Applied Mathematical Sciences

Laurent Younes

Shapes and Diffeomorphisms

Second Edition

Springer

Applied Mathematical Sciences

Volume 171

The mathematization of all sciences, the fading of traditional scientific boundaries, the impact of computer technology, the growing importance of computer modeling and the necessity of scientific planning all create the need both in education and research for books that are introductory to and abreast of these developments. The purpose of this series is to provide such books, suitable for the user of mathematics, the mathematician interested in applications, and the student scientist. In particular, this series will provide an outlet for topics of immediate interest because of the novelty of its treatment of an application or of mathematics being applied or lying close to applications. These books should be accessible to readers versed in mathematics or science and engineering, and will feature a lively tutorial style, a focus on topics of current interest, and present clear exposition of broad appeal. A compliment to the Applied Mathematical Sciences series is the Texts in Applied Mathematics series, which publishes textbooks suitable for advanced undergraduate and beginning graduate courses.

More information about this series at http://www.springer.com/series/34

Laurent Younes

Shapes and Diffeomorphisms

Second Edition

Springer

Laurent Younes
Center for Imaging Science
Johns Hopkins University
Baltimore, MD, USA

ISSN 0066-5452 ISSN 2196-968X (electronic)
Applied Mathematical Sciences
ISBN 978-3-662-58495-8 ISBN 978-3-662-58496-5 (eBook)
https://doi.org/10.1007/978-3-662-58496-5

Library of Congress Control Number: 2018964233

Mathematics Subject Classification (2010): 68T10, 53-01, 68-02, 92C55, 37E30, 37C10, 53A04, 53A05, 68U05

This Springer imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

To Geneviève, Hannah, Salomé, Simon

Introduction to the Second Edition

Besides correcting some of the typos and mistakes from the first edition and implementing a few changes to the notation to make it uniform, several significant additions appear in the second edition. Most notable is the introduction of a discussion of optimal control theory in an infinite-dimensional framework (Appendix D), which is then used in multiple places to enrich the presentation of diffeomorphic matching, and a new chapter on shape datasets (Chap. 13 of the second edition). A few other changes have been made. These are listed below, where chapters are referred to as x/y, the first and second numbers indicating the first and second editions, respectively.

- Chapter 1/1: The presentation of closed curves and integrals along them has been revised. Some sections have been reordered to improve readability.
- Chapter 2/2: No significant changes.
- Chapter 3 of the first edition has been removed.
- Chapter 4/3: No major changes were made. The material on discrete surfaces was moved to Chap. 5/4, in which a new discussion is added on consistent discrete-to-continuous approximation.
- Chapters 6/5 and 7/6 only had minor modifications.
- Chapter 8/7: A few additional results on compositions of diffeomorphisms and their derivatives have been included. Small modifications were made to the rest of the chapter, which also relies more directly on results from the appendix.
- Chapter 9/8: The discussion on invariant operators and kernels has been updated and clarified.
- Chapter 10/9: Only minor changes were made in most sections. Exceptions are (1) curve and surface matching, in which a discussion of varifold distances was added, and (2) frame matching, which has been rewritten with a slightly different action, and a more rigorous handling of the supports of the frame fields.
- Chapter 11/10: Some of the discussion now directly uses results from the appendix on ODEs and optimal control. The proof of the existence and uniqueness of solutions of the EPDiff equation has also been rewritten.
- Chapter 12/11 had only minor changes.

- Chapter 13/12 on metamorphosis has been extended, with more examples, including in particular a detailed discussion of metamorphosis applied to the tangent representation of curves.
- Chapter 13 in the second edition is new.
- Only minor changes were made in Appendix A.
- Appendix B was extended with a discussion of differential forms and some consequences of Stokes's theorem.
- Appendix C has essentially the same content, even though part of the write-up was revised.
- Appendix D now includes an introduction to optimal control theory.
- Appendices E and F were only slightly modified.

Introduction to the First Edition

Shape is a fascinating object of study. Understanding how a single shape can incur a complex range of transformations, while defining the same perceptually obvious figure, entails a rich and enticing collection of problems at the interface between applied mathematics, statistics, and computer science. Various applications in computer vision, object recognition, and medical imaging bring additional motivation for researchers to develop adequate theoretical background and methodology for solving these problems.

This book is an attempt at providing a description of the large range of methods that have been invented to represent, detect, or compare shapes (or more generally, deformable objects), together with the necessary mathematical background that they require. While certainly being a book on applied mathematics, it is also written in a way that will be of interest to an engineering- or computer-science-oriented reader, including in several places concrete algorithms and applicable methods, including experimental illustrations.

This book starts with a discussion of shape representation methods (Chapters 1–4), including classical aspects of the differential geometry of curves and surfaces, but borrowing also from other fields that have positively impacted the analysis of shape in practical applications, such as medial axes and discrete differential geometry.

The second part (Chapters 5–7) studies curve and surface evolution algorithms and how they relate to segmentation methods that can be used to extract shapes from images, using active contours or deformable templates. A reader with enough background in differential geometry may start reading this book at Chapter 6 or at Chapter 7 if the main focus of interest is on diffeomorphic registration and comparison methods.

In Chapters 7 and 8, basic concepts related to diffeomorphisms are introduced, discussing in particular how using ordinary differential equations associated with vector fields belonging to reproducing kernel Hilbert space provides a computationally convenient framework to handle them. Chapters 9 and 10 then focus on the registration of deformable objects using diffeomorphisms; in Chapter 9, we catalog a large collection of deformable objects and discuss matching functionals that can be used to compare them. Chapter 10 addresses diffeomorphic matching and

focuses in particular on methods that optimize a matching functional combined with a regularization term that penalizes the distance of a diffeomorphism to the identity within the group.

The next two chapters (11 and 12) discuss metric aspects of shape analysis, with a special focus on the relation between distances and group actions. Both the global and infinitesimal points of view are presented. The classical Kendall's metric over configurations of labeled points is included, as well as a short discussion of Riemannian metrics on plane curves. Chapter 12 provides a presentation of the theory of metamorphosis. Chapter 13 provides an introduction to the statistical analysis of shape data.

In the appendices are provided fundamental concepts that are needed in order to understand the rest of this book. The main items are some elements of functional analysis (Appendix A), of differential and Riemannian geometry (Appendix B), and of ordinary differential equations (Appendix C). Appendix D provides an introduction to optimization and optimal control. Appendix E focuses on principal component analysis and Appendix F on dynamic programming. In all cases, the appendices do not provide a comprehensive presentation of these theories, but simply what is needed in the particular context of this book.

Chapters 1 to 5, which are (with a few exceptions) rather elementary, provide an introduction to applied differential geometry that is suitable for an advanced undergraduate class. They can be combined with Chapter 6 to form a graduate-level class on the same subject. The first six chapters are written (with a few exceptions) in order to be accessible without using the more advanced features developed in the appendices. Chapters 8 to 13 represent specialized, advanced graduate topics.

I would like to thank my students and collaborators, who have helped to make the ideas that are developed in these notes reach their current state of maturation. I would like, in particular, to express my gratitude to Alain Trouvé and Michael Miller, whose collaboration over the last decade has been invaluable. Special thanks also to Darryl Holm, David Mumford, and Peter Michor. This book was written while the author was partially supported by the National Science Foundation, the National Institute of Health, and the Office for Naval Research.

Contents

List of Figures

Chapter 1
Parametrized Plane Curves

1.1 Definitions

We start with some definitions.

Definition 1.1 A (parametrized plane) curve is a continuous mapping $m : I \to \mathbb{R}^2$, where $I = [a, b]$ is an interval.

A curve m is closed if $m(a) = m(b)$.

A curve m is a Jordan curve if it is closed and has no self-intersection: $m(x) = m(y)$ only for $x = y$ or $\{x, y\} = \{a, b\}$.

A curve is piecewise C^1 if it has everywhere left and right derivatives, which coincide except at a finite number of points.

The range of a curve m is the set $m([a, b])$. It will be denoted by $\mathcal{R}_m$.

Notice that we have defined curves as functions over bounded intervals. Their range must therefore be a compact subset of $\mathbb{R}^2$ (this forbids, in particular, curves with unbounded branches).

A Jordan curve is what we can generally accept as a definition of the outline of a shape. An important theorem [292] states that the range of a Jordan curve partitions the plane $\mathbb{R}^2$ into two connected regions: a bounded one, which is the interior of the curve, and an unbounded one (the exterior). The proof of this rather intuitive theorem is quite complex (see, for example [184] for an argument using Brouwer's fixed point theorem).

However, requiring only continuity for curves allows for more irregularities than what we would like to handle. This is why we will always restrict ourselves to piecewise C^1, generally Jordan, curves. We will in fact often ask for more, and consider curves which are regular (or piecewise regular).

Definition 1.2 A C^1 curve $m : I \mapsto \mathbb{R}^2$ is a regular curve if $\partial m \neq 0$ for all $u \in I$. If m is only piecewise C^1, we extend the definition by requiring that all left and right derivatives are non-vanishing.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_1

Here, and in the rest of the book, we will use either ∂m or $\dot{m}$ to denote the derivative of a function $u \mapsto m(u)$.

The previous definition is fundamental. It avoids, in particular, curves which are smooth functions (C^∞, for example) but with a range having geometric singularities. Consider the following example: let

$$m(u) = \begin{cases} (\varphi(u), 0), & u \in [0, 1/2] \\ (1, \varphi(u - 1/2)), & u \in [1/2, 1] \end{cases}$$

with $\varphi(u) = 16u^2(1-u)^2$, $u \in [0, 1]$. It is easy to check that m is continuously differentiable, whereas the range of m is the corner $[0, 1] \times \{0\} \cup \{1\} \times [0, 1]$.

We will say that a curve $m : [a, b] \to \mathbb{R}^2$ is C^p if it is p times continuously differentiable, including all right derivatives at a and left derivatives at b up to order p. If the curve is closed, we will implicitly require that the derivatives at a and b coincide. More precisely, a closed curve is C^p if and only if m is C^p when restricted to the open interval (a, b), and so is the curve $\tilde{m}$ defined on (a, b) by $\tilde{m}(u) = m(u + \varepsilon)$ if $u \in (a, b - \varepsilon]$ and $\tilde{m}(u) = m(u + \varepsilon - b + a)$ if $u \in [b - \varepsilon, b)$ (for some $0 < \varepsilon < b - a$).

Alternatively (and more conveniently), closed curves can be handled by considering the interval $[a, b]$ closed onto itself after identifying a and b (which provides a one-dimensional torus). We will denote this torus by $[a, b]_*$ and let $a \sim b \in [a, b]_*$ denote a or b after the identification. For $u, v \in [a, b]_*$, we let

$$d_*(u, v) = \min(|u - v|, (b - a) - |u - v|). \tag{1.1}$$

This is a distance on $[a, b]_*$ (if u (or v) are equal to $a\tilde{b}$, the result does not depend on which value is chosen to compute the expression).

If $\delta \in \mathbb{R}$ and $u \in [a, b]_*$, we define $u +_* \delta \in [a, b]_*$ by $u + \delta - (u +_* \delta) = k(b - a)$ for some integer k (so that we consider addition modulo $b - a$). A function $f : [a, b]_* \to \mathbb{R}^d$ is continuous on $[a, b]_*$ if and only if, for all $u \in [a, b]_*$, $|f(u +_* \delta) - f(u)| \to 0$ when $\delta \to 0$, which is equivalent to $f(u) = \hat{f}(u)$ for some function $\hat{f}$ continuous on $[a, b]$ satisfying $\hat{f}(a) = \hat{f}(b)$. One defines derivatives of functions by

$$\partial f(u) = \lim_{\delta \to 0} \frac{f(u +_* \delta) - f(u)}{\delta}$$

when the right-hand side exists and higher derivatives are defined accordingly. With this notation, it is easy to see that C^p closed curves are functions $m : [a, b]_* \to \mathbb{R}^2$ with at least p continuous derivatives.

We also use integrals along $[a, b]_*$ as follows: if $u_0, u_1 \in [a, b]$ and $f : [a, b] \to \mathbb{R}^d$ is continuous, then

$$\int_{u_0}^{u_1} f(v)dv_* := \begin{cases} \displaystyle\int_{u_0}^{u_1} f(v)dv \text{ if } u_0 \le u_1 \\ \displaystyle\int_a^b f(v)dv - \int_{u_1}^{u_0} f(v)dv \text{ if } u_1 \le u_0. \end{cases} \tag{1.2}$$

This is the integral of f along the "positively oriented arc" going from u_0 to u_1 in $[a, b]_*$. Note that

$$\int_{u_0}^{u_1} f(v)dv_* \neq -\int_{u_1}^{u_0} f(v)dv_*$$

in general. For example $\int_b^a f(v)dv_* = 0$ for all f.

The length of the positively oriented arc going from u_0 to u_1 is

$$\ell_*(u_0, u_1) = \int_{u_0}^{u_1} dv_* := \begin{cases} |u_1 - u_0| \text{ if } u_0 \le u_1 \\ (b-a) - |u_1 - u_0| \text{ if } u_1 \le u_0. \end{cases}$$

With this notation, $d_*(u_0, u_1) = \min(\ell_*(u_0, u_1), \ell_*(u_1, u_0))$.

1.2 Reparametrization Equivalence

1.2.1 Open Curves

Definition 1.3 Let $m : I \to \mathbb{R}^2$ be a plane curve. A change of parameter for m is a function $\psi : I' \to I$ such that:
(i) I' is a bounded interval;
(ii) ψ is continuous, increasing (strictly) and onto.

From (ii), ψ is one-to-one and onto, hence invertible. Its inverse, ψ^{-1} is also a change of parameter (the proof being left to the reader). In particular, ψ is a *homeomorphism* (a continuous invertible function with a continuous inverse).

The new curve $\tilde{m} = m \circ \psi$ is called a reparametrization of m. The ranges $\mathcal{R}_m$ and $\mathcal{R}_{\tilde{m}}$ coincide.

When m belongs to a specific smoothness class, the same properties will be implicitly required for the change of parameter. For example, if m is (piecewise) C^1, ψ will also be assumed to be C^1 (in addition to the previous properties). When working with regular curves, the following assumption will be made.

Definition 1.4 If I, I' are bounded intervals, a regular change of parameter is a C^1 function $\psi : I' \to I$ which is onto and satisfies $\dot{\psi} > 0$ everywhere (including left and right limits at the bounds).

A piecewise regular change of parameter is continuous, piecewise C^1 and such that its left and right derivatives (which coincide everywhere except at a finite number of points) are all strictly positive.

It is easy to see that the property of two curves being related by a change of parameter is an equivalence relation. This is called "parametric equivalence." We will denote the parametric equivalence class of m by $[m]$. A property, or a quantity, which only depends on $[m]$ will be called parametrization-invariant. For example, the range of a curve is parametrization-invariant.

Note that the converse is not true. If two curves have the same range, they are not necessarily parametrically equivalent: the range of the piecewise C^1 curve defined on $I = [0, 1]$ by $m(t) = (2t, 0), t \in [0, 1/2]$ and $m(t) = (2 - 2t, 0),\ t \in [1/2, 1]$ is the segment $[0, 1] \times 0$, but this curve is not equivalent to $\tilde{m}(t) = (t, 0), t \in [0, 1]$, even though they have the same range (the first one travels back to its initial point after reaching the end of the segment). Also, if m is a curve defined on $I = [0, 1]$, then $\tilde{m}(t) = m(1 - t)$ has the same range, but is not equivalent to m, since we have required the change of parameter to be increasing (changes of orientation are not allowed).

Changes of parameter will always be assumed to match the class of curves that is being considered: (piecewise) regular reparametrizations for (piecewise) regular curves, or, when more regularity is needed, C^p regular reparametrizations for C^p regular curves.

1.2.2 Closed Curves

Changes of parameters for closed curves must be slightly more general than for open curves, because the starting point of the parametrization is not uniquely defined. Using representations over tori, we will say that a continuous mapping $\psi : [a', b']_* \to [a, b]_*$ is increasing if, for all u, one can write, for small enough δ,

$$\psi(u +_* \delta) = \psi(u) +_* \varepsilon(\delta),$$

where $\varepsilon : (-\delta_0, \delta_0) \to \mathbb{R}$ can be defined for some $\delta_0 > 0$ as a (strictly) increasing function such that $\varepsilon(\delta) \to 0$ if $\delta \to 0$. This says that ψ moves in the same direction as u.

A change of parameter is then a continuous, increasing, one-to-one transformation ψ from $[a', b']_*$ onto $[a, b]_*$ (and its inverse is then continuous too). The main difference with the open case is that such a transformation does not necessarily satisfy $\psi(a') = a$: it can start anywhere and wrap around to return to its initial point. We will then say that the change of parameter is regular if it is C^1 with $\dot{\psi} > 0$ everywhere, as in the open case.

Letting $c' = \psi^{-1}(b)$ (recall that $a = b \in [a, b]_*$) and taking $\hat{\psi} : [a', b'] \to [a, b]$ to be such that $\hat{\psi}(u') = \psi(u')$ if $u' \neq c'$ and $\hat{\psi}(u') = b$ otherwise, the definition is

equivalent to requiring that $\hat{\psi}$ is increasing over $[a', c']$ and over $(c', b']$, continuously differentiable over these intervals, with left and right derivatives coinciding at c', and the right derivative at a' coinciding with the left derivative at b'.

1.3 Unit Tangent and Normal

If $M \subset \mathbb{R}^d$ is an arbitrary set, we will say that a vector $v \in \mathbb{R}^d$ is tangent to M at a point p in M if one can find points x in M that are arbitrarily close to p and such that v is arbitrarily close to the half line $\mathbb{R}^+(x - p)$. This is formalized in the following definition (see [107]):

Definition 1.5 If $M \subset \mathbb{R}^d$, and $p \in M$, a vector $v \in \mathbb{R}^d$ is an oriented tangent to M at p if, for any $\varepsilon > 0$, there exist $x \in M$ and $r > 0$ such that $|x - p| < \varepsilon$ and $|v - r(x - p)| < \varepsilon$.

The set of oriented tangents to M at p will be denoted by $T_p^+ M$, and the set of (unoriented) tangents by $T_p M$, so that $v \in T_p M$ if either v or $-v$ belongs to $T_p^+ M$. Taking $x = p$, one sees that $v = 0$ always belong to $T_p^+ M$, which is therefore never empty.

Let $m : I \to \mathbb{R}^2$ be a regular curve (here, I can be a closed interval or a torus). The unit tangent at $u \in I$ is the vector

$$T_m(u) = \frac{\dot{m}(u)}{|\dot{m}(u)|}.$$

We then have

Proposition 1.6 *If $m : I \to \mathbb{R}^2$ is regular, and $p \in \mathcal{R}_m$, then*

$$T_p \mathcal{R}_m = \{\lambda T_m(u) : \lambda \in \mathbb{R}, m(u) = p\}.$$

Note that, when m is regular the set of parameters u such that $m(u) = p$ is necessarily finite. (Each such u is necessarily isolated because $\dot{m} \neq 0$ and any family of isolated points in a compact set must be finite.)

Proof Let $I = [a, b]$ (the case of a closed curve being addressed similarly). Take $p \in \mathcal{R}_m$ and u such that $p = m(u)$. Fix $\lambda \in \mathbb{R}$ and $\varepsilon > 0$. One has $m(u + \delta) - m(u) - \delta|\dot{m}(u)|T_m(u) = o(\delta)$. So, taking δ small enough so that $|m(u + \delta) - m(u)| < \varepsilon$ and

$$\left| \lambda T_m(u) - \frac{\lambda}{\delta|\dot{m}(u)|}(m(u + \delta) - m(u)) \right| < \varepsilon$$

one gets $|x - p| < \varepsilon$ and $|\lambda T_m(u) - r(x - p)| < \varepsilon$ with $x = m(u + \delta)$ and $r = \lambda/(\delta|\dot{m}|)$. If $u \in (a, b)$, one can ensure that $r > 0$ by choosing the sign of δ appropriately. If $u = a$ one must take $\delta > 0$ and $r > 0$ only if $\lambda > 0$. Similarly, if $u = b$,

one needs $\lambda < 0$. In any case, either $\lambda T_m(u)$ or $-\lambda T_m(u)$ belongs to $T_p^+\mathcal{R}_m$, so that $\lambda T_m(u) \in T_p\mathcal{R}_m$.

Conversely, if $v \in T_p\mathcal{R}_m$, there exists sequences (u_n) and (r_n) with $u_n \in I$ and $r_n > 0$ such that $|m(u_n) - p| < 1/n$ and $|v - r_n(m(u_n) - p)| < 1/n$. Taking a subsequence if needed, one can assume that $u_n \to u \in I$, necessarily with $p = m(u)$. If $u_n = u$ for an infinite number of v, then $|v| \leq 1/n$ for these n, which implies $v = 0$. Otherwise, remove these values from the sequence to ensure $u_n \neq u$ for all n and use the fact that

$$v - r_n(u_n - u)\frac{m(u_n) - p}{u_n - u} \to 0$$

with $(m(u_n) - p)/(u_n - u) \to \dot{m}(u) \neq 0$ to prove that $r_n(u_n - u)$ converges to some $\lambda \in \mathbb{R}$. We then have $v = \lambda \dot{m}_u$, which completes the proof. □

The unit normal is the unique vector $N_m(u)$ which extends $T_m(u)$ to a positively oriented orthonormal basis of $\mathbb{R}^2$: $(T_m(u), N_m(u))$ is orthonormal and $\det[T_m(u), N_m(u)] = 1$. The subscript m is generally dropped in the absence of ambiguity.

The frame (T, N) is parametrization-invariant in the following sense: if $\varphi : I \to \tilde{I}$ is a regular change of parameter, and $m = \tilde{m} \circ \varphi$, then $T_{\tilde{m}}(\varphi(u)) = T_m(u)$ and similarly for the normal.

1.4 Embedded Curves

Letting I be either an interval or a torus, a C^1 function $m : I \to \mathbb{R}^2$ such that $\dot{m}(u) \neq 0$ everywhere is a special case of an immersion (see Definition B.13), and regular curves are also sometimes called *immersed curves*. Among immersed curves, one also distinguishes *embedded curves* which are furthermore assumed to be non-intersecting (so that closed embedded curves are regular Jordan curves). For embedded curves, $T_m(u)$ is (up to a sign change) the only unit element of $T_{m(u)}M$. Moreover, if $m : I \mapsto \mathbb{R}^2$ is an embedding, the inverse map $m^{-1} : \mathcal{R}_m \to I$, which is well defined by assumption, is continuous: if $p_n \in \mathcal{R}_m$ is a sequence that converges to $p \in \mathcal{R}_m$, then, for some u_n and u, $p_n = m(u_n)$ and $p = m(u)$. Any limit v of a subsequence of u_n (recall that I is compact, so that any sequence has at least a convergent subsequence, and any limit of a subsequence belongs to I) must satisfy, by continuity, $m(v) = p$, which implies $v = u$. This implies that $m^{-1}(p) = u$.

If two embedded curves have the same range, they can be deduced from one another through a change of parameters, possibly after reorientation (this is not true for regular curves). Letting $m : I \mapsto \mathbb{R}^2$ and $m' : I' \mapsto \mathbb{R}^2$ be two such curves, $\psi = m^{-1} \circ m'$ is a homeomorphism (continuous, with a continuous inverse) between I' and I. If v is any point in the case of closed curves, or $v \in (a', b')$ for open curves, one can apply the implicit function theorem to the identity $m' = m \circ \psi$ to prove that ψ is differentiable with

$$\dot{\psi}\,\dot{m} \circ \psi = \dot{m}',$$

implying

$$\dot{\psi} = \frac{(\dot{m}')^T \, \dot{m} \circ \psi}{|\dot{m} \circ \psi|^2}.$$

For open curves, one shows that ψ has non-zero right and left derivatives at a' and b' by passing to the limit, the detailed argument being left to the reader. Because of this, any parametrization-invariant quantity only depends on the (oriented) range of the curve when restricted to embeddings.

With some abuse of terminology, we will say that a subset $\mathcal{R} \subset \mathbb{R}^2$ is an embedded curve if there exists an embedded curve m such that $\mathcal{R} = \mathcal{R}_m$. Such a curve m is then defined up to a change of parameter.

1.5 The Integral Along a Curve and Arc Length

Let $m : [a, b] \to \mathbb{R}^2$ be a parametrized curve. If $\sigma = (a = u_0 < u_1 < \cdots < u_n < u_{n+1} = b)$ is a subdivision of $[a, b]$, one can approximate m by the polygonal line m_σ with vertices $(m(u_0), \ldots, m(u_{n+1}))$. The length of m_σ is the sum of lengths of the segments that form it, namely

$$L_{m_\sigma} = \text{length}(m_\sigma) = \sum_{i=1}^{n+1} |m(u_i) - m(u_{i-1})|.$$

One then defines the length of m as

$$L_m = \sup_\sigma L_{m_\sigma},$$

where the supremum (which can be infinite) is over all possible subdivisions σ of $[a, b]$.

One then has the following proposition.

Proposition 1.7 *If $m : [a, b] \to \mathbb{R}^2$ is C^1, then*

$$L_m = \int_a^b |\dot{m}(t)| dt < \infty.$$

Proof The fact that the integral is finite results from the derivative being bounded on the compact interval $[a, b]$ (because the curve is C^1). If $\sigma = (a = u_0 < u_1 < \cdots < u_n < u_{n+1} = b)$ is a subdivision of $[a, b]$, then one has

$$|m(u_{i+1}) - m(u_i)| = \left| \int_{u_i}^{u_{i+1}} \dot{m}(t) dt \right| \leq \int_{u_i}^{u_{i+1}} |\dot{m}(t)| dt.$$

Summing over i yields the fact that $L_{m_\sigma} \leq \int_a^b |\dot{m}(t)|dt$ and taking the supremum on the right-hand side implies the same inequality for L_m.

On the other hand, given any σ, the finite increment theorem implies that, for all i, there exists $v_i \in (u_i, u_{i+1})$ such that $m(u_{i+1}) - m(u_i) = \dot{m}(v_i)(u_{i+1} - u_i)$. Using this, we see that

$$L_{m_\sigma} = \sum_{i=0}^{n} |\dot{m}(v_i)|(u_{i+1} - u_i),$$

which is a Riemann sum for $\int_a^b |\dot{m}(t)|dt$, and can therefore be made arbitrarily close to the integral by taking fine enough subdivisions. So for any ε, one can find σ such that

$$\int_a^b |\dot{m}(t)|dt \leq L_{m_\sigma} + \varepsilon$$

and since the upper-bound is less than $L_m + \varepsilon$, we find $\int_a^b |\dot{m}(t)|dt \leq L_m$ by letting ε tend to 0. This completes the proof of the proposition. □

If $f : I \to \mathbb{R}$ is a continuous function, one defines the integral of f along m by

$$\int_m f \, d\sigma_m = \int_a^b f(u) \, |\dot{m}(u)| \, du. \tag{1.3}$$

The definition is parametrization-independent: if $\psi : [a', b'] \to [a, b]$ is a change of parameters, then, using a change of variable,

$$\begin{aligned}\int_{a'}^{b'} f(\psi(u'))|\partial(m \circ \psi)(u')|du' &= \int_{a'}^{b'} f(\psi(u'))|\dot{m} \circ \psi(u')| \, \dot{\psi}(u') \, du' \\ &= \int_a^b f(u)|\dot{m}(u)|du\end{aligned}$$

so that

$$\int_{m\circ\psi} f \circ \psi \, d\sigma_{m\circ\psi} = \int_m f \, d\sigma_m.$$

The same result holds if $\psi : [a', b']_* \to [a, b]_*$ is a change of parameter between closed curves. In that case, taking c' such that $\psi(c') = a \sim b$ and letting $\psi(a' \sim b') = c$, we have

$$\begin{aligned}\int_{a'}^{b'} f \circ \psi(u')du' &= \int_{a'}^{c'} f \circ \psi(u')|\partial(m \circ \psi)(u')|du' \\ &\quad + \int_{c'}^{b'} f \circ \psi(u')|\partial(m \circ \psi)(u')|du' \\ &= \int_c^b f(u)|\dot{m}(u)|du + \int_a^c f(u)|\dot{m}(u)|du\end{aligned}$$

$$= \int_a^b f(u)|\dot{m}(u)|du.$$

These results imply that, when m is an embedding, the integral along m only depends on the range $\mathcal{R}_m$. This allows us to define the integral of a function over $\mathcal{R}_m$ by

$$\int_{\mathcal{R}_m} f\, d\sigma_{\mathcal{R}_m} = \int_m f\, d\sigma_m,$$

which does not depend on how $\mathcal{R}_m$ is parametrized.

We now give the following important definition.

Definition 1.8 Let $m : I \to \mathbb{R}^2$ be a (piecewise) C^1 curve, where I is either $[a, b]$ or $[a, b]_*$. A change of parameter $\sigma : I \to [0, L_m]$ (or $[0, L_m]_*$) is an arc-length reparametrization of m if

$$\dot{\sigma} = |\dot{m}|.$$

One says that m is parametrized by arc length if $m : [0, L_m] \to \mathbb{R}^2$ satisfies $|\dot{m}| = 1$.

If m is regular, then σ is a regular change of parameter and $m \circ \sigma^{-1}$ is an arc-length reparametrization of m. When $I = [a, b]$, the arc-length reparametrization is unique and given by

$$\sigma_m(u) = \int_a^u |\dot{m}(v)|dv. \tag{1.4}$$

When $I = [a, b]_*$, the parametrization is unique once the starting point $c = \sigma^{-1}(0)$ is chosen, and is given by (following (1.2))

$$\sigma_{m,c}(u) = \int_c^u |\dot{m}(v)|dv_*. \tag{1.5}$$

The arc length is parametrization-invariant: if m is a curve, with arc-length reparametrization σ, and $\tilde{m} = m \circ \psi$ is another parametrization of m, then $\sigma \circ \psi$ is an arc-length parametrization of $\tilde{m}$ (this is obvious, since $\tilde{m} \circ (\sigma \circ \psi)^{-1} = m \circ \sigma^{-1}$).

When a curve is parametrized by arc length, it is customary to denote its parameter by s instead of u, and we will follow this convention. From our definition of integrals, we clearly have in that case

$$\int_m f d\sigma_m = \int_0^{L_m} f(s)ds$$

(or ds_* in the case of closed curves).

We will also use the notion of derivative with respect to the arc length. For open curves, this corresponds to the limit of the ratio

$$\frac{g(u+\varepsilon)-g(u)}{\sigma_m(u+\varepsilon)-\sigma_m(u)}$$

as $\varepsilon \to 0$ (for closed curves, replace $+$ by $+_*$) therefore leading to the following definition.

Definition 1.9 Let $m : I \to \mathbb{R}^2$ be a C^1 regular curve. The operator ∂_{s_m} transforms a C^1 function g over I into the function $\partial_{s_m} g$, which is defined over I by

$$\partial_{s_m} g(u) = \frac{\dot{g}(u)}{|\dot{m}(u)|}. \tag{1.6}$$

We will write ∂_s if there is no ambiguity concerning the curve m. Note that, if m is parametrized by arc length (so that $u = s$), this notation coincides with the usual derivative with respect to s and therefore introduces no conflict.

The next proposition expresses that the derivative with respect to the arc length is parametrization invariant.

Proposition 1.10 *Let $m : I \to \mathbb{R}^2$ be a regular curve and $\psi : I' \to I$ be a change of parameter, with $\tilde{m} = m \circ \psi$. Then, for any C^1 function g defined on I,*

$$(\partial_{s_m} g) \circ \psi = \partial_{s_{\tilde{m}}}(g \circ \psi).$$

Proof This derives from the definition and from the chain rule, namely

$$\partial_{s_{\tilde{m}}}(g \circ \psi) = \frac{\partial_{\tilde{u}}(g \circ \psi)}{|\partial_{\tilde{u}}(m \circ \psi)|} = \frac{(\partial_u g) \circ \psi}{|(\partial_u m) \circ \psi|} = (\partial_{s_m} g) \circ \psi$$

(the positive term $\partial_{\tilde{u}} \psi$ cancels in the ratio). □

Note that, with this definition, one can rewrite the definition of the unit tangent as $T_m = \partial_{s_m} m$.

The following proposition shows how the arc length parametrization can be used to stitch several local parametrizations of a set to a global one forming an embedding.

Proposition 1.11 *Let $\mathcal{R} \subset \mathbb{R}^2$ be compact and connected. Assume that there exists a family $V_1, \ldots, V_n$ of open sets in $\mathbb{R}^2$, and a family $m_i : [a_i, b_i] \to \mathbb{R}^2$ of embeddings, such that $\mathcal{R} \subset \bigcup_{i=1}^n V_i$ and, for every $i = 1, \ldots, n$, $\mathcal{R} \cap V_i = m_i((a_i, b_i))$. Then, there exists a closed embedding $m : [a, b]_* \to \mathbb{R}^2$ such that $\mathcal{R} = \mathcal{R}_m$.*

Proof Note that, since $\mathcal{R}$ is compact (hence closed), it contains each extremity $m_i(a_i)$ or $m_i(b_i)$. Also, assume, without loss of generality, that each curve is parametrized by arc length so that $a_i = 0$ and $b_i = L_i$ (the length of m_i). Let $I^{(1)} = [0, L_1]$ and $m^{(1)} = m_1$, and define the following iterative construction.

Given the current interval $I_n = [0, \ell_n]$ and embedding $m : I_n \to \mathbb{R}^2$ such that $\mathcal{R}_m \subset \mathcal{R}$, choose an index j such that $m(\ell_n) \in \mathcal{R} \cap V_j$ and $\mathcal{R}_{m_j} \not\subset \mathcal{R}_m$. Let $\mathcal{R}_m^0 = m((0, \ell_n))$, the set $\mathcal{R}_m$ without its extremities.

Define $u_j \in (0, L_j)$ by $m_j(u_j) = m(\ell_n)$. Let A_j be the connected component of $m^{-1}(\mathcal{R}_m^0 \cap V_j)$ that contains u_j: A_j is a sub-interval of $(0, L_j)$ taking either the form (x_j, u_j), $0 < x_j < u_j$ or (u_j, y_j), $u_j < y_j < L_j$. Reorienting m_j if needed, assume that $\mathcal{R} \cap V_j = (x_j, u_j)$.

We now consider two cases.

(i) $m_j([u_j, L_j]) \cap \mathcal{R}_m = \emptyset$. Define $\ell_{n+1} = \ell_n + L_j - u_j$ and extend m to $[0, \ell_{n+1}]$ by $m(u) = m_j(u - \ell_n + u_j)$ for $u > \ell_n$. Then m is an embedding and the construction can continue.

(ii) $m_j([u_j, L_j]) \cap \mathcal{R}_m \neq \emptyset$. Let $v_j > u_j$ be the first parameter such that $m_j(v_j) \in \mathcal{R}_m$. If $m(v_j) \neq m_1(0)$, then, by construction, there exists a V_i, $i \neq j$ such that $m(v_j) \in \mathcal{R}_m \cap V_i$. This implies that m_j coincides with m_i in $m_j((u_j, v_j) \cap V_i$, but this contradicts the fact that v_i was the first point of self-intersection.
So we have $m(v_i) = m_1(0)$ and we conclude the construction with $\ell_{n+1} = \ell_n + v_j - u_j$, extending m to $[0, \ell_{n+1}]$ by $m(u) = m_j(u - \ell_n + u_j)$ for $u > \ell_n$.

Note that we always reach case (ii) (there are at most n steps). After case (ii) is completed, $\mathcal{R}_m$ is an embedded closed curve which is necessarily equal to $\mathcal{R}$, which is connected. □

1.6 Curvature

The curvature of a C^2 regular curve $m : I \to \mathbb{R}^2$ is a function $\kappa_m : I \to \mathbb{R}$, related to the arc-length derivative of the tangent through the formula:

$$\partial_{s_m} T_m = \kappa_m N_m. \tag{1.7}$$

Note that $T_m^T T_m = 1$ implies that $T_m^T \partial_{s_m} T_m = 0$ so that $\partial_{s_m} T_m$ is collinear to N_m and κ_m is the coefficient of collinearity. From the remark made at the end of the previous section, one also has

$$\kappa_m N_m = \partial_{s_m}^2 m, \tag{1.8}$$

the second derivative of the curve with respect to its arc length. This implies that

$$\kappa_m = N_m^T \partial_{s_m}^2 m = \det(T_m, \partial_{s_m}^2 m). \tag{1.9}$$

Assume that T_m can be expressed as $T_m(u) = (\cos\theta_m(u), \sin\theta_m(u))$ (so that $N_m = (-\sin\theta_m, \cos\theta_m)$) where θ is differentiable in u (we will show below that this is always true). Then, from a direct computation, $\partial_{s_m} T_m = \partial_{s_m} \theta N_m$, from which we deduce an alternative interpretation of κ_m:

$$\kappa_m(u) = \partial_{s_m} \theta_m(u), \tag{1.10}$$

where θ_m is a C^1 version of the angle between T_m and the "horizontal axis."

The same kind of easy computation yields

$$\partial_{s_m} N_m = -\kappa_m T_m \tag{1.11}$$

and Eqs. (1.7) and (1.11) together form what are called the *Frénet formulas* for the curve m.

Since it is defined as a double arc-length derivative, the curvature is parametrization invariant. Indeed, if $\tilde{m} = m \circ \psi$, then, applying Proposition 1.10 twice,

$$\partial^2_{s_{\tilde{m}}} \tilde{m} = \partial_{s_{\tilde{m}}}((\partial_{s_m} m) \circ \psi) = (\partial^2_{s_m} m) \circ \psi$$

so that $\kappa_{\tilde{m}} = \kappa_m \circ \psi$.

When $\kappa_m(u) \neq 0$, one defines the radius of curvature $\rho_m(u) = 1/|\kappa_m(u)|$ and the center of curvature $c_m(u) = m(u) + N_m(u)/\kappa_m(u)$. The circle with center $c_m(u)$ and radius $\rho_m(u)$ is called the *osculating circle* of the curve at $m(u)$.

We now prove the fact that a smooth version of the tangent angle θ exists as a consequence of the following lemma.

Lemma 1.12 *Let $I = [a, b]$ or $[a, b]_*$ and $f : I \to \mathbb{R}^2$ be a C^p function satisfying $|f(u)| = 1$ for all $u \in I$, with $p \geq 0$. Assume that for all $u \in I$, there is a small neighborhood $J_u \subset I$ and a C^p function $\tau_u : J_u \to \mathbb{R}$ such that $f(u') = (\cos \tau(u'), \sin \tau(u'))$ for $u' \in J_u$. Then there exists a C^p function $\tau : I \to \mathbb{R}$ such that $f = (\cos \tau, \sin \tau)$.*

Proof Since I is compact, we can find a finite number of $u_1, \ldots u_n$ such that $I = \bigcup_{i=1}^n J_{u_i}$. The result can then be proved by induction on n. There is nothing to prove if $n = 1$. Assume that $n > 1$ and that the result is true for $n - 1$. Then there must exist a subset J_{u_j} with $j \neq n$ such that $J_{u_j} \cap J_{u_n} \neq \emptyset$. Assume without loss of generality that $j = n - 1$. There must exist an integer k such that, for any u in this intersection, $\tau_{u_n}(u) = \tau_{u_{n-1}}(u) + 2k\pi$. Define $\tilde{J}_{u_{n-1}} = J_{u_{n-1}} \cup J_{u_n}$ and $\tilde{\tau}_{u_{n-1}}(u) = \tau_{u_{n-1}}(u)$ on $J_{u_{n-1}}$ and $\tilde{\tau}_{u_{n-1}}(u) = \tau_{u_n}(u) - 2k\pi$ on J_{u_n}, so that $\tilde{\tau}$ is C^p on $\tilde{J}_{u_{n-1}}$. Then we can apply the induction hypothesis to $J_{u_1}, \ldots, J_{u_{n-2}}, \tilde{J}_{u_{n-1}}$ with associated functions $\tau_{u_1}, \ldots, \tau_{u_{n-2}}, \tilde{\tau}_{u_{n-1}}$. □

To prove the existence of a differentiable $\theta(u)$, the lemma needs to be applied with $p = 1$, $f = T_m$, $\tau = \theta$ and $\tau_u(u') = \theta_0(u) + \arcsin(\det(T_m(u), T_m(u')))$.

