleration of the exact DLRA solution curve to Problem 6.1.

Proposition 6.6. *If a smooth curve on $\mathcal{M}_k$ is defined by $\dot{Y} = \Pi(Y)F(Y)$ for some $F \in \mathfrak{X}(\mathcal{E})$, then its intrinsic acceleration can be computed as*

$$\ddot{Y} = \Pi(Y)\mathrm{D}F(Y)[\Pi(Y)F(Y)] + \mathcal{W}_Y(\Pi(Y)F(Y), \Pi(Y)^\perp F(Y)).$$

Proof. By definition of tangential connection,

$$\ddot{Y} = \nabla_{\dot{Y}}\dot{Y} = \Pi(Y)\mathrm{D}_{\dot{Y}}(\Pi(Y)F(Y))$$

Using the product rule, $\Pi(Y)^2 = \Pi(Y)$ and Proposition 6.5, we evince

$$\Pi(Y)\mathrm{D}_{\dot{Y}}(\Pi(Y)F(Y)) = \Pi(Y)\left(\Pi(Y)\mathrm{D}F(Y)[\dot{Y}] + \Pi(Y)\mathrm{D}\Pi(Y)[\dot{Y}]F(Y)\right)$$

$$= \Pi(Y)\mathrm{D}F(Y)[\Pi(Y)F(Y)] + \Pi(Y)\mathrm{D}\Pi(Y)[\Pi(Y)F(Y)]\,\Pi(Y)^\perp F(Y)$$

$$= \Pi(Y)\mathrm{D}F(Y)[\Pi(Y)F(Y)] + \mathcal{W}_Y(\Pi(Y)F(Y), \Pi(Y)^\perp F(Y)).$$

$\square$

To mimic the Euclidean AFE update defined by equation (6.5), we need to construct a smooth manifold curve $Y^{\mathrm{AFE}}(\Delta t)$ such that

$$Y^{\mathrm{AFE}}(0) = Y_0, \quad \dot{Y}^{\mathrm{AFE}}(0) = \dot{Y}(0), \quad \ddot{Y}^{\mathrm{AFE}}(0) = \ddot{Y}(0).$$

Proposition 3.2 shows this is possible given any second-order retractions. Let $R^{\mathbb{II}}$ indicate any second-order retraction, then the manifold analogous of the AFE update (6.5) reads

$$Y^{\mathrm{AFE}}(\Delta t) = R_{Y_0}^{\mathbb{II}}\left(\Delta t \dot{Y}_0 + \Delta t^2 \ddot{Y}_0/2\right).$$

Then the AFE scheme for DLRA takes the form

$$Y_{i+1}^{\mathrm{AFE}} = R_{Y_i}^{\mathbb{II}}\left(\Delta t \Pi(Y_i)F(Y_i) + \frac{\Delta t^2}{2}\ddot{Y}_i\right) \tag{6.6}$$

with

$$\ddot{Y}_i = \Pi(Y_i)\mathrm{D}F(Y_i)[\Pi(Y_i)F(Y_i)] + \mathcal{W}_{Y_i}\left(\Pi(Y_i)F(Y_i), \Pi(Y_i)^{\perp}F(Y_i)\right). \tag{6.7}$$

All the retractions for the fixed-rank manifold presented in Section 2.4.4 as well as the KSL and KLS retractions introduced in Section 6.2 have the second-order property. In principle, all of them are suited to be used in (6.6), however, experiments reported in Section 6.6 suggest the orthographic retraction is the most convenient in terms of speed, accuracy and stability. In virtue of Proposition 3.22, the local truncation error associated with the one step of the AFE integration scheme is of order $O(\Delta t^3)$. A discussion on the stability and the global error convergence of the method is reported in the numerical experiments section. A possible limitation to the applicability of the AFE integration scheme is the need to compute at each step the differential of the forcing term. Unless F contains only linear terms and pointwise nonlinearities, the differential may not be readily available or be too costly to compute.

6.5   The Ralston-Hermite method

The retraction-based Hermite (RH) interpolant developed in Chapter 5 can be used to define a manifold curve with prescribed endpoints and endpoint velocities. In this section, we develop a numerical integration scheme using this manifold curve as a building block.

6.5.1   Extrapolation of the RH interpolant

When data is sampled from a smooth manifold curve, Proposition 5.8 states that the RH interpolant is well-defined between interpolation nodes provided samples are sufficiently close. The following proposition generalizes the well-posedness result in two ways. First, the interpolation data is not sampled from any curve but is just a pair of tangent bundle points $(p_0, v_0), (p_1, v_1) \in T\mathcal{M}$ that we wish to interpolate at instants t_0 and t_1, respectively. Second, we also argue that the curve can be evaluated in an interval larger than $[t_0, t_1]$. Whenever we need to highlight the dependence on the interpolation data, we

shall denote this RH interpolant $H(t) = H(t; (t_0, p_0, v_0), (t_1, p_1, v_1))$.

Proposition 6.7. *Consider tangent bundle data points $(p_0, v_0), (p_1, v_1) \in T\mathcal{M}$ and some parameters $t_0 < t_1$ and let R denote a retraction on $\mathcal{M}$. If p_0 and p_1 are close enough and $t_1 - t_0$ is sufficiently small, there exists a RH interpolant H satisfying $H(t_i) = p_i$ and $\dot{H}(t_i) = v_i$, for $i = 1, 2$, that is well-defined for any $t \in (t_0 - \varepsilon, t_1 + \varepsilon)$, for some $\varepsilon > 0$.*

Proof. By Corollary 5.6, we know the curve

$$H(t) = \alpha \left(\frac{t - t_0}{h}; p_0, hv_0, p_1, hv_1 \right), \quad \text{with } h = t_1 - t_0,$$

satisfies the interpolation conditions provided it is well-defined. This curve corresponds to the generalized de Casteljau curve given by Definition 5.4, with control points $b_0 = p_0$, $b_1 = R_{p_0}(hv_0/3)$, $b_2 = p_1$ and $b_3 = R_{p_1}(-hv_1/3)$. If $d(p_0, p_1) \leq \bar{\rho}(p_0)$ and h is small enough, all control points belong to a retraction-convex set centered at p_0. Hence the curve H is well-defined on $[t_0, t_1]$. The building blocks of H are the r-endpoint retraction curves given in Definition 3.4. Although their definition restricts them on $[0, 1]$, they actually can be evaluated on an open interval containing $[0, 1]$ by openness of the invertibility domain of the retraction. Hence, the curve H can be evaluated on an open interval containing $[t_0, t_1]$. $\square$

6.5.2 The integration scheme

As stated in Theorem 5.9, the retraction-based Hermite interpolant of a smooth manifold curve can achieve $O(\Delta t^4)$ approximation error on the interval $[t_i, t_i + \Delta t]$ as $\Delta t \to 0$. Assuming the interval over which the interpolant is well-defined and satisfies the error bounds extends to the interval $[t_i, t_i + 2\Delta t]$, we may conclude the following update rule

$$Y_{i+2} = H\left(t_{i+2}; \left(t_i, Y_i, \dot{Y}_i\right), \left(t_{i+1}, Y_{i+1}, \dot{Y}_{i+1}\right)\right), \tag{6.8}$$

where H is the Hermite interpolant as introduced in Proposition 6.7, produces a small local truncation error. Let us apply this scheme to the scalar ODE given by $y' = f(y)$ with by taking H as the unique polynomial Hermite interpolant of the data $(t_j, y_j, f_j := f(y_j))$, for $j = i, i + 1$. Then, the recursive relation (6.8) becomes

$$y_{i+2} = y_{i+1} + \Delta t \left(2f_i + 4f_{i+1} - 5 \left(\frac{y_{i+1} - y_i}{\Delta t} \right) \right).$$

As pointed out in [HNW93, §III.3], this scheme indeed has a local truncation error $O(\Delta t^4)$. However, it is not zero-stable and thus does not produce a convergent scheme of order 3.

Nevertheless, stability may be recovered by combing this update rule with a suitable chosen intermediate step. Consider the family of multistep methods parametrized by $\alpha \in (0, 1)$ obtained by concatenating a forward-Euler step of length $\alpha \Delta t$ and a Hermite

interpolation update of the form (6.8) up to t_{i+1}:

$$\begin{cases} y_{i+\alpha} = y_i + \alpha \Delta t f_i, \\ y_{i+1} = H(t_{i+1}; (t_i, y_i, f_i), (t_{i+\alpha}, y_{i+\alpha}, f_{i+\alpha})). \end{cases}$$

This family of schemes are all part of the family of the 2-stages explicit Runge-Kutta methods with Butcher table

$$\begin{array}{c|cc} 0 & & \\ \alpha & \alpha & \\ \hline & \frac{\alpha^2+\alpha-1}{\alpha^2} & \frac{1-\alpha}{\alpha^2} \end{array}$$

Since the scheme is explicit and has two stages we can aim for a scheme of order two by choosing α in order to satisfy the first two order conditions of the Runge-Kutta methods [HNW93, Theorem 1.6]. For any $\alpha \in (0,1)$, the scheme satisfies the first-order condition. Choosing α to satisfy also the second-order conditions narrows down the family to the scheme with $\alpha = 2/3$. This scheme is an explicit second-order RK method known as the Ralston scheme [Ral62] and is defined by the following Butcher table.

$$\begin{array}{c|cc} 0 & & \\ \frac{2}{3} & \frac{2}{3} & \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array}$$

This scheme can be easily generalized to a scheme for manifold ODE that uses only retractions thanks to Proposition 3.2 and Proposition 6.7. Let R denote any retraction and R^I denote any retraction whose local inverse can be computed efficiently. To indicate which retraction is used to construct the retraction-based Hermite interpolant H, we add it to its list of arguments. The Ralston-Hermite (RH) scheme for Problem 6.1 read as

$$\begin{cases} Y_{i+2/3} = R_{Y_i}\left(\frac{2}{3}\Pi(Y_i)F(Y_i)\right), \\ Y_{i+1}^{\mathrm{RH}} = H(t_{i+1}; (t_i, Y_i, \Pi(Y_i)F(Y_i)), \left(t_i + \frac{2}{3}\Delta t, Y_{i+2/3}, \Pi(Y_{i+2/3})F(Y_{i+2/3})\right), R^I) \end{cases}$$

A suitable candidate for both retractions R and R^I is the orthographic retraction defined in Section 2.2.2. As experiments in Section 6.6 suggest, this generalization to the manifold setting of the Ralston scheme maintains its second-order accuracy.

6.5.3 The ARH integration scheme

Given the two-steps structure of the RH scheme, we propose to investigate a third scheme which combines the AFE and the RH schemes. Replacing the intermediate forward-Euler step of the RH scheme with a AFE update defines what we will call the accelerated

Table 6.1: Average time in milliseconds per step for the experiments of Figure 6.2-(d).

PRK1	KSL	KLS	AFE	RH	ARH	PRK2	PRK3
5.27	5.41	5.88	7.71	12.62	17.38	10.41	14.16

Ralston-Hermite scheme (ARH). It is defined by the recursive relation

$$\begin{cases} Y_{i+2/3} = R_{Y_i}^{\mathrm{I\!I}}\left(\frac{2}{3}\Delta t \Pi(Y_i)F(Y_i) + \frac{4\Delta t^2}{18}\ddot{Y}_i \right), \\ Y_{i+1}^{\mathrm{RH}} = H(t_{i+1}; (t_i, Y_i, \Pi(Y_i)F(Y_i)), \left(t_i + \frac{2}{3}\Delta t, Y_{i+2/3}, \Pi(Y_{i+2/3})F(Y_{i+2/3}) \right), R^I), \end{cases}$$

where $\ddot{Y}_i$ is given by (6.7).

6.6 Numerical experiments

The following sections are dedicated to illustrating the performances of the accelerated forward Euler (AFE) method, the Ralston-Hermite (RH) method and the accelerated Ralston-Hermite (ARH) method. Experiments were executed with Matlab 2022b on a laptop computer with Intel i7 CPU (1.8GHz with single-thread mode) with 8GB of RAM, 1MB of L2 cache and 8MB of L3 cache. The implementation uses the differential geometry tools of the Manopt library [BMAS14]. The orthographic retraction is chosen for AFE, RH and ARH. An implementation of the KSL and KLS retractions as described by Algorithms 6.1 and 6.2 were added to the fixed-rank manifold factory. For the implementation of the projected Runge-Kutta method of [KV19], we also added an implementation of the truncated SVD extended retraction, accepting as inputs a list of s tangent vectors of other tangent spaces.

6.6.1 Differential Lyapunov equation

The modeling error (6.3) introduced by DLRA is associated with the normal component of the vector field of the original differential equation. The effect of the modeling error magnitude on the performance of DLRA integrators can be assessed by considering a class of matrix differential equations, going by the name of differential Lyapunov equations [UV20, §6.1], which take the form

$$\begin{cases} A' = LA + AL^\top + Q, & t \in [0, T], \\ A(0) = A_0, \end{cases} \tag{6.9}$$

for some $A_0, L, Q \in \mathbb{R}^{n \times n}$. If A_0 has rank exactly k and the matrix Q is zero, then $A(t)$ is also of rank k for every $t \in [0, T]$ [HM94, Lemma 1.22]. Therefore, the norm of Q is proportional to the modeling error.

In the following experiment, we take L as the discretization of a Laplacian operator in two dimensions, i.e. L is the tridiagonal matrix with -2 on the main diagonal and 1 on

(a) $\|Q\|_F = 0$.

(b) $\|Q\|_F = 0.01$.