1.7 Expression in Coordinates

1.7.1 Cartesian Coordinates

To provide explicit formulas for the quantities that have been defined so far, we introduce the space coordinates (x, y) and write, for a curve m: $m(u) = (x(u), y(u))$.

The first, second and higher derivatives of x will be denoted by $\dot{x}, \ddot{x}, x^{(3)}, \ldots$ and similarly for y. The tangent and the normal vector expressions in coordinates are

$$T = \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}, \quad N = \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \begin{pmatrix} -\dot{y} \\ \dot{x} \end{pmatrix}.$$

The arc length is $ds = \sqrt{\dot{x} + \dot{y}}\, du$ and the curvature is

$$\kappa = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}. \tag{1.12}$$

The last formula is proved as follows. Since $\kappa N = \partial_s T$, we have

$$\kappa = N^T \partial_s T.$$

Using $T = \partial_s m = \dot{m}/\dot{s}$, we have

$$N^T \partial_s T = \dot{s}^{-2} \ddot{m}^T N + \dot{s}^{-1} \partial \left(\dot{s}^{-1}\right) \dot{m}^T N.$$

The last term vanishes, and the first one gives (1.12) after introducing the coordinates.

1.7.2 Polar Coordinates

Let (Oxy) be a fixed frame. A point m in the plane can be characterized by its distance, r, to the origin, O, and by θ, the angle between the horizontal axis (Ox) and the half-line Om. (Notice that this is different from the angle of the tangent with the horizontal, for which we also used θ. Unfortunately, this is the standard notation in both cases.) The relation between the Cartesian coordinates (x, y) of m and its polar coordinates (r, θ) is $(x = r\cos\theta, y = r\sin\theta)$. This representation is unique, except for $m = O$, for which θ is undetermined.

A polar parametrization of a curve $u \mapsto m(u)$ is a function $u \mapsto (r(u), \theta(u))$. Often, the parameter u coincides with the angle θ and the parametrization reduces to a function $r = f(\theta)$. Some shapes have very simple polar coordinates, the simplest being a circle centered at O for which the equation is $r = \text{const}$.

Let us compute the Euclidean curvature from such a parametrization. Let $\tau = (\cos\theta, \sin\theta)$ and $\nu = (-\sin\theta, \cos\theta)$. We have $m = r\tau$, and

$$\dot{m} = \dot{r}\tau + r\dot{\theta}\nu\,,$$

$$\ddot{m} = (\ddot{r} - r\dot{\theta}^2)\tau + (2\dot{r}\dot{\theta} + r\ddot{\theta})\nu.$$

Therefore,

$$\kappa = \frac{\det[\dot{m}, \ddot{m}]}{|\dot{m}|^3} = \frac{r^2(\dot{\theta})^3 - r\ddot{r}\dot{\theta} + 2\dot{r}^2\dot{\theta} + r\dot{r}\ddot{\theta}}{(\dot{r}^2 + r^2\dot{\theta}^2)^{3/2}}.$$

When the curve is defined by $r = f(\theta)$, we have $\theta = u$, $\dot{\theta} = 1$ and $\ddot{\theta} = 0$, so that

$$\kappa = \frac{r^2 - r\ddot{r} + 2\dot{r}^2}{(\dot{r}^2 + r^2)^{3/2}}.$$

The polar representation does not have the same invariance properties as the arc length (see the next section), but still has some interesting features. Scaling by a factor λ simply corresponds to multiplying r by λ. Making a rotation with center O and angle α simply means replacing θ by $\theta + \alpha$. However, there is no simple relation for a translation. This is why a curve is generally expressed in polar coordinates with respect to a curve-dependent origin, such as its center of gravity.

1.8 Euclidean Invariance

The arc length and the curvature have a fundamental invariance property. If a curve is transformed by a rotation and translation, both quantities are invariant. The rigorous statement of this is as follows. Let R be a planar rotation and b a vector in $\mathbb{R}^2$. Define the transformation $g : \mathbb{R}^2 \to \mathbb{R}^2$ by $g(p) = Rp + b$. Then, if $m : I = [a, b] \to \mathbb{R}^2$ is a plane curve, one can define $g \cdot m : I \to \mathbb{R}^2$ by $(g \cdot m)(u) = g(m(u)) = Rm(u) + b$. Then, the statements are:

(i) $\sigma_{g\cdot m}(u) = \sigma_m(u)$, and in particular $L_{g\cdot m} = L_m = L$.
(ii) The curvatures κ_m and $\kappa_{g\cdot m}$, reparametrized over $[0, L]$ (as functions of the arc length), coincide.

The proof of (i) is straightforward from the definition of σ_m (see Eq. (1.4)). For (ii), use $\partial^2_{s_m}(g \cdot m) = R\partial^2_{s_m} m$, $N_{g\cdot m} = RN_m$ and (1.9).

Note that in this discussion we have taken $I = [a, b]$, an interval, for which the arc length reparametrization is uniquely defined by (1.4). If one wants to consider "wrapped intervals" $[a, b]_*$, arc lengths should be compared with the same inverse image of 0 (c in (1.5)).

We now state and prove the converse statement of (ii).

Theorem 1.13 (Characterization Theorem) *If two C^2 regular plane curves m and $\tilde{m}$ have the same curvature as a function of the arc length, denoted $\kappa : [0, L] \to \mathbb{R}$, then there exist R and b, and a change of parameter, ψ, such that $\tilde{m} = Rm \circ \psi + b$.*

With our notation, the assumption means that

$$\kappa = \kappa_m \circ \sigma_m^{-1} = \kappa_{\tilde{m}} \circ \sigma_{\tilde{m}}^{-1}$$

and implicitly implies that the lengths of the two curves coincide (with L).

Proof Let m^* and $\tilde{m}^*$ be m and $\tilde{m}$ reparametrized with arc length. We prove that

$$\tilde{m}^* = Rm^* + b$$

for some R and b, which implies the statement of the theorem after reparametrization. Equivalently, we assume without loss of generality that both m and $\tilde{m}$ are parametrized by arc length.

Now, let $\kappa : [0, L] \to \mathbb{R}$ be an integrable function. We build all possible curves m that are parametrized by arc length over $[0, L]$ and have κ as curvature and prove that they all differ by a rotation and translation. By definition, the angle θ_m, defined over $[0, L]$, must satisfy:

$$\dot{\theta}_m = \kappa \text{ and } \dot{m} = (\cos\theta_m, \sin\theta_m).$$

Let $\theta(s) = \int_0^s \kappa(u)du$. The first equality implies that, for some $\theta_0 \in [0, 2\pi)$, we have $\theta_m(s) = \theta(s) + \theta_0$ for all $s \in [0, L]$. The second implies that, for some $b \in \mathbb{R}^2$,

$$m(s) = \int_0^s (\cos(\theta(u) + \theta_0), \sin(\theta(u) + \theta_0))du + b.$$

Introduce the rotation $R = \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{pmatrix}$. From standard trigonometric formulas, we have

$$R\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} \cos(\theta + \theta_0) \\ \sin(\theta + \theta_0) \end{pmatrix}$$

so that, letting $\hat{m}(s) = \int_0^s (\cos\theta(u), \sin\theta(u))du$, we have $m = R\hat{m} + b$. Since $\hat{m}$ is uniquely defined by κ, we obtain the fact that m is uniquely defined up to a rotation and translation. □

1.9 The Frénet Frame

If m is a C^2 regular plane curve, its Frénet frame is defined by

$$F_m(u) = \begin{pmatrix} T_m(u) & N_m(u) \end{pmatrix}.$$

Considering T_m and N_m as column vectors, F_m is a rotation matrix satisfying $F_m^T F_m = \mathrm{Id}$ and $\det(F_m) = 1$. It is a *moving frame* along the curve.

Equations (1.7) and (1.11), which, put together, form the *Frénet formulas* for plane curves, can be summarized in matrix form as

$$\partial_{s_m} F_m = F_m S_m \tag{1.13}$$

with

$$S_m = \begin{pmatrix} 0 & -\kappa_m \\ \kappa_m & 0 \end{pmatrix}. \tag{1.14}$$

Note that, applying ∂_{s_m} to $F_m^T F_m = \mathrm{Id}$, we get

$$(\partial_{s_m} F_m)^T F_m + F_m^T \partial_{s_m} F_m = 0,$$

which states that the matrix $F_m^T \partial_{s_m} F_m$ (which is equal to S_m) must be skew-symmetric ($S_m^T = -S_m$). This implies that Eqs. (1.13) and (1.14) can be used as alternative definitions of the curvature, via the Frénet formulas.

The advantage of this construction is that it generalizes to arbitrary dimensions (cf. Sect. 3.1), and to more general forms of moving frames (like affine, or projective frames). It also leads to an alternative proof of the Characterization Theorem, as detailed below.

Proof (Alternative proof of Theorem 1.13) If one applies a rotation, R, and a translation to a curve m, the Frénet frame of the new curve, $\tilde{m}$, is $F_{\tilde{m}} = RF_m$, and using $R^T R = \mathrm{Id}$ and $\sigma_m = \sigma_{\tilde{m}}$, we have

$$S_{\tilde{m}} = F_{\tilde{m}}^T \partial_{s_{\tilde{m}}} F_{\tilde{m}} = F_m^T \partial_{s_m} F_m = S_m.$$

We therefore retrieve the fact that κ_m is invariant under rotation. The invariance by change of parameter is again a consequence of the invariance of the arc-length derivative.

We now prove the converse, assume that m and $\tilde{m}$ are such that $S_m = S_{\tilde{m}} =: S$ with both curves parametrized by arc length (as in the first proof of Theorem 1.13, it suffices to restrict to this case).

Let $G_m(s) = F_m(0)^T F_m(s)$ and $G_{\tilde{m}}(s) = F_{\tilde{m}}(0)^T F_{\tilde{m}}(s)$, so that

$$\begin{cases} \dot{G}_m = G_m S \\ \dot{G}_{\tilde{m}} = G_{\tilde{m}} S. \end{cases}$$

Both G_m and $G_{\tilde{m}}$ are therefore solutions of the differential equation $\dot{G} = GS$. We have, in addition $G_{\tilde{m}}(0) = G_m(0) = \mathrm{Id}$, and the theory of differential equations states that two functions that satisfy the same linear differential equation with the same initial condition must coincide. Thus $G_{\tilde{m}} = G_m$, which yields $F_{\tilde{m}} = RF_m$ with $R = F_{\tilde{m}}(0) F_m(0)^T$. This implies, in particular, that $T_{\tilde{m}} = RT_m$, and, since $T_m = \dot{m}_s$ for curves parametrized with arc length,

$$\tilde{m}(s) - \tilde{m}(0) = \int_0^s T_{\tilde{m}}(u)du = \int_0^s RT_m(u)du = Rm(s) - Rm(0)$$

so that $\tilde{m} = Rm + b$ with $b = \tilde{m}(0) - Rm(0)$. □

1.10 Enclosed Area and the Green (Stokes) Formula

When a closed curve m is embedded, its enclosed area can be computed with a single integral instead of a double integral. Let Ω_m be the bounded connected component of $\mathbb{R}^2 \setminus \mathcal{R}_m$. We assume that m is defined on $I = [a, b]_*$, and that *the curve is oriented so that the normal N points inward*, which means that for any $u \in [a, b]_*$, there exists an $\varepsilon > 0$ such that $m(u) + tN(u) \in \Omega_m$ for $0 < t < \varepsilon$. Since this is a convention that will be used repeatedly, we state it as a definition.

Definition 1.14 A closed regular curve oriented so that the normal points inward is said to be positively oriented.

For a circle, positive orientation corresponds to moving counter-clockwise.

We have the following proposition:

Proposition 1.15 *Using the notation above, and assuming that m is positively oriented, we have*

$$\text{Area}(\Omega_m) = \int_{\Omega_m} dx\, dy = -\frac{1}{2}\int_a^b N(u)^T m(u)\, |\dot{m}(u)|\, du. \tag{1.15}$$

Note that the last integral can also be written as $-(1/2)\int_m (N^T m)$, as defined in Sect. 1.5. We also have $N^T m = -\det(m(s), T(s))$ which provides an alternative expression. Indeed, we have

$$\begin{aligned} -(1/2)\int_m (N^T m)\, d\sigma_m &= (1/2)\int_m \det(m, T)\, d\sigma_m \\ &= (1/2)\int_a^b \det(m(u), T(u))|\dot{m}(u)|du \end{aligned}$$

so that, using $T(u) = \dot{m}(u)/|\dot{m}(u)|$,

$$\text{Area}(\Omega_m) = (1/2)\int_a^b \det(m(u), \dot{m}(u))du. \tag{1.16}$$

We will not prove Proposition 1.15, but simply remark that (1.15) is a particular case of the following important theorem.

Theorem 1.16 (Divergence theorem) *If $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a smooth function (a vector field), then*

$$\int_a^b N(u)^T f(m(u))\, |\dot{m}(u)|\, du = -\int_{\Omega_m} \text{div}\, f\, dx\, dy \tag{1.17}$$

where, letting $f(x, y) = (\alpha(x, y), \beta(x, y))$, one has $\text{div}\, f = \partial_1\alpha + \partial_2\beta$.

(Here ∂_i denotes the derivative with respect to the ith coordinate.)

Equation (1.17) is called Green's formula. To retrieve Eq. (1.15) from it, take $f(x, y) = (x, y)$, for which $\operatorname{div} f = 2$. Note that Green's formula is sometimes given with a plus sign, N being chosen as the *outward* normal.

Formula (1.15) can also be nicely interpreted as the limit of an algebraic sum of triangle areas. For this, consider a polygonal discretization, say $\tilde{m}$, of m with vertices $p_1, \ldots, p_N$. Let O be an arbitrary point in $\mathbb{R}^2$.

First consider the simple case in which the segment Op_k is included in the region $\Omega_{\tilde{m}}$ for all k (the polygonal curve is said to be star shaped with respect to O). In this case, the area enclosed by the polygon is the sum of the areas of the triangles. The area of (O, p_k, p_{k+1}) is $|\det(p_k p_{k+1}, Op_k)|/2$.[1] Assuming that the discretization is counterclockwise, which is consistent with the fact that the normal points inward, the vectors Op_k and $p_k p_{k+1}$ make an angle between 0 and π, which implies that their determinant is positive. We therefore get

$$\operatorname{Area}(\Omega_{\tilde{m}}) = \frac{1}{2} \sum_{k=1}^{N} \det[Op_k, p_k p_{k+1}]. \tag{1.18}$$

Since this can be written as $\frac{1}{2} \sum_{k=1}^{N} \det(Op_k, p_k p_{k+1}/|p_k p_{k+1}|)|p_k p_{k+1}|$, this is consistent with the continuous formula

$$\frac{1}{2} \int_a^b \det(Om(u), T(u))|\dot{m}(u)|du.$$

The interesting fact is that (1.18) is still valid for polygons which are not star shaped around the origin. In this case, the determinant may take negative values, which provides a necessary correction because, for general polygons, some triangles can intersect $\mathbb{R}^2 \setminus \Omega_m$.

Finally, we mention a classical inequality comparing the area and the perimeter of a simple closed curve.

Theorem 1.17 (Isoperimetric Inequality) *It m is a simple closed curve with perimeter L and area A, then*

$$4\pi A \leq L^2 \tag{1.19}$$

with equality if and only if m is a circle.

1.11 The Rotation Index and Winding Number

Let m be a closed, C^1, plane curve, defined on $I = [a, b]$. Express $T : [a, b] \to S^1$ (the unit circle) as a function $t \mapsto (\cos\theta(t), \sin\theta(t))$ where θ is a *continuous* function (cf. Lemma 1.12).

[1] The general expression of the area of a triangle (A, B, C) is $|\det(AB, AC)|/2$, half the area of the parallelogram formed by the two vectors.

Since m is closed, we must have $T(b) = T(a)$, which implies that $\theta(b) = \theta(a) + 2r_m\pi$, where r_m is an integer called the *rotation index* of the curve.

The rotation index is parametrization-invariant, since it is defined in terms of T, which is itself parametrization-invariant. If the curve is regular and C^2, then, taking the arc length parametrization, we find, using $\kappa = \dot{\theta}$,

$$\theta(L) - \theta(0) = \int_0^L \kappa(s)ds$$

or

$$r_m = \frac{1}{2\pi}\int_0^L \kappa(s)ds.$$

The rotation index provides an algebraic count of the number of loops in the curve: a loop is counted positively if it is parametrized counter-clockwise (normal inward), and negatively otherwise. The figure "8", for example, has a rotation index equal to 0. This also provides an alternative definition of a positively oriented curve: *a simple closed curve is positively oriented if and only if its rotation index is* $+1$.

A similar notion is the winding number of a curve. It depends on a reference point $p_0 \in \mathbb{R}^2$, and is based on the angle between $p_0m(t)/|p_0m(t)|$ and the horizontal axis, which is again assumed to be continuous in t. Denoting this angle by $\alpha_{p_0}(t)$, the winding number of m around p_0 is

$$w_{p_0}(m) = (\alpha_{p_0}(b) - \alpha_{p_0}(a))/2\pi.$$

It provides the number of times the curve loops around p_0. Again, it depends on the curve orientation.

If a curve is *simple* (i.e., it has no self-intersection), then it is intuitively obvious that it can loop only once. This is the statement of the theorem of turning tangents, which says that *the rotation index of a simple closed curve is either 1 or* -1. However, proving this statement is not so easy (even in the differentiable case we consider) – the reader may refer to [86] for a proof.

1.12 More on Curvature

There is an important relationship between positive curvature (for positively oriented curves) and convexity. One says that a simple closed curve is convex if the bounded region it outlines is convex (it contains all line segments between any two of its points). Another characterization of convexity is that the curve lies on a single side of any of its tangent lines. The relation between convexity and curvature is stated in the next theorem.

Theorem 1.18 *A positively oriented* C^2 *curve is convex if and only if its curvature is everywhere nonnegative.*

We only provide a partial justification of the only if part. Assume that m is positively oriented and that its interior, Ω_m, is convex. For a fixed arc length, s and ε small enough, we have (since m is positively oriented): $m(s) + \varepsilon N(s) \in \Omega_m$ if $\varepsilon > 0$ and $\in \overline{\Omega_m^c}$ if $\varepsilon < 0$. Now, using a second-order expansion around s, we get

$$\frac{1}{2}(m(s+h) + m(s-h)) = m(s) + \frac{h^2}{s}\kappa(s)N(s) + o(h^2)$$

and this point cannot be in Ω_m if h is small and $\kappa(s) < 0$.

The local extrema of the curvature are also of interest. They are called the vertices of the curve. The four-vertex theorem, which we also state without proof, is another classical result for plane curves [63, 212, 228].

Theorem 1.19 *Every simple closed C^2 curve has at least four vertices.*

1.13 Discrete Curves and Curvature

1.13.1 Least-Squares Approximation

Because it involves a ratio of derivatives, the numerical computation of the curvature is unstable (very sensitive to noise). We give here a brief account of how one can deal with this issue.

Assume that the curve is discretized as a finite sequence of points, say $m(1), \ldots, m(N)$. The usual finite-difference representation of derivatives are:

$$m'(k) = (m(k+1) - m(k-1))/2;$$
$$m''(k) = m(k+1) - 2m(k) + m(k-1).$$

The simplest formula for the approximate curvature is then

$$\kappa(k) = \frac{\det(m'(k), m''(k))}{|m'(k)|^3}.$$

This is however very sensitive to noise. A small variation in the position of $m(k)$ can have large consequences on the value of the estimated curvature. To be robust, curvature estimation has to include some kind of smoothing. As an example of such an approach, we describe a procedure in which one fits a curve of order 2 at each point.

Fix an approximation scale $\Delta \geq 1$, where Δ is an integer. For each k, compute three two-dimensional vectors $a(k), b(k), c(k)$ in order to have, for $-\Delta \leq l \leq \Delta$:

$$m(k+l) \simeq a(k)\frac{l^2}{2} + b(k)l + c(k).$$

Once this is done, $b(k)$ will be our approximation of the first derivative of m and $a(k)$ our approximation of the second derivative. The curvature will then be approximated by

$$\kappa(k) = \frac{\det[b(k), a(k)]}{|b(k)|^3}.$$

We will use least-squares estimation to compute a, b, c. First, build the matrix

$$A = \begin{pmatrix} \sum_{l=-\Delta}^{\Delta} \frac{l^4}{4} & 0 & \sum_{l=-\Delta}^{\Delta} \frac{l^2}{2} \\ 0 & \sum_{l=-\Delta}^{\Delta} l^2 & 0 \\ \sum_{l=-\Delta}^{\Delta} \frac{l^2}{2} & 0 & 2\Delta + 1 \end{pmatrix}$$

which is the matrix of second moments for the "variables" $l^2/2$, l and 1. They can be computed in closed form as a function of Δ, since

$$\sum_{l=-\Delta}^{\Delta} l^2 = \frac{\Delta}{3}(2\Delta^2 + 3\Delta + 1) \text{ and } \sum_{l=-\Delta}^{\Delta} l^4 = \frac{\Delta}{15}(6\Delta^4 + 15\Delta^3 + 10\Delta^2 - 1).$$

The second computation is, for all k:

$$z_0(k) = \sum_{l=-\Delta}^{\Delta} m(k+l),$$
$$z_1(k) = \sum_{l=-\Delta}^{\Delta} l m(k+l),$$
$$z_2(k) = \sum_{l=-\Delta}^{\Delta} \frac{l^2}{2} m(k+l).$$

Given this, the vectors $a(k), b(k), c(k)$ are provided by the row vectors of the matrix

$$A^{-1} \begin{pmatrix} z_2(k) \\ z_1(k) \\ z_0(k) \end{pmatrix}$$

where z_0, z_1, z_2 are also written as row vectors. As shown in Fig. 1.1, this method gives reasonable results for smooth curves. However, if the curve has sharp angles, the method will oversmooth and underestimate the curvature.

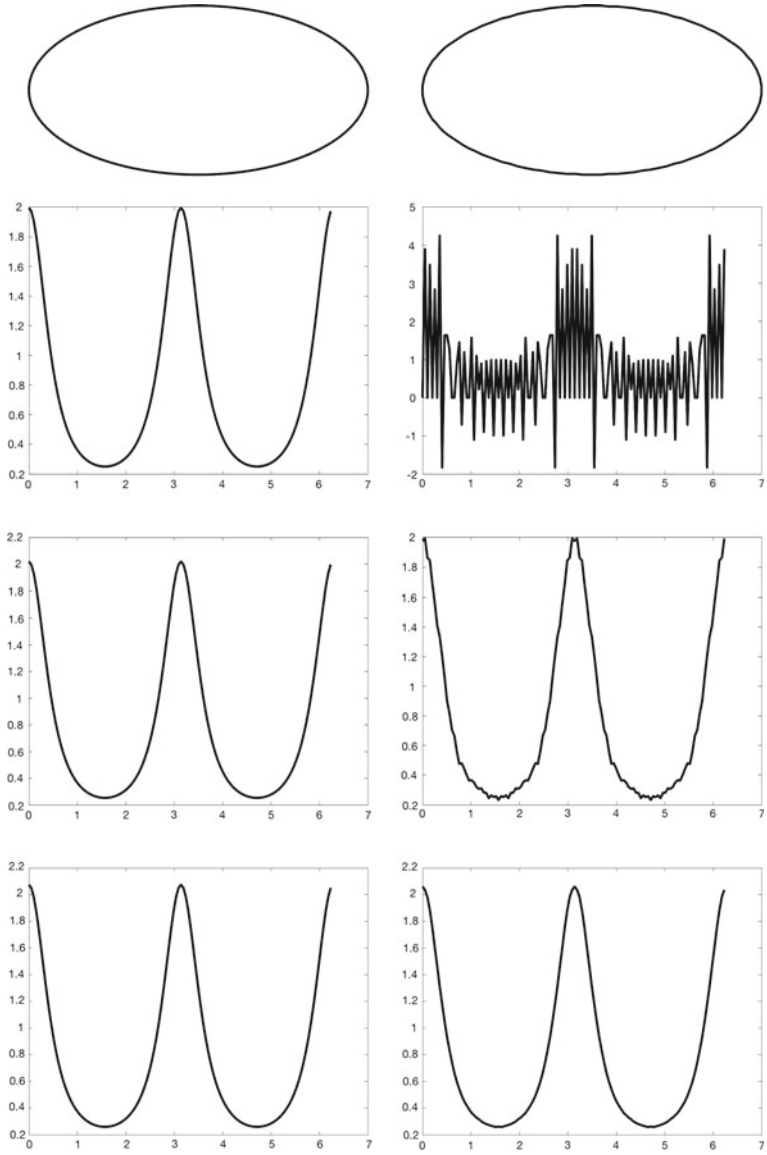

Fig. 1.1 Noise and curvature. The first curve on the left is an ellipse discretized over 125 points. The second on the right is the same ellipse, with coordinates rounded to two decimal points. The difference is almost imperceptible. However, the second row shows the result of estimating the curvature without smoothing, on the first and the second ellipse, with a very strong noise effect. The third (resp. fourth) row shows the result of the second-order approximation with $\Delta = 5$ (resp. $\Delta = 10$). The computed curvature for the truncated curve progressively improves while that of the original curve is minimally affected

1.13.2 Curvature and Distance Maps

If $A \subset \mathbb{R}^2$, one defines the distance map to A as

$$d_A(p) = \operatorname{dist}(p, \mathcal{A}) = \inf\{|p - q|, q \in \mathcal{A}\}.$$

If A is a closed set (which we will assume in the following), then for any $p \in \mathbb{R}^2$ there exists a $q \in A$ such that $d_A(p) = |p - q|$ (i.e., the infimum is a minimum). This is because any minimizing sequence q_n such that $|p - q_n| \to d_A(p)$ is necessarily bounded, and therefore has, according to the Heine–Borel theorem, a convergent subsequence, with limit $q \in A$ (because A is closed) and such that $|p - q| = d_A(p)$.

The optimal q is not always unique. For example, all points in a circle are closest to its center. The set of points $p \in \mathbb{R}^2$ for which there exists a unique $q \in A$ such that $|p - q| = d_A(p)$ will be denoted by $\mathcal{U}_A$, and we let $\pi_A : \mathcal{U}_A \to A$ be the projection, uniquely defined by $|p - \pi_A(p)| = d_A(p)$.

For $p \in \mathbb{R}^2$, we let $B(p, r) = \left\{q \in \mathbb{R}^2 : |p - q| < r\right\}$ denote the (open) disc with center p and radius r. For $q \in A$, define

$$r(A, q) = \sup\{r : B(q, r) \subset \mathcal{U}_A\}$$

and $r(A) = \inf\{r_A(q) : q \in A\}$, which is called the *reach* of A, and also has the following alternative definition.

Proposition 1.20

$$r(A) = \sup\{r : d_A(p) < r \Rightarrow p \in \mathcal{U}_A\}. \tag{1.20}$$

Proof Denote temporarily by $r'(A)$ the right-hand side of (1.20). Assume that $r \leq r'(A)$. If $q \in A$ and $p \in B(q, r)$, then $d_A(p) \leq |p - q| < r$ so that $p \in \mathcal{U}_A$ by definition of r'_A. Therefore, $B(q, r) \subset \mathcal{U}_A$ and $r \leq r(A, q)$ for all $q \in A$, which implies that $r \leq r(A)$. Taking the maximum in r, we get $r'(A) \leq r(A)$.

Assume now that $r \leq r(A)$. If $d_A(p) < r$, then $p \in B(\pi_A(p), r)$, and since $r(A) \leq r(A, \pi_A(p))$, we have $p \in \mathcal{U}_A$. This proves that $r \leq r'(A)$, and taking the maximum in r, we get $r(A) \leq r'(A)$, which concludes the proof. □

We have the following proposition.

Proposition 1.21 *The distance map is 1-Lipschitz, i.e., for all p, $p' \in \mathbb{R}^2$, one has*

$$|d_A(p) - d_A(p')| \leq |p - p'| \tag{1.21}$$

and the projection π_A is continuous on its domain.

Proof One has, for all p, $p' \in \mathbb{R}^2$ and $q \in A$,

$$d_A(p) \leq |p - q| \leq |p' - q| + |p - p'|.$$

Taking the inf of the right-hand side, we get $d_A(p) \le d_A(p') + |p - p'|$. By symmetry, we also have $d_A(p') \le d_A(p) + |p - p'|$ and (1.21) holds.

Now, take $p \in \mathcal{U}_A$ and a sequence $p_n \in \mathcal{U}_A$ such that $|p_n - p| \to 0$. Let $q_n = \pi_A(p_n)$, $q = \pi_A(p)$ and assume that there exists a subsequence of q_n (that we will still denote by q_n) and $\varepsilon > 0$ such that $|q_n - q| \ge \varepsilon$. Because $|p - q_n| \le |p - p_n| + d_A(p_n) \le 2|p - p_n| + d_A(p)$, which is bounded, q_n has a convergent subsequence (still called q_n), with limit $q' \in A$. But $|p - q'| = \lim_n |p_n - q_n| = \lim_n d_A(p_n) = d_A(p)$. Since $p \in \mathcal{U}_A$, this implies $q = q'$, a contradiction to the fact that $|q_n - q| \ge \varepsilon$ for all n. The latter condition being impossible implies that π_A is continuous. □

Proposition 1.22 *Assume that d_A is differentiable at $p \in \mathring{\mathcal{U}}_A$ (the interior of $\mathcal{U}_A$). Then, if $p \notin A$,*

$$\nabla d_A(p) = \frac{p - \pi_A(p)}{|p - \pi_A(p)|}. \tag{1.22}$$

Proof To see this, first note that, letting $q = \pi_A(p)$, one has $p_t := q + t(p - q) \in \mathcal{U}_A$ for all $t \in [0, 1]$, with $\pi_A(p_t) = q$. Indeed, if $q' \in A$, $q' \ne q$, one has $|p - q| < |p - q'| \le |p - p_t| + |p_t - q'| = (1 - t)|p - q| + |p_t - q'|$. This yields

$$|p_t - q| = t|p - q| < |p_t - q'|$$

so that $p_t \in \mathcal{U}_A$ with $q = \pi_A(p_t)$. This also implies that $d_A(p_t) = t|p - q|$ and taking the derivative with respect to t at $t = 1$, we get

$$\nabla d_A(p)^T (p - q) = |p - q|.$$

However, (1.21) implies that $|\nabla d_A(p)| \le 1$. This is only possible for $\nabla d_A(p)$ given by (1.22).

One can use the fact that the gradient of d_A is prescribed in $\mathring{\mathcal{U}}_A \setminus A$ whenever d_A is differentiable, in combination with Rademacher's theorem [107], which states that Lipschitz functions are differentiable *almost everywhere*, to prove that d_A is actually differentiable on the whole set $\mathring{\mathcal{U}}_A \setminus A$. Similarly, d_A^2 is differentiable on $\mathring{\mathcal{U}}_A$, with $\nabla(d_A^2)(p) = 2(p - \pi_A(p))$. This general fact is proved below in the special case $A = \mathcal{R}_m$, where m is a C^2, closed, regular curve with no self-intersection. Note that our definitions, so far, and Propositions 1.20–1.22 are valid for arbitrary closed sets, and in any dimension (and so is the differentiability of d_A on $\mathring{\mathcal{U}}_A \setminus A$).

We now specialize to the case $A = \mathcal{R}_m$, and we will write $d_m = d_{\mathcal{R}_m}, \mathcal{U}_m = \mathcal{U}_{\mathcal{R}_m}$, etc.

Proposition 1.23 *Let m be a simple closed C^2 regular curve. Then, we have the following statements.*

(i) If $|p - m(s)| = d_m(p)$, then $p = m(s) + tN_m(s)$ with $|t| = d_m(p)$ and $t\kappa_m(s) \le 1$.

(ii) Let

$$\rho_m = \max\left\{ \frac{2\,|(m(\tilde{s}) - m(s))^T N_m(s)|}{|m(\tilde{s}) - m(s)|^2} : s, \tilde{s} \in [0, L]_*, s \neq \tilde{s} \right\}. \qquad (1.23)$$

Then $\rho_m < \infty$ and $r(\mathcal{R}_m) \geq 1/\rho_m > 0$. In particular, $\mathring{\mathcal{U}}_m$ is not empty.

(iii) The distance map is differentiable on $\mathring{\mathcal{U}}_m$.

Proof Assume that m is parametrized by arc length over the wrapped interval $[0, L]_*$. The function $f : u \mapsto |p - m(s +_* u)|^2$ has by assumption a global minimum at $u = 0$. We therefore have $\dot{f}(0) = 0$ and $\ddot{f}(0) \geq 0$. Since $\dot{f}(0) = -2(p - m(s))^T T_m(s)$, we get the fact that $p - m(s)$ is normal to m, so that $p = m(s) + tN_m(s)$ with $|t| = d_m(p)$. We also have $\ddot{f}(0) = 2 - 2(p - m(s))^T N_m(s)\kappa_m(s) = 2(1 - t\kappa_m(s))$ yielding $t\kappa_m(s) \leq 1$. This proves (i).

We now prove that ρ_m is finite. If $m(s_n) \neq m(\tilde{s}_n)$ are such that

$$c_n := \frac{2\,|(m(\tilde{s}_n) - m(s_n))^T N_m(s_n)|}{|m(\tilde{s}_n) - m(s_n)|^2}$$

tends to infinity, then, necessarily, $m(\tilde{s}_n) - m(s_n) \to 0$. We can assume (taking subsequences if needed) that both s_n and $\tilde{s}_n$ converge, necessarily to the same limit (say s) because m is non-intersecting. Assume that $s \neq 0$ so that $s_n \in (0, L)$ for large enough n (otherwise, just reparametrize m with another starting point). We have, making a Taylor expansion,

$$m(\tilde{s}_n) = m(s_n) + (\tilde{s}_n - s_n)T_m(s_n) + \kappa(s_n)\frac{(\tilde{s}_n - s_n)^2}{2}N_m(s_n) + o((\tilde{s}_n - s_n)^2),$$

$$\left|(m(\tilde{s}_n) - m(s_n))^T N_m(s_n)\right| = |\kappa(s_n)|\frac{(\tilde{s}_n - s_n)^2}{2} + o((\tilde{s}_n - s_n)^2)$$

and

$$|m(\tilde{s}_n) - m(s_n)|^2 = (\tilde{s}_n - s_n)^2 + o((\tilde{s}_n - s_n)^2).$$

Thus $c_n \to |\kappa_m(s)|$, which is a contradiction, proving that ρ_m is finite. Note that the same limit argument also proves that $\rho_m \geq \|\kappa_m\|_\infty := \max_s |\kappa_m(s)|$.

Now, take $q \in \mathbb{R}^2$ with $d_m(q) = t < 1/\rho_m$ and assume that it has two closest points, so that there exists $s_0 \neq s_1$ such that $t = |q - m(s_0)| = |q - m(s_1)|$. Then $q = m(s_0) + t_0 N_m(s_0) = m(s_1) + t_1 N_m(s_1)$ with $|t_0| = |t_1| = t$. Moreover,

$$\begin{aligned} |m(s_1) - m(s_0)|^2 = |t_0 N_m(s_0) - t_1 N_m(s_1)|^2 &= 2t_0^2 - 2t_1 t_0 N_m(s_0)^T N_m(s_1) \\ &= 2t\,|t_0 - t_1 N_m(s_0)^T N_m(s_1)| \end{aligned}$$

and

$$|(m(s_1) - m(s_0))^T N_m(s_0)| = |t_0 - t_1 N_m(s_0)^T N_m(s_1)|.$$

We therefore get

$$\frac{2|(m(s_1)-m(s_0))^T N_m(s_0)|}{|m(s_1)-m(s_0)|^2} = \frac{1}{t}.$$

By definition, the right-hand side is less than or equal to ρ_m, which contradicts our assumption that $t < 1/\rho_m$. Therefore, $q \in \mathcal{U}_m$. This proves that $r(\mathcal{R}_m) \geq 1/\rho_m$.

Conversely, take $t < r(\mathcal{R}_m)$. By definition, we have $B(m(s), t+\varepsilon) \subset \mathcal{U}_A$ for all $s \in [0, L]_*$ and some ε such that $t + \varepsilon < r(\mathcal{R}_m)$. Therefore, letting $q_+ = m(s) + tN_m(s)$ and $q_- = m(s) - tN_m(s)$, we have $|m(\tilde{s}) - q_+| > t$ and $|m(\tilde{s}) - q_-| > t$ for all $\tilde{s} \neq s$. Developing the expression $|m(\tilde{s}) - m(s) \mp tN_m(s)| - t^2$ yields

$$|m(\tilde{s}) - m(s)|^2 \geq \pm 2t(m(\tilde{s}) - m(s))^T N_m(s)$$

so that

$$\frac{1}{t} \geq \frac{2\left|(m(\tilde{s}) - m(s))^T N_m(s)\right|}{|m(\tilde{s}) - m(s)|^2}$$

for all $s \neq \tilde{s}$, i.e., $t \leq 1\rho_m$. Taking the maximum in t implies $r(\mathcal{R}_m) \leq 1/\rho_m$.

We now prove (iii). Take $q \in \mathring{\mathcal{U}}_m$ and $m(s) = \pi_m(q)$. Write $q = m(s) + rN_m(s)$ with $|r| = d_m(q)$. Since $B(q, \varepsilon) \subset \mathring{\mathcal{U}}_m$ for ε small enough, we have $m(s) + tN(s) \in \mathring{\mathcal{U}}_m$ for $t \in (r-\varepsilon, r+\varepsilon)$, and $\pi_m(m(s) + tN(s)) = m(s)$. From (i), we get $\kappa_m(s)t \leq 1$ for $t \in (r-\varepsilon, r+\varepsilon)$, which implies $\kappa_m(s)r < 1$.

Take $\delta > 0$ such that $\kappa_m(u)t < 1$ if $|s-u| < \delta$ and $t \in (r-\varepsilon/2, r+\varepsilon/2)$. Consider the mapping $\varphi : (s-\delta, s+\delta) \times (r-\varepsilon/2, r+\varepsilon/2) \to \mathbb{R}^2$ defined by $\varphi(u,t) = m(u) + tN(u)$. Then $\partial_1\varphi(u,t) = (1 - t\kappa_m(u))T_m(u)$ and $\partial_2\varphi(u,t) = N_m(u)$ so that $\det(d\varphi) = 1 - t\kappa_m(u) \neq 0$. The inverse function theorem implies that φ (possibly restricted to a smaller open neighborhood of (s,r)) is invertible with a differentiable inverse. So, there exists a neighborhood of q in $\mathring{\mathcal{U}}_m$ such that $\varphi^{-1}(p) = (\pi_m(p), t(p))$ is differentiable with $t(p) = \pm d_m(p)$. Making sure that this neighborhood does not intersect m, we can ensure that the sign of $t(p)$ is constant so that d_m is differentiable in this neighborhood and, in particular, at q. □

Consider the mapping (a local version of which was introduced in the previous proof)

$$\begin{aligned} \varphi_m : \quad [0, L]_* \times (-r, r) &\to \quad \mathbb{R}^2 \\ (s, t) &\mapsto \quad m(s) + tN_m(s) \end{aligned}$$

for some $r < r(\mathcal{R}_m)$. As shown in the proof of Proposition 1.23, φ_m is locally invertible, but because it is also one-to-one, it provides a diffeomorphism from $[0, L]_* \times (-r, r)$ to the set

$$V_m(r) = \{q : d_m(q) < r\}.$$

Consider now the set $V_m^+(r) = \varphi_m([0, L]_* \times (0, r))$. We can write

$$\begin{aligned}\text{Area}(V_m^+(r)) &= \int_0^L \int_0^r \det d\varphi_m(s,t)\, ds\, dt \\ &= \int_0^L \int_0^r (1 - t\kappa_m)\, ds\, dt \\ &= Lr - \frac{r^2}{2}\int_0^L \kappa_m\, ds.\end{aligned}$$

The interesting conclusion is that the area is a second-degree polynomial in r. The first-degree coefficient is the curve's length and the second-degree coefficient is the integral of the curvature, i.e., the rotation index of the curve.

The formula can be localized without difficulty by restricting $V^+(m)$ to points s, t such that $s_0 < s < s_1$, the result being obviously

$$(s_1 - s_0)r - \frac{r^2}{2}\int_{s_0}^{s_1} \kappa_m\, ds = r\int_{s_0}^{s_1} (1 - \kappa_m r/2)\, ds.$$

The "infinitesimal limit" $r(1 - \kappa_m(s)r/2)ds$ provides the infinitesimal area of the set of points that are within distance r to the curve and project on $m(s)$ for some $s \in (s_0, s_1)$. This area is at first order given by the arc length times r, with a corrective term involving the curvature.

This computation is a special case of a very general construction of what are called curvature measures [106]. They can be defined for a large variety of sets, in any dimension. We will see a two-dimensional description of them when discussing surfaces.

Proposition 1.23 needs to be modified to apply to open curves. Consider such a curve, $m : [0, L] \to \mathbb{R}^2$. Then point (i) in the proposition remains true with a proper definition of a normal vector to $\mathcal{R}_m$: one says that N is a unit normal to $\mathcal{R}_m$ (or simply to m) at $m(s)$ if

$$\begin{cases} N = \pm N_m(s) & \text{if } s \in (0, L) \\ N = t_1 N_m(0) + t_2 T_m(0) & \text{if } s = 0 \\ N = t_1 N_m(L) - t_2 T_m(L) & \text{if } s = L \end{cases}$$

with $t_1^2 + t_2^2 = 1$, $t_2 > 0$. Denoting by $\mathcal{N}_m(s)$ the set of unit normals to m at $m(s)$, the first statement in (i) can be replaced by: $p = m(s) + d_m(p)N$ where $N \in \mathcal{N}(s)$. The fact that $\kappa_m(s)d_m(p) \leq 1$ holds for $s \in (0, L)$.

If one replaces the definition of ρ_m by

$$\rho_m = \max\left\{ \frac{2\,(m(\tilde{s}) - m(s))^T N}{|m(\tilde{s}) - m(s)|^2} : s, \tilde{s} \in [0, L]_*, s \neq \tilde{s}, N \in \mathcal{N}_m(s) \right\}, \qquad (1.24)$$

then (ii) remains true. Note that (1.24) boils down to (1.23) for closed curves, where $\mathcal{N}_m(s) = \{\pm N_m(s)\}$. Finally, (iii) is true.

The reader can try to prove these statements directly, or refer to [106], where these statements are proved for arbitrary closed sets, with a proper definition of the set of unit normal vectors, and without the finiteness of ρ_m, which does not hold in general. (It does not hold, for example, for polygonal curves.)

1.14 Implicit Representation

1.14.1 Introduction

Implicit representations can provide simple descriptions of relatively complex shapes and can in many cases be a good choice when designing stable shape processing algorithms. The zero level set of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is the set C_f of all $p \in \mathbb{R}^2$ such that $f(p) = 0$ (cf. Fig. 1.2). One says that f is regular if its derivative never vanishes on C_f, that is,

$$f(p) = 0 \Rightarrow \nabla f(p) \neq 0. \tag{1.25}$$

The set C_f can have several connected components, each of them being the image of a curve (level sets can therefore be used to represent multiple curves). Our first goal is to show how local properties of curves can be computed directly from the function f. We will always assume, in this chapter, that the function f tends to infinity as p tends to infinity. This implies that the zero level sets are bounded.

The implicit function theorem implies that, in a neighborhood of any regular point of f (such that $\nabla f(m) \neq 0$), the set C_f can be locally parametrized as a regular curve, for example by expressing one of the coordinates (x, y) as a function of the other. This fact and Proposition 1.11 implies that, if f is regular, each connected component of C_f can be parametrized as a regular curve. The existence of higher derivatives in f implies the same regularity for the parametrization.

Fix a connected component and assume that such a parametrization has been chosen. This results in a curve $m : I \to \mathbb{R}^2$ such that $m(0) = m_0$ and $f(m(u)) = 0$ for $u \in I$ ($\mathcal{R}_m$ coincides with the chosen connected component). From the chain rule, we have:

$$\nabla f(m)^T \partial_u m = 0.$$

This implies that $\nabla f(m)$ is normal to m.

Orientation. *We will say that f is positively oriented if $f < 0$ in the bounded connected components of $\mathbb{R}^2 \setminus C_f$ and $f > 0$ otherwise. If m is also positively oriented, then $\nabla f(m)$ points outward while the normal N to m points inward, so that $\nabla f(m) = -|\nabla f(m)|N$ (recall that (T, N) must have determinant 1, with $T = \dot{m}/|\dot{m}|$).*

Assuming positive orientation, we obtain

$$T = \frac{1}{|\nabla f|}\big(-\partial_2 f, \partial_1 f\big)$$

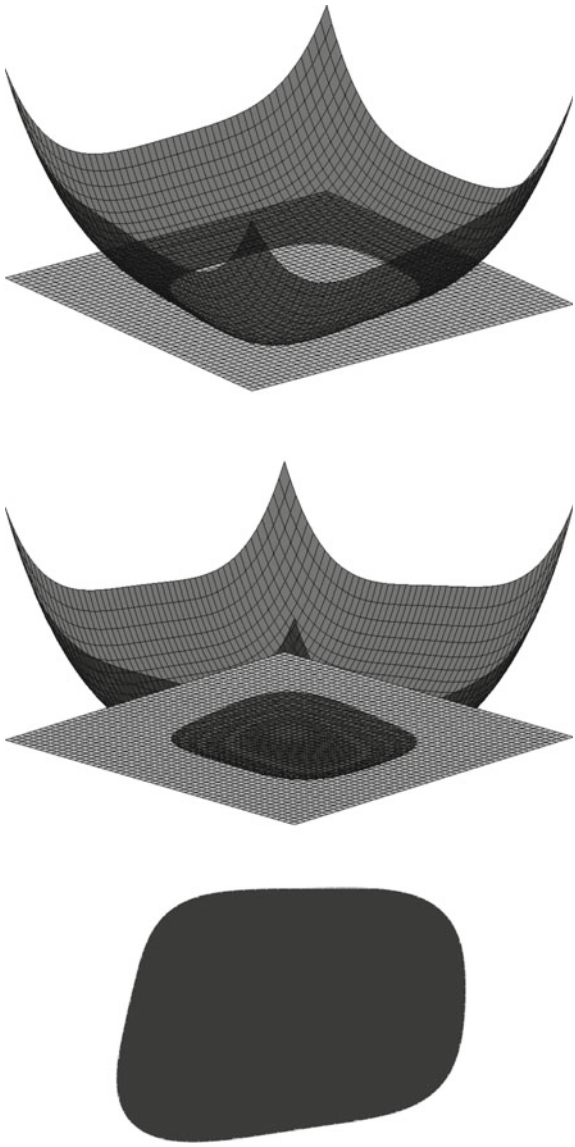

Fig. 1.2 Implicit representation: The first two panels provide two views of the graph of a function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ intersecting the plane $z = 0$. The third panel is the corresponding level set

Assume that f is twice differentiable. From the second derivative of the equation $f(m(u)) = 0$, we have

$$\dot{m}^T d^2 f(m) \dot{m} + \nabla f(m)^T \ddot{m} = 0 \,.$$

(recall that the second derivative of f is a 2 by 2 matrix).

Since $\nabla f(m) = -|\nabla f(m)|N$ and $\ddot{m}^T N = \kappa|\dot{m}|^2$, the previous equation yields (after division by $|\dot{m}|^2$),

$$T^T d^2 f(m)\, T - \kappa|\nabla f(m)| = 0\,.$$

so that

$$\kappa = \frac{T^T d^2 f\, T}{|\nabla f|} = \frac{\partial_1^2 f\, \partial_2 f^2 - 2\partial_1\partial_2 f\, \partial_1 f\, \partial_2 f + \partial_2^2 f\, \partial_1 f^2}{(\partial_1 f^2 + \partial_2 f^2)^{3/2}}.$$

This can also be written as (the computation being left to the reader)

$$\kappa = \operatorname{div} \frac{\nabla f}{|\nabla f|}. \tag{1.26}$$

1.14.2 Example: Implicit Polynomials

A large variety of shapes can be obtained by restricting the function f to be a polynomial of small degree [169], therefore involving a dependency on a small number of parameters. A polynomial in two variables and total degree less than n is given by the general formula

$$f(x, y) = \sum_{p+q \le n} a_{pq} x^p y^q\,.$$

The zero level set of f, $C_f = \{z = (x, y),\ f(x, y) = 0\}$, is called an algebraic curve. It can be a complicated object, with branches at infinity, self-intersections, or multiple loops.

The principal part of f is the homogeneous polynomial

$$g(x, y) = \sum_{k=0}^{n} a_{k,n-k} x^k y^{n-k}.$$

A sufficient condition for the compactness of C_f is that g has no non-trivial zeros, i.e., $g(x, y) = 0 \Rightarrow x = y = 0$. Adding our usual regularity condition, $f = 0 \Rightarrow \nabla f \neq 0$, ensures that C_f is a union of Jordan curves.

Figure 1.3 provides a few examples of zero level sets of implicit polynomials.

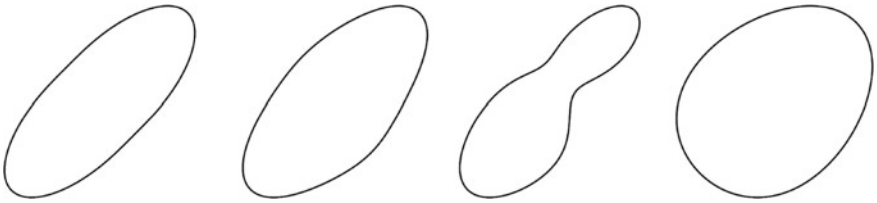

Fig. 1.3 Shapes generated by implicit polynomials of degree 4. The first curve is the level set of the polynomial $f(x, y) = x^4 + y^4 - xy - 0.1$. The other three are generated by adding a small noise to the coefficients (including zeros) of f

1.15 Invariance for Affine and Projective Transformations

Invariance, which searches for quantities that remain unchanged under certain classes of transformations, is a fundamental concept when dealing with shapes. So far, we have discussed two classes of transformations: parameter change and Euclidean motion (rotations, translations). We found in particular that Euclidean curvature was an invariant for these two classes together. We now consider additional invariants to complement these two.

We will start with transformation by scaling. This corresponds to replacing the curve m by $\tilde{m} = \lambda m$ where λ is a positive number. Visually, this corresponds to viewing the shape from a location that is closer or further away. Because of the renormalization, the unit tangent, normal and the angles θ_m are invariant. However, the length and arc length are multiplied by the constant factor λ. Finally, since the curvature is the rate of change of the angle as a function of arc length, it is divided by the same constant, $\kappa_{\tilde{m}} = \kappa_m/\lambda$.

It will also be interesting to consider invariants of affine transformations $m \mapsto Am + b$ where A is a 2 by 2 invertible matrix (a general affine transformation). Arc length and curvature are not conserved by such transformations, and there is no simple formula to compute their new value. This section describes how new quantities, which will be called affine arc length and affine curvature, can be introduced to obtain the same type of invariance.

However, a comprehensive study of the theory of differential invariants of curves [224] lies beyond the scope of this book. Here, we content ourselves with the computation in some particular cases. Although this repeats what we have already done with arc length and curvature, it will be easier to start with the simple case of rotation invariance. We know that s_m and κ_m are invariant under translation and rotation, and we now show how this can be obtained with a systematic approach that will in turn be applied to more general cases.

1.15.1 Euclidean Invariance

The generic approach to defining generalized notions of length and arc length is to look for a function Q which depends only on the derivatives of a curve at a given point, such that $Q(\dot{m}, \ddot{m}, \ldots)du$ provides the length of an element of the curve between u and $u + du$.

An arc length is then defined by

$$\sigma_m(u) = \int_0^u Q(\dot{m}, \ddot{m}, \ldots)dv.$$

The function Q will be designed to meet invariance properties. We will always require parametrization invariance, ensuring that $m = \tilde{m} \circ \varphi$ implies $\sigma_m = \sigma_{\tilde{m}} \circ \varphi$. Computing the derivative of this identity yields, in terms of Q:

$$Q(\dot{m}, \ddot{m}, \ldots) = \dot{\varphi} Q(\dot{\tilde{m}} \circ \varphi, \ddot{\tilde{m}} \circ \varphi, \ldots). \tag{1.27}$$

Now, for $m = \tilde{m} \circ \varphi$, we have

$$\begin{aligned} \dot{m} &= \dot{\varphi}\, \dot{\tilde{m}} \circ \varphi, \\ \ddot{m} &= \ddot{\varphi}\, \dot{\tilde{m}} \circ \varphi + \dot{\varphi}^2 \ddot{\tilde{m}} \circ \varphi, \end{aligned}$$

and so on for higher derivatives.

As a consequence, if Q only depends on the first derivative, we must have

$$Q(\dot{\varphi}\, \dot{\tilde{m}} \circ \varphi) = \dot{\varphi}\, Q(\dot{\tilde{m}} \circ \varphi).$$

This is true in particular when, for all $z_1 \in \mathbb{R}^2$, $\lambda_1 > 0$:

$$Q(\lambda_1 z_1) = \lambda_1 Q(z_1).$$

This is the order 1 condition for Q. It is sufficient by the discussion above, but one can show that it is also necessary. Similarly, the order 2 condition is that, for all $z_1, z_2 \in \mathbb{R}^2$, for all $\lambda_1 > 0$, $\lambda_2 \in \mathbb{R}$:

$$Q(\lambda_1 z_1, \lambda_2 z_1 + \lambda_1^2 z_2) = \lambda_1 Q(z_1, z_2).$$

This argument can be applied to any number of derivatives. The general expression (based on the *Faà di Bruno formula*) is quite heavy, and we will not need it for this discussion, but the trick for deriving new terms is quite simple. Think in terms of derivatives: the derivative of λ_k is λ_{k+1} and the derivative of z_k is $\lambda_1 z_{k+1}$; then apply the product rule. For example, the second term is the derivative of the first term, $\lambda_1 z_1$, and therefore:

$$\begin{aligned}(\lambda_1 z_1)' &= (\lambda_1)' z_1 + \lambda_1 (z_1)' \\ &= \lambda_2 z_1 + \lambda_1^2 z_2,\end{aligned}$$

which is what we found by direct computation. The constraint with three derivatives would be

$$Q(\lambda_1 z_1, \lambda_2 z_1 + \lambda_1^2 z_2, \lambda_3 z_1 + 3\lambda_2 \lambda_1 z_2 + \lambda_1^3 z_3) = \lambda_1 Q(z_1, z_2, z_3).$$

The second type of constraint which is required for Q is invariance under some class of transformations of the plane. If A is such a transformation, and $\tilde{m} = Am$, the requirement is $\sigma_{\tilde{m}} = \sigma_m$, or

$$Q(\dot{m}, \ddot{m}, \ldots) = Q(\partial(Am), \partial^2(Am), \ldots). \tag{1.28}$$

We consider affine transformations (the results will be extended to projective transformations at the end of this section). The equality is always true for translations $Am = m + b$, since Q only depends on the derivatives of m, and therefore we can assume that A is purely linear. Equality (1.28) therefore becomes: for all $z_1, z_2, \ldots \in \mathbb{R}^2$,

$$Q(z_1, z_2, \ldots) = Q(Az_1, Az_2, \ldots).$$

We now specialize to rotations. We will favor the lowest complexity for Q, and therefore first study whether a solution involving only one derivative exists. In this case, Q must satisfy: for all $\lambda_1 > 0$, for all $z_1 \in \mathbb{R}^2$ and for any rotation A,

$$Q(Az_1) = Q(z_1) \text{ and } Q(\lambda_1 z_1) = \lambda_1 Q(z_1)\,.$$

Let $e_1 = (1, 0)$ be the unit vector in the x-axis. Since one can always use a rotation to transform any vector z_1 into $|z_1|e_1$, the first condition implies that $Q(z_1) = Q(|z_1|e_1)$, which is equal to $|z_1|Q(e_1)$ from the second condition. We therefore find that $Q(z_1) = c|z_1|$ for some constant c, yielding $Q(\dot{m}) = c|\dot{m}| = c\sqrt{\dot{x}^2 + \dot{y}^2}$. We therefore retrieve the previously defined arc length up to a multiplicative constant c. The choice $c = 1$ is quite arbitrary, and corresponds to the condition that e_1 provides a unit speed: $Q(e_1) = 1$. We will refer to this σ_m as the Euclidean arc length, since we now consider other choices to obtain more invariants.

1.15.2 Scale Invariance

Let us now add scale to translation and rotation. Since it is always possible to transform any vector z_1 into e_1 with a rotation and scaling, considering only one derivative

is not enough anymore.[2] We need at least two derivatives and therefore consider z_1 and z_2 with the constraints

$$Q(Az_1, Az_2) = Q(z_1, z_2) \text{ and } Q(\lambda_1 z_1, \lambda_2 z_1 + \lambda_1^2 z_2) = \lambda_1 Q(z_1, z_2)\,.$$

Similar to rotations, the first step is to use the first condition to place z_1 and z_2 into a canonical position. Consider the combination of rotation and scaling which maps e_1 to z_1. The first column of its matrix must therefore be z_1, but, because combinations of rotation and scaling have matrices of the form $S = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, we see that, letting $z_1 = (x_1, y_1)$, the obtained matrix is

$$S_{z_1} = \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix}.$$

Now take $A = S_{z_1}^{-1}$ to obtain, from the first condition:

$$Q(z_1, z_2) = Q(e_1, S_{z_1}^{-1} z_2)\,.$$

A direct computation yields

$$S_{z_1}^{-1} z_2 = \frac{1}{x_1^2 + y_1^2} \begin{pmatrix} x_1 x_2 + y_1 y_2 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

So far, we have obtained the fact that Q must be a function F of the quantities $a = (z_1^T z_2)/|z_1|^2$ and $b = \det(z_1, z_2)/|z_1|^2$.

Now consider the second condition. The transformation $z_1 \to \lambda_1 z_1$ and $z_2 \to \lambda_1^2 z_2 + \lambda_2 z_1$ takes a to $\lambda_1 a + \lambda_2/\lambda_1$ and b to $\lambda_1 b$. Thus, if $Q(z_1, z_2) = F(a, b)$, we must have

$$F(\lambda_1 a + \lambda_2/\lambda_1, \lambda_1 b) = \lambda_1 F(a, b)$$

for all real numbers a, b, λ_2 and $\lambda_1 > 0$. Given a, b we can take $\lambda_2 = -\lambda_1^2 a$ and $\lambda_1 = 1/|b|$, at least when $b \neq 0$. This yields, for $b \neq 0$:

$$F(a, b) = |b| F(0, \text{sign}(b)).$$

For $b = 0$, we can take the same value for λ_2 to obtain $F(0, 0) = \lambda_1 F(a, 0)$ for every λ_1 and a, which is only possible if $F(a, 0) = 0$ for all a. Thus, in full generality, the function Q must take the form

[2] This would give $Q(z_1) = Q(e_1) = \text{const}$ and $Q(\lambda_1 z_1) = \lambda_1 Q(z_1) = Q(z_1)$ for all $\lambda_1 > 0$, yielding $Q = 0$.

$$Q(z_1, z_2) = \begin{cases} c_+ |\det(z_1, z_2)|/|z_1|^2 \text{ if } \det(z_1, z_2) > 0, \\ 0 \text{ if } \det(z_1, z_2) = 0, \\ c_- |\det(z_1, z_2)|/|z_1|^2 \text{ if } \det(z_1, z_2) < 0, \end{cases}$$

where c_0, c_+, c_- are positive constants. To ensure invariance by a change of orientation, however, it is natural to choose $c_+ = c_-$. Taking this value equal to 1 yields

$$Q(z_1, z_2) = |\det(z_1, z_2)|/|z_1|^2.$$

We obtain the definition of the arc length for similitudes[3]:

$$d\sigma^{sim} = \frac{|\dot{x}\ddot{y} - \ddot{x}\dot{y}|}{\dot{x}^2 + \dot{y}^2} du\,. \tag{1.29}$$

1.15.3 Special Affine Transformations

We now consider the case of area-preserving, or special affine transformations. These are affine transformations A such that $\det(A) = 1$. As before, we need two derivatives, and the first step is again to normalize $[z_1, z_2]$ using a suitably chosen matrix A. Here, the choice is natural and simple, at least when z_1 and z_2 are independent: take A to be the inverse of $[z_1, z_2]$, normalized to have determinant 1, namely

$$A = \begin{cases} \sqrt{\det(z_1, z_2)}[z_1, z_2]^{-1} \text{ if } \det(z_1, z_2) > 0, \\ \sqrt{\det(z_2, z_1)}[z_2, z_1]^{-1} \text{ if } \det(z_1, z_2) < 0. \end{cases}$$

When $\det(z_1, z_2) > 0$, this yields

$$Q(z_1, z_2) = Q(\sqrt{\det(z_1, z_2)}e_1, \sqrt{\det(z_1, z_2)}e_2)$$

so that Q must be a function F of $\sqrt{\det(z_1, z_2)}$. Applying the parametrization invariance condition, we find

$$F(\lambda_1^{3/2}\sqrt{\det(z_1, z_2)}) = \lambda_1 F(\sqrt{\det(z_1, z_2)}),$$

which implies, taking $\lambda_1 = (\det(z_1, z_2))^{-1/3}$, that

$$Q(z_1, z_2) = F(1)(\det(z_1, z_2))^{1/3}.$$

The same result is true for $\det(z_1, z_2) < 0$, yielding

[3]To complete the argument, one needs to check that the required conditions are satisfied for the obtained Q; this is indeed the case, although we skip the computation.

$$Q(z_1, z_2) = \tilde{F}(1)(\det(z_2, z_1))^{1/3}.$$

with a possibly different constant $\tilde{F}(1)$. Again, for orientation invariance, it is natural to define the area-preserving arc length by

$$d\sigma^{s.aff} = |\ddot{x}\dot{y} - \ddot{y}\dot{x}|^{1/3} du.$$

We have left aside the case $\det(z_1, z_2) = 0$. In this case, assume that $z_2 = \alpha z_1$. The second condition implies, taking $\lambda_2 = -\lambda_1^2 \alpha$:

$$\lambda_1 Q(z_1, \alpha z_1) = Q(\lambda_1 z_1, \lambda_1^2 \alpha z_1 + \lambda_2 z_1) = Q(\lambda_1 z_1, 0),$$

but we can always find an area-preserving transformation which maps $\lambda_1 z_1$ to e_1 so that $\lambda_1 Q(z_1, \alpha z_1) = Q(e_1, 0)$ is true for every $\lambda_1 > 0$ only if $Q(z_1, \alpha z_1) = 0$. This is consistent with the formula obtained for $\det(z_1, z_2) \neq 0$.

Computations are also possible for the full affine group and also for the projective group, but they require us to deal with four and more derivatives and are quite lengthy. They will be provided at the end of this section. The reader may refer to Sect. B.4 for a quick introduction to groups of linear transformations and their actions.

1.15.4 Generalized Curvature

In addition to arc length, new definitions of curvature can be adapted to more invariance constraints. One way to understand the definition is to return to the rotation case, and our original definition of curvature.

We have interpreted the curvature as the speed of rotation of the tangent with respect to arc length. Consider the matrix $P_m = [T_m, N_m]$ associated to the tangent and normal to m. Because (T_m, N_m) is an orthonormal system, this matrix is a rotation, called a moving frame [55, 104, 108, 109], along the curve. The rate of variation of this matrix is defined by

$$W_m = P_m^{-1} \partial_s P_m.$$

In the Euclidean case, it is

$$W_m = \partial_s \theta_m \begin{pmatrix} \cos\theta_m & \sin\theta_m \\ -\sin\theta_m & \cos\theta_m \end{pmatrix} \begin{pmatrix} -\sin\theta_m & -\cos\theta_m \\ \cos\theta_m & -\sin\theta_m \end{pmatrix} = \kappa_m(s) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This illustrates the moving frame method, which provides here the Euclidean curvature. It can be shown to always provide a function which is invariant under the considered transformations and change of parametrization. More precisely, we have the following definition. For a group G with associated arc length $d\sigma = Q du$, we will use the notation

$$\partial_\sigma = \partial_u / Q,$$

which generalizes the arc-length derivative defined in the Euclidean case. The following discussion concerns curves such that $Q \neq 0$, which generalizes the notion of regular curves.

Let $J_k(G)$ be the set of vectors $(z_0, z_1, \ldots, z_k) \in (\mathbb{R}^2)^{k+1}$ such that there exists a curve m such that $z_k = \partial_\sigma^k m$. That this condition induces restrictions on $z_1, \ldots, z_k$ is already clear in the case of rotations, for which one must have $|z_1| = 1$.

Definition 1.24 Let G be a group acting on $\mathbb{R}^2$ (e.g., a subgroup of $\mathrm{GL}_2(\mathbb{R})$). A G-moving frame of order k is a one-to-one function $P_0 : J_k(G) \to G$ with the following property. For all curves $m : I \to \mathbb{R}^2$ with $Q \neq 0$ on m, define $P_m : I \to G$ by

$$P_m = P_0(m, \partial_\sigma m, \ldots, \partial_\sigma^k m).$$

Then, one must have $P_{gm} = g P_m$ for all $g \in G$.

We now consider affine transformations, with group G a subgroup of $\mathrm{GL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$ (cf. Sect. B.4.3). An element of G is represented by a pair (A, b) for a linear map A and $b \in \mathbb{R}^2$. We will therefore write $P_0 = (A_0, b_0)$, $P_m = (A_m, b_m)$. We denote by G_0 the linear part of G, i.e., $(A, b) \in G \Rightarrow A \in G_0$. The invariance condition in Definition 1.24 yields, for all $U \in G_0$, $h \in \mathbb{R}^2$,

$$\begin{aligned} A_0(Uz_0 + h, Uz_1, Uz_2, \ldots, Uz_k) &= U A_0(z_0, z_1, z_2, \ldots, z_k), \\ b_0(Uz_0 + h, Uz_1, Uz_2, \ldots, Uz_k) &= U b_0(z_0, z_1, z_2, \ldots, z_k) + h. \end{aligned} \tag{1.30}$$

We have the following result, which generalizes Theorem 1.13. We here use the same notation as in Sect. B.5.

Theorem 1.25 (Moving Frame: affine case) *Let $G = G_0 \ltimes \mathbb{R}^2$ be a subgroup of $\mathrm{GL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$. If $P_0 = (A_0, b_0)$ is a G-moving frame, then, for any plane curve m*

$$\bar{W}_m = A_m^{-1} \partial_\sigma P_m = (A_m^{-1} \partial_\sigma A_m, A_m^{-1} \partial_\sigma b_m)$$

is invariant under change of parametrization and under the action of G. It moreover characterizes the curve up to the action of G: if $\bar{W}_{m^*} = \bar{W}_{\tilde{m}^*}$, where m^* and $\tilde{m}^*$ are respectively the arc-length reparametrization of m and $\tilde{m}$, then $\tilde{m} = gm \circ \psi$ for some $g \in G$ and a change of parameter ψ.

Proof Invariance by change of parametrization relies on the fact the arc length is, by construction, invariant and the details are left to the reader. If $\tilde{m} = Um + h$, then $P_{\tilde{m}} = (UA_m, Ub_m + h)$ and

$$\bar{W}_{\tilde{m}} = A_m^{-1} U^{-1} (U \partial_\sigma A_m, U \partial_\sigma b_m) = P_m^{-1} \partial_\sigma P_m = \bar{W}_m,$$

which proves G-invariance.

Conversely, assume that $\bar{W}_{\tilde{m}} = \bar{W}_m = W$, and assume, without loss of generality, that they are both parametrized by arc length. Let $g = (U, h) = P_{\tilde{m}}(0)P_m(0)^{-1}$. The proof that $\tilde{m} = gm$ for some g derives from the uniqueness theorem for ordinary differential equations (cf. Appendix C); $P_m = (A_m, b_m)$ and $P_{\tilde{m}} = (A_{\tilde{m}}, b_{\tilde{m}})$ are both solutions of the equation $\partial_\sigma(A, b) = AW$, and gP_m is another solution, as can easily be checked. Since $gP_m(0) = P_{\tilde{m}}(0)$ by definition of g, we have

$$P_0(\tilde{m}, \dot{\tilde{m}}, \ldots, \tilde{m}^{(k)}) = gP_0(\dot{m}, \ldots, m^{(k)}) = P_0(gm, U\dot{m}, \ldots, Um^{(k)}).$$

Because P_0 is assumed to be one-to-one, we have $\tilde{m} = gm$, which proves the theorem. □

For affine groups, we select a moving frame P_0 of the form $P_0(z_0, z_1, \ldots, z_k) = (A_0(z_1, \ldots, z_k), z_0)$. This implies that

$$\bar{W}_m = \left(A_m^{-1}\partial_\sigma A_m, A_m^{-1}\partial_\sigma m\right).$$

We will mainly focus on the first term, which we denote by

$$W_m = A_m^{-1}\partial_\sigma A_m.$$

The choice made for rotations corresponds to $A_0(z_1) = [z_1, Rz_1]$, R being the $(\pi/2)$-rotation. It is obviously one-to-one and satisfies the invariance requirements. The second term in $\bar{W}_m$ is constant, namely $A_m^{-1}\partial_\sigma m = (1, 0)$.

It can be shown that W_m can lead to only one, "fundamental", scalar invariant. All other coefficients are either constant, or can be deduced from this fundamental invariant. This invariant will be called the curvature associated to the group.

Consider this approach applied to similitudes. Assume that the curve is parametrized by the related arc length, σ. The frame, here, must be a similitude, A_m, and, as above, we take

$$A_m = \begin{pmatrix} \dot{x} & -\dot{y} \\ \dot{y} & \dot{x} \end{pmatrix}.$$

Define $W_m = A_m^{-1}\partial_\sigma A_m$, so that

$$\begin{aligned} W_m &= \frac{1}{\dot{x}^2 + \dot{y}^2} \begin{pmatrix} \dot{x} & \dot{y} \\ -\dot{y} & \dot{x} \end{pmatrix} \begin{pmatrix} \ddot{x} & -\ddot{y} \\ \ddot{y} & \ddot{x} \end{pmatrix} \\ &= \frac{1}{\dot{x}^2 + \dot{y}^2} \begin{pmatrix} \ddot{x}\dot{x} + \ddot{y}\dot{y} & \ddot{x}\dot{y} - \ddot{y}\dot{x} \\ -\ddot{x}\dot{y} + \ddot{y}\dot{x} & \ddot{x}\dot{x} + \ddot{y}\dot{y} \end{pmatrix}. \end{aligned}$$

When the curve is parametrized by arc length, we have

$$\frac{|\dot{x}\ddot{y} - \ddot{x}\dot{y}|}{\dot{x}^2 + \dot{y}^2} = 1$$

along the curve. Therefore

$$W_m(\sigma) = \begin{pmatrix} \frac{\ddot{x}\dot{x}+\ddot{y}\dot{y}}{\dot{x}^2+\dot{y}^2} & \mp 1 \\ \pm 1 & \frac{\ddot{x}\dot{x}+\ddot{y}\dot{y}}{\dot{x}^2+\dot{y}^2} \end{pmatrix}.$$

(σ being the similitude arc length). The computation exhibits a new quantity, which is

$$K = \frac{\ddot{x}\dot{x} + \ddot{y}\dot{y}}{\dot{x}^2 + \dot{y}^2}. \tag{1.31}$$

This is the curvature for the group of similitudes: it is invariant under translation, rotation and scaling, and characterizes curves up to similitudes.

We now consider special affine transformations (affine with determinant 1). Assume that the curve is parametrized by the corresponding arc length, σ, i.e.,

$$|\ddot{x}\dot{y} - \ddot{y}\dot{x}|^{1/3} = 1.$$

One can choose $A_m = \begin{pmatrix} \dot{x} & \ddot{x} \\ \dot{y} & \ddot{y} \end{pmatrix}$, which has determinant 1. Since $A_m(1,0)^T = \dot{m}$, the term $A_m^{-1}\dot{m}$ is trivial. We have

$$A_m^{-1}\partial A_m = \begin{pmatrix} \ddot{y} & -\ddot{x} \\ -\dot{y} & \dot{x} \end{pmatrix} \begin{pmatrix} \ddot{x} & x^{(3)} \\ \ddot{y} & y^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & \ddot{y}x^{(3)} - \ddot{x}y^{(3)} \\ 1 & -\dot{y}x^{(3)} + \dot{x}y^{(3)} \end{pmatrix}.$$

Since $\partial(\ddot{x}\dot{y} - \ddot{y}\dot{x}) = \dot{y}x^{(3)} - \dot{x}y^{(3)} = 0$, the only non-trivial coefficient is $\ddot{y}x^{(3)} - \ddot{x}y^{(3)}$, which can be taken (up to a sign change) as a definition of the *special affine curvature*:

$$K = \det(\ddot{m}, m^{(3)}). \tag{1.32}$$

Again, this is expressed as a function of the affine arc length and is invariant under the action of special affine transformations.

The local invariants with respect to rotation, similitude and the special affine group probably reach the limits of numerical feasibility, based on the number of derivatives they require. Going further involves even higher derivatives, and has only theoretical interest. However, we include here, for completeness, the definition of the affine and projective arc lengths and curvatures. This section can be safely skipped. In discussing the projective arc lengths, we will use a few notions that are related to Lie groups and manifolds. The reader can refer to Appendix B for more details.

1.15.5 *Affine Arc Length*

We first introduce new parameters which depend on the sequence $z_1, \ldots, z_n$ that describes the first derivatives. We assume that $\det(z_1, z_2) \neq 0$ and let

$$\alpha_k = \alpha_k(z_1, \ldots, z_n) = \frac{\det(z_k, z_2)}{\det(z_1, z_2)}$$
$$\text{and } \beta_k = \beta_k(z_1, \ldots, z_n) = \frac{\det(z_1, z_k)}{\det(z_1, z_2)}.$$

These are defined so that

$$z_k = \alpha_k z_1 + \beta_k z_2, \tag{1.33}$$

which also yields

$$\begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} = [z_1, z_2]^{-1} z_k.$$

In particular, we have $\alpha_1 = \beta_2 = 1$, $\alpha_2 = \beta_1 = 0$.

Assuming affine invariance, we must have

$$Q(z_1, \ldots, z_n) = Q([z_1, z_2]^{-1} z_1, \ldots, [z_1, z_2]^{-1} z_n),$$

which implies that Q must be a function of the α_k's and β_k's. We see also that we must have at least $n = 3$ to ensure a non-trivial solution. In fact, we need to go to $n = 4$, as will be shown by the following computation.

For $n = 4$, the parametric invariance constraint yields: for all $\lambda_1 > 0$, λ_2, λ_3, λ_4,

$$Q(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \tilde{z}_4) = \lambda_1 Q(z_1, z_2, z_3, z_4)$$

with $\tilde{z}_1 = \lambda_1 z_1$, $\tilde{z}_2 = \lambda_2 z_1 + \lambda_1^2 z_2$, $\tilde{z}_3 = \lambda_3 z_1 + 3\lambda_2\lambda_1 z_2 + \lambda_1^3 z_3$ and

$$\tilde{z}_4 = \lambda_4 z_1 + (3\lambda_2^2 + 4\lambda_3\lambda_1) z_2 + 6\lambda_1^2\lambda_2 z_3 + \lambda_1^4 z_4.$$

We now make specific choices for λ_1, λ_2, λ_3 and λ_4 to progressively reduce the functional form of Q. We will abuse the notation by keeping the letter Q to design the function at each step. Our starting point is $Q = Q(\alpha_3, \beta_3, \alpha_4, \beta_4)$.

(i) We start by taking $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 0$, yielding $\tilde{z}_1 = z_1$, $\tilde{z}_2 = z_2$, $\tilde{z}_3 = z_3$ and $\tilde{z}_4 = z_4 + \lambda_4 z_1$. Denote by $\tilde{\alpha}_k$, $\tilde{\beta}_k$ the α_k, β_k coefficients associated to the $\tilde{z}$'s. For the considered variation, the only coefficient that changes is α_4, which becomes $\tilde{\alpha}_4 = \alpha_4 + \lambda_4$. This implies that

$$Q(\alpha_3, \beta_3, \alpha_4, \beta_4) = Q(\alpha_3, \beta_3, \alpha_4 + \lambda_4, \beta_4).$$

Taking $\lambda_4 = -\alpha_4$, we see that Q does not depend on α_4, yielding the new functional form $Q = Q(\alpha_3, \beta_3, \beta_4)$.

(ii) Let's now consider $\lambda_1 = 1, \lambda_2 = \lambda_4 = 0$. In this case, z_1, z_2 remain unchanged, and z_3 and z_4 become $\tilde{z}_3 = z_3 + \lambda_3 z_1$, $\tilde{z}_4 = z_4 + 4\lambda_3 z_2$. This implies $\tilde{\alpha}_3 = \alpha_3 + \lambda_3$, $\tilde{\beta}_3 = \beta_3$ and $\tilde{\beta}_4 = \beta_4 + 4\lambda_3$. Taking $\lambda_3 = -\alpha_3$ yields the new functional form $Q = Q(\beta_3, \beta_4 - 4\alpha_3)$.

(iii) Now, take $\lambda_1 = 1$, $\lambda_2 = \lambda_4 = 0$, yielding $\tilde{z}_1 = z_1$, $\tilde{z}_2 = z_2 + \lambda_2 z_1$, $\tilde{z}_3 = z_3 + 3\lambda_2 z_2$ and $\tilde{z}_4 = z_4 + 6\lambda_2 z_3 + 3\lambda_2^2 z_2$, so that $\tilde{\beta}_3 = \beta_3 + 3\lambda_2$, $\tilde{\alpha}_3 = \alpha_3 - 3\lambda_2^2 - \lambda_2\beta_3$ and $\tilde{\beta}_4 = \beta_4 + 3\lambda_2^2 + 6\lambda_2\beta_3$. In particular,

$$\tilde{\beta}_4 - 3\tilde{\alpha}_3 = \beta_4 - 4\alpha_3 + 15\lambda_2^2 + 10\lambda_2\beta_3.$$

Taking $\lambda_2 = -\beta_3/3$ yields $Q = Q(\beta_4 - 4\alpha_3 - 5\beta_3^2/3)$.

(iv) Finally, take $\lambda_2 = \lambda_2 = \lambda_4 = 0$ yielding $\tilde{\beta}_3 = \lambda_1\beta_3$ $\tilde{\beta}_4 = \lambda_1^2\beta_4$ and $\tilde{\alpha}_3 = \lambda_1^2\alpha_3$. This gives

$$Q(\lambda_1^2(\beta_4 - 4\alpha_3 - 5\beta_3^2/3)) = \lambda_1 Q(\beta_4 - 4\alpha_3 - 5\beta_3^2/3).$$

Taking $\lambda_1 = 1/\sqrt{|\beta_4 - 4\alpha_3 - 5\beta_3/3|}$, assuming this expression does not vanish, yields

$$Q(\beta_4 - 4\alpha_3 - 5\beta_3^2/3) = \begin{cases} Q(1)\sqrt{|\beta_4 - 4\alpha_3 - 5\beta_3^2/3|} \text{ if } \beta_4 - 4\alpha_3 - 5\beta_3^2/3 > 0, \\ Q(-1)\sqrt{|\beta_4 - 4\alpha_3 - 5\beta_3^2/3|} \text{ if } \beta_4 - 4\alpha_3 - 5\beta_3^2/3 < 0. \end{cases}$$

Here again, it is natural to ensure an invariance by a change of orientation and let $Q(1) = Q(-1) = 1$ so that

$$Q(z_1, z_2, z_3, z_4) = \sqrt{|\beta_4 - 4\alpha_3 - 5\beta_3^2/3|}.$$

This provides the affine-invariant arc length.