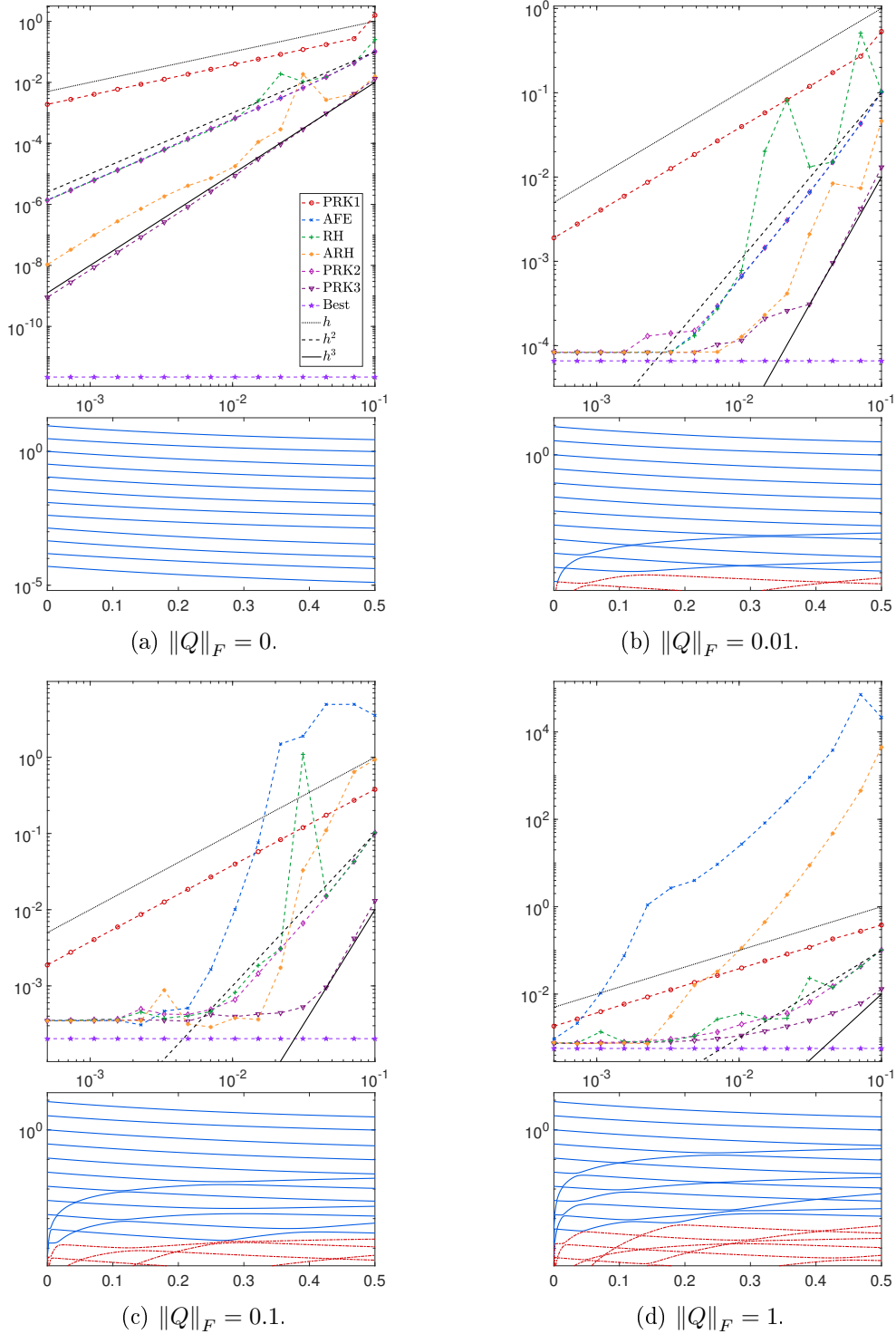(c) $\|Q\|_F = 0.1$.

(d) $\|Q\|_F = 1$.

Figure 6.2: Convergence of the error at final time for different DLRA integration schemes applied to the Lyapunov equation (6.9) with sources terms of different norms. The top plot in each panel is the final error $\|Y_{\Delta t}(T) - A(T)\|_2$ versus the step size Δt, where $Y_{\Delta t}$ is the approximation of A obtained with a step size Δt. The bottom plot reports the evolution of the singular values of the reference solution over time. The red dashed curves correspond to discarded singular values.

the first off-diagonals. For the source term, we take $Q = \eta\tilde{Q}/\|\tilde{Q}\|_F$ for some $\eta > 0$, where $\tilde{Q}$ is a full rank matrix generated from its singular value decomposition with randomly chosen singular vectors and prescribed singular values, decaying as $\sigma_i(\tilde{Q}) = 10^{2-i}$, for $i = 1, \ldots, n$. The initial condition is taken to be of rank exactly k and is also assembled from a randomly generated singular value decomposition with a prescribed geometric decay of singular values: $\sigma_i(A_0) = 3^{2-i}$, for all $i = 1, \ldots, k$, and $\sigma_i(A_0) = 0$, for all $i = k+1, \ldots, n$.

In Figure 6.2, we report the results with $n = 100$ and $k = 12$ of the following experiment. For different values of η, we numerically integrate the rank-k DLRA differential equation (6.2) applied to (6.9) with different numerical schemes and different times steps up to $T = 0.5$. A reference solution to the ambient equation (6.9) is found using the MATLAB routine `ode45` between each time step, for a time step that is the smallest among those considered for the numerical integrators. We then plot as a function of the step size the 2-norm discrepancy between the reference solution at final time and its approximation obtained by numerical integration. The schemes presented in this chapter are compared with the projected Runge-Kutta method (PRK) of [KV19] of order 1,2 and 3. The numerical results for the KSL and the KLS scheme were very similar to the ones of PRK1. Hence, they were omitted not to overcrowd the plots.

The panels of Figure 6.2 correspond to the cases (a) $\eta = 0$, (b) $\eta = 0.01$, (c) $\eta = 0.1$, (d) $\eta = 1.0$. When the source term is zero, the reference solution is also of rank exactly k, as can be seen from the value of the best approximation error in panel (a). In this regime, the AFE and the RH scheme fulfill the promise to exhibit $O(\Delta t^2)$ error convergence, while the ARH scheme seem to reach an asymptotic $O(\Delta t^3)$ trend. The trade-off between accuracy and computational effort that can be seen in Figure 6.3-(a) shows that in this simple setting, the RH, AFE and ARH schemes have comparable performances to PRK2. Turning on the source term determines a non-negligible best approximation error due to the growth of singular values that were initially zero, as can be seen in the bottom plots of panels (b), (c) and (d). The larger the source term's norm, the faster and the greater these singular values grow. Then, the numerical integrators converge to the exact solution of the projected system and so the error with respect the ambient solutions stagnates at a value slightly higher than the best 2-norm approximation. While the RH scheme preserve the $O(\Delta t^2)$ trend up to some oscillations as η increases, the AFE and ARH schemes seem to suffer instability when the normal component of the vector field is too large. A satisfactory explanation for this behavior remained elusive. In this more realistic scenario where the normal component of the vector field is non-negligible, only the RH scheme remains comparable to PRK2 in terms of the trade-off between accuracy and effort, see Figure 6.3-(b) and Table 6.1

6.6.2 Robustness to small singular values

A fundamental prerequisite for competitive DLRA integrators is to be resilient to the presence of small singular values in the solution. A detailed discussion on the topic

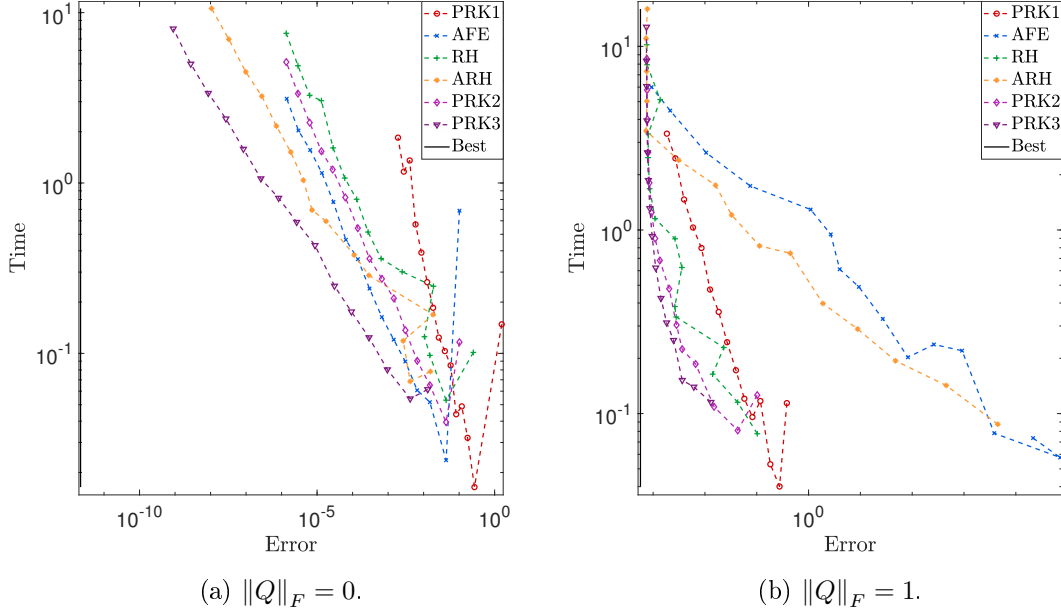(a) $\|Q\|_F = 0$.   (b) $\|Q\|_F = 1$.