We can take the formal derivative in (1.33), yielding

$$z_{k+1} = \alpha_k' z_1 + \alpha_k z_2 + \beta_k' z_2 + \beta_k z_3 = (\alpha_k' + \beta_k\alpha_3)z_1 + (\alpha_k + \beta_k' + \beta_k\beta_2)z_2,$$

so that $\alpha_{k+1} = \alpha_k' + \beta_k\alpha_3$ and $\beta_{k+1} = \beta_k' + \alpha_k + \beta_k\beta_3$. This implies that higher-order coefficients can always be expressed in terms of α_3, β_3 and their (formal) derivatives, which are represented using prime exponents. In particular, using $\beta_4 = \beta_3' + \alpha_3 + \beta_3^2$, we get

$$Q(z_1, z_2, z_3, z_4) = \sqrt{|\beta_3' - 3\alpha_3 - 2\beta_3^2/3|}. \tag{1.34}$$

Returning to parametrized curves, let $\alpha_{m,k}$ and $\beta_{m,k}$ be the coefficients α_k, β_k in which $(z_1, z_2, \ldots)$ are replaced by their corresponding derivatives $(\dot{m}_u, \ddot{m}_{uu}, \ldots)$, so that

$$m^{(k)} = \alpha_{m,k}\dot{m} + \beta_{m,k}\ddot{m}.$$

We want to express the affine arc length in terms of the Euclidean curvature. Assuming that m is parametrized by Euclidean arc length, we have $\ddot{m} = \kappa R\dot{m}$, where R is the $\pi/2$ rotation. Taking one derivative yields (using $R^2 = -\mathrm{Id}$)

$$m^{(3)} = \kappa R\ddot{m} + \dot{\kappa}R\dot{m} = -\kappa^2\dot{m} + (\dot{\kappa}/\kappa)\ddot{m}.$$

This implies that $\alpha_{m,3} = -\kappa^2$ and $\beta_{m,3} = \dot{\kappa}/\kappa$; thus, (1.34) implies that the affine arc length, σ, and the Euclidean arc length are related by

$$d\sigma = \sqrt{|\partial(\dot{\kappa}/\kappa) + 3\kappa^2 - 2(\dot{\kappa}/\kappa)^2/3|}ds.$$

1.15.6 Projective Arc Length

The problem is harder to address for the projective group (see Sect. B.4.3 for a definition) because of the non-linearity of the transformations. We keep the same notation for α_k and β_k as in the affine case (since the projective group includes the affine group, we know that the function Q will have to depend on these reduced coordinates).

Before the computation, we need to express the effects that a projective transformation has on the derivative of the curve. We still let the symbol z_k hold for the kth derivative. A projective transformation applied to a point $z \in \mathbb{R}^2$ takes the form $g : z \mapsto (Uz + b)/(w^T z + 1)$ for a 2 by 2 matrix U, and vectors $b, w \in \mathbb{R}^2$. Let $\gamma_0 = (w^T z_0 + 1)^{-1}$ so that z_0 is transformed as $\tilde{z}_0 = \gamma_0(Uz_0 + b)$. We need to express the higher derivatives $\tilde{z}_1, \tilde{z}_2, \ldots$ as functions of the initial $z_1, z_2, \ldots$ and the parameters of the transformations. Letting γ_k represent the kth derivative of γ_0, the rule for the derivation of a product (Leibniz's formula) yields

$$\tilde{z}_k = \gamma_k(Uz_0 + b) + \sum_{q=1}^{k} \binom{k}{q} \gamma_{k-q} U z_q. \tag{1.35}$$

This provides a group action, which will be denoted $\tilde{z} = g \star z$. Our goal is to find a function Q such that $Q(z_1, z_2, \ldots, z_k) = Q(\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_k)$, and which is also invariant under the transformations induced by a change of variables. It will be necessary to go to $k = 5$ for the projective group.

We first focus on projective invariance, and make an analysis equivalent to the one that allowed us to remove z_0, z_1 and z_2 in the affine case. More precisely, we

show that U, b, and w can be found such that $\tilde{z}_0 = 0$, $\tilde{z}_1 = e_1$, $\tilde{z}_2 = e_2$ and $\tilde{z}_3 = 0$, with $e_1 = (1, 0)$ and $e_2 = (0, 1)$.

First note that $\gamma_1 = -w^T z_1 \gamma_0^2$ and $\gamma_2 = -w^T z_2 \gamma_0^2 + 2(w^T z_1)^2 \gamma_0^3$. Take $b = -Uz_0$ to ensure $\tilde{z}_0 = 0$. We have $\tilde{z}_1 = \gamma_0 U z_1$, $\tilde{z}_2 = 2\gamma_1 U z_1 + \gamma_0 U z_2$ and

$$\tilde{z}_3 = 3\gamma_2 U z_1 + 3\gamma_1 U z_2 + \gamma_0 U z_3.$$

We therefore need

$$Uz_1 = e_1/\gamma_0, Uz_2 = e_2/\gamma_0 - (2\gamma_1/\gamma_0^2)e_1 = e_2/\gamma_0 + 2w^T z_1 e_1$$

and (after some algebra)

$$\begin{aligned} Uz_3 &= -3(\gamma_2/\gamma_0)Uz_1 - 3(\gamma_1/\gamma_0)Uz_2 \\ &= 3w^T z_2 e_1 + 3w^T z_1 e_2. \end{aligned}$$

Using the decomposition $z_k = \alpha_k z_1 + \beta_k z_2$, we also have $Uz_3 = \alpha_3(e_1/\gamma_0) + \beta_3(e_2/\gamma_0 - (2\gamma_1/\gamma_0^2)e_1)$, which yields the identification

$$w^T z_1 = \beta_3/(3\gamma_0) \text{ and } w^T z_2 = (3\alpha_3 + 2\beta_3^2)/9.$$

Using the definition of γ_0, this can be written as

$$\begin{cases} w^T(z_1 - \beta_3/3z_0) = \beta_3/3 \\ w^T(z_2 - (\alpha_3/3 + 2\beta_3/9)z_0) = (\alpha_3/3 + 2\beta_3/9), \end{cases}$$

which uniquely defines w, under the assumption (which we make here) that z_0, $(3/\beta_3)z_1$, $(9/(3\alpha_3 + 2\beta_3^2))z_2$ forms an affine frame. Given W, we can compute b and U. We have in particular, using the decomposition of z_k:

$$Uz_k = (\alpha_k/\gamma_0 + 2\beta_k\beta_3/(3\gamma_0))e_1 + (\beta_k/\lambda)e_2.$$

Similarly, we have

$$w^T z_k = \alpha_k\beta_3/(3\gamma_0) + \beta_k(3\alpha_3 + 2\beta_3^2)/9.$$

With this choice of U, w and b, the resulting expressions of $\tilde{z}_3$, $\tilde{z}_4$ and $\tilde{z}_5$ can be obtained. This is a heavy computation for which the use of a mathematical software is helpful; the result is that the projective invariance implies that the function Q must be a function of the following four expressions:

$$A = \alpha_4 - \frac{8}{3}\alpha_3\beta_3 - \frac{8\beta_3^3}{9} + \frac{2}{3}\beta_3\beta_4$$

$$B = \alpha_5 - \frac{10}{3}\alpha_4\beta_3 + \frac{40}{9}\alpha_3\beta_3^3 + \frac{40\beta_3^4}{27} - \frac{5}{3}\alpha_3\beta_4 - \frac{20}{9}\beta_3^2\beta_4 + \frac{2}{3}\beta_3\beta_5$$
$$C = -2\alpha_3 - \frac{4\beta_3^2}{3} + \beta_4$$
$$D = -\frac{10}{3}\alpha_3\beta_3 - \frac{5}{3}\beta_3\beta_4 + \beta_5.$$

Given this, it remains to carry out the reductions associated to the invariance by change of parameter. This is done as in the affine case, progressively selecting the coefficients λ_i to eliminate one of the expressions and modify the others, the difference being that there is one extra constraint here associated to the fifth derivative. Note that with five constraints, we would normally be short of one expression, but one of the invariances is (magically) satisfied in the reduction process, which would otherwise have required using six derivatives. We spare the reader the details, and directly provide the final expression for Q, which is

$$Q = \left| \frac{40\beta_3^3}{9} + \beta_5 - 5\beta_3(\beta_4 - 2\alpha_3) - 5\alpha_4 \right|^{1/3}.$$

As before, this can be expressed in terms of the formal derivatives of α_3 and β_3, yielding

$$Q = \left[\beta_3'' - 3\alpha_3' - 2\beta_3\beta_3' + 2\beta_3\alpha_3 + (4/9)\beta_3^3\right]^{1/3}. \tag{1.36}$$

1.15.7 Affine Curvature

We can apply the moving frame method described in Sect. 1.15.4 to obtain the affine curvature of a curve m. We assume here that m is parametrized by affine arc length, σ. A moving frame on m is immediately provided by the matrix $A_m = [\dot{m}_\sigma, \ddot{m}_{\sigma\sigma}]$, or, with our z notation, $A_0 = [z_1, z_2]$. By definition of α_3 and β_3, the matrix $W_m = A_m^{-1}\partial_\sigma A_m$ is equal to

$$W_m = \begin{pmatrix} 0 & \alpha_{m,3} \\ 1 & \beta_{m,3} \end{pmatrix}.$$

Since the curve is parametrized by affine arc length, we have $Q = 1$, where Q is given by $\sqrt{|\dot{\beta}_{m,3} - 3\alpha_{m,3} - 2\beta_{m,3}^2/3|}$. This implies that $\alpha_{m,3}$ is a function of $\beta_{m,3}$ and $\dot{\beta}_{m,3}$ along the curve; the moving frame therefore only depends on $\beta_{m,3}$ and its derivatives, which indicates that $\beta_{m,3}$ is the affine curvature. Thus, when a curve is parametrized by affine arc-length, σ, its curvature is given by

$$\kappa_m(\sigma) = \frac{\det(\dot{m}, m^{(3)})}{\det(\dot{m}, \ddot{m})}.$$

If the curve now has an arbitrary parametrization, the curvature is obtained by using $d\sigma = Q du$, where Q is given by (1.34). This yields the following expression:

$$\kappa_m(s) = \frac{1}{Q}\frac{\det(\dot{m}, m^{(3)})}{\det(\dot{m}, \ddot{m})} - \frac{3\dot{Q}}{Q}.$$

1.15.8 Projective Curvature

In the projective case, the moving frame method cannot be used exactly as described in Sect. 1.15.4, because of the non-linearity of the transformations. The moving frame is still associated to a one-to-one function $P_0(z_0, \ldots, z_k) \in G = \mathrm{PGL}_2(\mathbb{R})$. The invariance property in this case gives, with the definition of the action $z \mapsto g \star z$ given in (1.35), $P_0(g \star z) = gP_0(z)$. For Theorem 1.25 to make sense, we must use the differential of the left translation $L_g : h \mapsto hg$ on $\mathrm{PGL}_2(\mathbb{R})$, and define

$$\bar{W}_m = dL_{P_m}(\mathrm{Id})^{-1}\partial_\sigma P_m,$$

which belongs to the Lie algebra of $\mathrm{PGL}_2(\mathbb{R})$. This is the general definition of a moving frame on a Lie group [108], and coincides with the definition that has been given for affine groups, for which we had $dL_g = A$ when $g = (A, b)$.

We first need to build the matrix A_0. For this, using as before the notation (e_1, e_2) for the canonical basis of $\mathbb{R}^2$, we define a projective transformation that takes the family $\omega = (0, e_1, e_2, 0)$ to the family $z = (z_0, z_1, z_2, z_3)$, i.e., we want to determine g such that $g \star \omega = z$ (we showed that its inverse exists in Sect. 1.15.6, but we need to compute it explicitly). Since this provides eight equations for eight dimensions, one can expect that a unique such transformation exists; this will be our $A_0(z)$.

Assuming that this existence and uniqueness property is satisfied, such a construction ensures the invariance of the moving frame under the group action. Indeed, letting z be associated to a curve m and $\tilde{z}$ to $\tilde{m} = g(m)$ for some $g \in \mathrm{PGL}_2(\mathbb{R})$, we have $\tilde{z} = g \star z$. Since $A_0(z)$ is defined by $A_0(z) \star \omega = z$, the equality $A_0(\tilde{z}) \star \omega = \tilde{z}$ is achieved by $A_0(\tilde{z}) = gA_0(z)$, which is the required invariance. (Indeed, because $\star$ is a group action, we have $(gA_0(z)) \star \omega = g \star (A_0(z)\omega) = g \star z = \tilde{z}$.)

We now proceed to the computation. The first step is to obtain the expression of $g \star z$ for $z = (z_0, z_1, z_2, z_3)$. We do this in the special case in which g is given by:

$$g(m) = (Um + b)/(1 + w^T m),$$

w and b being two vectors in $\mathbb{R}^2$ and $U \in \mathrm{GL}_2(\mathbb{R})$. Define $g \star (\tilde{z}_0, \tilde{z}_1, \tilde{z}_2, \tilde{z}_3) = (z_0, z_1, z_2, z_3)$. From $(1 + w^T\tilde{z}_0)z_0 = U\tilde{z}_0 + b$, we obtain

$$\begin{cases} (1+w^T\tilde{z}_0)z_0 = U\tilde{z}_0 + b \\ (1+w^T\tilde{z}_0)z_1 + w^T\tilde{z}_1 z_0 = U\tilde{z}_1 \\ (1+w^T\tilde{z}_0)z_2 + 2w^T\tilde{z}_1 z_1 + w^T\tilde{z}_2 z_0 = U\tilde{z}_2 \\ (1+w^T\tilde{z}_0)z_3 + 3w^T\tilde{z}_1 z_2 + 3w^T\tilde{z}_2 z_1 + w^T\tilde{z}_3 z_0 = U\tilde{z}_3. \end{cases} \tag{1.37}$$

Taking $\tilde{z} = \omega$, we get

$$\begin{cases} z_0 = b \\ z_1 + w_1 z_0 = u_1 \\ z_2 + 2w_1 z_1 + w_2 z_0 = u_2 \\ z_3 + 3w_1 z_2 + 3w_2 z_1 = 0, \end{cases} \tag{1.38}$$

where $w = (w_1, w_2)$, $Ue_1 = u_1$ and $Ue_2 = u_2$. The third equation yields

$$z_3 = -3w_2 z_1 - 3w_1 z_2. \tag{1.39}$$

We will assume that z_1 and z_2 are linearly independent, so that $w = (w_1, w_2)$ is uniquely defined by this equation, and therefore $U = [u_1, u_2]$ by the middle equations of (1.38). Using again the notation $z_3 = \alpha_3 z_1 + \beta_3 z_2$, we get

$$\begin{cases} w_1 = -\beta_3/3 \\ w_2 = -\alpha_3/3. \end{cases}$$

This fully defines our moving frame $A_0(z)$.

Recall that the formal derivative of a quantity M that depends on $z_0, \ldots, z_3$ is given, in our notation, by $M' = \sum_{k=0}^{3} (\partial M/\partial z_k) z_{k+1}$. Since $b = z_0$, we have $b' = z_1$; from $u_1 = z_1 + w_1 z_0$, we get

$$u_1' = w_1 z_1 + z_2 + w_1' z_0\,,$$

and from (1.39) and $u_2 = z_2 + 2w_1 z_1 + w_2 z_0$,

$$\begin{aligned} u_2' &= z_3 + 2w_1 z_2 + (2w_1' + w_2) z_1 + w_2' z_0 \\ &= (-2w_2 + 2w_1') z_1 - w_1 z_2 + w_2' z_0. \end{aligned}$$

We have $w_1' = -\beta_3'/3$ and $w_2' = -\alpha_3'/3$, which are therefore directly computable along the curve.

By taking the representation of a projective transformation by the triplet (U, b, w), we have chosen a local chart on $\mathrm{PGL}_2(\mathbb{R})$ which obviously contains the identity represented by $(\mathrm{Id}, 0, 0)$. To be able to compute the differential of the left translation $L_{A(z)}$, we need to express the product in this chart. One way to do this efficiently is to observe that, by definition of the projective group, products in $\mathrm{PGL}_2(\mathbb{R})$ can be deduced from matrix products in $\mathrm{GL}_3(\mathbb{R})$, up to a multiplicative constant. A function g with coordinates (U, b, w) in the chart is identified (up to multiplication by a scalar) with the matrix

$$\begin{pmatrix} U & b \\ w^T & 1 \end{pmatrix}$$

and the product of $\tilde{g} = (\tilde{U}, \tilde{b}, \tilde{w})$ and $\bar{g} = (\bar{U}, \bar{b}, \bar{w})$ is therefore identified with the product of the associated matrices, which is

$$\begin{pmatrix} \bar{U} & \bar{b} \\ \bar{w}^T & 1 \end{pmatrix} \begin{pmatrix} \tilde{U} & \tilde{b} \\ \tilde{w}^T & 1 \end{pmatrix} = \begin{pmatrix} \bar{U}\tilde{U}' + \bar{b}\tilde{w}^T & \bar{U}\tilde{b} + \bar{b} \\ \bar{w}^T\tilde{U} + \tilde{w}^T & \bar{w}^T\tilde{b} + 1 \end{pmatrix},$$

which yields the chart representation for the product

$$\bar{g}\tilde{g} = \Big((\bar{U}\tilde{U} + \bar{b}\tilde{w}^T)/(1 + \bar{w}^T\tilde{b}), \\ (\bar{U}\tilde{b} + \bar{b})/(1 + \bar{w}^T\tilde{b}), (\tilde{U}^T\bar{w} + \tilde{w}')/(1 + \bar{w}^T\tilde{b})\Big).$$

To compute the differential of the left translation in local coordinates, it suffices to take $\tilde{U} = \mathrm{Id} + \varepsilon H$, $\tilde{b} = \varepsilon\beta$ and $\tilde{w} = \varepsilon\gamma$, and compute the first derivative of the product with respect to ε at $\varepsilon = 0$. This yields

$$d_{\mathrm{Id}}L_{\bar{g}}(H, \beta, \gamma) = (\bar{U}H + \bar{b}\gamma^T - \bar{w}^T\beta\bar{U}, \bar{U}\beta - \bar{w}^T\beta\bar{b}, \gamma + H^T\bar{w} - \bar{w}^T\beta\bar{w}).$$

We need to compute the inverse of this linear transformation, and therefore solve

$$\begin{cases} \bar{U}H + \bar{b}\gamma^T - \bar{w}^T\beta\bar{U} = \tilde{H} \\ \bar{U}\beta - \bar{w}^T\beta\bar{b} = \tilde{\beta} \\ \gamma + H^T\bar{w} - \bar{w}^T\beta\bar{w} = \tilde{\gamma}. \end{cases}$$

The second equation yields $\beta = (\bar{U} - \bar{b}\bar{w}^T)^{-1}\tilde{\beta}$. Substituting γ in the first by its expression in the third yields

$$\tilde{H} = (\bar{U} - \bar{b}\bar{w}^T)H + \bar{b}\tilde{\gamma}^T + (\bar{w}^T\beta)\bar{b}\bar{w}^T - \bar{w}^T\beta\bar{U}$$

so that

$$H = (\bar{U} - \bar{b}\bar{w}^T)^{-1}(\tilde{H} - \bar{b}\tilde{\gamma}^T) + (\bar{w}^T\beta)\mathrm{Id}.$$

Finally, we have

$$\gamma = \tilde{\gamma} - H^T\bar{w} + \bar{w}^T\beta\bar{w}.$$

$\bar{W}$ is obtained by applying these formulae to $\bar{g} = A(z) = (U, b, w)$ and $\tilde{H} = (\theta_1, \theta_2)$ with

$$\begin{cases} \theta_1 = u_1' = w_1z_1 + z_2 + w_1'z_0 \\ h_2' = u_2' = (2w_1' - 2w_2)z_1 - w_1z_2 + w_2'z_0 \\ \tilde{\beta} = z_1 \\ \tilde{\gamma} = w'. \end{cases}.$$

Note that, since $(A - bw^T)h = Ah - w^T hb$, the identity $z_1 = u_1 - w_1 b$ implies

$$\beta = (U - bw^T)^{-1} z_1 = e_1 .$$

Similarly, from $u_2 - w_2 b = z_2 + 2w_1 z_1$, we get

$$(U - bw^T)^{-1} z_2 = e_2 - 2w_1 e_1 .$$

We have, using $b = z_0$ and $\tilde{\gamma} = w'$,

$$\tilde{H} - b\tilde{\gamma}^T = \big(w_1 z_1 + z_2,\ (w_1' - 2w_2)z_1 - w_1 z_2\big).$$

We therefore obtain

$$h_1 = (U - bw^T)^{-1}(w_1 z_1 + z_2) + w_1 e_1 = w_1 e_1 + e_2 - 2w_1 e_1 + w_1 e_1 = e_2 ,$$

$$\begin{aligned} h_2 &= (U - bw^T)^{-1}((2w_1' - 2w_2)z_1 - w_1 z_2) + w_1 e_2 \\ &= (2w_1' - 2w_2)e_1 - w_1(e_2 - 2w_1 e_1) + w_1 e_2 \\ &= (2w_1' + 2w_1^2 - 2w_2)e_1 . \end{aligned}$$

With $c = w_1' + w_1^2 - w_2$, we have obtained $W = \begin{pmatrix} 0 & 2c \\ 1 & 0 \end{pmatrix}$. Moreover, we have

$$\gamma = w' - H^T w + w_1 w = \begin{pmatrix} w_1' - w_2 + w_1^2 \\ w_2' - 2cw_1 + w_1 w_2 \end{pmatrix}.$$

Because we assume that $\left[\beta_3'' - 3\alpha_3' - 2\beta_3\beta_3' + 2\beta_3\alpha_3 + (4/9)\beta_3^3\right]^{1/3} = 1$, we see that $w_2' = -\alpha_3'/3$ can be expressed as a function of α_3 and the derivatives of β_3 (up to the second one), while c is equal to $-(\beta_3' - \beta_3^2/3 - \alpha_3)/3$. The invariant of smallest degree can therefore be taken to be $\beta_3' - \beta_3^2/3 - \alpha_3$ (in fact, $w_2' - 2cw_1 + w_1 w_2 = -c'/6$). The projective curvature can therefore be taken as (assuming a curve parametrized by projective arc length)

$$\kappa_m(\sigma) = \partial\left(\frac{\det(\dot{m}, m^{(3)})}{\det(\dot{m}, \ddot{m})}\right) - \frac{\det(m^{(3)}, \ddot{m})}{\det(\dot{m}, \ddot{m})} + \frac{1}{3}\left(\frac{\det(\dot{m}, m^{(3)})}{\det(\dot{m}, \ddot{m})}\right)^2 .$$

The computation of the expression of the curvature for an arbitrary parametrization is left to the reader. It involves the second derivative of the arc length, and therefore the seventh derivative of the curve.

1.16 Non-local Representations

1.16.1 Semi-local Invariants

The invariants that we have defined so far depend on derivatives that can be difficult to estimate in the presence of noisy data (as seen in Fig. 1.1). Semi-local invariants attempt to address this issue by replacing derivatives by estimates depending on nearby, but not coincident, points. They provide new curve "signatures", different from the one associated to the curvature.

A general recipe for building semi-local invariants can be described as follows [48]. For a given integer, k, one needs to provide:

1. An algorithm to select k points on the curve, relative to a single point $m(u)$.
2. A formula to compute a signature based on the k selected points.

We introduce some notation. First, let S_m represent the selection of k points along m. If $p = m(u)$ is a point on m, we let $S_m(p) = (p_1, \ldots, p_k)$. Second, let F be the signature function: it takes $p_1, \ldots, p_k$ as input and returns a real number.

We need to enforce invariance at both steps of the method. Reparametrization invariance is implicitly enforced by the assumption that S_m only depends on $p = m(u)$ (and not on u). Consider now the issue of invariance with respect to a class G of affine transformations. For A in this class, we want that:

1. The point selection process "commutes": if $S_m(p) = (p_1, \ldots, p_k)$, then $S_{Am}(Ap) = (Ap_1, \ldots Ap_k)$.
2. The function F is invariant: $F(Ap_1, \ldots, Ap_k) = F(p_1, \ldots, p_k)$.

Enforcing Point 2 becomes easy if one introduces a transformation A which places the first points in $S_m(p)$ in a generic position, leading to a normalization of the function F. We clarify this operation with examples. Assume that the class of transformations being considered are translations and rotations. Then, there is a unique such transformation that displaces p_1 on O and p_2 on $|p_1 - p_2|e_1$, where e_1 is the unit vector of the horizontal axis. Denote this transformation by A_{p_1,p_2}. Then, we must have

$$\begin{aligned} F(p_1, p_2, \ldots, p_k) &= F(A_{p_1,p_2}p_1, A_{p_1,p_2}p_2, \ldots, A_{p_1,p_2}p_k) \\ &= F(0, |p_1 - p_2|e_1, A_{p_1,p_2}p_3, \ldots, A_{p_1,p_2}p_k). \end{aligned}$$

Conversely, it is clear that any function F of the form

$$F(p_1, p_2, \ldots, p_k) = \tilde{F}(|p_1 - p_2|, A_{p_1,p_2}p_3, \ldots, A_{p_1,p_2}p_k)$$

is invariant under rotation and translation. The transformation $A_{p_1p_2}$ can be made explicit: skipping the computation, this yields ((x_i, y_i) being the coordinates of p_i)

$$A_{p_1,p_2}p_j = \frac{1}{|p_2 - p_1|}\begin{pmatrix}(x_2 - x_1)(x_j - x_1) + (y_2 - y_1)(y_j - y_1)\\ (x_2 - x_1)(y_j - y_1) - (y_2 - y_1)(x_j - x_1)\end{pmatrix}.$$

Thus, with three selected points, the general form of F is

$$F(p_1, p_2, p_3) = \tilde{F}\Big(|p_2 - p_1|, \frac{(p_2 - p_1)^T(p_3 - p_1)}{|p_2 - p_1|}, \frac{\det(p_2 - p_1, p_3 - p_1)}{|p_2 - p_1|}\Big).$$

If scaling is added to the class of transformations, the same argument shows that the only choice with three points is:

$$F(p_1, p_2, p_3) = \tilde{F}\left(\frac{(p_2 - p_1)^T(p_3 - p_1)}{|p_2 - p_1|^2}, \frac{\det(p_2 - p_1, p_3 - p_1)}{|p_2 - p_1|^2}\right).$$

Similar computations hold for larger classes of transformations.

There are several possible choices for point selection (Step 1). One can use the arc length (relative to the class of transformations) that we have defined in the previous sections, and choose $p_1, \ldots, p_k$ symmetrically around p, with fixed relative arc lengths $\sigma_m(p_1) - \sigma_m(p), \ldots, \sigma_m(p_k) - \sigma_m(p)$. For example, letting $\delta_i = \sigma_m(p_i) - \sigma_m(p)$, and if $k = 2l + 1$, one can take $\delta_1 = -l\varepsilon$, $\delta_2 = -(l - 1)\varepsilon, \ldots, \delta_k = l\varepsilon$.

However, the arc length requires using curve derivatives, and this is precisely what we wanted to avoid. Some purely geometric constructions can be used instead. For rotations, for example, we can choose $p_1 = p$, and p_2 and p_3 to be the two intersections of the curve m with a circle of radius ε centered at p (taking the ones closest to p on the curves) with ε small enough. For scale and rotation, consider again circles, but instead of fixing the radius in advance, adjust it so that $|p_2 - p_3|$ becomes smaller that $1 - \varepsilon$ times the radius of the circle. This is always possible, unless the curve is a straight line.

Considering the class of special affine transformations [48], one can choose p_1, p_2, p_3, p_4 such that the line segments (p_1, p_2) and (p_3, p_4) are parallel to the tangent at p, and the areas of the triangles (p_0, p_1, p_2) and (p_0, p_3, p_4) are respectively given by ε and 2ε.

1.16.2 The Shape Context

The shape context [33] represents a shape by a collection of histograms along its outline. Here we give a presentation of this concept in the continuum and do not discuss discretization issues.

Let $s \mapsto m(s)$ be a parametrized curve, defined on some interval I. For $s, t \in I$, let $v(s, t) = m(t) - m(s)$. Fixing t, the function $s \mapsto v(s, t)$ takes values in $\mathbb{R}^2$.

Consider a density kernel, i.e, a function $K : \mathbb{R}^2 \to \mathbb{R}^2$ such that, for fixed x, $K(x, \cdot)$ is a probability density on $\mathbb{R}^2$, usually symmetric around x. The typical example is

$$K(x, y) = \exp(-|x - y|^2/(2\sigma^2))/(2\pi\sigma^2). \tag{1.40}$$

Using this kernel, let, for $s \in I$

$$f^{(m)}(s, y) = \int_I K(y, v(s, t))dt.$$

The density $f^{(m)}(s, \cdot)$ is the shape context of the curve at s and the bivariate function $f^{(m)}$ is the shape context of the whole curve. To discuss some invariance properties of this representation, we assume that the curve is parametrized by arc length (and therefore focus on translation and rotations), and that K is radial, i.e., $K(x, y)$ only depends on $|x - y|$, which is true for (1.40).

A translation applied to the curve has no effect on $v(s, t)$ and therefore leaves the shape context invariant. A rotation R transforms v into Rv, and we have $f^{(Rm)}(s, Ry) = f^{(m)}(s, y)$. The representation is not scale-invariant, but can be made so with an additional normalization (e.g., by forcing the mean distance between different points in the shape to be equal to 1, cf. [33]).

The shape context is a global representation, since it depends for any point on the totality of the curve. To some extent, however, it shares the property of local representations that small variations of the contour will have a small influence on the shape context of other points, by only slightly modifying the density $f(s, \cdot)$.

1.16.3 Conformal Welding

Conformal welding is a complex analysis operation that provides a representation of a curve by a diffeomorphism of the unit circle. While a rigorous description of the method requires advanced mathematical concepts (compared to the rest of this book), the resulting representation is interesting enough to justify the effort.

We will identify $\mathbb{R}^2$ with $\mathbb{C}$, via the usual correspondence $(x, y) \to x + iy$, and add to $\mathbb{C}$ a point at infinity that will confer the structure of a two-dimensional sphere to it. This can be done using the mapping

$$F(re^{i\theta}) = \Big(\frac{2r\cos\theta}{r^2+1}, \frac{2r\sin\theta}{r^2+1}, \frac{r^2-1}{r^2+1}\Big).$$

This mapping can be interpreted as identifying parallel circles on the sphere with zero-centered circles on the plane; zero is mapped to the south pole, the unit disc is mapped to the equator, and the representation tends to the north pole as $r \to \infty$. With this representation, the interior and the exterior of the unit disc are mapped to hemispheres and therefore play a symmetric role. We will let $\bar{\mathbb{C}}$ denote $\mathbb{C} \cup \infty$.

The complex derivative of a function is defined as the limit of $(f(z+h)-f(z))/h$ as $h \to 0$ in $\mathbb{C}$.

Two domains $\Omega_1, \Omega_2 \subset \bar{\mathbb{C}}$ are said to be conformally equivalent if there exists a function $f : \Omega_1 \to \Omega_2$ such that f is onto and one-to-one and the complex derivative $f'(z)$ exists for all $z \in \Omega_1$, with $f'(z) \neq 0$. Such a function has the property of conserving angles, in the sense that the angle made by two curves passing by z remains unchanged after a transformation by f.

The Riemann mapping theorem [249] states that any simply connected domain (i.e., any domain within which any simple closed curve can be continuously deformed into a point) is conformally equivalent to the unit disc. This domain may or may not include a point at infinity and therefore may or may not be bounded. For example, the transformation $z \mapsto 1/z$ maps the interior of the unit disc to its exterior and vice-versa. This conformal transformation is obviously unique up to any conformal mapping of the unit disc onto itself. It can be shown that the latter transformations must belong to a three-parameter family (a sub-class of the family of Möbius transformations of the plane), containing functions of the form

$$z \mapsto e^{i\alpha} \frac{z^{i\beta} + r}{rze^{i\beta} + 1} \tag{1.41}$$

with $r < 1$. We let M_1 be the set of such transformations (which forms a three-parameter group of diffeomorphisms of the unit disc). A transformation in M_1 can be decomposed into three steps: a rotation $z \mapsto ze^{i\beta}$ followed by the transformation $z \mapsto (z+r)/(zr+1)$, followed again by a rotation $z \mapsto ze^{i\alpha}$.

The Riemann mapping theorem can be applied to the interior and to the exterior of any Jordan curve γ. Letting Ω_γ represent the interior, and $\overline{\Omega}_\gamma^c$ the exterior (the notation holding for the complement of the closure of Ω_γ), and D being the open unit disc, we therefore have two conformal transformations $\Phi_- : \Omega_\gamma \to D$ and $\Phi_+ : \overline{\Omega}_\gamma^c \to D$. These two maps can be extended to the boundary of Ω_γ, i.e., the range R_γ of the curve γ, and the extension remains a homeomorphism. Restricting Φ^+ to R_γ yields a map $\varphi^+ : R_\gamma \to S^1$ (where S^1 is the unit circle) and similarly $\varphi^- : R_\gamma \to S^1$. In particular, the mapping $\varphi = \varphi^- \circ (\varphi^+)^{-1}$ is a homeomorphism of S^1 onto itself. It is almost uniquely defined by γ. In fact Φ^+ and Φ^- are both unique up to composition (on the left) by a Möbius transformation, as given by (1.41), so that φ is unique up to a Möbius transformation applied on the left or on the right. The indeterminacy on the right can be removed by the following normalization; one can constrain Φ^+, which associates two unbounded domains, to transform the point at infinity into itself, and be such that its differential at this point has a positive real part and a vanishing imaginary part. Under this constraint, φ is unique up to the left action of Möbius transformations.

In mathematical terms, we obtain a representation of (smooth) Jordan plane curves by the set of diffeomorphisms of S^1 (denoted $\mathrm{Diff}(S^1)$) modulo the Möbius transformations (denoted $PSL_2(S^1)$), writing

$$\text{2D shapes} \sim \mathrm{Diff}(S^1)/PSL_2(S^1).$$

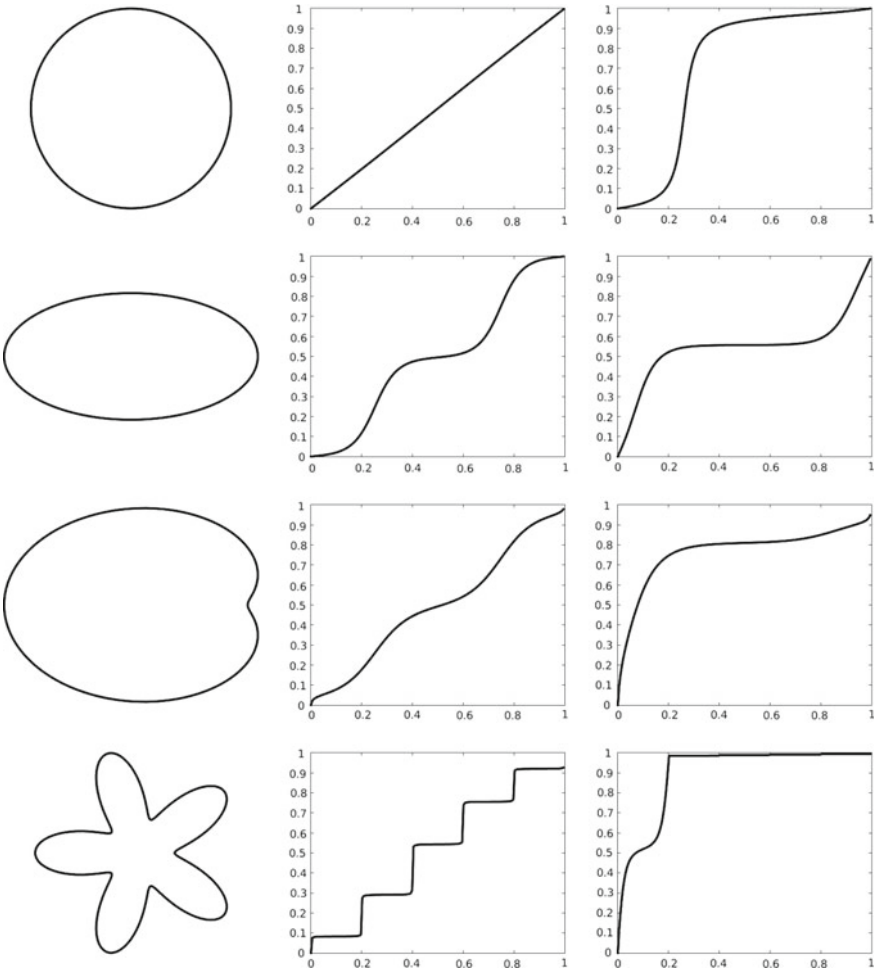

Fig. 1.4 Conformal disparity between the interior and exterior of four planar curves. *First column:* original curves; *second and third columns:* two representations of the curve signature rescaled over the unit interval, related by a Möbius transformation, illustrating the fact that these signatures are equivalent classes of diffeomorphisms of the unit disc

We now describe the two basic operations associated to this equivalence, namely computing this representation from the curve, and retrieving the curve from the representation. The first operation requires computing the trace of the conformal maps of the interior and exterior of the curve. Several algorithms are available to compute conformal maps. The plots provided in Fig. 1.4 were obtained using the Schwarz–Christoffel toolbox developed by T. Driscoll.

The solution to the second problem (going from the representation to the curves) is described in [260, 261] (Fig. 1.5). It is proved in [261] that, if φ is the mapping

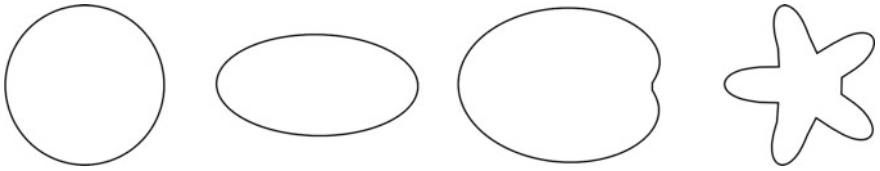

Fig. 1.5 Reconstruction of the curves in Fig. 1.4 from their signatures

above, and $\psi = \varphi^{-1}$, the corresponding shape (defined up to translation, rotation and scaling) can be parametrized as $\theta \mapsto F(\theta) \in \mathbb{C}$, $\theta \in [0, 2\pi]$, where F is the solution of the integral equation

$$K(F)(\theta) + F(\theta) = e^{i\theta},$$

where $K(F)(\theta) = \int_0^{2\pi} K(\theta, \tilde{\theta}) F(\tilde{\theta}) d\tilde{\theta}$, and the kernel K is given by

$$K(\theta, \tilde{\theta}) = \frac{i}{2} \operatorname{ctn}\left(\frac{\theta - \tilde{\theta}}{2}\right) - \frac{i}{2} \dot{\psi}(\tilde{\theta}) \operatorname{ctn}\left(\frac{\psi(\theta) - \psi(\tilde{\theta})}{2}\right)$$

which has limit $i\ddot{\psi}(\theta)/4\dot{\psi}(\theta)$ as $\theta \to \tilde{\theta}$. The inverse representation can then be computed by solving, after discretization, a linear equation in F. More precisely, assume that $((\theta_k, \varphi_k), i = 0, \ldots, N)$ is a discretization of φ (with $\varphi_N = \varphi_0 + 2\pi$ and $\theta_N = \theta_0 + 2\pi$). Following [261], one then makes the approximation

$$\begin{aligned} \int_0^{2\pi} \operatorname{ctn}\left(\frac{\varphi - \tilde{\varphi}}{2}\right) F(\tilde{\varphi}) d\tilde{\varphi} &\simeq \sum_{k=1}^{N} F_k \int_{\varphi_{k-1}}^{\varphi_k} \operatorname{ctn}\left(\frac{\varphi - \tilde{\varphi}}{2}\right) d\tilde{\varphi} \\ &= 2 \sum_{k=1}^{N} F_k \log \frac{|\sin((\varphi - \varphi_k)/2)|}{|\sin((\varphi - \varphi_{k-1})/2)|}, \end{aligned}$$

where we have set $F_k = F((\varphi_k + \varphi_{k-1})/2)$. Similarly, letting $\theta = \psi(\varphi)$,

$$\begin{aligned} \int_0^{2\pi} \operatorname{ctn}\left(\frac{\theta - \psi(\tilde{\varphi})}{2}\right) F(\tilde{\varphi}) \dot{\psi}(\tilde{\varphi}) d\tilde{\varphi} &\simeq \sum_{k=1}^{N} F_k \int_{\theta_{k-1}}^{\theta_k} \operatorname{ctn}\left(\frac{\theta - \tilde{\theta}}{2}\right) d\tilde{\theta} \\ &= 2 \sum_{k=1}^{N} F_k \log \frac{|\sin((\theta - \theta_k)/2)|}{|\sin((\theta - \theta_{k-1})/2)|}. \end{aligned}$$

Letting $\bar{\varphi}_l = (\varphi_l + \varphi_{l-1})/2$ and $\bar{\theta}_l = (\theta_l + \theta_{l-1})/2$, one obtains a discretization $((\bar{\varphi}_l, F_l), l = 1, \ldots, N)$ of F by solving the equation

$$F_l + i \sum_{k=1}^{N} K_{lk} F_k = e^{i\bar{\varphi}_l}, \quad l = 1, \ldots, N$$

with

$$K_{lk} = \log \frac{|\sin((\bar{\varphi}_l - \varphi_k)/2) \sin((\bar{\theta}_l - \theta_{k-1})/2)|}{|\sin((\bar{\varphi}_l - \varphi_{k-1})/2) \sin((\bar{\theta}_l - \theta_k)/2)|}.$$

Chapter 2
The Medial Axis

2.1 Introduction

The medial axis (or skeleton) of a shape is the set of centers of discs of maximal radii inscribed in the shape. It provides a skeleton-like structure, which, when the associated maximal radii are also stored (providing the medial axis transform [37]) uniquely encodes the shape geometry.

To be specific, represent a shape by an open connected bounded set in the plane, denoted Ω. Let $B(p, r)$ denote the open disc of center $p \in \mathbb{R}^2$ and radius $r > 0$. One says that such a disc is maximal in Ω if and only if it is included in Ω, and no disc in which it is (strictly) contained is included in Ω. The skeleton of Ω, denoted $\Sigma(\Omega)$, is the set of all p such that $B(p, r)$ is maximal in Ω for some $r > 0$, i.e., $\Sigma(\Omega)$ is the set of loci of the centers of maximal discs. We shall also denote by $\Sigma^*(\Omega)$ the set of pairs (p, r) such that $B(p, r)$ is maximal. This is the medial axis transform (MAT). We have the following proposition.

Proposition 2.1 *The medial axis transform, $\Sigma^*(\Omega)$, uniquely characterizes Ω.*

Proof Let

$$\tilde{\Omega} = \bigcup_{(p,r)\in\Sigma^*(\Omega)} B(p, r).$$

By definition of Σ^*, we have $\tilde{\Omega} \subset \Omega$ and we want to prove the reverse inclusion (therefore proving that $\Sigma^*(\Omega)$ characterizes Ω).

For $x \in \Omega$, let

$$r_x = \text{dist}(x, \Omega^c) = \min\{d(x, y), y \notin \Omega\}.$$

One has $B(x, r_x) \subset \Omega$. Define

$$G_x = \left\{y \in \Omega : B(y, r_y) \supset B(x, r_x)\right\}.$$

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_2

Let $r_x^* = \sup\{r_y : y \in G_x\}$. By definition, there exists a sequence (y_n) such that $r_{y_n} \to r_x^*$, and, because Ω is bounded, we can assume (replacing y_n by a subsequence if needed) that $y_n \to y^* \in \overline{\Omega}$. Obviously, y^* cannot belong to $\partial\Omega$ because this would imply $r_{y_n} \to 0$ while $r_x^* \geq r_x > 0$ (since $x \in G_x$). Also, because $B(y_n, r_{y_n}) \subset \Omega$, we have at the limit $B(y^*, r_x^*) \subset \overline{\Omega}$, which implies $B(y^*, r_x^*) \subset \Omega$ because Ω is open. Similarly, passing to the limit in the inclusion $B(x, r_x) \subset B(y_n, r_{y_n})$ implies $B(x, r_x) \subset B(y^*, r_x^*)$.

We now show that $B(y^*, r_x^*)$ is maximal. If $B(y^*, r_x^*)$ is included in some ball $B(y, r) \in \Omega$, it will be *a fortiori* included in $B(y, r_y)$ and since $B(x, r_x) \subset B(y^*, r_x^*)$, we see that y must be in G_x with $r_y > r_x^*$, which is a contradiction.

We have therefore proved that every $x \in \Omega$ belongs to a maximal disk, therefore proving that $\Omega \subset \tilde{\Omega}$.

2.2 The Structure of the Medial Axis

We assume that Ω is the interior of a piecewise smooth Jordan curve. Some structural properties of the skeleton can be obtained under some assumptions on the regularity of the curve [65]. The assumption is that the smooth arcs are analytic[1] everywhere except at a finite number of points; for these exceptional points, it is required that m has both left and right tangents. The simplest example of a curve satisfying this assumption is a polygon.

For such a curve, it can be shown that all but a finite number of points in the skeleton are such that the maximal disc $B(p, r)$ meets the curve m at exactly two points. Such points on the skeleton are called regular. Non-regular points separate into three categories.

The first one is when the maximal disc, $B(m, r)$, meets the curve in more than two connected regions. Such points are *bifurcation points* of the skeleton. The second possibility is when there is only one connected component; then, there are two possibilities: either m is the center of an osculating circle to the curve, or there exists a concave angle at the intersection of the curve and the maximal disc. The third possibility is when there are two connected components, but one of them is a sub-arc of the curve. This happens only when the curve has circular arcs.

The skeleton itself is connected, and it is composed of a finite number of smooth curves.

2.3 The Skeleton of a Polygon

Consider a closed polygon, without self-intersections. Denote its vertices by $m_1, \ldots, m_N, m_{N+1} = m_1$. Let s_i denote the ith edge, represented by the open line segment (m_i, m_{i+1}), with $i = 1, \ldots, N$. A maximal disc $B(m, r)$ within the polygon

[1] A function m is analytic at t if it is infinitely differentiable at this point, and equal to the limit of its infinite Taylor series.

has to meet the boundary at at least two points. We separate the cases depending on whether these points are on edges or vertices.

Assume that $B(m, r)$ is tangent to s_i at some point $p \in s_i$. Let $T_i = (m_{i+1} - m_i)/|m_{i+1} - m_i|$ be the unit tangent to s_i and N_i the unit normal. We assume that the orientation is such that N_i points inward. We must have

$$p = m - rN_i \text{ and } p = m_i + tT_i$$

for some $t \in (0, |m_{i+1} - m_i|)$. Taking the dot product of both equations with T_i and computing the difference yields

$$t = (m - m_i)^T T_i.$$

We therefore obtain the fact that $B(m, r)$ is tangent to s_i if and only if

$$m - rN_i = m_i + ((m - m_i)^T T_i)T_i \qquad \text{with } 0 \leq (m - m_i)^T T_i \leq |m_{i+1} - m_i|.$$

We can distinguish three types of maximal discs:

1. Bitangents: there exists $i \neq j$ with

$$m = m_i + ((m - m_i)^T T_i)T_i + rN_i = m_j + ((m - m_j)^T T_j)T_j + rN_j \text{ and}$$
$$0 \leq (m - m_i)^T T_i \leq |m_{i+1} - m_i|, 0 \leq (m - m_j)^T T_j \leq |m_{j+1} - m_j|.$$

2. Discs that meet the boundary at exactly one edge and one vertex: there exists $i \neq j$ such that

$$\begin{aligned} m \quad &= \quad m_i + ((m - m_i)^T T_i)T_i + rN_i, \\ &\qquad 0 \leq ((m - m_i)^T)T_i \leq |m_{i+1} - m_i| \\ &\text{and} \quad |m - m_j| = r. \end{aligned}$$

3. Discs that meet the boundary at two vertices: there exists $i \neq j$ such that $|m - m_i| = |m - m_j| = r$.

Note that a maximal ball can meet a vertex only if this vertex points inward (concave vertex). In particular, with convex polygons, only the first case can happen.

An interesting consequence of this result is that the skeleton of a polygon is the union of line segments and parabolic arcs. To see this, consider the equations for the three previous cases. For bitangents, we have

$$r = (m - m_i)^T N_i = (m - m_j)^T N_j$$

which implies

$$(m - m_i)^T (N_j - N_i) = (m_j - m_i)^T N_j.$$

If $N_i \neq N_j$, this is the equation of a line orthogonal to $N_i - N_j$. The case $N_i = N_j$ can never occur because the normals have to point to the interior of maximal balls and therefore coincide only if $s_i = s_j$.

For the second case, we have

$$m - m_i = ((m - m_i)^T)T_i T_i + |m - m_j|N_i,$$

which yields

$$(m - m_i)^T N_i = |m - m_j|.$$

This is the equation of a parabola. To see why, express m as $m = m_i + \alpha T_i + \beta N_i$. The previous equations yield $\beta \geq 0$ and

$$\beta^2 = (\alpha - (m_j - m_i)^T T_i)^2 + (\beta - (m_j - m_i)^T N_i)^2$$

or

$$2(m_j - m_i)^T N_i \beta = (\alpha - (m_j - m_i)^T T_i)^2 + ((m_j - m_i)^T N_i)^2.$$

Finally, in the last case, the skeleton coincides with the line of points which are equidistant from the two vertices. We have therefore proved the following fact (which comes in addition to the properties discussed in Sect. 2.2).

Proposition 2.2 *The skeleton of a polygonal curve is a union of line segments and parabolic arcs. For a convex polygon, the skeleton only contains line segments.*

2.4 Voronoï Diagrams

2.4.1 Voronoï Diagrams of Families of Closed Sets

The previous computation and the most efficient algorithms to compute skeletons are related by the theory of Voronoï diagrams. We start with their definition:

Definition 2.3 Let $F_1, \ldots F_N$ be closed subsets of $\mathbb{R}^2$. The associated Voronoï cells are the sets $\Omega_1, \ldots, \Omega_N$ defined by

$$x \in \Omega_i \Leftrightarrow d(x, F_i) < \min_{j \neq i} d(x, F_j).$$

The union of the boundaries, $\bigcup_{i=1}^N \partial\Omega_i$, forms the Voronoï diagram associated to $F_1, \ldots, F_N$.

In the case of a polygonal curve, the skeleton is included in the Voronoï diagram of the closed line segments that form the curve. Indeed, a maximal disc has to meet at least two segments (sometimes at their common vertices), and is at a strictly larger

distance from the segments it does not intersect. It therefore belongs to the boundary of the cells. The converse is false: a point from the diagram is not necessarily in the skeleton (some points may correspond to external disks).

There exist very efficient algorithms to compute these diagrams. We shall not detail them here, but references can be found in [222, 233].

The notion of Voronoï diagrams for a polygon can be extended to a general curve. The question is to find sub-arcs $F_1, \ldots, F_N$ of the curve with the property that their diagram contains the curve's skeleton. What we have said concerning polygons applies, except in one case: when a maximal disc meets an arc at two distinct points. This could not happen with straight lines, and a condition ensuring that this does not happen for a given arc is as follows [170]. Recall that a vertex of a smooth curve m is a local extremum of the curvature.

Theorem 2.4 *A sub-arc of a C^2 closed curve which has two points belonging to a maximal disc necessarily contains a vertex.*

Therefore, it suffices to cut the curve at vertices to be sure that the obtained arcs cannot hold two contacts with maximal discs.

2.4.2 Voronoï Diagrams of Discretized Boundaries

The medial axis can, in some sense, also be understood as deriving from the Voronoï diagram of the infinite family of points in the boundary of Ω. The Voronoï cell associated to a point $x \in \partial\Omega$ is the set

$$V_x = \left\{ y \in \Omega : |x - y| < |x' - y| \text{ for all } x' \in \partial\Omega \setminus \{x\} \right\}.$$

(The set of points y such that x is closer to y than any other point on the boundary.) When the boundary is smooth, V_x is normal to the boundary and extends to a point y_0 that belongs to the medial axis.

Consider now a finite family of points sampling the boundary, i.e., a finite subset $F \subset \partial\Omega$. Let $d = \max(d(x, F), x \in \partial\Omega)$ measure the density of F in $\partial\Omega$. The cells of the Voronoï diagram of $\{\{x\}, x \in F\}$ are possibly unbounded polygonal domains, and those among their edges (boundaries) that are included in Ω will provide a good approximation of the skeleton of Ω (see Fig. 2.1). One way to understand this is through the relationship between Voronoï diagrams and Delaunay triangulations. A triangulation of the finite set F is a family of triangles with vertices in F covering the convex hull of F (the smallest convex set containing F) such that the intersection of any two triangles is either empty, or a common vertex, or a common edge to the two triangles. Such a triangulation is a Delaunay triangulation if in addition, the circumcenters of any triangle contain no other point in F. Because of this property, these circumcenters, which are equidistant to three vertices, and closer to them than to any other, form vertices of the Voronoï diagram of F. Line segments joining the

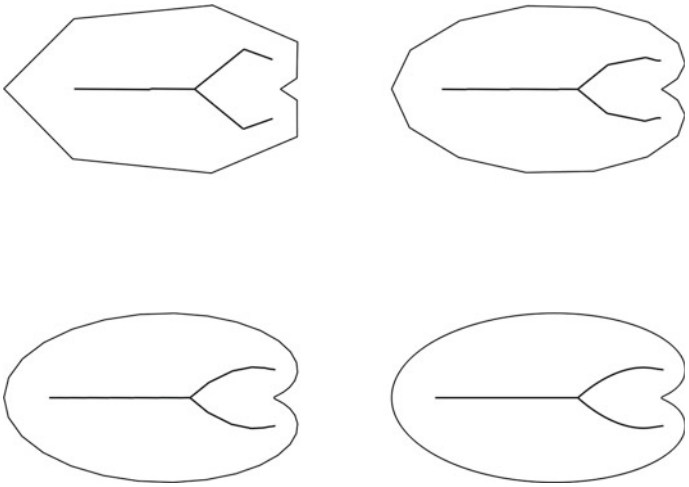

Fig. 2.1 Comparison of medial axes computed using Voronoï diagrams with different degrees of discretization

circumcenters of two adjacent triangles (sharing an edge) form the edges of the diagram. For the same reason, assuming that d is small enough, the circumcircles can be seen as approximations of maximal balls with three contact points in Ω, and those of them that belong to Ω therefore provide approximations of vertices of the medial axis, the edges that connect them in the Voronoï diagram providing approximate lines in the skeleton. Note that some of the circumcenters may fall outside of Ω (and must therefore be excluded from the medial axis) and that the Voronoï diagram of F contains, in addition, half-lines stemming from centers of boundary triangles and through midpoints of segments of the boundary, which must also be excluded. There is an extensive literature on the computation of Delaunay triangulation and Voronoï diagrams, which constitute fundamental algorithms in computational geometry [39, 226, 233], and numerous implementations in various softwares.

2.5 Thinning

There exist other approaches aiming at defining skeletal structures, which slightly differ from the medial axis that we have described so far (sometimes called the Blum medial axis, because it was introduced by Blum in [37]). Thinning algorithms, in particular, create their own kind of skeleton which does not necessarily correspond to centers of maximal discs. They are, however, quite efficient and generally easy to implement. The principle is to progressively "peel off" the boundary of the region until only a skeletal structure remains. One of the first methods, defined for discrete binary images, is the Hilditch algorithm [147], in which a sequence of simple tests are performed to decide whether a pixel must be removed or not from the region.

A more formal definition of thinning, which is briefly described below, can be based on the erosion operation in mathematical morphology [254].

Define a structuring element B to be a symmetric subset of $\mathbb{R}^2$ (for example a small disc centered at 0). Using B, we define a sequence of operators that apply to a set X and create a new set:

$$\begin{aligned} E_B(X) &= \{x : x + B \subset X\} \text{ (erosion)}, \\ D_B(X) &= \{x : (x + B) \cap X \neq \emptyset\} \text{ (dilation)}, \\ O_B(X) &= D_B \circ E_B(X) \text{ (opening)}, \\ L_B(X) &= X \setminus O_B(X). \end{aligned}$$

Erosion is like peeling X with a knife shaped like B. Dilation spreads matter around X, adding around each point some material once again shaped like B. Opening is an erosion followed by a dilation, which essentially puts back what the erosion has removed, except the small structures that have completely been removed and cannot be recovered (since there is nothing left to spread on). The last operation, L_B, precisely collects these lost structures (called linear parts), and is the basic operator for the morphological skeleton which is defined by

$$S(X) = \bigcup_{n=1}^{\infty} L_B(E_{nB}(X)).$$

This is the union of the linear parts of X after successive erosions. Note that, for bounded X, this union is actually finite since $E_{nB}(X) = \emptyset$ for large enough n.

2.6 Sensitivity to Noise

One of the main issues with the medial axis transform is its lack of robustness to noise. Figure 2.2 provides an example of how small variations at the boundary of a shape can result in dramatic changes in the skeleton. In fact, we have seen in our discussion of polygons that the addition of a convex vertex automatically results in a branch of the skeleton reaching it.

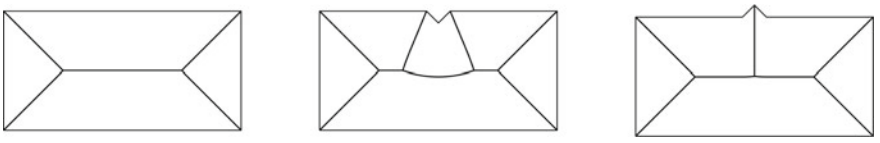

Fig. 2.2 Effect of a small shape change in the boundary on the skeleton of a rectangular shape

Because of this, many skeletonization algorithms come with a way to prune the skeleton of spurious branches. There are two ways to do this.

- *Prior smoothing of the curve.* One can apply curve smoothing algorithms (which will be described in Chap. 5). For polygons, smoothing can be done by removing small structures or flattening vague angles. It is interesting to note that smoothing curves does not always result in simplifying the skeleton (see [25] for a discussion).
- *Pruning.* Branches can be removed after the computation of the skeleton. This can be based on several principles using the fact that skeleton branches resulting from small incidents at the boundary can be characterized. We refer to [124] for more details.

2.7 Recovering the Initial Curve from the MAT

Given a parametrized sub-arc of the medial axis transform, one can explicitly reconstruct the part of the boundary $\partial\Omega$ which is associated with it (the contact points on $\partial\Omega$ of the maximal balls centered on the subarc). Let γ be such a parametrization, defined on $(a, b) \subset \mathbb{R}$, with values in $\Sigma^*(\Omega)$, so that $\gamma(u) = (m(u), r(u))$ for some functions m and r that we will assume to be at least C^2.

Without loss of generality, assume that $u \mapsto m(u)$ is an arc length parametrization ($|\dot{m}_u| = 1$). Assume also that $B(m(u), r(u))$ has exactly two contacts with $\partial\Omega$ (this is typically true on all $\Sigma^*(\Omega)$ except at a finite number of points). If $x \in \partial\Omega \cap B(m(u), r(u))$, then $|x - m(u)| = r(u)$ and, for all $\varepsilon \neq 0$, $|x - m(u + \varepsilon)| \geq r(u + \varepsilon)$ (because $B(m(u + \varepsilon), r(u + \varepsilon)) \subset \Omega$). Thus, letting $f(\varepsilon) = |x - m(u + \varepsilon)|^2 - r(u + \varepsilon)^2$, we have, because $\varepsilon = 0$ is a minimizer: $f(0) = \dot{f}(0) = 0$, with

$$\dot{f}(0) = -2\langle x - m(u)\,,\, \dot{m}(u)\rangle + 2r(u)\dot{r}(u).$$

Solving this equation in x, we obtain two solutions given by

$$x_+(u) = m(u) + r(u)\left[-\dot{r}(u)\dot{m}(u) + \sqrt{1 - \dot{r}(u)^2}q(u)\right],$$
$$x_-(u) = m(u) + r(u)\left[-\dot{r}(u)\dot{m}(u) - \sqrt{1 - \dot{r}(u)^2}q(u)\right],$$

with $q(u) \perp \dot{m}(u)$, $|q(u)| = 1$. Note that $|\dot{r}| < 1$ is a necessary condition for the existence of two distinct solutions.

The curvature of the boundary can also be related to the medial axis via an explicit formula. Let ρ_+ (resp. ρ_-) be the vector $-\dot{r}\dot{m} + \sqrt{1 - \dot{r}^2}\, q$ (resp. $-\dot{r}\dot{m} - \sqrt{1 - \dot{r}^2}\, q$) so that $x_+ = m + r\rho_+$ and $x_- = m + r\rho_-$. The following discussion holds for both arcs and we temporarily drop the $+$ and $-$ indices in the notation.

We have $x = m + r\rho$; ρ is a unit vector, and since the maximum disc is tangent to the curve at x, ρ is normal to the curve. Since r is positive and ρ is a radial vector for a maximal disc, ρ points outward from the curve at x and therefore is oriented

in the opposite direction to the normal (assuming that the boundary is positively oriented). Introduce the vector $h = \dot{m} + \dot{r}\rho$. We have $h^T\rho = -\dot{r} + \dot{r} = 0$ so that h is orthogonal to ρ. Since $|\rho| = 1$, $\dot{\rho}$ is also orthogonal to ρ and there exists a number c such that $\dot{\rho} = -ch$ (we have $|h|^2 = 1 - \dot{r}^2 > 0$ so that $h \neq 0$). Since $\rho = -N$, we also have

$$\dot{\rho} = \partial_s \rho \frac{ds}{du} = \kappa \frac{ds}{du} T,$$

where κ is the curvature of the considered arc of curve. Likewise, $\dot{x} = (ds/du)T$ so that $\dot{\rho} = \kappa\dot{x}$. We now use these identities to compute κ: we have $\dot{x} = \dot{m} + \dot{r}\rho + r\dot{\rho} = (1 - cr)h$. This implies

$$\kappa = -c/(1 - cr),$$

which provides a very simple relation between c and the curvature.

To be complete, it remains to compute c. From $\dot{\rho} = -c(\dot{m} + \dot{r}\rho)$, we get

$$\dot{\rho}^T\dot{m} = -c(1 + \dot{r}\rho^T\dot{m}) = -c(1 - \dot{r}^2).$$

We also have

$$-\ddot{r} = \partial(\rho^T\dot{m}) = \dot{\rho}^T\dot{m} + \rho^T\ddot{m} = \dot{\rho}^T\dot{m} + K\rho^T q,$$

where K is the curvature of the skeleton. Writing $\rho^T q = \varepsilon\sqrt{1 - \dot{r}^2}$ with $\varepsilon = \pm 1$, we get the equation:

$$\dot{\rho}^T\dot{m} = -\ddot{r} - \varepsilon K\sqrt{1 - \dot{r}^2},$$

which yields (reintroducing the $+$ and $-$ subscripts for each contact) $c_+ = \ddot{r}/(1 - \dot{r}^2) + K/\sqrt{1 - \dot{r}^2}$ and $c_- = \ddot{r}/(1 - \dot{r}^2) - K/\sqrt{1 - \dot{r}^2}$.

2.8 Generating Curves from Medial and Skeletal Structures

The previous section described how to retrieve a curve once its medial axis transform has been computed. Here we want to discuss the issue of using the medial axis transform as a modeling tool, i.e., of specifying a curve by starting from a medial axis transform.

This is a more difficult problem, because not any combination of curves and radii is a valid medial axis. Even when the skeleton consists of only one curve, we have already seen conditions in the above section, like $|\dot{r}_u| < 1$ at all points in the interior of the medial curve, that are required in the skeletal representation. We must also ensure that the specified curve is regular on both sides of the axis, which, since $\dot{x} = (1 - cr)h$, must ensure that $1 - cr$ does not vanish along the curve. In fact, $1 - cr$ must be positive. To see this, note that we have proved that $1 - cr = (1 - r\kappa)^{-1}$. At a convex point ($\kappa > 0$), r must be smaller than the radius of curvature $1/\kappa$ so that

$1 - r\kappa > 0$. Since points of positive curvature always exist, we see that $1 - cr$ must remain positive along the curve in order to never be zero. Using the expression for c found in the previous section, this provides a rather complex condition:

$$1 - \frac{r\ddot{r}}{1 - \dot{r}^2} > \frac{|K|r}{\sqrt{1 - \dot{r}^2}}. \tag{2.1}$$

To ensure continuity of the reconstructed curve when boundary branches meet, we need $|\dot{r}| = 1$ at terminal points of the medial axis. Also, if the medial axis has multiple branches, the corresponding parts of the curve must have the same limits on both sides. More conditions are needed to ensure that the contacts at these points are smooth. This provides a rather complicated set of constraints that must be satisfied by a generative medial axis model. This can be made feasible, however, in some simple cases, as shown in the following examples.

2.8.1 Skeletons with Linear Branches

Let us consider the situation in which each branch of the medial axis is a line segment, i.e., $K = 0$. The constraints on r are then $r > 0$, $\dot{r}^2 < 1$ and $r\ddot{r} + \dot{r}^2 < 1$. The last inequality comes from the fact that $cr < 1 \Leftrightarrow r\ddot{r} < 1 - \dot{r}^2$. Introducing $z = r^2/2$, this can also be written $\ddot{z} < 1$. Assume that $\ddot{z} = -f$ with $f > -1$. Integrating twice, we find

$$\begin{aligned} \dot{z}(u) &= \dot{z}(0) - \int_0^u f(t)dt \\ z(u) &= z(0) + u\dot{z}(0) - \int_0^u (u - t) f(t)dt. \end{aligned} \tag{2.2}$$

We now analyze the other conditions in some special cases.

Shapes with a Single Linear Branch

We start with the simplest situation in which the medial axis is composed of a single segment, say $m(u) = (u, 0)$, $u \in [0, 1]$. Since $|\dot{r}| = 1$ at the extremities and the medial axis cannot cross the boundary curve, we need $\dot{r}(0) = 1$ and $\dot{r}(1) = -1$. Define

$$\begin{aligned} M_0(u) &= \int_0^u f(t)dt \\ M_1(u) &= \int_0^u tf(t)dt. \end{aligned}$$

Using the identities $\dot{z}(0) = r(0)$, $\dot{z}(1) = -r(1)$, $z(0) = r(0)^2/2$ and $z(1) = r(1)^2/2$, we can solve (2.2) with respect to $r(0)$ and $r(1)$ to obtain:

$$r(0) = \frac{M_0(1) + M_0(1)^2/2 - M_1(1)}{1 + M_0(1)}$$
$$r(1) = M_0(1) - r(0) = \frac{M_0(1)^2/2 + M_1(1)}{1 + M_0(1)}.$$

These quantities must be positive, and we will assume that f is chosen with this property (note that the denominator is always positive since $f > -1$). These equations imply that z, and therefore m, are uniquely determined by f.

Consider now the remaining constraints, which are (in terms of z) $z > 0$ and $\dot{z}^2 < 2z$ on $(0, 1)$. Since the latter implies the former, we can concentrate on it, and introduce the function $h(u) = 2z(u) - \dot{z}(u)^2$. We have $h(0) = r(0)^2$ and $h(1) = r(1)^2$. Moreover,

$$\dot{h} = 2\dot{z}(1 - \ddot{z}) = 2\dot{z}(1 + f).$$

Since $1 + f > 0$, $\dot{h}$ vanishes for $\dot{z} = 0$, or $M_0(u) = r(0)$. Note that $\dot{h}(0) = 2r(0)(1 + f) > 0$ and $\dot{h}(1) = -2r(1)(1 + f) < 0$ so $\dot{h}$ changes signs over $(0, 1)$.

Also, since the extrema of h only occur when $\dot{z} = 0$ (and $h = 2z$ at these points), h will be positive under any condition that ensures that $z > 0$ when $\dot{z} = 0$, which reduces to $r(0)^2/2 + M_1(u) > 0$ whenever $M_0(u) = r(0)$.

There is an easy case: if $f > 0$, then $M_1(u) > 0$ and the condition is satisfied. Moreover, if $f > 0$, then $M_1(1) \leq M_0(1)$ so that $r(0)$ and $r(1)$ are positive. However, as Fig. 2.3 shows, interesting shapes are obtained when $f < 0$ is allowed.

Shapes with Three Intersecting Linear Branches

Let's now consider a slightly more complex example with one multiple point and three linear branches. So we have three lines, ℓ_1, ℓ_2, ℓ_3, starting from a single point p_0. Let $\ell_i = \{p_0 + uw_i, u \in [0, s_i]\}$, where w_1, w_2, w_3 are unit vectors. Let q_i be a unit vector completing w_i in a positively oriented orthonormal frame. Finally, let $r^{(1)}, r^{(2)}$ and $r^{(3)}$ be the radii along each of these lines and $z^{(i)} = (r^{(i)})^2/2$. Assume that $\ddot{z}^{(i)} = -f^{(i)}(u/s_i)$ for $u \in (0, s_i)$, where $f^{(i)} > -1$ as before, and is defined over $[0, 1]$.

We need to work out the compatibility conditions for the $r^{(i)}$ at the intersection point, $u = 0$. Assume that the branches are ordered so that (w_1, w_2), (w_2, w_3) and (w_3, w_1) are positively oriented. The compatibility conditions are

$$x_+^{(1)}(0) = x_-^{(2)}(0),\ x_+^{(2)}(0) = x_-^{(3)}(0),\ x_+^{(3)}(0) = x_-^{(1)}(0).$$

Identifying the norms, we see that the radii must coincide: $r^{(1)}(0) = r^{(2)}(0) = r^{(3)}(0) := r_0$. So, defining h_1, h_2, h_3 by

$$h_1 = \rho_+^{(1)}(0) = \rho_-^{(2)}(0), h_3 = \rho_+^{(2)}(0) = \rho_-^{(3)}(0), h_2 = \rho_+^{(3)}(0) = \rho_-^{(1)}(0),$$

we see that the triangle $(p_0 + h_1, p_0 + h_2, p_0 + h_3)$ has p_0 as circumcenter, and the lines defining the axis are the perpendicular bisectors of its edges.

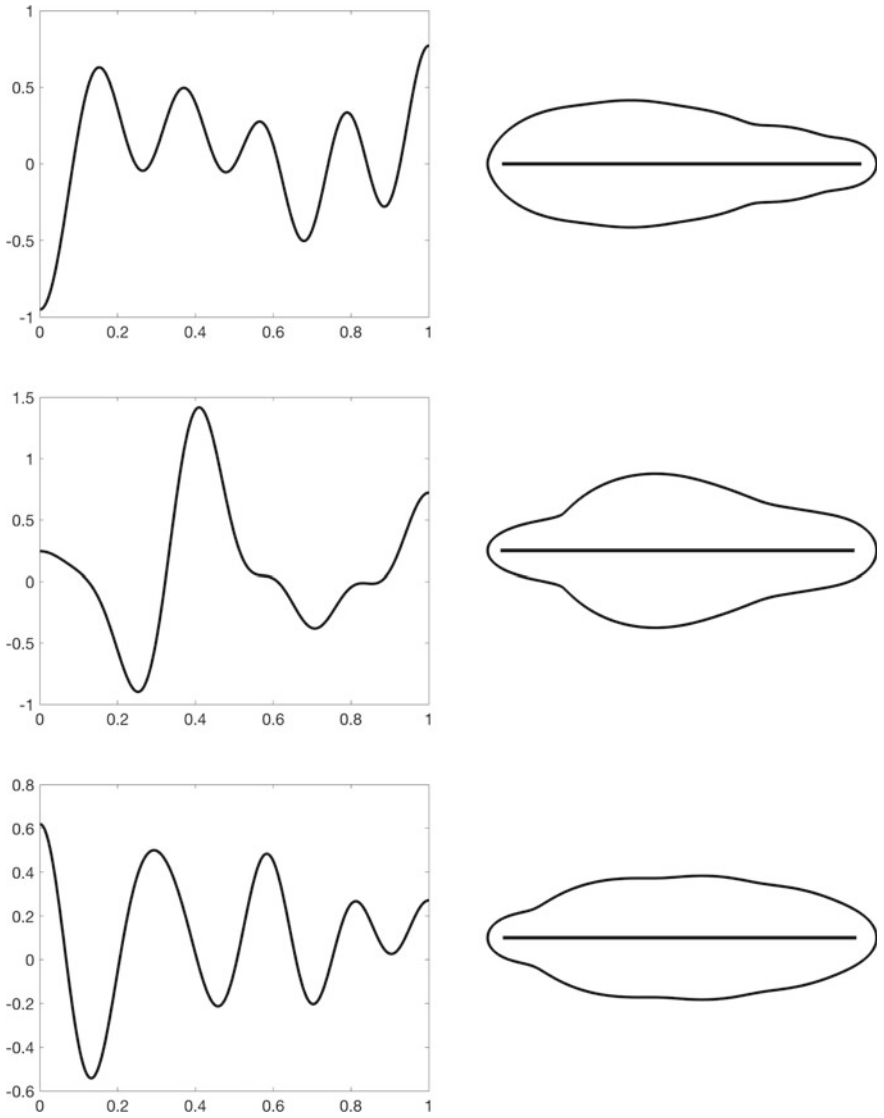

Fig. 2.3 Shapes with horizontal medial axes. The shapes (right column) are obtained with $\partial^2_{uu}(r^2) = -2f$; f is shown in the left column

Given the above, it is easier to organize the construction by first specifying p_0 and the three directions h_1, h_2, h_3. This specifies the vectors w_1, w_2, w_3: given $i \in \{1, 2, 3\}$, denote the other two indices by j and j'. Then

$$w_i = (h_j + h_{j'})/|h_j + h_{j'}|$$

and, from the expression of ρ, we see that this also specifies $\dot{r}^{(i)}(0)$, with

$$\dot{r}^{(i)}(0) = -z_i^T h_j = -\frac{1}{\sqrt{2}}\sqrt{1 + h_j^T h_{j'}} = -\cos(\theta_i/2),$$

where θ_i is the angle between h_j and h'_j.

This gives, for $u \in [0, s_i]$

$$\dot{z}^{(i)}(u) = \dot{z}^{(i)}(0) - \int_0^u f^{(i)}(t/s_i)dt = -r_0 \cos\frac{\theta_i}{2} - s_i M_0^{(i)}(u/s_i)$$

and

$$z^{(i)}(u) = \frac{r_0^2}{2} - r_0 u \cos\frac{\theta_i}{2} - s_i u M_0^{(i)}(u/s_i) + s_i^2 M_1^{(i)}(u/s_i).$$

Since we need $\dot{r}^{(i)}(s_i) = -1$, we have $z^{(i)}(s_i) = r^{(i)}(1)^2/2$ and $\dot{z}^{(i)}(s_i) = -r^{(i)}(1)$. Identifying $r^{(i)}(1)^2$ in the two equations above yields

$$r_0^2 \cos^2\frac{\theta_i}{2} + 2r_0 s_i \cos\frac{\theta_i}{2} M_0^{(i)}(1) + s_i^2 M_0^{(i)}(1)^2 = r_0^2 - 2r_0 s_i \cos\frac{\theta_i}{2} - s_i^2 M_0^{(i)}(1) + s_i^2 M_1^{(i)}(1)$$

or

$$\left(M_0^{(i)}(1)^2 + 2M_0^{(i)}(1) - 2M_1^{(i)}(1)\right)\frac{s_i^2}{r_0^2} + 2\cos\frac{\theta_i}{2}\left(1 + M_0^{(i)}(1)\right)\frac{s_i}{r_0} - \left(1 - \cos^2\frac{\theta_i}{2}\right) = 0. \quad (2.3)$$

Assuming that $f^{(i)}$ satisfies

$$M_0^{(i)}(1)^2/2 + M_0^{(i)}(1) - M_1^{(i)}(1) > 0,$$

which is a condition already encountered in the previous case, this equation has a unique solution, specifying s_i. The curve is then uniquely defined by p_0, h_1, h_2, h_3, $f^{(1)}$, $f^{(2)}$, $f^{(3)}$, with constraints on the $f^{(i)}$'s similar to those obtained in the one-branch case. Examples are provided in Fig. 2.4.

Note that this construction does not freely specify the medial axis, but only the orientation of its branches (since the s_i's are constrained by the rest of the parameters). One possibility to deal with this is to relax the specification of the $f'_i s$ by adding a factor α_i, using

$$\ddot{z}^{(i)} = -\alpha_i f^{(i)}.$$

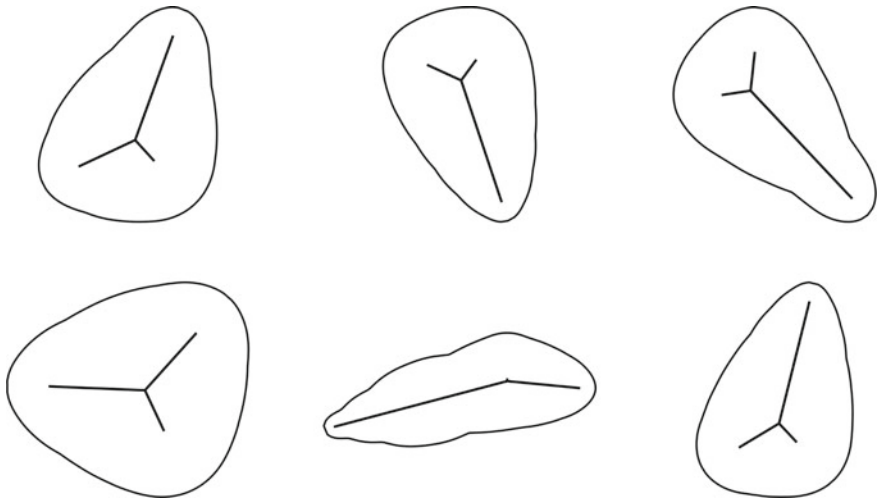

Fig. 2.4 Shapes generated from a medial axis with three linear branches

This implies that $M_0^{(i)}$ and $M_1^{(i)}$ must be replaced by $\alpha_i M_0^{(i)}$ and $\alpha_i M_1^{(i)}$ in the computation above, and Eq. (2.3), with fixed s_i, becomes a second-degree equation in α_i. The consistency conditions (existence of a solution to this equation, requirement that $\alpha_i f^{(i)} > -1$, etc.) are, however, harder to work out in this case.

Shapes with Generic Linear Branches

Conceptually, the above construction can be generalized to any skeleton with a ternary tree structure and linear branches. Indeed, the derivative $\dot{r}$ is uniquely specified at the extremities of any branch: it is -1 if the branch ends and $-\cos\theta/2$ at an intersection, where θ is specified by the branch geometry as above. Also, the radii at all branching points are uniquely specified as soon as one of them is (the constraint propagates along the tree). Of course, as before, the fact that the solution is uniquely defined does not guarantee consistency, which become harder to specify when the medial axis gets more complex. Finally, it is important to note that, for all the previous methods, even if the consistency conditions are satisfied, there is still a possibility for the shape to self-intersect non-locally (without singularity).

2.8.2 Skeletal Structures

One way to simplify the construction of a shape from a skeleton is to relax some of the conditions that are associated with medial axes. Skeletal structures, which we briefly describe now, have been introduced by J. Damon in [76–78] with this idea in mind.

There are two parts in Damon's skeletal structure. The first one is the skeletal set (the skeleton), which is a union of smooth open curves that meet at singular points (branching points or end-points) with well-defined tangents at their extremities.