Figure 6.3: Computational effort in terms of wall clock time against the error with respect to the reference solution achieve by different numerical integration schemes and with different step sizes. The results were collected from the same experiment of panels (a) and (d) of Figure 6.2.

can be found in [KLW16]. In applications, very often the ambient solution admits an exponential decay of singular values. Hence, a good low-rank approximation is possible but the occurrence of small singular values is inevitable for DLRA to be accurate: a rank-k approximation of the solution must match the k-th singular value of the ambient solution, which is small if the approximation error is small.

The smaller the singular values of the solution, the greater the stiffness of the DLRA differential equation (6.2): the Lipschitz constant of the vector field F gets multiplied by the Lipschitz constant of the tangent space projection, which is inversely proportional to the smallest non-zero singular value [KL07, Lemma 4.2] of the base point. Therefore, standard numerical integration methods fail to provide a good approximation unless the step size is taken to be very small. Projector-splitting integrators for DLRA do not require such step size restrictions and the results on the convergence of the error are independent of the smallest non-zero singular value of the approximation. These scheme are commonly qualified as robust to small singular values. The robustness property was shown for the KSL scheme [KLW16, Theorem 2.1] and the KLS scheme [CL22, Theorem 4]. The PRK method also enjoy the robustness property [KV19, Theorem 6]. In the following we experimentally study the robustness of the AFE, RH and ARH integration schemes to the presence of small singular values.

The typical setting to assess the stability to small singular values of a given integration

scheme, considers a matrix curve $t \in [0, T] \to A(t) \in \mathbb{R}^{n \times n}$ of the form

$$A(t) = U(t)\Sigma(t)V(t)^\top \tag{6.10}$$

with

$$U(t) := \exp\left(t\Omega_U\right), \quad \Sigma(t) := \exp(t)D, \quad V(t) := \exp(t\Omega_V),$$

for some $n \times n$ skew-symmetric matrices Ω_U, Ω_V and a diagonal matrix $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$, for a positive and geometrically decaying sequence σ_i. A rank k approximation of this curve is reconstructed by numerically integrating with the given scheme the DLRA equation (6.2) where the scalar field F is replaced by the exact derivative of the ambient curve (6.10):

$$A'(t) = U(t)\left(\Omega_U\Sigma(t) + \Sigma(t) + \Sigma(t)\Omega_V^\top\right)V(t)^\top. \tag{6.11}$$

The approximation error at final time is constituted mainly of the integration error which can be reduced by decreasing the step size, and the modeling error affected only by the choice of k. A scheme is said to be robust to small singular values, if the integration error is independent of the choice of k. In practice, one must observe that the trend of the error as a function of the step size is unaffected by the choice of k for step sizes where modeling error is negligible compared to the integration error.

Figure 6.4 presents the results for the experiment described in the previous paragraph on a curve (6.10) with randomly generated Ω_U and Ω_V, initial singular values $\sigma_i = 2^{-i}$ and $n = 100$. The panels from left to right corresponds respectively to the AFE, the RH and the ARH schemes. Note that for the AFE and the ARH schemes, we use the exact expression for the second derivative of (6.10) given by

$$\begin{aligned}
A''(t) = U(t)\Big(&\Omega_U^2\Sigma(t) + \Sigma(t) + \Sigma(t)(\Omega_V^2)^\top \\
&+ 2\Omega_U\Sigma(t) + 2\Omega_U\Sigma(t)\Omega_V^\top + 2\Sigma(t)\Omega_V^\top\Big)V(t)^\top.
\end{aligned}$$

The results for AFE show the ideal outcome: the error curves for increasing values of k are superimposed until the modeling error plateau determined by the value of k is reached. These results empirically suggest that the AFE integration scheme is robust to small singular values. On the other hand, the RH and ARH scheme which rely on retraction-based Hermite interpolation suffer from small singular values. Panels (b) and (c) of Figure 6.4 exhibit the same oscillatory convergence trend that could be observed for both schemes in the experiments on the differential Lyapunov equation in Section 6.6.1. A partial explanation for this behavior comes from studying robustness of the retraction-based Hermite interpolant defined in Proposition 6.7 to the presence of small singular values at the interpolation points. Consider the following experiment. Take $Y_0 \in \mathcal{M}_r \in \mathbb{R}^{m \times n}$ with $m = n = 100$, $r = 12$ defined by

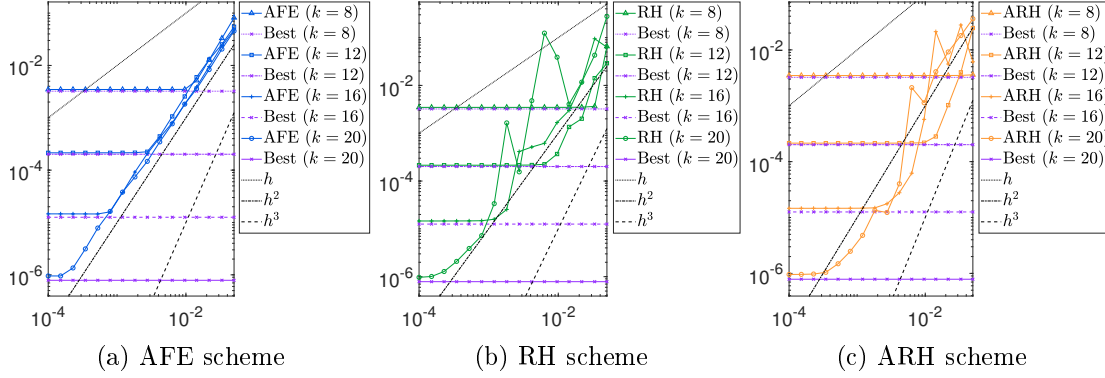$$Y_0(\sigma_r) = U_0\mathrm{diag}(1, \ldots, \sigma_r)V_0^\top$$

(a) AFE scheme	(b) RH scheme	(c) ARH scheme

Figure 6.4: Convergence of the error at final time $\|Y_{\Delta t}(T) - A(T)\|_2$ versus the step size Δt of DLRA integration applied to (6.11) to reconstruct the curve (6.1) for different values of rank.

for some randomly generated orthogonal matrices U_0 and V_0 and with σ_i logarithmically spaced on the interval $[\sigma_r, 1]$, for some $\sigma_r \leq 1$. The second interpolation point is found by first moving away from Y_0 along the orthographic retractions along a random tangent vector $Z \in T_Y \mathcal{M}_r$ such that $\|Z\|_F = 1$ to get $\tilde{Y}_1 = R_{Y_0(\sigma_r)}(Z)$. The second interpolation point Y_1 is obtained from $\tilde{Y}_1$ by replacing its singular values with

$$\sigma_i(Y_1) = \sigma_i(Y_0)(1 + \xi_i),$$

for some random ξ_i drawn from a uniform distribution on $[1/2, 2]$. This way, the singular values decay of both Y_0 and Y_1 mimic a situation encountered in one step of the RH and ARH integration schemes, when the smallest singular value of the current approximation is of the order of σ_r. Then, we randomly generate $Z_0 \in T_{Y_0}\mathcal{M}_r$ and $Z_1 \in T_{Y_1}\mathcal{M}_r$ with $\|Z_0\| = \|Z_1\| = 1$ and form the retraction based interpolant

$$H(\tau) = H(\tau; (0, Y_0, Z_0), (1, Y_1, Z_1)), \quad \tau \in [0, 1]. \tag{6.12}$$