The second part of the skeletal structure is formed by the vectors that correspond to $r\rho$ in our previous notation, with some smoothness and consistency conditions; referring to [76] for details, here are the most important ones. Like with the medial axis, each point in the smooth curves of the skeletal set carries two of these vectors (one on each side of the curve), and singular points can carry one vector (at end-points) or more than two (at branching points). When one continuously follows one of these vectors along a smooth branch until a branching point, it must have a limit within the set of vectors at this point, and all vectors at this point can be obtained by such a process. At end-points, there is a unique vector which is tangent to the skeletal curve.

To summarize, a skeletal structure requires a skeletal set, say S, and, at each point p in the skeletal set, a set $U(p)$ of vectors that point to the generated curve, subject to the previous conditions. The generated curve itself is simply

$$C = \{p + U(p),\, p \in S\}\,.$$

The medial axis transform does induce a skeletal structure, but has additional properties, including the facts that, at each p, all vectors in $U(p)$ must have the same norm, and if p is on a smooth curve, the difference between the two vectors in $U(p)$ must be perpendicular to the curve. These properties are not required for skeletal structures.

Most of the analysis done in the previous section on the regularity of the generated curve with the medial axis transform can be carried over to skeletal structures. Along any smooth curve in the skeletal structure, one can follow a smooth portion of the generated curve, writing

$$x(u) = m(u) + r(u)\rho(u)$$

and assuming an arc-length parametrization in $m(u)$. Letting $c = -\dot{m}^T\dot{\rho}$, one can write, for some $\alpha \in \mathbb{R}$,

$$\dot{\rho} = -c\dot{m} + \alpha\rho$$

because ρ is assumed to be non-tangent to the skeletal set (except at its end-points). This definition of c generalizes the one given for the medial axis, in which we had $\dot{\rho} = -ch = -c\dot{m} + c\dot{\rho}\rho$. Since we have $\dot{x} = (1 - cr)\dot{m} + (\alpha + \dot{r})\rho$, we see that $cr < 1$ is here also a sufficient condition for the regularity of the curve.

We need to check that different pieces of curves connect smoothly at branching points. With the medial axis, a first-order contact (same tangents) was guaranteed by the fact that the generated curve was everywhere perpendicular to ρ. With skeletal structures, we have (since $\dot{\rho}^T\rho = 0$)

$$\dot{x}^T\rho = \dot{m}^T\rho + \dot{r}.$$

So, a sufficient condition for smooth contacts at branching points and at end-points is that $\dot{r} + \dot{m}^T \rho$ vanishes at the extremities of the smooth curves that form S (while this quantity vanishes everywhere with the medial axis transform).

Obviously, these conditions are much less constraining than those associated with the medial axis transform. One can start fixing ρ, which defines c, then r such that $rc < 1$, with a few end-point conditions that must be satisfied. The simplification that is brought to curve generation, however, comes at a price, which is that a skeletal structure is not uniquely specified by a given curve, as the medial axis transform was. It is not a one-to-one curve representation.

Chapter 3
Local Properties of Surfaces

In this chapter, we start discussing representations that can be associated with three-dimensional shapes, where surfaces now replace curves. We begin with some basic definitions and results on the theory of surfaces in $\mathbb{R}^3$. Although some parts are redundant with the abstract discussion of submanifolds that is provided in Appendix B, we have chosen to give a more elementary presentation here, very close to [86], to ensure that this important section can be read independently.

3.1 Curves in Three Dimensions

Before addressing surfaces, we extend our developments on plane curves to the three-dimensional case. A three-dimensional parametrized curve is a function $\gamma : [a, b] \mapsto \mathbb{R}^3$. It is regular if it is C^1 and $|\dot{\gamma}| \neq 0$ for all $t \in [a, b]$. For regular curves, the unit tangent is defined by $T = \dot{\gamma}/|\dot{\gamma}|$ and the arc length is $ds = |\dot{\gamma}(t)|dt$.

Assume that γ is C^2 and parametrized by arc length. One then defines the *curvature* of γ at s by $\kappa(s) = |\dot{T}|$. This differs from the planar case, for which a sign was attributed to the curvature: here, the curvature is always non-negative.

One says that the γ is bi-regular if $\kappa(s) \neq 0$ for all s. In this case, one uniquely defines a unit vector N by the relation $\dot{T} = \kappa N$; N is perpendicular to T because T has unit norm. Finally, the binormal is the unique unit vector B which completes (T, N) into a positive orthonormal basis of[1] $\mathbb{R}^3$: $B = T \times N$. The frame (T, N, B) is called the Frénet frame, and the plane passing through $\gamma(t)$ and generated by T and N is called the osculating plane.

[1] If $h = (a, b, c)$ and $k = (a', b', c')$ are three-dimensional vectors, their cross product $h \times k$ is defined by

$$h \times k = (bc' - cb', a'c - ac', ab' - a'b).$$

It is orthogonal to both h and k and vanishes if and only if h and k are collinear. Moreover, for any third vector l: $(h \times k)^T l = \det(h, k, l)$.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_3

The derivative of the normal is orthogonal to N and can therefore be written $\dot{N} = aT + bB$. We have $N^T T = 0$ so that $a = \dot{N}^T T = -N^T \dot{T} = -\kappa$. The *torsion* of the curve is given by $-b$ by definition and denoted τ so that $\dot{N} = -\kappa T - \tau B$. Using the fact that $\dot{B}^T T = -B^T \dot{T} = -\kappa B^T N = 0$ and $\dot{B}^T N = -\dot{N}^T B = \tau$, we have $\dot{B} = \tau N$, which provides the third equation of Frénet's formulas for three-dimensional curves:

$$\begin{cases} \partial_s T = \kappa N, \\ \partial_s N = -\kappa T - \tau B, \\ \partial_s B = \tau N. \end{cases} \tag{3.1}$$

(These equations are valid for any parametrization if one defines $\partial_s f = \dot{f}/|\dot{m}|$.) Note that, if F is the 3 by 3 rotation matrix associated with the Frénet frame, i.e., $F = [T, N, B]$, then the Frénet formulas can be written as

$$\partial_s F = F S_m, \tag{3.2}$$

where S_m is the skew-symmetric matrix

$$S_m = \begin{pmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}.$$

There is a three-dimensional version of Theorem 1.13. The proof, based on Eq. (3.2), is identical to the alternative proof given in two dimensions in Sect. 1.9.

Theorem 3.1 *Two C^2 curves γ and $\tilde{\gamma}$ have the same curvature and torsion as functions of their arc length if and only if there exist a rotation R, a vector b and a change of parameter ϕ such that $\tilde{\gamma} = R\gamma \circ \phi + b$.*

3.2 Regular Surfaces

Curves being represented by one parameter, one may think of surfaces as bi-parametrized objects, i.e., functions $(u, v) = m(u, v)$ defined on some subset of $\mathbb{R}^2$.

Definition 3.2 A C^p parametrized (regular) surface is a C^p map $m : U \mapsto \mathbb{R}^3$, where U is an open subset of $\mathbb{R}^2$, such that:

1. m is one-to-one and its inverse, $m^{-1} : V = m(U) \to U$ is continuous (m is a homeomorphism between U and V), i.e., if a sequence u_n is such that $m(u_n)$ converges to $p = m(u) \in V$, then u_n converges to u.
2. For all $q \in U$, the differential $dm(q)$ is one-to-one.

The last statement is equivalent to the fact that the 3 by 2 matrix of partial derivatives $[\partial_1 m, \partial_2 m]$ has rank 2. It is a direct generalization of regularity for curves.

We did not assume that curve parametrizations were one to one, but this assumption provides an important simplification for surfaces. The second part of Condition 1 also prevents situations in which the boundary of some part of the surface intersects another part (see examples).

Finally, letting $S = m(U)$ be the *range* of m, we will often abuse the terminology by saying that S (the geometric object) is a parametrized surface. However, for many interesting surfaces, it is generally impossible (or simply not convenient) to find a parametrization which satisfies the previous requirement and covers the whole surface. This is a fundamental difference with the theory of plane curves. To be able to handle interesting cases, we need to limit our requirement for parametrizations to hold only within *patches* that together cover the surface, with additional conditions ensuring that the surface is smooth and non-intersecting, and that the patches fit well together.

Definition 3.3 A subset $S \subset \mathbb{R}^3$ is a C^k regular surface if, for each $p \in S$, there exists an open set V in $\mathbb{R}^3$, with $p \in V$, and a C^k parametrization of the surface patch $V \cap S$. The local parametrizations are also called local charts.

This definition requires more than just M being covered with parametrized patches. These patches must be obtained from intersections of S with three-dimensional open sets. In particular, this prevents non-local self-intersection, since, along such an intersection, the surface would contain two local patches and would not be locally parametrizable. Figure 3.3 provides an illustration of how local parametrized patches can be combined to cover a surface.

If $m : U \to V \cap S$ is as specified in the definition, for any p in $V \cap S$, there exist parameters $(u(p), v(p))$ in U such that $m(u(p), v(p)) = p$. The functions $p \mapsto u(p)$ and $p \mapsto v(p)$ are called the local coordinates on $V \cap S$.

3.2.1 Examples

Graphs of Functions

The simplest example of a parametrized surface is the graph of a C^1 function $f : U \subset \mathbb{R}^2 \to \mathbb{R}$. The parametrization is then $m(u, v) = (u, v, f(u, v))$. Since the inverse of (u, v, z) on the surface is (u, v), this is a homeomorphism, and the differential is

$$(u, v) \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \partial_1 f & \partial_2 f \end{pmatrix}$$

which has rank 2.

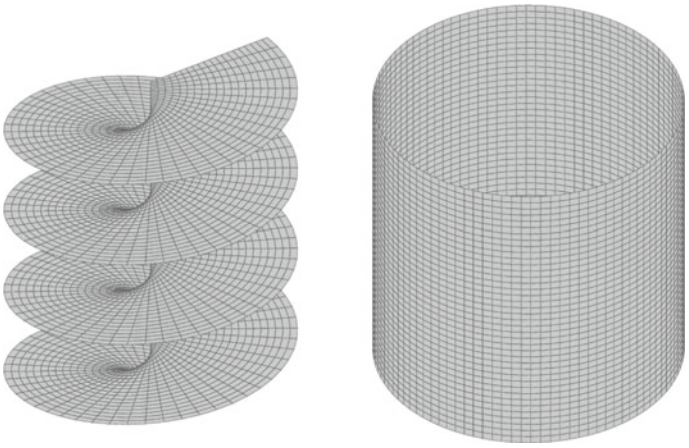

Fig. 3.1 Helicoid (left) and cylinder (right)

Helicoid

A parametrized surface does not have to be a graph. An example is the helicoid (Fig. 3.1), defined over $(0, a) \times \mathbb{R}$ by

$$m(u, v) = (u\cos(v), u\sin(v), \lambda v)$$

for some $a, \lambda > 0$.

Cylinder

The cylinder, which can be defined by the set of $m(u, v) = (\cos u, \sin u, v)$, for $u \in [0, 2\pi)$ and $v \in (-1, 1)$ (Fig. 3.1), is an example of a surface which, according to our definition, cannot be globally parametrized. This map is one-to-one and in fact a homeomorphism, and the only reason why this is not a parametrization is that we have required parametrizations to be defined on open sets ($[0, 2\pi) \times (-1, 1)$ is not open). The cylinder is a regular surface, by considering patches for the same map m, defined on $(0, 2\pi) \times (-1, 1)$ and say $(-\pi, \pi) \times (-1, 1)$.

Sphere

Consider now the example of the unit sphere (Fig. 3.2), which is denoted

$$S^2 = \left\{p \in \mathbb{R}^3, |p| = 1\right\}.$$

Like the cylinder, this surface cannot be globally parametrized. The simplest choice of local charts are the projections: $(u, v) \mapsto (u, v, \sqrt{1 - u^2 - v^2})$ and $(u, v) \mapsto (u, v, -\sqrt{1 - u^2 - v^2})$, both defined for $u^2 + v^2 < 1$, the open unit disc. The two maps cover the whole sphere, except the equator for which the third coordinate is 0. One can add other projections, like $(u, v) \mapsto (u, \pm\sqrt{1 - u^2 - v^2}, v)$,

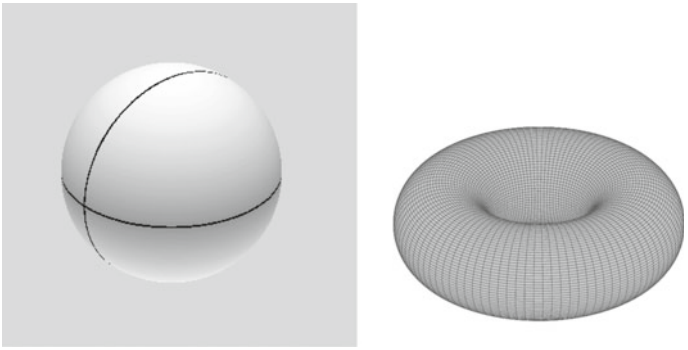

Fig. 3.2 Sphere (left) and torus (right)

$(u, v) \mapsto (\pm\sqrt{1-u^2-v^2}, u, v)$ to cover everything, or use cylindrical-like charts close to the equator.

Another useful coordinate system for the sphere is the (properly named) spherical coordinate system: $(u, v) \mapsto (\cos u \cos v, \sin u \cos v, \sin v)$. These coordinates cover the whole sphere when (u, v) varies in $[0, 2\pi) \times [-\pi/2, \pi/2]$ but they do not provide a local parametrization, since this set is not open (and the map is not one-to-one for $v = -\pi/2$ and $v = \pi/2$). Restricting to the open intervals requires using other charts to cover the meridian $u = 0$, for example the same coordinates on $(-\pi, \pi) \times (-\pi/2, \pi/2)$ which now only leave the poles uncovered. A neighborhood of the poles can be covered by the previous projection maps.

Torus

The torus (a surface with a shape like a donut, see Fig. 3.2) can be represented as the image of $[0, 2\pi) \times [0, 2\pi)$ under the map

$$m(u, v) = ((R + r\cos v)\cos u, (R + r\cos v)\sin u, r\sin v),$$

where $0 < r < R$, which is one-to-one but once again not defined on an open set. The whole torus can be covered by considering this map restricted to open subsets of $[0, 2\pi) \times [0, 2\pi)$. Let us check that the rank of the differential of m is always 2. We have

$$dm = \begin{pmatrix} -(R + r\cos v)\sin u & -r\sin v\cos u \\ (R + r\cos v)\cos u & -r\sin v\sin u \\ 0 & r\cos v \end{pmatrix}.$$

The determinant of the first two rows is $-r\sin v(R + r\cos v)$. Since $r < R$, it can only vanish when $\sin v = 0$. For the remaining two determinants, which are $r(R + r\cos v)\sin u\cos v$ and $r(R + r\cos v)\cos u\cos v$, to vanish together, one needs $\cos v = 0$. So at least one of the three two-by-two determinants does not vanish (Fig. 3.3).

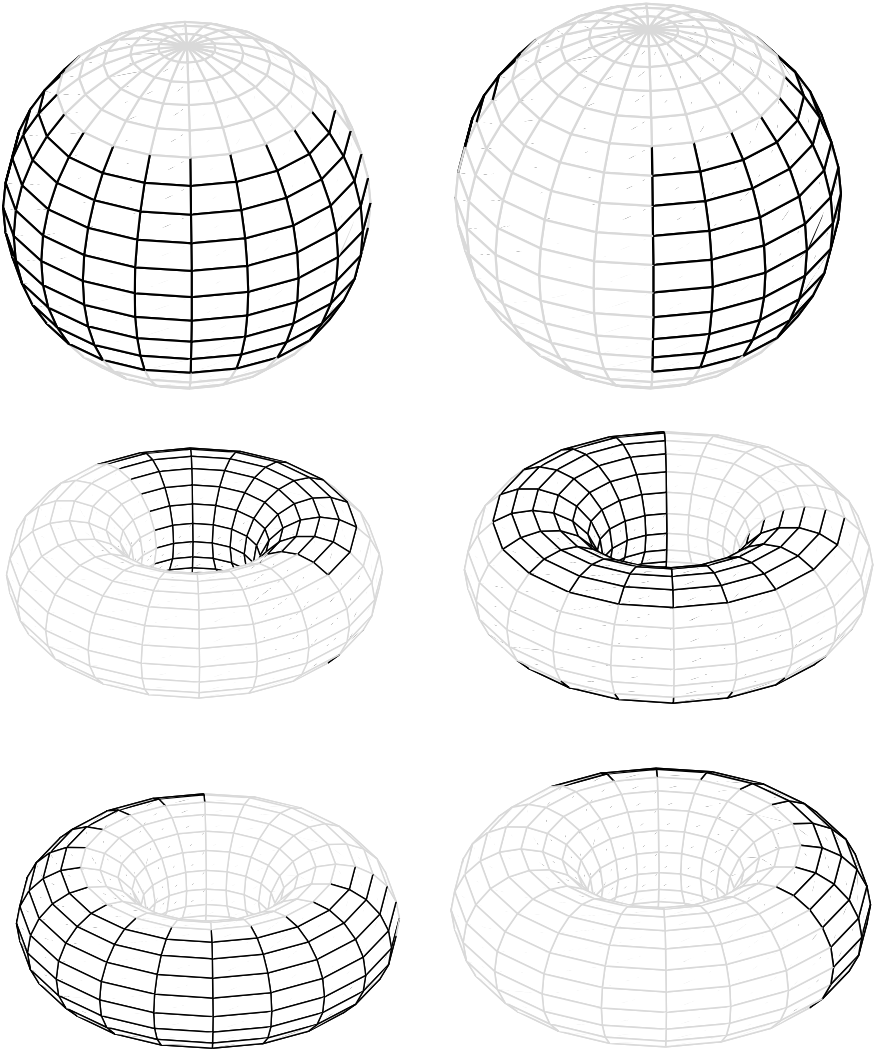

Fig. 3.3 Examples of decompositions in local charts for the sphere and the torus. Parametrizations are represented by grids over the surface, black inside local patches and gray outside

A Non-regular Surface

As a last example, consider the set S defined by

$$S = \{m(u, v) : (u, v) \in (-1, 1) \times (-\pi/2, \pi)\}$$

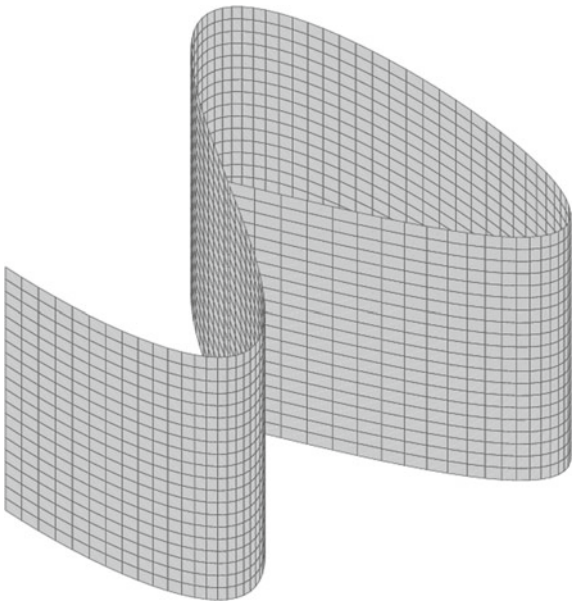

Fig. 3.4 A non-regular surface

with

$$m(u, v) = \begin{cases} \cos v \\ \sin 2v \\ u \end{cases}$$

(see Fig. 3.4). The parametrization is defined over an open set, it is one to one and

$$dm = \begin{pmatrix} 0 & -\sin v \\ 0 & 2\cos 2v \\ 1 & 0 \end{pmatrix}$$

has rank two everywhere. S is not a parametrized surface, however, because

$$\lim_{v \to -\pi/2} m(0, v) = m(0, \pi/2),$$

which contradicts the assumption that m^{-1} is continuous. The same contradiction can be obtained for $S \cap V$ where V is any open subset of $\mathbb{R}^3$ that contains 0, so that S is not a regular surface either.

3.2.2 Changing Coordinates

As we have seen, several different valid parametrizations can be defined at a single point of a surface. Like for curves, "geometric" properties or quantities should not depend on the parametrization. We will define a few of them in the following: normals, curvature, length, area, etc.

It can be deduced from the requirements in Definition 3.2 that changes of coordinates are C^1 homeomorphisms. To be more specific, assume that in a neighborhood V of a point p on S, there exist two parametrizations $m : U \to V$ and $\tilde{m} : \tilde{U} \to V$. Then, because of the invertibility of the parametrization, one can go from U to V via m, then from V to $\tilde{U}$ via the inverse of $\tilde{m}$. The resulting map, $\varphi = \tilde{m}^{-1} \circ m : U \to \tilde{U}$, is called a change of coordinates, and is a *diffeomorphism* between U and $\tilde{U}$ (it is C^1, invertible, with a C^1 inverse). This consequence of Definition 3.2 can be proved using the inverse mapping theorem.

3.2.3 Implicit Surfaces

An implicit surface is defined by an equation of the form $f(p) = 0$ where $f : \mathbb{R}^3 \to \mathbb{R}$ is a scalar function which is such that $\nabla f(p) \neq 0$ if $f(p) = 0$. In this case, the set

$$S = \left\{ p \in \mathbb{R}^3,\, f(p) = 0 \right\}$$

is a regular surface. (This is a consequence of the implicit function theorem.)

3.3 Tangent Planes and Differentials

3.3.1 Tangent Planes

For a curve, we were able to define a unique unit tangent, but this is obviously no longer possible for surfaces. Still, curves provide a simple way to define tangent vectors to surfaces.

A curve $m : I \to \mathbb{R}^3$ is supported by a surface S if and only if, for all $t \in I$, one has $m(t) \in S$. We have the following definition:

Definition 3.4 Let S be a regular surface. A vector $T \in \mathbb{R}^3$ is tangent to S at a point $p \in S$ if and only if, for some $\varepsilon > 0$, there exists a C^1 curve $\gamma : (-\varepsilon, \varepsilon) \to S$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = T$.

Assume, in the previous definition, that ε is chosen small enough so that the curve γ is completely inscribed in a parametrized patch of the surface S. Let $m : (u, v) \mapsto m(u, v)$ be the parametrization. Since m is one-to-one, one can express

$\gamma(t) = m(u(t), v(t))$. The plane curve $t \mapsto (u(t), v(t))$ is the expression of γ in the local coordinates. From the chain rule, we have

$$\dot{\gamma} = \dot{u}\partial_1 m + \dot{v}\partial_2 m.$$

Thus, $\dot{\gamma}$ must be a linear combination of the two independent vectors $\partial_1 m$ and $\partial_2 m$. Conversely, if $p = m(u_0, v_0)$, then, for any $\alpha, \beta \in \mathbb{R}$, the vector $\alpha\partial_1 m + \beta\partial_2 m$ is the derivative of $t \mapsto m(u_0 + \alpha t, v_0 + \beta t)$ and is therefore tangent to S at p. This proves the following proposition:

Proposition 3.5 *Let S be a regular surface, $p \in S$ and $m : U \to S$ a parametrization of S in a neighborhood of p. The set of tangent vectors to S at p is the plane generated by $\partial_1 m$ and $\partial_2 m$.*

The tangent plane to S at p will be denoted $T_p S$. Although the generating vectors $\partial_1 m$ and $\partial_2 m$ depend on the local parametrization m, the plane itself does not, because we gave a parametrization-independent definition of tangent vectors.

If S is defined implicitly by $f(p) = 0$, the tangent plane at p is characterized by the equation $\nabla f(p)^T T = 0$ (recall that f must be such that $\nabla f(p) \neq 0$ if $f(p) = 0$). Indeed, if γ is a curve on S, then $f \circ \gamma(t) = 0$ for all t, and the chain rule implies: $\nabla f(\gamma(0))^T \dot{\gamma}(0) = 0$. This implies that $T_p S \subset (\nabla f(p))^\perp$. Because $T_p S$ and $(\nabla f(p))^\perp$ have the same dimension (two), they coincide.

3.3.2 Differentials

Differentials describe how measurements made on a surface vary locally. Consider a scalar function $f : S \to \mathbb{R}$ and take a local parametrization on S, $m : U \to V \cap S$. For $(u, v) \in U$, we can define the function $f_m(u, v) = f(m(u, v))$; this is a function from an open subset of $\mathbb{R}^2$ to $\mathbb{R}$, which provides the expression of f in the local system of coordinates: we have $f(p) = f_m(u(p), v(p))$. We have the following definition:

Definition 3.6 Let S be a regular surface. A function $f : S \to \mathbb{R}$ is C^1 at $p \in S$ if and only if, for some local parametrization m on S around p, the function f_m is C^1 at $m^{-1}(p)$.

We say that f is C^1 on S if it is C^1 at all $p \in S$.

(Because changes of coordinates are C^1, the definition does not depend on the choice of local parametrization at p.)

We now want to evaluate the effect that small variations in p have on the function f, i.e., we want to define the derivative of f. Usually, a first-order variation of $p \in \mathbb{R}^3$ in the direction h is represented by $p + \varepsilon h$, with small ε. This cannot be applied to S, since there is no reason for $p + \varepsilon h$ to belong to S if p does. It is reasonable, and rather intuitive, to define a first-order variation of p as an element of a curve on S containing p. This leads to:

Definition 3.7 Let S be a regular surface and $p \in S$. A first-order variation of p in the direction $h \in \mathbb{R}^3$ is a C^1 curve $\gamma : (-\varepsilon, \varepsilon) \to S$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = h$.

Note that, from this definition, first-order variations on S can only arise in directions which are tangent to S.

Now, we can define the differential of a scalar function f defined on S as the limit (if it exists) of the ratio $(f(\gamma(\delta)) - f(p))/\delta$ as δ tends to 0, γ being a first-order variation of p. This will be denoted $df(p)h$, with $h = \dot{\gamma}(0)$. Implicit in this notation is the fact that this limit only depends on $\dot{\gamma}(0)$, which is true if f is C^1 as stated in the next proposition.

Proposition 3.8 *Let f be a C^1 scalar function on a regular surface S. Then, for any $p \in S$, and $h \in T_pS$, the differential of f at p in the direction h exists, and is equal to the limit of the ratio $(f(\gamma(\delta)) - f(p))/\delta$ for any C^1 curve γ on S with $\gamma(0) = p$ and $\dot{\gamma}(0) = h$.*

Proof What we need to prove is that the limit of the ratio exists for any γ and only depends on h. Take a local parametrization m around p. We know that the function $f(m(u, v))$ is C^1, and letting $\gamma(t) = m(u(t), v(t))$, we have

$$\begin{aligned}\lim_{\delta \to 0} \frac{f(\gamma(\delta)) - f(p)}{\delta} &= \lim_{\delta \to 0} \frac{f_m(u(\delta), v(\delta)) - f_m(u(0), v(0))}{\delta} \\ &= \partial_1 f_m \dot{u}(0) + \partial_2 f_m \dot{v}(0).\end{aligned}$$

This proves the existence of the limit. We have $h = \dot{\gamma}(0) = \dot{u}(0)\partial_1 m + \dot{v}(0)\partial_2 m$: since $(\partial_1 m, \partial_2 m)$ has rank 2, $\dot{u}(0)$ and $\dot{v}(0)$ are uniquely specified by h and thus the limit above only depends on h. The notation $df(p)h$ is therefore valid. □

Note that the expression provided in this proof shows that $df(p)h$ is linear with respect to h. In other terms, $df(p)$ is a linear form from T_pS to $\mathbb{R}$. Most of the time, the computation of $df(p)$ is easy, because f can be expressed as the restriction to S of a differentiable function which is defined on $\mathbb{R}^3$. In this case, $df(p)h$ coincides with the usual differential of f, but restricted to the two-dimensional plane T_pS.

The proof above also provides a simple way to compute differentials in local charts: let $f : S \to \mathbb{R}$ be C^1, $p \in S$ and m be a local parametrization around p. Then, if $h = \alpha\partial_1 m + \beta\partial_2 m$, we have

$$df(p)h = \alpha\partial_1 f_m + \beta\partial_2 f_m. \tag{3.3}$$

When f is a vector-valued function ($f : S \to \mathbb{R}^d$), the differential $df(p)$ is defined in the same way, and is also vector-valued. It is a linear map from T_pS to $\mathbb{R}^d$.

The simplest examples of differentiable maps are the coordinates: if $m : U \to V \cap S$ is a local chart, the function $f = m^{-1}$ is such that $f_m(u, v) = (u, v)$, which is the identity map, and therefore differentiable. In particular, the coordinates: $p \mapsto u(p)$ and $p \mapsto v(p)$ are scalar differentiable maps. If $T = \alpha\partial_1 m + \beta\partial_2 m$, we have $du(p)T = \alpha$, $dv(p)T = \beta$ and $df(p)T = (\alpha, \beta)$.

Consider now the example of the sphere S^2. The tangent plane is easy to describe if one uses the fact that S^2 can be defined by the implicit equation $|p|^2 = 1$. If $\phi(p) = |p|^2$, we have $\nabla\phi(p)^T h = 2p^T h$ so that h is tangent to S^2 at p if and only if $p^T h = 0$ (h is perpendicular to p).

Fix a vector $p_0 \in S^2$ and consider the function $f(p) = p^T p_0$. Then, since f is well-defined on $\mathbb{R}^3$, we can use its restriction, which yields $df(p)h = h^T p_0$. This was an easy result, but for illustration purposes, let us retrieve it via local charts, which will require a little more computation.

Consider the parametrization $m(u, v) = (\cos u \cos v, \sin u \cos v, \sin v)$. Then,

$$\partial_1 m = (-\sin u \cos v, \cos u \cos v, 0) \text{ and}$$

$$\partial_2 m = (-\cos u \sin v, -\sin u \sin v, \cos v).$$

A straightforward computation shows that both $\partial_1 m$ and $\partial_2 m$ are orthogonal to $m(u, v)$. In the chart, letting $p_0 = (a, b, c)$, the function f_m is

$$f_m(u, v) = a \cos u \cos v + b \sin u \cos v + c \sin v.$$

Obviously, $\partial_1 f_m = {p_0}^T \partial_1 m$ and $\partial_2 f_m = {p_0}^T \partial_2 m$, so that, if $h = \alpha\partial_1 m + \beta\partial_2 m$, we get, by Eq. (3.3),

$$df(p)h = \alpha\partial_1 f_m + \beta\partial_2 f_m = {p_0}^T h.$$

3.4 Orientation and Normals

Let S be a surface and m a local parametrization on S. The vector $\partial_1 m \times \partial_2 m$ is non-vanishing and orthogonal to both $\partial_1 m$ and $\partial_2 m$. Since $\partial_1 m$ and $\partial_2 m$ generate $T_p S$ at $p = m(u, v)$, $\partial_1 m \times \partial_2 m$ is normal to the tangent plane at p.

In particular, the vector $N = \partial_1 m \times \partial_2 m / |\partial_1 m \times \partial_2 m|$ is a unit normal to the tangent plane. One also says that N is *normal to the surface* S. Since unit normals to a plane are defined up to a sign change, the one obtained from another parametrization must be either N or $-N$. This leads to the following definition:

Definition 3.9 Two local parametrizations, m and $\tilde{m}$, on a regular surface S have the same orientation at a given point at which they are both defined if

$$\frac{\partial_1 m \times \partial_2 m}{|\partial_1 m \times \partial_2 m|} = \frac{\partial_1 \tilde{m} \times \partial_2 \tilde{m}}{|\partial_1 \tilde{m} \times \partial_2 \tilde{m}|}$$

and have opposite orientation otherwise.

The surface S is said to be orientable if it can be covered by local parametrizations that have the same orientation wherever they intersect.

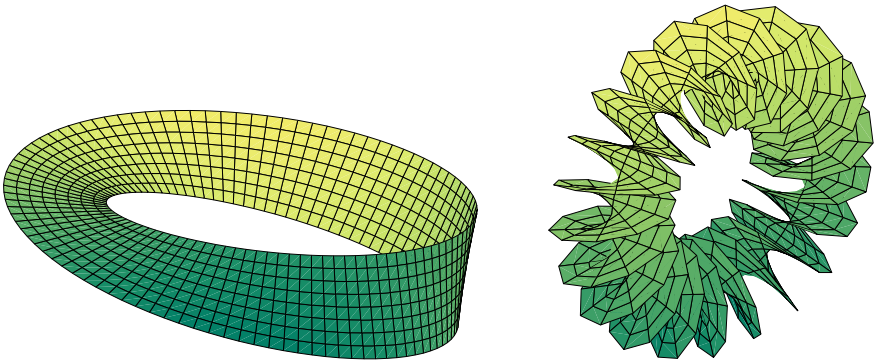

Fig. 3.5 Two examples of non-orientable surfaces. On the left is the Möbius band; the surface on the right is similar, with an odd number of twists

A surface is therefore orientable if there is a consistent (continuous) definition of a normal all over it. Not all surfaces are orientable (Fig. 3.5). A typical example is a twisted ring (the Möbius band).

3.5 Integration on an Orientable Surface

Let S be an orientable surface and $f : S \to \mathbb{R}$ be a continuous function. We want to compute the integral of f over S. We first define it within a local chart. Let $m : U \to V \cap S$ be a parametrized patch of the surface S. To motivate the definition, let U be divided into small rectangular cells (neglecting boundary issues). Consider a cell of the form $(u_0 - \varepsilon/2, u_0 + \varepsilon/2) \times (v_0 - \varepsilon/2, v_0 + \varepsilon/2)$. In this cell, we can make a first-order expansion of m in the form

$$m(u, v) = m(u_0, v_0) + (u - u_0)\partial_1 m(u_0, v_0) + (v - v_0)\partial_2 m(u_0, v_0) + o(\varepsilon)$$

so that, at first order, the image of the rectangular cell by m is a parallelogram in space, centered at $p_0 = m(u_0, v_0)$, namely

$$\sigma_0 = \{p_0 + \alpha\partial_1 m + \beta\partial_2 m, \alpha \in (-\varepsilon/2, \varepsilon/2), \beta \in (-\varepsilon/2, \varepsilon/2)\}.$$

Its area is given by $\varepsilon^2|\partial_1 m \times \partial_2 m|$, and the integral of a function f over this parallelogram can legitimately be estimated by $\varepsilon^2 f(p_0)|\partial_1 m \times \partial_2 m|$. Summing over cells and letting ε tend to 0 leads to the following definition:

Definition 3.10 Let f be a function defined on a regular surface S, and $m : U \to V \cap S$ a regular patch on S. The integral of f on $V \cap S$ is defined and denoted by

$$\int_{V\cap S} f(p)d\sigma_S(p) = \int_U f_m(u,v)\,|\partial_1 m \times \partial_2 m|\,du\,dv\,. \tag{3.4}$$

The integral of f over the whole surface S is defined as the sum of such integrals over non-overlapping local patches that cover S (maybe leaving out a finite number of curves or points on S). It is denoted

$$\int_S f(p)\,d\sigma_S(p) \text{ or } \int_S f\,d\sigma_S.$$

This can be shown to be independent of the chosen family of patches. The notation $d\sigma_S$ refers to the area form on S, defined on a local chart by $d\sigma_S = |\partial_1 m \times \partial_2 m|\,du\,dv$.

Note that the area form that we have defined here is a special case of a volume form in an arbitrary finite-dimensional manifold. For this reason, it is also often called the volume form of S (even though it measures areas).

Another (equivalent) way to globally define the integral is to use partitions of unity. Given a family $((U_i, m_i), i = 1, \ldots, n)$ of local parametrizations which cover the surface (so that $\bigcup_i m_i(U_i) = S$), but may overlap, one defines a partition of unity as a family of continuous functions $(\omega_i, i = 1, \ldots, n)$ where each ω_i is defined on S and takes values in $[0, 1]$, with $\omega_i(p) = 0$ if $p \notin m_i(U_i)$, and for all $p \in S$,

$$\sum_{i=1}^{n} \omega_i(p) = 1.$$

Such partitions of unity always exist, and one can define

$$\int_S f(p)\,d\sigma_S(p) = \sum_{i=1}^{N} \int_{U_i} \omega_i(m_i(u,v))\,f_{m_i}(u,v)|\partial_1 m_i \times \partial_2 m_i|\,du\,dv\,.$$

Here also, the result does not depend on the local parametrizations, or on which partition of unity is chosen.

That the right-hand side of (3.4) does not depend on the chosen parametrization should be clear from the approximation process which led to its definition (which was purely geometric), and can be checked directly as follows. Let $\tilde{m} : \tilde{U} \to V \cap S$ be another parametrization of the same patch. For $p \in V \cap S$, the equation $p = m(u,v) = \tilde{m}(\tilde{u}, \tilde{v})$ provides a relation between homologous coordinates given by

$$\begin{cases} \partial_1 m = \partial_1 \tilde{u}\partial_1 \tilde{m} + \partial_1 \tilde{v}\partial_2 \tilde{m} \\ \\ \partial_2 m = \partial_2 \tilde{u}\partial_1 \tilde{m} + \partial_2 \tilde{v}\partial_2 \tilde{m}. \end{cases}$$

The left-hand sides are computed at (u, v) and the right-hand sides at $(\tilde{u}, \tilde{v})$. This implies

$$\partial_1 m \times \partial_2 m = (\partial_1 \tilde{u} \partial_2 \tilde{v} - \partial_1 \tilde{v} \partial_2 \tilde{u})\, \partial_1 \tilde{m} \times \partial_2 \tilde{m}\,.$$

Letting ϕ be the change of variables ($\phi(u, v) = (\tilde{u}, \tilde{v})$), this is $\partial_1 m \times \partial_2 m = (\det \phi)(\partial_1 \tilde{m} \times \partial_2 \tilde{m}) \circ \phi$. Therefore

$$\begin{aligned}\int_U f(m(u, v))\, |\partial_1 m \times \partial_2 m|\, du\, dv \\ = \int_U f(\tilde{m} \circ \phi(u, v))\, |\partial_1 m \times \partial_2 m| \circ \phi(u, v)\, |\det \phi(u, v)|\, du\, dv \\ = \int_{\tilde{U}} f(\tilde{m}(\tilde{u}, \tilde{v}))\, |\partial_1 \tilde{m} \times \partial_2 \tilde{m}|\, d\tilde{u}\, d\tilde{v}.\end{aligned}$$

As an example, we compute the area of the unit sphere, which can be parametrized (poles excepted) by $m(u, v) = (\cos u \cos v, \sin u \cos v, \sin v)$. Then

$$\begin{aligned}\partial_1 m &= (-\sin u \cos v, \cos u \cos v, 0), \\ \partial_2 m &= (-\cos u \sin v, -\sin u \sin v, \cos v)\end{aligned}$$

and $|\partial_1 m \times \partial_2 m|^2 = \cos^2 v$, so that

$$\int_{S^2} d\sigma = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos v\, du\, dv = 2\pi [\sin v]_{-\pi/2}^{\pi/2} = 4\pi.$$

3.6 Regular Surfaces with Boundary

Consider the surface S defined by $x^2 + y^2 < 1$, $z = 0$, which is the unit disc in the horizontal plane. It is natural to define the boundary of S to be the circle $x^2 + y^2 < 1$, $z = 0$. Such a definition cannot coincide with the topological boundary in $\mathbb{R}^3$, $\partial S = \bar{S} \setminus \mathring{S}$, which would be the unit disc $\bar{S}$ defined by $x^2 + y^2 \leq 1$, $z = 0$ (because S has an empty interior in $\mathbb{R}^3$). Because of this, one defines the boundary of a surface S by $\partial S = \bar{S} \setminus S$ (and never use the topological boundary).

For a regular surface to be a "regular surface with boundary", some additional requirements are made to ensure that the boundary is locally a smooth curve.

Definition 3.11 Let S be a regular surface. One says that $p \in \partial S$ is a regular boundary point if there exists a parametrized surface $m : U \to \mathbb{R}^3$, where U is open in $\mathbb{R}^2$ with $0 \in U$ such that $m(0, 0) = p$, and, if $U^+ = \{(u, v) \in U : v > 0\}$ and $U^0 = \{(u, v) \in U : v = 0\}$, one has $m(U^+) = m(U) \cap S$ and $m(U^0) = m(U) \cap \partial S$.

One says that S is a regular surface with boundary if every point $p \in \partial S$ is regular.

Equivalently, $p \in \partial S$ is regular if there exists a regular surface $\tilde{S}$ such that $S \subset \tilde{S}$, $p \in \tilde{S}$ and $\partial S \cap \tilde{S}$ is a C^1 regular curve on $\tilde{S}$.

With this definition, the horizontal unit disc considered above is a regular surface with boundary, and its boundary is the horizontal unit circle. If one removes the origin from this unit disc, one still has a regular surface, but not "with boundary" anymore, because 0, which is now in the boundary, is not a regular point. Similarly, a triangular region, such as $\{x > 0, y > 0, x + y < 1, z = 0\}$, is not a regular surface with boundary, because its boundary has angles.

Let S be a regular surface with boundary and assume that S is oriented. Let $p \in \partial S$ and m be a local parametrization such as the one defined in Definition 3.11. Assume that $m : U^+ \to S$ is positively oriented (otherwise, take its composition with the transformation $(u, v) \mapsto (-u, v)$). Then one defines the unit tangent and normal to ∂S at p by $T_{\partial S}(p) = \partial_1 m(0, 0)/|\partial_1 m(0, 0)|$ and $N_{\partial S}(p) = N_S(p) \times T_{\partial S}(p)$ where $N_S(p)$, defined by

$$N_S(p) = \frac{\partial_1 m(0, 0) \times \partial_2 m(0, 0)}{|\partial_1 m(0, 0) \times \partial_2 m(0, 0)|},$$

extends the normal to S to its boundary. We let the reader check that this definition does not depend on the chosen parametrization m (or refer to the general argument made in Sect. B.7.3). With this definition, $N_{\partial S}$ is the inward pointing normal to ∂S in the tangent plane to S.

Note that the term "boundary" is not the only difference between the terminology used for surfaces and the one for standard topology. Here is another example.

Definition 3.12 One says that a regular surface S is a "closed surface", or a "surface without boundary", if and only if S is a compact subset of $\mathbb{R}^3$.

With this definition, a sphere and a torus are closed surfaces. However, the horizontal plane $z = 0$ is a closed subset of $\mathbb{R}^3$ and a regular surface, but not a closed surface according to this definition.

3.7 The First Fundamental Form

3.7.1 Definition and Properties

Let S be a regular surface. When h and k are two tangent vectors at $p \in S$, their dot product in $\mathbb{R}^3$ will be denoted $\langle h\ ,\ k\rangle_p$. It is simply the usual dot product, the sum of products of coordinates, but gets a specific notation because it is restricted to T_pS. The associated quadratic form is called *the first fundamental form*, and denoted

$$I_p(h) := |h|_p^2 . \tag{3.5}$$

This form is the key instrument for metric measurements on surfaces. Although its definition is straightforward, one must remember that surfaces are mostly described by local charts, and the expression of the form in such charts is not the standard

norm anymore. Indeed, let m be a local parametrization around p, and $h = \alpha\partial_1 m + \beta\partial_2 m \in T_p S$. Then

$$\begin{aligned} I_p(h) &= \alpha^2 \langle \partial_1 m\,,\, \partial_1 m \rangle_p + 2\alpha\beta \langle \partial_1 m\,,\, \partial_2 m \rangle_p + \beta^2 \langle \partial_2 m\,,\, \partial_2 m \rangle_p \\ &= \alpha^2 E + 2\alpha\beta F + \beta^2 G \end{aligned}$$

with the notation

$$E = \langle \partial_1 m\,,\, \partial_1 m \rangle_p,\ F = \langle \partial_1 m\,,\, \partial_2 m \rangle_p,\ G = \langle \partial_2 m\,,\, \partial_2 m \rangle_p. \tag{3.6}$$

E, F and G are the coefficients of the first fundamental form in the chart. They depend on the parameters u, v.

The following proposition allows one to use convenient local charts around a given point.

Proposition 3.13 *If S is a regular surface and $p \in S$, there exists a local parametrization $m : U \mapsto S$ around p such that $(\partial_1 m, \partial_2 m)$ is orthogonal on U.*

(Note that this proposition does not hold if 'orthogonal' is replaced with 'orthonormal'.)

3.7.2 Geodesics

The first fundamental form provides all the information required to compute lengths of curves on S: let γ be such a curve; assuming that γ is contained in a parametrized patch and letting $\gamma(t) = m(u(t), v(t))$, we have

$$|\dot\gamma|^2 = |\dot u \partial_1 m + \dot v \partial_2 m|^2 = \dot u^2 E + 2\dot u \dot v F + \dot v^2 G$$

so that the length of the curve from its expression in local coordinates is provided by

$$\text{length}(\gamma) = \int_a^b \sqrt{\dot u^2 E(u,v) + 2\dot u \dot v F(u,v) + \dot v^2 G(u,v)}\, dt.$$

Similarly, one defines the energy of a curve γ by

$$\text{energy}(\gamma) = \frac{1}{2}\int_a^b |\dot\gamma|^2 dt = \frac{1}{2}\int_a^b \left(\dot u^2 E(u,v) + 2\dot u \dot v F(u,v) + \dot v^2 G(u,v)\right) dt.$$

Curves of minimal energy on a surface are called *minimizing geodesics*, as formalized by the following definition.

Definition 3.14 Given two points p and p' on a surface M, a curve γ on M achieving the minimum energy among all piecewise C^1 curves on M linking p and p' is called a (minimizing) geodesic.

In addition to minimizing the energy, it can be shown that geodesics are curves of minimal length between two points [86, 87]. Moreover, if we define

$$d_M(p, p') = \inf \left\{ \text{length}(\gamma) : \gamma \text{ piecewise } C^1 \text{ on } M \right\}$$

then d_M is a distance on M, called the geodesic distance. Therefore if γ is a minimizing geodesic between p and p' then $\text{length}(\gamma) = d_M(p, p')$.

Minimizing geodesics between two given points do not always exist, however. Let M be, for example, the plane $z = 0$ with the point $(0, 0, 0)$ removed, which is a regular surface. Then the geodesic distance between p and $-p$ in M is $2|p|$, but this distance cannot be achieved because the optimal curve must be a straight line containing 0. We however have the following theorem, which is an application of the standard Hopf–Rinow theorem (see [86], for example).

Theorem 3.15 *If M is a (topologically) closed surface, then there exists a minimizing geodesic connecting any pair of its points.*

If γ is a minimizing geodesic between p and p', and $h(t)$ is for all t a vector tangent to the surface at $\gamma(t)$, one can define, for small ε, a one-parameter family of curves $\tilde{\gamma}(t, \varepsilon)$ such that $\tilde{\gamma}(t, 0) = \gamma(t)$ and $\partial_\varepsilon \tilde{\gamma}(t, 0) = h(t)$. Since γ is minimizing, the function $\varepsilon \mapsto \text{energy}(\tilde{\gamma}(\cdot, \varepsilon))$ has a vanishing derivative at $\varepsilon = 0$. This derivative is given by

$$\int_a^b \dot{\gamma}^T \dot{h} dt = -\int_a^b \ddot{\gamma}^T h dt$$

by integration by parts. The fact that this expression vanishes for any h tangent to the surface along γ implies that the "acceleration" $\ddot{\gamma}$ is normal to the surface. By extension, curves satisfying this property are also called geodesics. They generalize the notion of straight lines in a plane.

Definition 3.16 A C^2 regular curve γ on M is called a geodesic if its second derivative $\ddot{\gamma}$ is always normal to M.

Note that, using $\partial |\dot{\gamma}|^2 = 2\dot{\gamma}^T \ddot{\gamma} = 0$ for geodesics, one finds immediately that such curves have "constant speed": $|\dot{\gamma}| = \text{const}$.

Let us compute the geodesics of the unit sphere. Such geodesics must satisfy $|\gamma(t)| = 1$ for all t and, in order to be normal,

$$\ddot{\gamma}(t) = \lambda(t)\gamma(t)$$

for some real-valued function λ. On the sphere, we can write, since $\gamma^T \dot{\gamma} = 0$,

$$0 = \partial_t \gamma^T \dot{\gamma} = |\dot{\gamma}|^2 + \lambda(t)|\gamma|^2,$$

which implies (because $\dot{\gamma}$ is constant along geodesics and $|\gamma| = 1$) that λ is constant. So geodesics must satisfy the equation $\ddot{\gamma} = \lambda\gamma$. By making a constant time change, we can assume that $|\dot{\gamma}| = -\lambda(t) = 1$, and that γ is parametrized by arc length. Since $\partial\ddot{\gamma} = \dot{\gamma}$, we see that the curve has unit curvature and zero torsion and therefore coincides with a portion of unit circle. The only unit circles included in the sphere must be centered at 0, and constitute the great circles on the sphere. So we find that geodesics on the sphere are great circles parametrized at constant speed.

Finally, we note that the first fundamental form also determines the area form used in the computation of integrals over the surface. Indeed, one can easily check that $|\partial_1 m \times \partial_2 m| = \sqrt{EG - F^2}$ (both terms are equal to $|\partial_1 m|\,|\partial_2 m|\,|\sin\theta|$ where θ is the angle between the two tangent vectors) so that

$$d\sigma_S = \sqrt{EG - F^2}\, du\, dv. \tag{3.7}$$

3.7.3 The Divergence Theorem on Surfaces

A vector field on S is a function $h : S \to \mathbb{R}^3$ such that, for all p, $h(p) \in T_pS$. We start with a simple definition of the divergence of a C^1 vector field.

Definition 3.17 Let h be a C^1 vector field on a regular surface S. The divergence of h on S is defined by

$$\mathrm{div}_S h(p) = {e_1}^T dh(p) e_1 + {e_2}^T dh(p) e_2 \tag{3.8}$$

whenever e_1, e_2 is a positively oriented orthonormal basis of T_pM (the result being independent of the choice made for e_1, e_2).

In this definition, $dh(p)$ is a linear transformation between T_pS and $\mathbb{R}^3$. If h is defined on S and takes values in $\mathbb{R}^3$ (not necessarily in TS), the definition remains meaningful. We will use the notation $\mathrm{div}'_S(h)$ for the left-hand side of (3.8) in that case. In fact, if h decomposes as $h = h_T + \mu N$ where h_T is a vector field on S, we have

$$\mathrm{div}'_S(h) = \mathrm{div}_S(h_T) + \mu \mathrm{div}'_S(N). \tag{3.9}$$

Another way of understanding the definition is by introducing the orthogonal projection on T_pS (denoted π_{T_pS}) and the operator

$$\nabla_S h(p) = \pi_{T_pS} \circ dh(p) : T_pS \to T_pS. \tag{3.10}$$

This operator is the covariant derivative on S, as described in Appendix B, and Definition 3.17 simply says that

$$\mathrm{div}_S h(p) = \mathrm{trace}(\nabla_S h(p)). \tag{3.11}$$

Note that we have, for $\xi \in T_pS$

$$\nabla_S h(p)\xi = dh(p)\xi - ((dh(p)\xi)^T N)N.$$

This definition can be made explicit in a chart. This yields the following proposition (the proof, which is just a computation, is left to the reader):

Proposition 3.18 *If m is a local chart on S and the C^1 vector field h decomposes as $h = \alpha\partial_1 m + \beta\partial_2 m$ in this chart, we have*

$$\operatorname{div}_S h = \partial_1\alpha + \partial_2\beta + (\alpha\partial_1\rho + \beta\partial_2\rho)/\rho, \tag{3.12}$$

where $\rho = |\partial_1 m \times \partial_2 m| = \sqrt{EG - F^2}$.

We also have the nice formula, still valid in a chart, that says that

$$\partial_1 h \times \partial_2 m + \partial_1 m \times \partial_2 h = \rho(\operatorname{div}_S h)N. \tag{3.13}$$

This result is a direct consequence of the following simple computation in linear algebra, the proof of which is left to the reader.

Lemma 3.19 *Let A be a linear operator from M, an oriented two-dimensional linear subspace of $\mathbb{R}^3$, to $\mathbb{R}^3$. Let n be the unit normal to M. Define, for $e_1, e_2 \in M$,*

$$\phi_A(e_1, e_2) = (Ae_1)^T(e_2 \times n) + (Ae_2)^T(n \times e_1).$$

Then, there exists a real number $\rho(A)$ such that

$$\phi_A(e_1, e_2) = \rho(A)\det(e_1, e_2, n),$$

which is also equal to $\rho(A)|e_1 \times e_2|$ if e_1, e_2 are positively oriented. Moreover, we have

$$\rho(A) = \operatorname{trace}((\mathrm{Id} - nn^T)A), \tag{3.14}$$

where $(\mathrm{Id} - nn^T)A$ (which is A followed by the projection on M) is considered as an operator from M to itself.

Equation (3.13) just comes by applying Lemma 3.19 with $M = T_pM$, $A = dh(p)$, $e_1 = \partial_1 m$ and $e_2 = \partial_2 m$.

We now give the divergence theorem on a surface, which is a direct generalization of the one we saw on $\mathbb{R}^2$ (Theorem 1.16):

Theorem 3.20 *Let S be an oriented regular surface, and h a smooth vector field on S. Then, if $\Sigma \subset S$ is a bounded subdomain of S with a regular boundary, we have*

$$\int_{\partial\Sigma} h^T N_{\partial\Sigma}\, d\sigma_{\partial\Sigma} = -\int_{\Sigma} \operatorname{div}_S(h)\, d\sigma_S,$$

where the first integral is a line integral over the curve $\partial\Sigma$, and $N_{\partial\Sigma}$ is the inward normal to Σ (normal to $\partial\Sigma$ and tangent to S).

The proof (which we skip) is an application of Green's formula in $\mathbb{R}^2$ combined with a decomposition in local coordinates.

In addition to the divergence, one can define the gradient operator on a surface S, which applies to scalar-valued functions.

Definition 3.21 Let $f : S \to \mathbb{R}$ be C^1. The gradient of f at $p \in S$ is denoted $\nabla_S f(p)$ and defined by $\nabla_S f(p) \in T_p S$ and

$$\forall \xi \in T_p S, \quad \left\langle \nabla_S f(p)\,,\, \xi \right\rangle_p = df(p)\,\xi. \tag{3.15}$$

Note that, even if they are using the same symbol ∇_S, the covariant derivative introduced in (3.10) and the gradient in (3.15) are similar, but different notions, since the former applies to vector fields on S and the latter to scalar functions. Their similarity (and some justification for the notation conflict) is supported by the following observation: if f is the restriction to S of a differentiable function $\hat{f}$ defined on $\mathbb{R}^3$, then (3.15) implies that $\nabla_S f$ is the orthogonal projection of $\nabla \hat{f}$ (the usual gradient in $\mathbb{R}^3$) on the tangent plane to S, namely

$$\nabla_S f(p) = \pi_{T_p S} \nabla \hat{f}(p) = \nabla \hat{f}(p) - (N(p)^T \nabla \hat{f}(p)) N(p). \tag{3.16}$$

In a chart $(u, v) \mapsto m(u, v)$, we have

$$\nabla_S f = \frac{G\partial_1 f - F\partial_2 f}{EG - F^2}\partial_1 m + \frac{E\partial_2 f - F\partial_1 f}{EG - F^2}\partial_2 m. \tag{3.17}$$

The usual formula, $\mathrm{div}(fh) = \nabla f^T h + f\,\mathrm{div} h$, extends to surfaces with

$$\mathrm{div}_S(fh) = \nabla_S f^T h + f\,\mathrm{div}_S h \tag{3.18}$$

for a scalar function f and a vector field h on S.

The generalization of the Laplacian on $\mathbb{R}^2$ is the Laplace–Beltrami operator on S. It is defined as follows:

Definition 3.22 The Laplace–Beltrami operator on a regular surface S associates to a scalar function f on S the scalar function $\Delta_S f$ defined by

$$\Delta_S f = \mathrm{div}_S \nabla_S f. \tag{3.19}$$

The Laplace–Beltrami operator in a chart is therefore given by the combination of (3.17) and (3.12), which yields a formula notably more complex than the ordinary Laplacian.

Theorem 3.20 relates surface integrals to linear integrals over the surface. Surface integrals can also be related to three-dimensional integrals, if the surface is closed, via the three-dimensional divergence theorem.

Theorem 3.23 *Let Ω be a bounded domain in $\mathbb{R}^3$ and assume that $S = \partial\Omega$ is a regular surface. If v is a C^1 vector field on $\mathbb{R}^3$, we have*

$$\int_\Omega \operatorname{div} v \, dx = -\int_S v^T(m) N(m) \, d\sigma_S(m), \tag{3.20}$$

where $N(m)$ is the inward normal to S at m.

From this theorem, we can derive an expression of the volume of Ω as an integral over its boundary, namely (taking $v(x, y, z) = (x, y, z)$ in (3.20))

$$\text{volume}(\Omega) = -\frac{1}{3}\int_S Om^T N(m) \, d\sigma_S(m). \tag{3.21}$$

3.8 Curvature and the Second Fundamental Form

Let S be a C^2 orientable regular surface, and N be its unit normal. The function N can be seen as a map defined on S with values in $\mathbb{R}^3$ (in fact in the unit sphere S^2 since $|N| = 1$), which is called the Gauss map. It therefore has a differential, dN. For any $p \in S$, $dN(p)$ is a linear map from T_pS to $\mathbb{R}^3$. The fact that $|N|^2 = 1$ implies that $(dN(p)h)^T N(p) = 0$ for all $h \in T_pS$ so that the range of $dN(p)$ is orthogonal to $N(p)$ and therefore coincides with T_pS. We can therefore consider $dN(p)$ as an endomorphism (a linear map from a vector space into itself)

$$dN(p) : T_pS \to T_pS.$$

This endomorphism (also called the *shape operator*) is essential for describing the curvature of the surface, which measures how the surface bends in a neighborhood of a point p. It has the interesting property of being symmetric:

Proposition 3.24 *Let S be a regular surface and $p \in S$: for any $h, k \in T_pS$, we have*

$$\langle dN(p)h \, , \, k \rangle_p = \langle h \, , \, dN(p)k \rangle_p.$$

Proof It suffices to show this for a basis of T_pS. Let us take the one provided by a local parametrization around p: $h = \partial_1 m$ and $k = \partial_2 m$. Let $N_m = N \circ m$ be the expression of N as a function of the parameters, so that

$$dN(p)(\alpha\partial_1 m + \beta\partial_2 m) = \alpha\partial_1 N_m + \beta\partial_2 N_m.$$

In particular, $dN(p)\partial_1 m = \partial_1 N_m$ and $dN(p)\partial_2 m = \partial_2 N_m$, and what we need to show is

$$\langle \partial_1 N_m \, , \, \partial_2 m \rangle_p = \langle \partial_1 m \, , \, \partial_2 N_m \rangle_p.$$

But, from $(\partial_1 m)^T N_m = 0$, we get $\langle \partial_1 m\ ,\ \partial_2 N_m \rangle_p = (\partial_1 m)^T \partial_2 N_m = -(\partial_2 \partial_1 m)^T N_m$. Similarly, $\langle \partial_2 m\ ,\ \partial_1 N_m \rangle_p = -(\partial_1 \partial_2 m)^T N_m$. Since partial derivatives commute, the two quantities are equal, yielding the required identity. □

Let γ be a curve on S, and assume that γ is parametrized by arc length. Let $T^{(\gamma)}$ be the unit tangent of γ, $\kappa^{(\gamma)}$ its curvature and $N^{(\gamma)}$ its unit normal, such that $\dot{T}^{(\gamma)} = \kappa^{(\gamma)} N^{(\gamma)}$. The normal $N^{(\gamma)}$ does not coincide with N in general, and we define the normal curvature of γ by the (algebraic) normal part of $\dot{T}^{(\gamma)}$ to the surface S. The interesting point is that it only depends on γ via $T^{(\gamma)}$.

Definition 3.25 The normal curvature at p of an arc length parametrized curve γ on a regular surface S is $\kappa_N^{(\gamma)}(s) = (\dot{T}^{(\gamma)}(s))^T N(\gamma(s))$, where $T^{(\gamma)} = \dot{\gamma}$.

The fact that the normal curvature only depends on $T^{(\gamma)}$ can be proved as follows: let γ be a curve on S such that $\dot{\gamma}(0) = T^{(\gamma)}$. For all s, we have $(T^{(\gamma)})^T N = 0$ since $T^{(\gamma)}$ is tangent to S. Computing the derivative with respect to arc length and applying the chain rule yields

$$(\dot{T}^{(\gamma)})^T N \circ \gamma + (T^{(\gamma)})^T dN(\gamma) T^{(\gamma)} = 0$$

so that

$$\kappa_N^{(\gamma)} = -(T^{(\gamma)})^T dN(\gamma) T^{(\gamma)}. \tag{3.22}$$

One also defines the *geodesic curvature* of γ at s_0 by the curvature (at s_0) of the projection of γ on the tangent plane to S at $\gamma(s_0)$, which is

$$\bar{\gamma}(s) = \gamma(s) - (\gamma(s) - \gamma(s_0))^T N(s_0)\, N(s_0).$$

Computing first and second derivatives in s and computing them at $s = s_0$ yields $\dot{\bar{\gamma}}(s_0) = \dot{\gamma}(s_0)$ and

$$\ddot{\bar{\gamma}}(s_0) = \ddot{\gamma}(s_0) - \kappa_N^{(\gamma)}(s_0) N(s_0).$$

Denoting the geodesic curvature by $\kappa_g^{(\gamma)}(s_0)$, we find (using the definition of the (signed) curvature for plane curves in the oriented tangent plane) that

$$\kappa_g^{(\gamma)} = \det(\dot{\gamma}, \ddot{\gamma}, N) = \ddot{\gamma}^T (N \times \dot{\gamma}),$$

where $N \times \dot{\gamma}$ is the unit normal to γ that belongs to $T_\gamma M$ and complements $\dot{\gamma}$ in a positively oriented basis of the tangent plane. Writing $\ddot{\gamma} = (\ddot{\gamma}^T (N \times \dot{\gamma}))(N \times \dot{\gamma}) + (\ddot{\gamma}^T N)$, one also gets the identity

$$(\kappa_g^{(\gamma)})^2 + (\kappa_N^{(\gamma)})^2 = (\kappa^{(\gamma)})^2,$$

the squared curvature of γ.

This expression in Eq. (3.22) involves another important quantity on S, its second fundamental form.

Definition 3.26 Let S be a regular surface and $p \in S$. The second fundamental form at p is the quadratic form defined on T_pS by

$$II_p(h) = -\langle h\,,\, dN(p)h\rangle_p.$$

In particular, we have the expression of the normal curvature of an arc length parametrized curve γ:

$$\kappa_N^{(\gamma)} = II_\gamma(\dot{\gamma}).$$

Because $dN(p)$ is symmetric, it can be diagonalized in an orthonormal basis of T_pS: let (e_1, e_2) be such a basis, with corresponding eigenvalues $-\kappa_1$ and $-\kappa_2$ such that $\kappa_1 \geq \kappa_2$. The numbers κ_1 and κ_2 are called the *principal curvatures* of the surface at p. The reason for this terminology is that any unit vector in T_pS can be written, for some θ, in the form $h = \cos\theta e_1 + \sin\theta e_2$ and

$$II_p(h) = -\langle h\,,\, dN(p)h\rangle = \kappa_1\cos^2\theta + \kappa_2\sin^2\theta.$$

This implies that $\kappa_2 \leq II_p(h) \leq \kappa_1$, the lower bound being attained for $h = e_2$ and the upper bound for $h = e_1$: κ_1 and κ_2, respectively, are the maximum and minimum normal curvatures of curves passing through p.

Definition 3.27 If κ_1 and κ_2 are the principal curvatures of a surface S at $p \in S$, one defines the mean curvature at p by $H(p) = (\kappa_1 + \kappa_2)/2$, and the Gauss curvature by $K(p) = \kappa_1\kappa_2$. They respectively coincide with the trace of $-dN(p)/2$ and the determinant of $dN(p)$.

From this definition, we can also write

$$2H = -\mathrm{div}'_S(N) \tag{3.23}$$

and rewrite (3.9) as (for $h = h_T + \mu N$)

$$\mathrm{div}'_S(h) = \mathrm{div}_S(h_T) - 2\mu H. \tag{3.24}$$

3.9 Curvature in Local Coordinates

In this section, we give the expression of the curvature in local coordinates, as functions of the coefficients of the first and second fundamental forms. Recall the notation (3.6) for the first fundamental form and a local parametrization m. We introduce a similar notation for the second form, letting

$$II_p(\alpha\partial_1 m + \beta\partial_2 m) = \alpha^2 e + 2\alpha\beta f + \beta^2 g$$

and

$$e = -\partial_1 m^T \partial_1 N = \partial_1^2 m^T N,\ f = -\partial_1 m^T \partial_2 N = \partial_1 \partial_2 m^T N,$$
$$g = -\partial_2 m^T \partial_2 N = \partial_2^2 m^T N. \tag{3.25}$$

Let $dN = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ in the basis $(\partial_1 m, \partial_2 m)$ (the matrix is not necessarily symmetric since the basis is not assumed to be orthonormal). We find:

$$\begin{aligned} -e &= \partial_1 m^T dN \partial_1 m = aE + bF \\ -f &= \partial_2 m^T dN \partial_1 m = aF + bG \\ -f &= \partial_1 m^T dN \partial_2 m = cE + dF \\ -g &= \partial_2 m^T dN \partial_2 m = cF + dG \end{aligned}$$

which yields, in matrix form: $-\begin{pmatrix} e & f \\ f & g \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. This implies that, in the basis $(\partial_1 m, \partial_2 m)$, dN is given by the matrix

$$-\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} e & f \\ f & g \end{pmatrix}.$$

From this, it can be deduced that

$$K = \frac{eg - f^2}{EG - F^2}$$

because it is just the ratio of the determinants. Also, after computation, one finds

$$H = \frac{eG - 2fF + gE}{2(EG - F^2)}.$$

The principal curvatures are then given by $\kappa = H \pm \sqrt{H^2 - K}$.

3.10 Implicit Surfaces

Assume that a surface is defined implicitly by

$$S = \left\{ p \in \mathbb{R}^3,\ f(p) = 0 \right\},$$

where f is a C^2 function from $\mathbb{R}^3$ to $\mathbb{R}$ with $\nabla f \neq 0$ on S. We have already noticed that the tangent plane to S is orthogonal to ∇f, and therefore $N(p) = -\nabla f(p)/|\nabla f(p)|$ is a smooth unit normal to S which therefore is orientable (and we take the orientation provided by this choice of N).

The interesting feature in this representation is that, since f is defined on $\mathbb{R}^3$, the function N can be extended to $\mathbb{R}^3$ (denote the extension by $\hat{N}$) so that $dN(p)$

is simply the restriction to T_pS of $d\hat{N}(p)$, In particular, the trace of $dN(p)$ is, by definition, $\langle e_1\,,\, dN(p)e_1\rangle_p + \langle e_2\,,\, dN(p)e_2\rangle_p$ for an arbitrary orthonormal basis of T_pS. It therefore suffices to add $(d\hat{N}N)^T N$ to obtain the trace of $d\hat{N}$, but this added quantity vanishes because $|\hat{N}|^2 = 1$ implies that $d\hat{N}N$ is perpendicular to N. Thus, we have, for the mean curvature:

$$H = -\text{trace}(d\hat{N})/2 = \frac{1}{2}\text{div}\frac{\nabla f}{|\nabla f|}. \tag{3.26}$$

(This is the usual divergence on $\mathbb{R}^3$, not to be confused with the S-divergence in Definition 3.17.)

Let P_N be the projection on $\hat{N}^\perp$: $P_N = \text{Id}_{\mathbb{R}^3} - \hat{N}\hat{N}^T$. The Gauss curvature can be computed after diagonalizing the matrix $P_N d\hat{N} P_N = d\hat{N} P_N$, which is symmetric and coincides with dN on T_pS. Using $\hat{N} = -\nabla f/|\nabla f|$, we get

$$\begin{aligned}(d\hat{N}P_N h)^T P_N k &= -\frac{1}{|\nabla f|}(d^2 f P_N h)^T P_N k + \frac{1}{|\nabla f|}\left((d^2 f P_N h)^T \hat{N}\right)\left((P_N k)^T N\right)\\ &= -\frac{1}{|\nabla f|}(P_N h)^T d^2 f P_N k,\end{aligned}$$

which is symmetric in h and k. The matrix $P_N d^2 f P_N/|\nabla f|$ has one vanishing eigenvalue since $P_N N = 0$, and the other two are the principal curvatures of S. Their product provides the Gauss curvature.

The Delta-Function Trick

When a surface or a curve is defined implicitly, integrals over its interior can be described in a straightforward way using the Heaviside function. Assume that S is the set $f(p) = 0$ for some smooth function f, and let Ω be its interior, defined by $f < 0$. Introduce the Heaviside function H_0 defined on $\mathbb{R}$ by $H_0(x) = 1$ if $x \geq 0$ and $H_0(x) = 0$ otherwise. Then, clearly, for any function V on $\mathbb{R}^3$, we have

$$\int_\Omega V(x)dx = \int_{\mathbb{R}^3}(1 - H_0(f(x)))V(x)dx. \tag{3.27}$$

Contour or surface integrals can be defined via a level-set representation, albeit requiring passing to a limit. For this, we need to replace H_0 by a smooth approximation denoted H_ε, which must be an increasing function that tends to H_0 as ε tends to 0. A possible example is (cf. [227, 316]) $H_\varepsilon(x) = 0$ for $x < -\varepsilon$, $H_\varepsilon(x) = 1$ for $x > \varepsilon$ and, on $[-\varepsilon, \varepsilon]$:

$$H_\varepsilon(x) = \frac{1}{2}\left(1 + \frac{x}{\varepsilon} + \frac{1}{\pi}\sin\left(\frac{\pi x}{\varepsilon}\right)\right). \tag{3.28}$$

Alternatively [58], one can take, for all $x \in \mathbb{R}$:

$$H_\varepsilon(x) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{x}{\varepsilon}\right)\right). \tag{3.29}$$

This choice being made, let δ_ε denote the derivative of H_ε. The function δ_ε can be considered as a smooth approximation of the Dirac function δ_0, in the sense that, for any bounded function u on $\mathbb{R}$ which is continuous at $t = 0$, one has

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}} \delta_\varepsilon(t)u(t)dt = u(0). \tag{3.30}$$

We leave the easy proof to the reader (simply divide the integral over domains around 0 or away from 0).

We now describe how surface integrals over implicitly defined surfaces can be approximated using δ_ε.

Proposition 3.28 *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a C^2 function with $\nabla f \neq 0$ if $f = 0$, and such that the implicit surface $S = f^{-1}(\varepsilon)$ is bounded in a neighborhood of 0. Then, if $V : \mathbb{R}^3 \to \mathbb{R}$ is continuous, we have*

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^3} \delta_\varepsilon \circ f(x)\, V(x)\, |\nabla f(x)|\, dx = \int_S V(m)\, d\sigma_S(m). \tag{3.31}$$

The same proposition holds for curves, with $f : \mathbb{R}^2 \to \mathbb{R}$ and the surface integral replaced by the integral along the curve.

Proof Let's consider the surface case (the case of curves is similar and simpler). We also assume that δ_ε is supported in $[-\varepsilon, \varepsilon]$, like for (3.28) (the general case requiring only minor modifications). Consider a local chart $(u, v) \mapsto m(u, v)$ on $S = f^{-1}(0)$. Consider the equation

$$f(m(u, v) + tN(u, v)) = y,$$

which we want to solve for t as a function of (u, v, y) in a neighborhood of some $u = u_0$, $v = v_0$ and $y = 0$. From the implicit function theorem, this is possible, because

$$\partial_t f(m + tN) = \nabla f^T N = -|\nabla f|,$$

which is not zero by assumption. Using the compactness of S, we can find a finite number of points $p_0 = m(u_0, v_0)$ and domains around $(u_0, v_0, 0) \in \mathbb{R}^3$ over which a function $t(m(u, v), y)$ such that $f(m + tN) = y$ is well-defined and such that the union of these domains forms an open set in $\mathbb{R}^3$ that contains S, and more generally contains the set $|f(p)| < y_0$ for y_0 small enough.

Taking $\varepsilon < y_0$, we can write

$$\int_{\mathbb{R}^d} \delta_\varepsilon \circ f(x)\, V(x)\, |\nabla f(x)|\, dx = \int_{|f|<y_0} \delta_\varepsilon \circ f(x)\, V(x)\, |\nabla f(x)|\, dx.$$

(Not assuming δ_ε to be compactly supported would add a small error to this identity, which is easily shown to be negligible when $\varepsilon \to 0$.)

We can decompose the integral over a partition of unity, which reduces the problem to the situation in which V is supported by one of the domains above. Working under this assumption, we make the change of variables $x(u, v, y) = m(u, v) + t(m(u, v), y)N(u, v)$ in this domain and let $J(u, v, t)$ be the associated Jacobian determinant, so that

$$\int_{|f|<y_0} \delta_\varepsilon \circ f(x) V(x) \, |\nabla f(x)| \, dx = \int_{|y|<y_0} \delta_\varepsilon(y) \, V(x(u, v, y)) \, |\nabla f(x(u, v, y))| \, J(u, v, y) \, du \, dv \, dy.$$

Our assumptions ensure that the integral

$$u(y) = \int V(x(u, v, y)) \, |\nabla f(x(u, v, y))| \, J(u, v, y) \, du \, dv$$

is continuous in y so that,

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} \delta_\varepsilon \circ f(x) \, V(x) \, |\nabla f(x)| \, dx = u(0).$$

Now,

$$J(u, v, 0) = |\det(\partial_1 m, \partial_2 m, \partial_3 t N)| = |\partial_1 m \times \partial_2 m| / |\nabla f(m)|$$

because $y = f(m + tN)$ implies $1 = \partial_3 t \, \nabla f^T N = -\partial_3 t |\nabla f|$. This implies that the $|\nabla f|$ terms cancel in the expression of $u(0)$, which is equal to

$$u(0) = \int V(m(u, v)) \, |\partial_1 m \times \partial_2 m| \, du \, dv = \int_S V \, d\sigma_S,$$

which concludes the proof. □

The theorem is particularly important for numerical computations, because it replaces computations over a surface with computations over a grid that contains the surface.

The left-hand side of (3.31) is often written using the symbolic notation

$$\int_{\mathbb{R}^2} \delta_0 \circ f(x) \, V(x) \, |\nabla f(x)| \, dx.$$

The assumption that V is continuous is important (of course, we only need continuity near $f^{-1}(0)$). Take the following simple example with curves; let $f(u, v) = u^2 + v^2 - 1$, so that $f^{-1}(0) = S^1$, the unit circle and let $V(u, v) = 1$ if $u^2 + v^2 \leq 1$ and 0 otherwise. Then

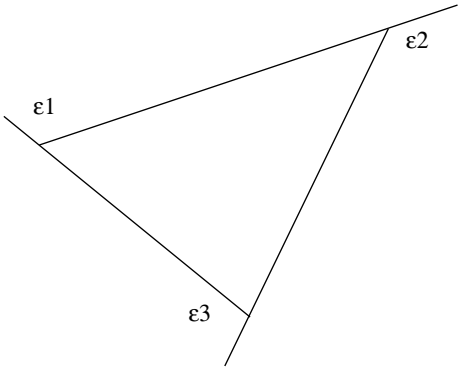

Fig. 3.6 The Gauss–Bonnet theorem in $\mathbb{R}^2$ reduces to the well-known property that the sum of consecutive angles in a polygon is 2π

$$\int_{S^1} V dl = 2\pi$$

but

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}^2} \delta_\varepsilon \circ f(x) V(x) |\nabla f(x)| dx = \pi$$

(both integrals being easily computed in radial coordinates).

3.11 The Gauss–Bonnet Theorem

The average of the Gauss curvature over a domain with piecewise geodesic boundary is provided by the Gauss–Bonnet formula [86]:

Theorem 3.29 *Let S be a regular surface and A be a domain on M such that ∂A is the union of N geodesics $\gamma^{(1)}, \ldots, \gamma^{(N)}$. Let $\varepsilon_i, i = 1, \ldots N$ be the sequence of consecutive angles between the curves at their intersection. Then*

$$\int_A K d\sigma = 2\pi - \sum_{i=1}^{N} \varepsilon_i. \tag{3.32}$$

For example, when $N = 3$ (∂A is a "geodesic triangle"), we obtain the fact that the sum of the angles of a triangle is 2π minus the integral of the Gauss curvature over its interior. This is consistent with the sum being 2π in the plane, which has zero Gauss curvature (Fig. 3.6).

Chapter 4
Computations on Triangulated Surfaces

4.1 Triangulated Surfaces

4.1.1 Definition and Notation

Triangulated surfaces provide a three-dimensional generalization of polygons in two dimensions. Surfaces are usually stored on computers in this form, and these are the kinds of objects that must be handled in practical applications.

In full generality, a triangulated surface is a set of *vertices* $\mathcal{V} = \{v_1, \ldots, v_M\}$ with a family of 3-tuples of indices $\mathcal{F} = \{f_1, \ldots, f_K\}$, where each f_k takes the form $f_k = (j_{k1}, j_{k2}, j_{k3}) \in \{1, \ldots, M\}$. One associates to f_k the triangle (or face) in the triangulation defined by $F_k = (v_{k1}, v_{k2}, v_{k3})$, using the abbreviated notation $v_{kl} := v_{j_{kl}}$. The set of edges of the triangulation is the family of unordered pairs of vertices which belong to the same face and will be denoted by $\mathcal{E} = \{e_1, \ldots, e_Q\}$.

The order of the vertices in each face is important and defines its orientation, which is invariant up to a cyclic permutation of the vertices. We will only consider *regular* triangulations, which are such that the intersection of two faces is either empty or an edge. This excludes those situations in which the contact between two faces occurs at a vertex only, or in which some vertex belongs to the interior of an edge. The number

$$\chi = |\mathcal{V}| - |\mathcal{E}| + |\mathcal{F}|$$

is a topological invariant of the surface called the Euler characteristic.

For a vertex v_i, we let $\mathcal{F}_i$ denote the set of indexes of faces that contain it, and $\mathcal{E}_i$ the set of indexes of edges that contain it. We also let $\mathcal{V}_i$ denote the set of indexes of vertices (distinct from v_i) that belong to one of the 3-tuples in $\mathcal{F}_i$. $(\mathcal{V}_i, \mathcal{E}_i, \mathcal{F}_i)$ represents the neighborhood of v_i in the triangulation.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_4

The triangulation is said to be consistent if, whenever two faces intersect, their common edge is ordered in different directions in the two faces. A consistent triangulation is the equivalent of an oriented surface. We only consider consistent triangulations in the following.

4.2 Estimating the Curvatures

Given a triangulated surface, the next step is to compute differential descriptors, and in particular discrete forms of the curvatures. We address this problem in this section, focusing on a few important methods that have recently emerged in the literature.

4.2.1 Taylor Expansions

The unit normal to an oriented triangle (v_1, v_2, v_3) is the vector

$$N = \frac{(v_2 - v_1) \times (v_3 - v_1)}{|(v_2 - v_1) \times (v_3 - v_1)|}.$$

Each face F_k in the triangulation therefore carries a uniquely defined normal, N_k^f. We can associate the normal to a specific point inside the face, for example its centroid $(v_1 + v_2 + v_3)/3$. (There are several possible definitions of the center of a triangle, however, including the circumcenter, which is the center of the circumscribed circle, the incenter, which is the center of the inscribed center, or the orthocenter, the intersection of the lines passing through the vertices and orthogonal to the opposite edge.)

In many cases, one also wants to define normals at the vertices. This can be done using a weighted average of the normals at the neighboring faces. If v_i is a vertex, define

$$N_i^v = \frac{\sum_{k \in \mathcal{F}_i} w_i(F_k) N_k^f}{|\sum_{k \in \mathcal{F}_i} w_i(F_k) N_k^f|},$$

where $w_i(F_k)$ gives a measure of the "importance" of face F_k relative to vertex v_i. The simplest definition is the area, $\text{area}(F_k)$, independent in this case of the chosen vertex. In [194], it is suggested to use the area of the part of the face which is closer to v_i than to any of the other two vertices. This is the intersection of the face F_k with the region delimited by the following four points: v_i, the two midpoints of the edges of F_k that contain v_i and the circumcenter of F_k. Notice that the circumcenter lies outside of F_k if the triangle is obtuse, as illustrated in Fig. 4.1. Such regions form Voronoï cells. Let F_{ki} denote the part of face F_k which is associated to v_i in this way. One can use $w_i(F_k) = \text{area}(F_{ki})$.