For different values of the smallest singular value σ_r, we measure the discrepancies $\|Z_0 - \dot{H}(0)\|_F$ and $\|Z_0 - \dot{H}(0)\|_F$, where derivatives of H are obtained by the finite differences formula (5.17). The experiment is repeated for each σ_r on 100 randomly generated instances and the error distribution is plotted against σ_r in Figure 6.5. These results unequivocally indicate the fragility of retraction-based Hermite interpolant on the fixed-rank manifold when small singular values are present in the interpolation points. As σ_r decrease, the velocity error in $\tau = 0$ increases, and even more severely in $\tau = 1$. The fact that the error is non-negligible even for moderately small values of σ_r suggests the RH and ARH integration schemes may occasionally employ very badly behaved retraction-based Hermite interpolants. This may contribute to the oscillatory behavior of the error observed RH and ARH in Figures 6.2 and 6.4.
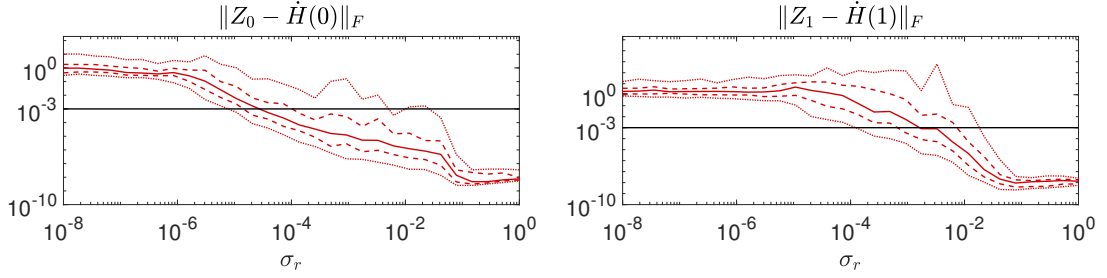
Figure 6.5: Robustness to small singular values of the retraction based Hermite interpolant (6.12). The solid line is the median of the error over 100 randomly generated instance for each value of σ_r while the dashed and dotted lines corresponds to the percentiles $[0.05, 0.25, 0.75, 0.95]$ of the sampled error.

6.7 Conclusions

This chapter contributes in strengthening the connection between retractions and numerical integration methods for manifold ODEs and in particular DLRA techniques. We derive two numerical integration schemes expressed in terms of retractions and showcase their performance on classic problem instances of DLRA. The derivation and the numerical results suggest that the methods can achieve second-order error convergence with respect to the time integration step. However, at the current stage, the two methods have shown mixed results and do not seem to offer computational advantages for the considered problems. While the AFE scheme exhibits instability in the presence of large normal components of the ambient vector field, the RH scheme appears more resilient to this aspect. On the other hand, the occurrence of small singular values in the approximation had no apparent effect on the performance of AFE. Concerning the RH method, small singular values may explain occasional deviations from the conjectured second-order convergence behavior. Further analysis would be required to better understand the drawbacks of both methods and may indicate a remedy or provide precise conditions under which the proposed methods can be reliable and competitive with respect to existing methods.

Another contribution to the connection between DLRA and retractions is the interpretation of the KLS integration scheme as a second-order retraction which approximates up to high-order terms the orthographic retraction. It remains an open question whether the same observation can be made for the recently proposed parallelized version of KLS [CKL23].

For other low-rank tensor formats, such as the Tucker or the tensor-train formats, retractions have also been proposed [KSV14, Ste16]. However, to the best of our knowledge no retraction with an efficiently computable inverse retraction is known and the orthographic retraction has remained elusive due to the complexity of the normal space structure for these manifolds. Yet, the KLS scheme has been extended to low-rank tensor manifolds [CL22, §5]. Hence, assuming the connection with the orthographic retraction carries over to the tensor setting, it may be possible to retrieve the orthographic retrac-

tion for those tensor manifolds as a small perturbation of the KLS update. Then, the possibility to easily compute the inverse orthographic retraction would enable using the retraction-based endpoint curves and Hermite interpolant presented in Chapters 3 and 5 for low-rank tensor manifolds.

Conclusions

The present thesis supports in several ways the claim that retractions as popularized by Riemannian optimization are a convenient tool for building and analyzing numerical methods in the broader context of problems involving a manifold constraint. Indeed, retractions can be effectively interpreted as a general purpose manifold curve generating device. This viewpoint allowed us to propose retraction-based numerical methods for the generalization to a manifold setting of the following problems: homotopy continuation, Hermite interpolation and the integration of ordinary differential equations. Let us briefly summarize the main contributions of the thesis while highlighting some open questions and possible directions of further inquiry.

Riemannian continuation, improvements and extensions

In Chapter 4, we have presented an extension of path-following numerical continuation to track solutions to parameter dependent Riemannian optimization problems. A retraction features in the prediction step of the proposed iterative predictor-corrector procedure, as well as in the correction step, where a locally superlinearly convergent Riemannian optimization method is used. The retraction is also central in the proposed adaptive step size variant of the algorithm. With the construction of suitable homotopies, fixed-step size continuation has demonstrated, up to manual tuning of the number of continuation steps, to allow for the reduction of total computation time in comparison with direct optimization.

The possibility to automatically balance the number of continuation steps and the cost of each correction would require more investigations. In fact, the currently proposed step size adaptation strategy, which aims at eliminating stagnation in the initial iterations of the correction phase, produces too stringent conditions on the step size preventing to strike the optimal balance. Also, the procedure to estimate the step size at each iteration comes at a relatively high cost. Alternative approaches may include generalizing to a manifold setting the other proposed step size selection strategies discussed in [AG90, §6], of which we have considered only the first. More heuristic approaches such as a backtracking line search strategy to limit the actual gradient norm at the initial condition of the subsequent correction step may also prove competitive, given the relatively low

operative cost of computing the gradient norm compared to computing the Hessian.

The effectiveness of embedding a given Riemannian optimization problem into a parameter dependent family of Riemannian optimization to be solved with the proposed continuation algorithm was empirically observed to greatly depend on the choice of parametric family. A more systematic study of the properties associated with particular strategies to build homotopies appears difficult in a general setting, but may be possible for specific Riemannian optimization problems of particular interest, for instance low-rank matrix completion. The intuition which sometimes guides the construction of initial conditions to be used in direct optimization routines may be applied to build homotopies that prove very effective in combination with our continuation algorithm.

As a generalization of the Riemannian optimization problems with a single parameter dependence that were considered in this chapter, one could tackle the case where the problem depends on multiple scalar parameters. The problem has been addressed for parametrized nonlinear equation on an Euclidean space [AG90, §15], leading to algorithms such as the moving frames algorithm which can be thought as a multidimensional generalization of the predictor-corrector approach we have considered. A local model of the solution manifold is constructed at a collection of sample points in parameter space and used to produce an estimate of the solution at new parameter values. Multiparametric interpolation on manifolds as recently developed in [ZB22] may be helpful in building such local models in the generalization of multiparametric continuation methods to a manifold setting.

Retraction-based Hermite interpolation, further theoretical investigations

In Chapter 5 we have proposed a generalization of the de Casteljau algorithm using retractions that can be used to solve the Hermite interpolation problem for a manifold curve. The method has a broader applicability than previously existing method in that it does not require the Riemannian exponential and logarithmic maps, the computation of which may be too expensive for their practical use. A retraction whose local inverse is conveniently computable is nevertheless required, but this requirement can be met for a large class of manifolds of practical interest, such as the fixed-rank matrix manifold.

The approximating power of the retraction-based Hermite interpolant is proved in Theorem 5.9, under the assumption that the derivatives of order two, three and four of the interpolant are bounded, see (5.8). While the proposed interpolant constructed with the particular choice (5.4) was experimentally demonstrated to fulfill this condition in Section 5.5.1, the a priori conditions which guarantee this have remained elusive. This is in part due to the layered complexity of the algorithm, already apparent in the proof of Lemma 5.11 proving the Lipschitz continuity of the interpolant. Studying the algorithm for specific manifolds, such as the sphere, for which the explicit expression of the retrac-

tion can be easily manipulated, may be helpful in formulating a general explanation.

The novel concept of retraction-convexity presented in Section 3.3 has proved to be a valuable tool of analysis. For instance, the existence of retraction-convex sets granted the well-posedness of the proposed retraction-based Hermite interpolant. That being said, a crucial assumption required to use retraction-convexity in this case is Assumption 3.11, which ensures that the retraction-convexity radius function introduced in Section 3.3.2 is continuous. In turn, this is used to provide a strictly positive lower bound to the retraction-convexity radius on the image of the curve, or more generally on a compact set of the manifold. The smoothness of the retraction seems sufficient to intuitively confirm the validity of this assumption, yet providing a rigorous argument would require more effort.

On the specific use of the proposed retraction-based Hermite interpolant on the manifold of fixed-rank matrices, we stress the concern raised in the experiments at the end of Section 6.6.2 about the robustness to small singular values of the interpolant. The capability of the interpolant to match a prescribed velocity at a given point seems to decline as the smallest singular value of the interpolated point decreases. On the one hand, this is not surprising given the numerical ill-conditioning that small singular values usually produce when dealing with fixed-rank manifolds. Matrices whose last non-zero singular value is small are located in regions of the fixed-rank manifold where the curvature is high. On the other hand, as done for the case of robust DLRA integrators, it may be possible to cure the instability by a careful reformulation of the inner steps of the algorithm.

Retraction-based time-stepping algorithms, analysis and extension to tensors

In Chapter 6, we have introduced two novel time-stepping schemes expressed in terms of retractions for the numerical integration of manifold-constrained differential equations. The construction of the methods has been guided by the practical and theoretical tools developed in Chapter 3 and Chapter 5 and thereby argued to achieve second-order accuracy. The methods have been experimentally demonstrated to attain the conjectured error convergence rate on simple instances of dynamical low-rank approximation and to exhibit comparable performances in terms of accuracy-effort ratio with existing methods in the simpler settings. However, the performance of the methods on more realistic yet still academic problem instances of DLRA are not entirely satisfactory. The presence of a large normal component in the ambient scalar field impaired the stability of the accelerated Forward Euler method, while the occurrence of small singular values of the ambient solution diminished the reliability of the Ralston-Hermite method. These drawback are potentially embedded in the methods themselves and a more in depth analysis could shed light on this aspect.

A close connection between the KLS unconventional integrator [CL22] and the orthographic retraction was established in Section 6.2.1. On the one hand, the geometric nature of the orthographic retraction provides a geometric interpretation for the KLS

scheme. On the other hand, if this connection carried over to the case of low-rank tensor manifolds for which the KLS integrators has been generalized, it may give a way to construct the orthographic retraction for these tensor manifolds. In turn, this would further broaden the applicability of the retraction-based curves of Chapters 3 and 5 to such tensor manifolds.

Bibliography

[AAM14]   P.-A. Absil, L. Amodei, and G. Meyer. Two Newton methods on the manifold of fixed-rank matrices endowed with Riemannian quotient geometries. *Comput. Statist.*, 29(3-4):569–590, 2014.

[ADM$^+$02]   R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub. Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.*, 22(3):359–390, 2002.

[AEM07]   D. Attali, H. Edelsbrunner, and Y. Mileyko. Weak witnesses for delaunay triangulations of submanifolds. In *Proceedings of the 2007 ACM Symposium on Solid and Physical Modeling*, SPM '07, page 143–150, New York, NY, USA, 2007. Association for Computing Machinery.

[AG90]   E. L. Allgower and K. Georg. *Numerical continuation methods*, volume 13 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1990.

[AGSW16]   P.-A. Absil, P.-Y. Gousenbourger, P. Striewski, and B. Wirth. Differentiable piecewise-Bézier surfaces on Riemannian manifolds. *SIAM J. Imaging Sci.*, 9(4):1788–1828, 2016.

[AM12]   P.-A. Absil and J. Malick. Projection-like retractions on matrix manifolds. *SIAM J. Optim.*, 22(1):135–158, 2012.

[AMHR13]   A. H. Al-Mohy, N. J. Higham, and S. D. Relton. Computing the Fréchet derivative of the matrix logarithm and estimating the condition number. *SIAM J. Sci. Comput.*, 35(4):C394–C410, 2013.

[AMS08]   P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008.

[Ams10]   D. Amsallem. *Interpolation on Manifolds of Cfd-Based Fluid and Finite Element-Based Structural Reduced-Order Models for On-Line Aeroelastic Predictions*. ProQuest LLC, Ann Arbor, MI, 2010. Thesis (Ph.D.)–Stanford University.

Bibliography

[AMT13]   P.-A. Absil, R. Mahony, and J. Trumpf. An extrinsic look at the Riemannian Hessian. In *Geometric science of information*, volume 8085 of *Lecture Notes in Comput. Sci.*, pages 361–368. Springer, Heidelberg, 2013.

[AO15]   P.-A. Absil and I. V. Oseledets. Low-rank retractions: a survey and new results. *Comput. Optim. Appl.*, 62(1):5–29, 2015.

[BA15]   N. Boumal and P.-A. Absil. Low-rank matrix completion via preconditioned optimization on the Grassmann manifold. *Linear Algebra Appl.*, 475:200–239, 2015.

[BCC21]   A. Bloch, M. Camarinha, and L. J. Colombo. Dynamic interpolation for obstacle avoidance on Riemannian manifolds. *Internat. J. Control*, 94(3):588–600, 2021.

[Bha07]   R. Bhatia. *Positive definite matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007.

[BKSL17]   J. Batista, K. Krakowski, and F. Silva Leite. Exploring quasi-geodesics on stiefel manifolds in order to smooth interpolate between domains. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 6395–6402, 2017.

[BMAS14]   N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15(42):1455–1459, 2014.

[Bou23]   N. Boumal. *An introduction to optimization on smooth manifolds*. Cambridge University Press, 2023.

[CKL23]   G. Ceruti, J. Kusch, and C. Lubich. A parallel rank-adaptive integrator for dynamical low-rank approximation. ArXiv preprint: https://arxiv.org/pdf/2304.05660.pdf, 2023.

[CKSL99]   P. Crouch, G. Kun, and F. Silva Leite. The de Casteljau algorithm on Lie groups and spheres. *J. Dynam. Control Systems*, 5(3):397–429, 1999.

[CL22]   G. Ceruti and C. Lubich. An unconventional robust integrator for dynamical low-rank approximation. *BIT*, 62(1):23–44, 2022.

[CL23]   A. Charous and P. F. J. Lermusiaux. Dynamically Orthogonal Runge–Kutta Schemes with Perturbative Retractions for the Dynamical Low-Rank Approximation. *SIAM J. Sci. Comput.*, 45(2):A872–A897, 2023.

[CSLC95]   M. Camarinha, F. Silva Leite, and P. Crouch. Splines of class C^k on non-Euclidean spaces. *IMA J. Math. Control Inform.*, 12(4):399–410, 1995.

178

[CT10] E. J. Candès and T. Tao. The power of convex relaxation: near-optimal matrix completion. *IEEE Trans. Inform. Theory*, 56(5):2053–2080, 2010.

[Dav53] D. F. Davidenko. On a new method of numerical solution of systems of nonlinear equations. *Doklady Akad. Nauk SSSR (N.S.)*, 88:601–602, 1953.

[dC59] P. de Casteljau. Outillages méthode calcul. Technical report, André Citroën Automobiles, Paris, 1959.

[dC92] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[DE99] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM J. Matrix Anal. Appl.*, 20(3):800–819, 1999.

[Deu11] P. Deuflhard. *Newton methods for nonlinear problems*, volume 35 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2011.

[DO05] D. Dunlavy and D. O'Leary. Homotopy optimization methods for global optimization. Technical report, Sandia National Laboratories, 12 2005.

[EAS99] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353, 1999.

[EM06] D. B. A. Epstein and A. Marden. Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces. In *Fundamentals of hyperbolic geometry: selected expositions*, volume 328 of *London Math. Soc. Lecture Note Ser.*, pages 117–266. Cambridge Univ. Press, Cambridge, 2006.

[Far02] G. Farin. *Curves and Surfaces for CAGD: A Practical Guide*. Computer graphics and geometric modeling. Elsevier Science, 2002.

[FL18] F. Feppon and P. F. J. Lermusiaux. A geometric approach to dynamical model order reduction. *SIAM J. Matrix Anal. Appl.*, 39(1):510–538, 2018.

[Gab82] D. Gabay. Minimizing a differentiable function over a differential manifold. *J. Optim. Theory Appl.*, 37(2):177–219, 1982.

[GHL04] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third edition, 2004.

[GK73] K. Grove and H. Karcher. How to conjugate C^1-close group actions. *Math. Z.*, 132:11–20, 1973.

Bibliography

[GSA14] P.-Y. Gousenbourger, C. Samir, and P.-A. Absil. Piecewise-Bézier C^1 interpolation on Riemannian manifolds with application to 2D shape morphing. In *Proceedings - International Conference on Pattern Recognition*, pages 4086–4091, 08 2014.

[GWZ84] J. Guddat, H. Wacker, and W. Zulehner. On imbedding and parametric optimization—a concept of a globally convergent algorithm for nonlinear optimization problems. *Math. Programming Stud.*, 21:79–96, 1984.

[HAG15] W. Huang, P.-A. Absil, and K. A. Gallivan. A Riemannian symmetric rank-one trust-region method. *Math. Program.*, 150(2, Ser. A):179–216, 2015.

[HAG18] W. Huang, P.-A. Absil, and K. A. Gallivan. A Riemannian BFGS method without differentiated retraction for nonconvex optimization problems. *SIAM J. Optim.*, 28(1):470–495, 2018.

[Hai01] E. Hairer. Geometric integration of ordinary differential equations on manifolds. *BIT*, 41:996–1007, 2001.

[HGA15] W. Huang, K. A. Gallivan, and P.-A. Absil. A Broyden class of quasi-Newton methods for Riemannian optimization. *SIAM J. Optim.*, 25(3):1660–1685, 2015.

[Hig08] N. J. Higham. *Functions of Matrices: Theory and Computation.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[HJ13] R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, second edition, 2013.

[HLW10] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics.* Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.

[HM94] U. Helmke and J. B. Moore. *Optimization and dynamical systems.* Communications and Control Engineering Series. Springer-Verlag London, Ltd., London, 1994. With a foreword by R. Brockett.

[HNW93] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, second edition, 1993. Nonstiff problems.

[JVV12] B. Jeuris, R. Vandebril, and B. Vandereycken. A survey and comparison of contemporary algorithms for computing the matrix geometric mean. *Electron. Trans. Numer. Anal.*, 39:379–402, 2012.

[Kar77] H. Karcher. Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.*, 30(5):509–541, 1977.

[KDLS21] K.-R. Kim, I. L. Dryden, H. Le, and K. E. Severn. Smoothing splines on Riemannian manifolds, with applications to 3D shape space. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 83(1):108–132, 2021.

[KEC23] J. Kusch, L. Einkemmer, and G. Ceruti. On the stability of robust dynamical low-rank approximations for hyperbolic problems. *SIAM J. Sci. Comput.*, 45(1):A1–A24, 2023.

[KFT13] T. Kaneko, S. Fiori, and T. Tanaka. Empirical arithmetic averaging over the compact Stiefel manifold. *IEEE Trans. Signal Process.*, 61(4):883–894, 2013.

[KH84] M. Kojima and R. Hirabayashi. *Continuous deformation of nonlinear programs*, pages 150–198. Springer Berlin Heidelberg, Berlin, Heidelberg, 1984.

[KL07] O. Koch and C. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.

[KLW16] E. Kieri, C. Lubich, and H. Walach. Discretized dynamical low-rank approximation in the presence of small singular values. *SIAM J. Numer. Anal.*, 54(2):1020–1038, 2016.

[KMSLB17] K. A. Krakowski, L. Machado, F. Silva Leite, and J. Batista. A modified Casteljau algorithm to solve interpolation problems on Stiefel manifolds. *J. Comput. Appl. Math.*, 311:84–99, 2017.

[KP13] S. G. Krantz and H. R. Parks. *The implicit function theorem.* Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. History, theory, and applications, Reprint of the 2003 edition.

[KSV14] D. Kressner, M. Steinlechner, and B. Vandereycken. Low-rank tensor completion by Riemannian optimization. *BIT*, 54(2):447–468, 2014.

[KV19] E. Kieri and B. Vandereycken. Projection methods for dynamical low-rank approximation of high-dimensional problems. *Comput. Methods Appl. Math.*, 19(1):73–92, 2019.

[Lee03] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2003.

[Lee13] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[Lee18] J. M. Lee. *Introduction to Riemannian manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, Cham, 2018.

[Lim21] L.-H. Lim. Tensors in computations. *Acta Numer.*, 30:555–764, 2021.

Bibliography

[LO14]   C. Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT*, 54(1):171–188, 2014.

[Lue72]   D. G. Luenberger. The gradient projection method along geodesics. *Management Sci.*, 18:620–631, 1972.

[LX15]   Q. Lin and L. Xiao. An adaptive accelerated proximal gradient method and its homotopy continuation for sparse optimization. *Comput. Optim. Appl.*, 60(3):633–674, 2015.

[Man12]   J. Manton. Optimisation geometry. ArXiv preprint: https://arxiv.org/pdf/1212.1775.pdf, 2012.

[MMH$^+$22]   A. Musolas, E. Massart, J. M. Hendrickx, P.-A. Absil, and Y. Marzouk. Low-rank multi-parametric covariance identification. *BIT*, 62(1):221–249, 2022.

[MMN$^+$20]   A. Maass, C. Manzie, D. Nesic, J. Manton, and I. Shames. Online zeroth-order optimisation on Riemannian manifolds. ArXiv preprint: https://arxiv.org/pdf/2010.00211.pdf, 2020.

[Moa05]   M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):735–747, 2005.

[NKS19]   L. Nguyen, J. Kim, and B. Shim. Low-rank matrix completion: A contemporary survey. *IEEE Access*, PP:1–1, 07 2019.

[NYP13]   E. Nava-Yazdani and K. Polthier. De Casteljau's algorithm on manifolds. *Comput. Aided Geom. Design*, 30(7):722–732, 2013.

[PN07]   T. Popiel and L. Noakes. Bézier curves and C^2 interpolation in Riemannian manifolds. *J. Approx. Theory*, 148(2):111–127, 2007.

[PR95]   F. C. Park and B. Ravani. Bézier curves on Riemannian manifolds and Lie groups with kinematics applications. *Journal of Mechanical Design*, 117(1):36–40, 03 1995.

[QSS07]   A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*, volume 37 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, second edition, 2007.

[Ral62]   A. Ralston. Runge-Kutta methods with minimum error bounds. *Math. Comp.*, 16:431–437, 1962.

[RSLJ05]   R. C. Rodrigues, F. Silva Leite, and J. Jakubiak. A new geometric algorithm to generate smooth interpolating curves on Riemannian manifolds. *LMS J. Comput. Math.*, 8:251–266, 2005.

[RW12] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM J. Optim.*, 22(2):596–627, 2012.

[San16] O. Sander. Geodesic finite elements of higher order. *IMA J. Numer. Anal.*, 36(1):238–266, 2016.

[SCK23] A. Séguin, G. Ceruti, and D. Kressner. Second-order time integrators for dynamical low-rank approximation using the orthographic retraction. in preparation, 2023.

[SH15] S. Sra and R. Hosseini. Conic geometric optimization on the manifold of positive definite matrices. *SIAM J. Optim.*, 25(1):713–739, 2015.

[Shu86] M. Shub. Some remarks on dynamical systems and numerical analysis. In *Dynamical systems and partial differential equations (Caracas, 1984)*, pages 69–91. Univ. Simon Bolivar, Caracas, 1986.

[SK22a] A. Séguin and D. Kressner. Continuation methods for Riemannian optimization. *SIAM J. Optim.*, 32(2):1069–1093, 2022.

[SK22b] A. Séguin and D. Kressner. Hermite interpolation with retractions on manifolds. ArXiv preprint: https://arxiv.org/abs/2212.12259, submitted for publication, 2022.

[Ste16] M. Steinlechner. Riemannian optimization for high-dimensional tensor completion. *SIAM J. Sci. Comput.*, 38(5):S461–S484, 2016.

[UV15] A. Uschmajew and B. Vandereycken. Greedy rank updates combined with Riemannian descent methods for low-rank optimization. *2015 International Conference on Sampling Theory and Applications, SampTA 2015*, pages 420–424, 07 2015.

[UV20] A. Uschmajew and B. Vandereycken. Geometric methods on low-rank matrix and tensor manifolds. In *Handbook of variational methods for nonlinear geometric data*, pages 261–313. Springer, Cham, [2020] ©2020.

[Van13] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM J. Optim.*, 23(2):1214–1236, 2013.

[vdV09] H. A. van der Vorst. *Iterative Krylov methods for large linear systems*, volume 13 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2009.

[Wat01] L. T. Watson. Theory of globally convergent probability-one homotopies for nonlinear programming. *SIAM J. Optim.*, 11(3):761–780, 2000/01.

[Wat79] L. T. Watson. A globally convergent algorithm for computing fixed points of C^2 maps. *Appl. Math. Comput.*, 5(4):297–311, 1979.

Bibliography

[WYZ12] Z. Wen, W. Yin, and Y. Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Math. Program. Comput.*, 4(4):333–361, 2012.

[ZB22] R. Zimmermann and R. Bergmann. Multivariate Hermite interpolation of manifold-valued data. ArXiv preprint: https://arxiv.org/abs/2212.07281, 2022.

[Zim17] R. Zimmermann. A matrix-algebraic algorithm for the Riemannian logarithm on the Stiefel manifold under the canonical metric. *SIAM J. Matrix Anal. Appl.*, 38(2):322–342, 2017.

[Zim20] R. Zimmermann. Hermite interpolation and data processing errors on Riemannian matrix manifolds. *SIAM J. Sci. Comput.*, 42(5):A2593–A2619, 2020.

[ZN19] E. Zhang and L. Noakes. The cubic de Casteljau construction and Riemannian cubics. *Comput. Aided Geom. Design*, 75:101789, 16, 2019.

Axel Séguin

E-mail: axel.seguin@epfl.ch. *Telephone number:* +41 76 801 0329

Nationalities: Italian, Canadian. *Date of birth:* 29/03/1995

Education

PhD thesis in applied mathematics   *June 2019 - September 2023*
Ecole Polytechnique Fédérale de Lausanne (advisor: Daniel Kressner)

"Retraction-based numerical methods for continuation, interpolation and time integration on manifolds".

Master in computational sciences and engineering   *September 2016 - February 2019*
Ecole Polytechnique Fédérale de Lausanne (final grade: 5.76/6)

Bachelor in mathematics for engineering   *September 2013 - July 2016*
Politecnico di Torino (final grade: 109/110)

Baccalauréat scientifiques spécialité mathématiques   *June 2013*
Lycée Jean Giono de Turin (final grade: 18.3/20)

Work experience

Software development for grinding trajectory computation   *September 2018 - ongoing*
Rollomatic SA

- Development of new mathematical model for grinding trajectory computation.
- C# implementation of the new framework (3D geometry, automatic differentiation, optimization).
- Internship and master project design and supervision.

Teaching assistantship in mathematics   *February 2017 - December 2022*
Ecole Polytechnique Fédérale de Lausanne

- Functional Analysis – 3rd year bachelor course in mathematics (4 times).
- Continuous optimization – 3rd year bachelor course in mathematics.
- Analysis II and IV – 1st and 2nd year bachelor course in engineering.
- Numerical analysis – 2nd year bachelor course in engineering.

Shape optimization internship   *February 2018 - July 2018*
Rollomatic SA

- Development of method to optimize grinding wheel geometry given wheel trajectory.
- Minimum viable prototype in MATLAB.

Programming skills

MATLAB	Expert
C#	Advanced
C++, Python, Mathematica	Basic

Languages

French	Native
Italian	Native
English	C1 level
Spanish	B2 level

List of publications

Second-order time integrators for DLRA using the orthographic retraction
A. Séguin, G. Ceruti, and D. Kressner, in preparation, 2023.

Hermite interpolation with retractions on manifolds
A. Séguin and D. Kressner, ArXiv preprint, submitted for publication, 2022.

Continuation methods for Riemannian optimization
A. Séguin and D. Kressner, SIAM Journal on Optimization, 2022.

List of further projects

Master thesis in Riemannian optimization   *Advisor: Daniel Kressner*
An application of Riemannian Optimization for SE(3) equations.

Master semester project in computational fluid dynamics   *Advisor: Marco Picasso*
Free surface flow of a viscoelastic fluid for simulation elastic solid dynamics.

Master semester project in reduced order modeling for PDE   *Advisors: Alfio Quarteroni*
A non-intrusive Reduced Basis method using RBF interpolation.   *Andrea Manzoni*

Bachelor thesis in compuational differential geometry   *Advisor: Emilio Musso*
Costa's Minimal Surface: derivation and a Mathematica implementation.

Conference contributions

11th Conference on Applied Mathematics and Scientific Computing   *September 2022*
Hermite interpolation with retractions on manifold (poster).   *Brijuni, Croatia*

- Awarded honorable mention.

SIAM Conference on Optimization (OP23)   *June 2023*
Continuation methods for Riemannian optimization (talk).   *Seattle, United States*