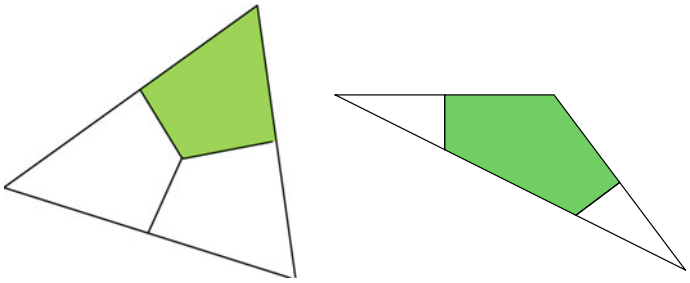

Fig. 4.1 Decomposition of triangles into Voronoï cells when the circumcenter is interior to the triangle (left) and when it is exterior (right)

Similarly, we can define a normal along an edge e to be a weighted average of the normals to the faces that intersect at e, using, for example, the areas of the faces as weights.

Having an estimation of the normal at each vertex allows for the approximation of the normal curvature of a curve on the surface passing through this vertex, which yields the second fundamental form. If $j \in \mathcal{V}_i$, the two-point path (v_i, v_j) provides a discrete curve fragment passing through v_i. Define the tangent vector $T_{ij} = (v_j - v_i)/|v_j - v_i|$. Using Definition 3.26, one possible approximation of the second fundamental form at the midpoint between v_i and v_j in the direction T_{ij} (which is also the normal curvature of the curve fragment at the midpoint) is

$$II_{ij} := -T_{ij}^T \left(\frac{N_j^v - N_i^v}{|v_j - v_i|} \right) = -\frac{(N_j^v - N_i^v)^T (v_j - v_i)}{|v_j - v_i|^2}.$$

Also, using a Taylor expansion (assuming that N^v is the restriction to the vertices of a smooth function), one can prove (the justification being left to the reader) that

$$\frac{(N_j^v + N_i^v)^T (v_j - v_i)}{|v_j - v_i|^2} = O(|v_j - v_i|),$$

and adding this expression to the previous estimate II_{ij} yields the alternative formula [273]

$$II_{ij} := \frac{2(N_i^v)^T (v_j - v_i)}{|v_j - v_i|^2}.$$

Because the matrix dN is symmetric in the tangent plane, it is described by three parameters in any orthonormal basis. Since each computation of the discrete second fundamental form yields one linear equation involving dN, this requires at least three edges for its estimation, which is the minimum number provided by the triangulation. One possible way to estimate curvatures is to select an arbitrary basis (a_i, b_i) of the tangent plane to the surface at v_i, $T_{v_i} M$, which is by definition the

plane perpendicular to N_i^v (for example, assuming that N_i^v is not parallel to the x-axis, take $a_i = (1, 0, 0)^T \times N_i^v$ and $b_i = N_i^v \times a_i$); then, compute, for each $j \in \mathcal{V}_i$, the coordinates (x_{ij}, y_{ij}) of the normalized orthogonal projection of T_{ij} onto this basis. We have

$$x_{ij} = a_i^T T_{ij} / \sqrt{(a_i^T T_{ij})^2 + (b_i^T T_{ij})^2}$$
$$\text{and } y_{ij} = b_i^T T_{ij} / \sqrt{(a_i^T T_{ij})^2 + (b_i^T T_{ij})^2}.$$

Then, letting $dN_i = \begin{pmatrix} \alpha_i & \gamma_i \\ \gamma_i & \beta_i \end{pmatrix}$ in this basis, we have the system of linear equations

$$\alpha_i x_{ij}^2 + 2\gamma_i x_{ij} y_{ij} + \beta_i y_{ij}^2 = -II_{ij}, \ \ j \in \mathcal{V}_i.$$

This is an over-constrained system, for which one can compute a least-squares solution. Once dN_i is computed, its trace, determinant and eigenvalues provide an estimation of the mean, Gaussian and principal curvatures.

A more direct approach to estimating the curvature from the second fundamental form has been proposed in [273]. Introduce, for continuous surfaces, the matrix (defined at a point p in the surface)

$$\Sigma_p = \frac{1}{2\pi} \int_0^{2\pi} \kappa_N(T_\theta) T_\theta T_\theta^T d\theta,$$

where T_θ is the rotation (within the tangent plane) of an arbitrary reference vector $T \in T_pM$ by an angle θ. A direct computation of this integral (using the basis $(T, T_{\pi/2})$) shows that

$$\Sigma_p = \frac{3}{8} dN(p) - \frac{1}{8} \det(dN(p)) dN(p)^{-1},$$

the last term being the adjugate matrix of $dN(p)$ (therefore also defined when $dN(p)$ is singular).

This implies that the eigenvalues of Σ_p are $\lambda_1 = -(3\kappa_1(p) - \kappa_2(p))/8$ and $\lambda_2 = -(3\kappa_2(p) - \kappa_1(p))/8$ (which can be used to compute the curvatures), and that the eigenvectors of Σ_p coincide with those of $dN(p)$ and therefore provide the principal directions.

Returning to the discrete case, the curvatures at vertex v_i can therefore be estimated from an approximation Σ_i of Σ_{v_i}. Such an approximation is provided by the simple formula

$$\Sigma_i = -\sum_{j \in \mathcal{V}_i} w_{ij} II_{ij} \, T_{ij} T_{ij}^T / \sum_{j \in \mathcal{V}_i} w_{ij},$$

where $w_{ij} = (w_i(F_j^+) + w_i(F_j^-))/2$, F_j^+ and F_j^- being the faces that contain the edge $\{v_i, v_j\}$.

4.2.2 Gauss–Bonnet and Area Minimization

In the previous section, the curvature computations were based on Taylor expansions of formulas that apply on smooth surfaces. More recently [85], an increased focus has been made on obtaining expressions that derive from intrinsic properties of surfaces that can be extended to polyhedral surfaces.

The right-hand side of Eq. (3.32) in the Gauss–Bonnnet theorem can still be defined on polyhedral surfaces. This fact is used in [194] to provide an approximation of the Gauss curvature, using, for a vertex v_i in the triangulation, the region A_i formed by the union of the Voronoï cells around v_i (Fig. 4.1). The expression is very simple because, in both cases in Fig. 4.1, the sum of the (one or two) exterior angles in each part of A_i coincides with the angle of the corresponding face at v_i. For $k \in G_i$, denoting by θ_{ik} the angle of the face F_k at v_i, we see that the right-hand side of (3.32) is given by $2\pi - \sum_{k\in\mathcal{F}_i} \theta_{ik}$. Approximating K by a constant over A_i, we get the formula

$$K_i = \frac{1}{|A_i|}\Big(2\pi - \sum_{k\in\mathcal{F}_i} \theta_{ik}\Big).$$

The area, $|A_i|$, can be computed in closed form. It is the sum of the areas of the shaded regions in Fig. 4.1, over all faces that contain v_i. Let as above θ_{ik} be the angle at v_i for a face F_k. Let v'_{ik} and v''_{ik} be the other two vertices of F_k so that v_i, v'_{ik} and v''_{ik} are ordered consistently with the orientation of F_k. Let e'_{ik} be the edge opposite v'_{ik} in F_k and e''_{ik} the edge opposite v''_{ik} (we will later denote by e_{ik} the edge opposite v_i). Finally, let θ'_{ik} and θ''_{ik} respectively denote the angles at v'_{ik} and v''_{ik}. Then, the area, a_{ik}, of the shaded region in Fig. 4.1 in the acute case is given by

$$a_{ik} = \frac{1}{8}(|e'_{ik}|^2 \operatorname{ctn}(\theta'_{ik}) + |e''_{ik}|^2 \operatorname{ctn}(\theta''_{ik})).$$

In the obtuse case, and if θ_{ik} is the obtuse angle,

$$a_{ik} = \frac{1}{2}|e'_{ik}|\,|e''_{ik}| \cos\theta_{ik} - \frac{1}{8}(|e'_{ik}|^2 \cos\theta''_{ik} + |e''_{ik}|^2 \cos\theta'_{ik}).$$

Finally, still in the obtuse case, and when θ_{ik} is one of the acute angles,

$$\frac{1}{8}|\tilde{e}_{ik}|^2 \cos\theta_{ik}$$

where $\tilde{e}_{ik}$ is the side opposed to the other acute vertex.

Given this, $|A_i|$ is the sum of these areas over $k \in G_i$. When there is no obtuse triangle around v_i, the area $|A_i|$ has another simple expression [194]. For $l \in \mathcal{E}_i$ (edges stemming from v_i), let α_{il} and β_{il} be the angles at vertices opposed to e_l in the triangles that intersect at e_l. Then

$$|A_i| = \frac{1}{8} \sum_{l \in \mathcal{E}_i} (\operatorname{ctn} \alpha_{il} + \operatorname{ctn} \beta_{il}) |e_l|^2.$$

To address the mean curvature, we first use an important interpretation of it as a "gradient" of the surface area. Let S be a surface and $h : S \to \mathbb{R}^3$ be a (smooth) vector field on S. Assume that $h = 0$ on the boundary of S (if S has one). Define the surface S_ε as the one obtained by displacing each $p \in S$ along the vector $\varepsilon h(p)$. Then (this will be proved in Proposition 5.4)

$$\partial \left(\operatorname{area}(S_\varepsilon) \right)_{|_{\varepsilon=0}} = 2 \int_S H(p)\, h(p)^T N(p)\, d\sigma_S(p).$$

One can make the same construction with a discrete surface Σ by associating to each vertex v_i a small displacement $\varepsilon h_i \in \mathbb{R}^3$, and computing the derivative of the area of the obtained surface Σ_ε. Approximating the right-hand side in the above formula, we will then identify:

$$\partial \left(\operatorname{area}(\Sigma_\varepsilon) \right)_{|_{\varepsilon=0}} = 2 \sum_{i=1}^{M} h_i^T (H_i N_i) |A_i|, \tag{4.1}$$

where A_i is the neighborhood attributed to v_i and $H_i N_i$ can then be interpreted as the discretized product of the mean curvature with the normal at v_i.

Given that the area of a triangle with vertices v_1, v_2, v_3 is given by the half-norm of the cross product $(v_2 - v_1) \times (v_3 - v_1)$, the left-hand side in (4.1) is

$$\begin{aligned} &\frac{1}{2} \sum_{k=1}^{K} ((h_{2k} - h_{1k}) \times (v_{k3} - v_{k1}) + (v_{k2} - v_{k1}) \times (h_{k3} - h_{k1}))^T N_k^f \\ &= \frac{1}{2} \sum_{k=1}^{K} (h_{k1} \times (v_{k2} - v_{k3}) + h_{k2} \times (v_{k3} - v_{k1}) + h_{k3} \times (v_{k2} - v_{k1}))^T N_k^f, \end{aligned}$$

where h_{k1}, h_{k2} and h_{k3} are the displacements associated with the vertices of F_k and N_k^f is the normal to F_k.

For $k \in \mathcal{F}_i$, let e_{ik} be the oriented edge opposite v_i. Using the relation $(x \times y)^T z = x^T (y \times z)$ and reordering the sums, we can write

$$\frac{d}{d\varepsilon} \operatorname{area}(\Sigma_\varepsilon)_{|_{\varepsilon=0}} = \frac{1}{2} \sum_{i=1}^{M} h_i^T \left(\sum_{k \in \mathcal{F}_i} e_{ik} \times N_k^f \right).$$

This provides a definition of the discrete mean curvature at v_i:

$$H_i N_i = \frac{1}{4|A_i|} \sum_{k \in \mathcal{F}_i} e_{ik} \times N_k^f.$$

Reordering this sum over edges and explicitly computing the cross product leads to the equivalent expression [194]

$$H_i N_i = \frac{1}{4|A_i|} \sum_{l \in \mathcal{E}_i} (\operatorname{ctn} \alpha_{il} + \operatorname{ctn} \beta_{il}) e_l,$$

where α_{il} and β_{il} are, as before, the angles at the vertices opposite to e_l in each of the faces that contain e_l (e_l being oriented from v_i to the other vertex).

Note that this computation provides an estimate of the normal and the mean curvature together.

4.2.3 Curvature Measures

The Smooth Case

There is another way to interpret curvature on a surface that can be generalized to the non-smooth case, leading to another formula for curvature approximation on triangulated surfaces. On smooth surfaces, this is related to the volume of so-called parallel sets. We first show that smooth surfaces have positive reach, in a discussion that parallels the one in Sect. 1.13.2. We use the same notation as in that section, letting, for a surface M,

$$d_M(p) = \operatorname{dist}(p, M) = \inf \{|p - q| : q \in M\},$$

$\mathcal{U}_M$ be the set of points p that have a unique closest point, $\pi_M(p)$, on M, $r(M, q)$ be the supremum of the radii of balls centered at q included in $\mathcal{U}_M$ and $r(M)$ be their minimum over $q \in M$ (the reach of M). Propositions 1.20, 1.21 and 1.22 remain true in the present case, as does the fact that d_M is differentiable on $\mathring{\mathcal{U}}_M \setminus M$. We prove a version of Proposition 1.23 for surfaces.

Proposition 4.1 *Let M be a closed C^2 regular surface. Then, we have the following statements.*

(i) If $|p - q| = d_M(p)$ $(q \in M)$ then $p = q + tN_M(q)$ with $|t| = d_M(p)$ and $\max(t\kappa_1(q), t\kappa_2(q)) \leq 1$, where κ_1 and κ_2 are the principal curvatures.

(ii) Let

$$\rho_M = \max \left\{ \frac{2\,|(\tilde{q} - q)^T N_M(q)|}{|\tilde{q} - q|^2} : q, \tilde{q} \in M, q \neq \tilde{q} \right\}. \tag{4.2}$$

Then $\rho_M < \infty$ and $r(M) \geq 1/\rho_M > 0$. In particular, $\mathring{\mathcal{U}}_M$ is not empty.

(iii) The distance map is differentiable on $\mathring{\mathcal{U}}_M \setminus M$.

Proof The proof is similar to that of Proposition 1.23 and we only highlight the differences. To prove (i), take an arc-length parametrized curve γ on M such that $\gamma(0) = q$

and let $f(t) = |p - \gamma(t)|^2$. Then, $\dot{f}(0) = -2(p-q)^T\dot{\gamma}(0)$, which much vanish for all γ, showing that $p - q$ is perpendicular to T_qM and therefore proportional to $N_M(q)$, i.e., $p = q + tN_M(q)$ for some t. Taking the second derivative of f (which must be non-negative at 0) yields $\ddot{f}(0) = 2 - 2(p-q)^T\ddot{\gamma}(0) = 2(1 - t\kappa_N^{(\gamma)}(0))$, so that $t\kappa_N^{(\gamma)}(0) \leq 1$. It then suffices to use the fact that $\kappa_1(q) \leq \kappa_N^{(\gamma)}(0) \leq \kappa_2(q)$.

To prove that ρ_M is finite, take $q_n, \tilde{q}_n$ such that

$$c_n := \frac{2\,|(\tilde{q}_n - q_n)^T N_M(q_n)|}{|\tilde{q}_n - q_n|^2}$$

tends to infinity, which implies that subsequences can be taken such that $q_n, \tilde{q}_n \to q$. Take a local chart around q such that $q_n = m(u_n, v_n)$ and $\tilde{q}_n = m(\tilde{u}_n, \tilde{v}_n)$. Let γ_n be a minimizing geodesic such that $\gamma_n(0) = q_n$ and $\gamma(s_n) = \tilde{q}_n$. Then the same argument as that of Proposition 1.23 can be used to prove that $|\tilde{q}_n - q_n|^2 = s_n^2 + o(s_n^2)$ and

$$|(\tilde{q}_n - q_n)^T N_M(q_n)| = |\kappa_N^{(\gamma_n)}(0)|s_n^2 + o(s_n^2),$$

contradicting the assumption that $c_n \to \infty$. The rest of the proof of (ii) is identical to Proposition 1.23.

For (iii), one shows that, if $p \in \mathring{\mathcal{U}}_M \setminus M$ and $q = \pi_M(p)$, then $1 - \max(t\kappa_1(q), t\kappa_2(q)) > 0$ with the same argument as in Proposition 1.23. Take a positively oriented chart $(u, v) = m(u, v)$ around q and consider the mapping $\varphi_m(u, v, t) = m(u, v) + tN_M(m(u, v))$. Then, letting $N_m = N_M \circ m$,

$$\partial_1\varphi_m = \partial_1 m + t\partial_1 N_m,\ \partial_2\varphi_m = \partial_2 m + t\partial_2 N_m,\ \partial_3\varphi_m = N_m$$

so that

$$\begin{aligned}\det(d\varphi_m) &= (\partial_1\varphi_m \times \partial_2\varphi_m)^T\partial_3\varphi_m \\ &= (\partial_1 m \times \partial_2 m)^T N_m + t(\partial_1 m \times \partial_2 N_m + \partial_1 N_m \times \partial_2 m)^T N_m \\ &\quad + t^2(\partial_1 N_m \times \partial_2 N_m)^T N_m.\end{aligned} \tag{4.3}$$

We have $\partial_1 m \times \partial_2 m = |\partial_1 m \times \partial_2 m|\, N_m$. Moreover, for any linear operator A on $\mathbb{R}^3$ and any basis (u_1, u_2, u_3) in $\mathbb{R}^3$, we have (the proof being left to the reader)

$$(u_1 \times u_2)^T Au_3 + (u_2 \times u_3)^T Au_1 + (u_3 \times u_1)^T Au_2 = \det(u_1, u_2, u_3)\mathrm{trace}(A).$$

Applying this to

$$(\partial_1 m \times \partial_2 N_m + \partial_1 N \times \partial_2 m)^T N_m = (N_m \times \partial_1 m)^T\partial_2 N_m + (\partial_2 m \times N_m)^T\partial_1 N_m$$

with $\partial_1 N_m = dN_m\,\partial_1 m$, $\partial_2 N_m = dN_m\,\partial_2 m$, taking $A = dN_m$ on T_pM and $AN_m = 0$, we get

$$(\partial_1 m \times \partial_2 N_m + \partial_1 N_m \times \partial_2 m)^T N_m = |\partial_1 m \times \partial_2 m| \operatorname{trace}(dN_m) \\ = -2|\partial_1 m \times \partial_2 m| H(m),$$

where H is the mean curvature. Moreover, since $\partial_1 N_m$ and $\partial_2 N_m$ are tangent to M at m, $(\partial_1 N_m \times \partial_2 N_m)^T N_m$ is the two-dimensional determinant of $[dN_m\, \partial_1 m, dN_m\, \partial_2 m]$, therefore equal to $K(m)|\partial_1 m \times \partial_2 m|$. We therefore have

$$\det(d\varphi_m) = (1 - 2tH(m) + t^2 K(m))|\partial_1 m \times \partial_2 m| \\ = (1 - t\kappa_1(m))(1 - t\kappa_2(m))|\partial_1 m \times \partial_2 m|. \quad (4.4)$$

The determinant is therefore positive, and the differentiability of d_M at p can be obtained using the inverse function theorem, as done in the proof of Proposition 1.23 □

Proposition 4.1 ensures that the mapping

$$\begin{aligned} \varphi_M : \quad M \times (-r, r) &\to \mathbb{R}^3 \\ (q, t) &\mapsto q + tN_M(q) \end{aligned}$$

is one-to-one for $r < r(M)$ onto $V_M(r) = \{p : d_M(p) < r\}$. More generally, for $B \subset M$, consider the sets $V_r(M, B) = \varphi_M(B \times (-r, r))$ and $V_r^+(M, B) = \varphi_M(B \times (0, r))$. Using the fact that φ_m introduced in the proof of Proposition 4.1 is such that $\varphi_m(u, v, t) = \varphi_M(m(u, v), t)$ and assuming that M is entirely covered by a local chart, Eq. 4.4 implies that

$$\begin{aligned} &\operatorname{Vol}(V_r^+(M, B)) \\ &= \int_0^r \int_{m^{-1}(B)} (1 - 2tH(m(u, v)) + t^2 K(m(u, v)))|\partial_1 m \times \partial_2 m|\, du\, dv \\ &= r\operatorname{Area}(B) - r^2 \int_B H\, d\sigma_M + \frac{r^3}{3} \int_B K\, d\sigma_M, \end{aligned} \quad (4.5)$$

where σ_M is the volume measure on M. The last expression for $\operatorname{Vol}(V_r^+(M, B))$ remains true even when B is not completely covered by a local chart, as can easily be proved by using partitions of unity, or covering B by a union of local charts up to a set of measure zero.

This discussion can also be extended to compact surfaces with boundary. Similar to the case of curves, we need to generalize the definition of normal vectors to boundary points. Let $p \in \partial M$ and $N_{\partial M}(p) \in T_p M$ denote the unit normal to the boundary at p pointing inward (toward M). Then, a vector N is normal to M at p if it can be written in the form

$$N = t_1 N_{\partial M}(p) + t_2 N_M(p)$$

with $t_1 \leq 0$. We let $\mathcal{N}_M(p)$ denote the set of unit normals at p. Then the statements of Proposition 4.1 remain true, provided that (4.2) is replaced by

$$\rho_M = \max\left\{\frac{2\,(\tilde{q}-q)^T N}{|\tilde{q}-q|^2} : q, \tilde{q} \in M, q \neq \tilde{q}, N \in \mathcal{N}_M(q)\right\}.$$

Equation (4.5) remains true whenever $B \subset M$. For $B \subset \partial M$, the computation must be modified. Represent $B \subset \partial M$ as a parametrized curve $\gamma : (0, L) \to \mathbb{R}^3$ and define

$$\varphi(s, t_1, t_2) = \gamma(s) + t_1 N_{\partial M}(\gamma(s)) + t_2 N_M(\gamma(s)).$$

Assume that γ is parametrized by arc length and oriented so that $\dot{\gamma} \times N_{\partial M} = N_M$. Consider the set $V_r(M, B) = \varphi(B \times \Gamma_r)$ where Γ_r is the half disc $\{(t_1, t_2) : t_1^2 + t_2^2 < r^2, t_1 \leq 0\}$, so that $V_r(M, B)$ is the set of points in $\mathbb{R}^3$ that have closest point to M in B at distance less than r. We have

$$\begin{aligned}\det(d\varphi(s, t_1, t_2)) &= \det(\dot{\gamma}(s), N_{\partial M}(\gamma(s)), N_M(\gamma(s)))\\ &\quad + t_1 \det(\partial_s N_{\partial M}, N_{\partial M}(\gamma(s)), N_M(\gamma(s)))\\ &\quad + t_2 \det(\partial_s N_M, N_{\partial M}(\gamma(s)), N_M(\gamma(s)))\\ &= 1 - t_1 \kappa_g^{(\gamma)}(s) - t_2 \kappa_N^{(\gamma)}(s).\end{aligned}$$

Indeed, we have $N_{\partial M} \times N_M = \dot{\gamma}$ and

$$\det(\partial_s N_{\partial M}, N_{\partial M}(\gamma(s)), N_M(\gamma(s))) = \partial_s N_{\partial M}^T \dot{\gamma} = -N_{\partial M}^T \ddot{\gamma} = -\kappa_g^{(\gamma)}.$$

Similarly

$$\det(\partial_s N_M, N_{\partial M}(\gamma(s)), N_M(\gamma(s))) = \partial_s N_M^T \dot{\gamma} = -\kappa_N^{(\gamma)}.$$

We can now compute (introducing radial coordinates)

$$\begin{aligned}\mathrm{Vol}(V_r(M, B)) &= \int_0^L \int_{t_1^2+t_2^2<r^2, t_1<0} (1 - t_1 \kappa_g^{(\gamma)}(s) - t_2 \kappa_N^{(\gamma)}(s))\, d\theta\, d\rho\, ds\\ &= \int_0^L \int_0^r \int_{-\pi/2}^{\pi/2} (1 + \rho \cos\theta \kappa_g^{(\gamma)}(s) - \rho \sin\theta \kappa_N^{(\gamma)}(s))\, d\theta\, d\rho\, ds\\ &= r\,\mathrm{length}(B) + r^2 \int_0^L \kappa^{(\gamma)}(s)\, ds.\end{aligned}$$

Applied to $r < r(M)$, Eq. (4.5) provides a new interpretation of the integrals of mean and Gauss curvatures over subsets of M in terms of the volumes of the sets $V_r^+(M, B)$ (which are often called *parallel sets* along the surface M). These parallel sets may be defined for sets that are much more general than smooth surfaces and their volumes then lead to generalized versions of the curvature. We now see how these ideas can be applied to triangulated surfaces.

The Discrete Case

We first generalize the definition of $V_B(r)$ to the non-smooth case, for which we will need a generalized definition of the set of unit normals to M at a given point. This can be done in two equivalent ways. Here M is an arbitrary closed set with positive reach, i.e., such that $r(M) > 0$.

The first point of view is to let, for $r < r(M)$,

$$V_r(M, B) = \{p : 0 < d_M(p) < r \text{ and } \pi_M(p) \in B\}.$$

When M is the boundary of a compact set Ω (e.g., when M is a closed surface), we can define

$$V_r^+(M, B) = V_r(\Omega, B),$$

still for $B \subset M$.

Federer [106] has proved that (4.5) can be generalized to sets of positive reach, so that $\mathrm{Vol}(V_B^+(r))$ is polynomial in r, taking the form

$$\mathrm{Vol}(V_r^+(M, B)) = r\mu_0(M, B) - r^2\mu_1(M, B) + \frac{r^3}{3}\mu_2(M, B). \tag{4.6}$$

In particular, $\mu_1(M, B)$ and $\mu_2(M, B)$ are generalizations of the integrals of the curvatures on B, and are called the mean and Gauss *curvature measures* on M. They have the important property of being additive, satisfying in particular

$$\mu_i(M \cup M', B) = \mu_i(M, B) + \mu_i(M', B) - \mu_i(M \cap M', B). \tag{4.7}$$

Although it already is a rich class of sets (including, for example, all convex sets), sets of positive reach do not include non-convex polyhedrons, so the construction cannot be immediately extended to triangulated surfaces. But formula (4.7) provides the key for this extension. Indeed, one can define a union of sets of positive reach as a set M that can be decomposed into

$$M = \bigcup_{j \in J} M_j,$$

where each M_j has positive reach and any nonempty intersection of M_j's has positive reach [314]. Then, iterating (4.7) (using the inclusion-exclusion formula), we can set

$$\mu_k(M, B) = \sum_{I \subset J} (-1)^{|I|-1} \mu_k\Big(\bigcap_{j \in I} M_j, B\Big), \tag{4.8}$$

the left-hand side being well-defined from the hypotheses. This is a valid definition of the right-hand side because it can be shown that the result does not depend on the chosen decomposition of M, which is not unique. This extension now includes all polyhedrons (and interiors of compact triangulated surfaces).

The second point of view gives an alternative interpretation of the curvature measures, based on the *normal bundle* to a set M. This requires a general definition of tangent and normal vectors to an arbitrary set $M \subset \mathbb{R}^3$ [107]. We already gave a general definition of tangent vectors in Definition 1.5, which defined the tangent set to M at $p \in M$ as the set $T_p M$ of vectors $v \in \mathbb{R}^d$ such that, for any $\varepsilon > 0$, there exist $x \in M$ and $r > 0$ such that $|x - p| < \varepsilon$ and $|v - r(x - p)| < \varepsilon$.

We will later use the fact that, when M is included in the boundary of an open set Ω, $T_m \Omega$ for $m \in M$ contains vectors in $T_m M$ and all vectors v such that $m + \varepsilon v \in \Omega$ for small ε (vectors that point to the interior of Ω). The special cases that follow will be important when studying triangulated surfaces.

Single points: Assume that $M = \{a\}$. It is clear from the definition that any tangent vector to M must satisfy $|v| < \varepsilon$ for any $\varepsilon > 0$ so that $T_a M = \{0\}$.

End-points of curves: Let $\gamma : [0, 1] \to \mathbb{R}^3$ be a smooth regular curve, $M = \gamma([0, 1])$ and $a = \gamma(0)$. Then, any $x \in M$ close to a is equal to $\gamma(u)$ for $u \simeq 0$, and a tangent vector v at a must be such that $v \simeq r(\gamma(u) - \gamma(0))$ with $r > 0$, so that $T_a M$ is the half-line $\mathbb{R}^+ \dot{\gamma}(0)$. By the same argument, if $b = \gamma(1)$, $T_b M = \mathbb{R}^- \dot{\gamma}(1)$. (Of course, the tangent space at interior points is the full line generated by the tangent vector $\dot{\gamma}$.)

Triangles: If M a triangle (including its interior) and a is on its boundary, then $T_a M$ is simply the set of vectors v such that $a + v$ points towards the interior of M.

Normals can now be derived from tangents, as stated in the next definition.

Definition 4.2 Let $M \subset \mathbb{R}^3$. For $p \in M$, the normal vectors to M at p form the set $\mathcal{N}_p M$, containing all vectors n such that $n^T v \leq 0$ for $v \in T_p M$.

The normal bundle of M is the set $\mathcal{N}M \subset \mathbb{R}^3 \times \mathbb{R}^3$ defined by

$$\mathcal{N}M = \left\{(p, n),\, p \in M, n \in \mathcal{N}_p M,\, |n| = 1\right\}.$$

When $M \subset \partial\Omega$, we can also consider

$$\mathcal{N}M^+ = \left\{(p, n),\, p \in M, n \in \mathcal{N}_p \Omega,\, |n| = 1\right\}.$$

This corresponds to normals to M pointing outward from Ω.

The normal bundle is the structure on which the new curvature measures will be defined. Let us describe it for the previous examples. First, if M is a smooth closed oriented surface in $\mathbb{R}^3$, $\mathcal{N}M$ is simply the set $\{(p, N(p)),\, p \in M\} \cup \{(p, -N(p)),\, p \in M\}$. The set $\mathcal{N}M^+$ only contains elements $(p, -N(p))$ (assuming that M is positively oriented).

If M is a closed curve, with regular parametrization $s \mapsto \gamma(s)$, then $\mathcal{N}M = \left\{(\gamma(s), n) : n^T \dot{\gamma} = 0, |n| = 1\right\}$ (this can be thought of as a tube centered around γ).

If $M = \{a\}$, then (since $T_a M = \{0\}$), $\mathcal{N}M = \{a\} \times S^2$, where S^2 is the unit two-dimensional sphere.

When M is an open curve, parametrized by γ, the set $\mathcal{N}_a M$ for $a = \gamma(0)$ is the half-sphere $S^2 \cap \{n^T \dot{\gamma}(0) \leq 0\}$, while, for $b = \gamma(1)$, it is $\mathcal{N}_b M = S^2 \cap \{n^T \dot{\gamma}(1) \geq 0\}$. The whole set $\mathcal{N}M$ can be thought of as a "sausage" around γ.

Finally, if M is a triangle, and a is on an edge, but not at a vertex, $\mathcal{N}_a M$ is a half-circle, being the intersection of the unit circle orthogonal to the edge and the half-space $\{n^T \nu \geq 0\}$, where ν is normal to the edge, in the triangle plane, and pointing towards the interior of the triangle. If a is a vertex, the set $\mathcal{N}_a M$ is the "spherical corner" formed by the intersection of the unit sphere and the two half-planes $\{n^T e \leq 0\}$ and $\{n^T e' \leq 0\}$, where e and e' are the two edges stemming from a (oriented from a to the other vertex).

The interesting fact in the previous examples is that, in each case, $\mathcal{N}M$ was a 2-D structure, i.e., it could always be parametrized by two parameters. In fact, $\mathcal{N}M$ is a two-dimensional continuous surface in a space of dimension 6.

For a smooth surface, we have introduced the function $\varphi_M(p, t) = p + tN_M(p)$, defined on $M \times (0, r_0)$. We now want to consider it as a function $\psi_M(p, n, t) = p + tn$ defined on $\mathcal{N}M \times (0, r_0)$. For smooth oriented surfaces, this is equivalent if the new definition of φ is restricted to the positive normal, i.e., $\mathcal{N}M^+$, since the latter is uniquely determined by m. But we have seen cases for which the normal was just partially specified by p, and this new definition of φ also applies to such situations. From now on, we assume that $M = \partial\Omega$ is the boundary of an bounded open subset of $\mathbb{R}^3$ and that (u, v) are local coordinates on $\mathcal{N}M^+$, so that we have a map $(u, v) \in U \mapsto (m(u, v), n(u, v))$ that locally parametrizes $\mathcal{N}M^+$ as a subset $\mathbb{R}^6$, U being an open subset of $\mathbb{R}^2$. We will let

$$A = \{(m(u, v), n(u, v)) : (u, v) \in U\}$$

denote the corresponding patch in $\mathcal{N}M^+$. We will assume that this map is differentiable in (u, v) (which may require the exclusion of some exceptional (negligible) sets from the integral that will be computed below. Our goal is to compute the volume of $V_r^+(M, A) = \psi_M(A \times (0, r))$.

The area form on $\mathcal{N}M^+$ is given by $g(u, v)\, du\, dv$ where

$$g(u, v)^2 = (|\partial_1 m|^2 + |\partial_1 n|^2)(|\partial_2 m|^2 + |\partial_2 n|^2) - (\partial_1 m^T \partial_2 m + \partial_1 n^T \partial_2 n)^2.$$

(In this expression, we have used $g(u, v) = \sqrt{EF - G^2}$, using (3.7), which still holds for two-dimensional surfaces in higher-dimensional spaces.)

One can check (we skip the proof) that the ratio $Q = |\det(\partial_1 m + t\partial_1 n_1, \partial_2 m + t\partial_2 n, n)|/g(u, v)$ is invariant under a change of local chart on $\mathcal{N}M^+$, allowing us to consider it as a function $Q(m, n)$ defined over A. When M is a smooth surface, this ratio can easily be computed, since we can assume that $\partial_1 m$ and $\partial_2 m$ correspond to the principal directions, yielding

$$Q = \frac{(1 + t\kappa_1)(1 + t\kappa_2)}{\sqrt{(1 + \kappa_1^2)(1 + \kappa_2^2)}}.$$

Returning to the general case, we have, by definition of Q:

$$\mathrm{Vol}(V_r^+(M,A)) = \int_U |\det(\partial_1 m + t\partial_1 n, \partial_2 m + t\partial_2 n, n)| du dv = \int_A Q d\sigma.$$

Assume that r can be chosen so that Q does not vanish for $t \in (0, r)$. (That such an r exists relates to the assumption of M having positive reach.) In this case, $\det(\partial_1 m + t\partial_1 n, \partial_2 m + t\partial_2 n, n)$ has constant sign, and one can expand Q in powers of t, yielding, for some functions S_0, S_1, S_2

$$\int_0^r \int_A Q d\sigma = r \int_A S_0 d\sigma - r^2 \int_A S_1 d\sigma + \frac{r^3}{3} \int_A S_2 d\sigma.$$

The functions S_k therefore provide densities for "generalized" curvature measures, defined on $\mathcal{N}M^+$ (instead of on M).

Recall that we assume that $M \subset \partial\Omega$ for some open set $\Omega \subset \mathbb{R}^2$. We can restrict the generalized curvature measures to M, letting for $B \subset \mathbb{R}^3$,

$$\mu_k(M, B) = \int_{\mathcal{N}M^+} \chi_B(m) S_k(m, n) d\sigma.$$

We now consider the case in which M is a triangulated surface and discuss the expression of this integral based on the location of the set B.

Face interiors: Let B be included in the interior of a face. Since M coincides there with a smooth surface with zero curvature, we have $\mu_0(M, B) = \mathrm{Area}(B)$, $\mu_1(M, B) = \mu_2(M, B) = 0$.

Convex edges: Now let B be included in the interior of a convex (salient) edge, e. At $p \in B$, normal vectors to Ω form the arc of the unit circle perpendicular to the edge delimited by the two outward normals to the neighboring faces. Fix an orientation of e and define $\beta_N(e)$ as the angle from the outward normal on the right to the one on the left of e (the saliency assumption implies that this angle is non-negative.) Now, on B, the normal bundle can be parametrized by $p = u(e/|e|) + n(v)$, where $n(v)$ is the normal to p that makes an angle v with one of the face normals. Using the fact that e and n are orthogonal, one finds $d(u, v) = 1$ and $|\det(\partial_u m, t\dot{n}_v, n)| = t$. This implies that $\mu_0 = \mu_2 = 0$ and $\mu_1 = \beta_N(e)\mathrm{length}(B)$.

Concave edges: The case of B included in a concave edge requires a little more care because Ω does not have positive reach on B. One can however split Ω into two parts on each side of the bisecting angle between the faces at e and apply the formula (letting Ω_1 and Ω_2 be the two sections)

$$\begin{aligned}\mu_k(\Omega, B) &= \mu_k(\Omega_1, B) + \mu_k(\Omega_2, B) - \mu_k(\Omega_1 \cap \Omega_2, B) \\ &= ((\pi - \beta_N)/2 + (\pi - \beta_N(e))/2 - \pi)\mathrm{length}(B) \\ &= -\beta_N \mathrm{length}(B),\end{aligned}$$

where $\beta_N(e)$ is again the absolute value of the angle between the normal to the faces meeting at e (taken between 0 and π).

Vertices: Let $B = \{v\}$, where v is a vertex of the triangulation. First note that, when M has positive reach, the volume of $V_r^+(M, \{v\})$ cannot be larger than that of the ball centered at v with radius r and is therefore an $O(r^3)$. From formula (4.8), this is also true when M is a triangulated surface. This implies that only the last term (the Gauss curvature measure) can be non-zero. The computation of this term is simplified if we also observe that the Gauss curvature measure does not change if we replace (locally at the vertex) Ω by $\Omega^c \cap \partial\Omega$, which corresponds to changing the orientation on M. This is because the function S_2 is an even function of the normal. Using this property, we get

$$2\mu_2(\Omega, B) = \mu_2(\Omega, B) + \mu_2(\Omega \cup \Omega^c, B) = \mu_2(\mathbb{R}^3, B) + \mu_2(M, B).$$

Since $\mu_2(\mathbb{R}^3, B) = 0$, it remains to compute $\mu_2(M, B)$. Let $F_1, \ldots, F_q$ represent the faces containing v. We want to apply the inclusion-exclusion formula to

$$\mu_2(M, B) = \mu_2\Big(\bigcup_{i=1}^{q} F_i, B\Big).$$

For this, we need to compute the Gauss curvatures in three special cases covering possible intersections between faces. The simplest case is $\mu_2(\{v\}, B)$. In this case, we can parametrize the normal bundle by $(m(u, u'), n(u, u')) = (v, n(u, u'))$, where (u, u') is a parametrization of the unit sphere, for which we assume that $\dot{n}_u$ and $\dot{n}_{u'}$ are orthogonal with unit norm. In this case, $d(u, u') = 1$ and $|\det(t\partial_1 n, t\partial_2 n, n)| = t^2$. This implies that $S_2 = 1$ and $\mu_2 = 4\pi$ (μ_2 is three times the volume of the sphere).

Now, let e be a segment having v as one of its extremities. Assume without loss of generality that e is supported by the first coordinate axis and $v = 0$. We can parametrize $\mathcal{N}e$ at v with $(u, u') \mapsto (v, n(u, u'))$, where $n(u, u')$ parametrizes the half-sphere that is normal to M at v. This provides $\mu_2(e, B) = 2\pi$.

The last case is a triangle F with vertex v. Let θ be the angle at v. Here, the normal bundle at v is the part of the unit sphere which is contained between the two hyperplanes normal to each edge of F incident at v, for which the volume is $2(\pi - \theta)/3$, so that $\mu_2(F, B) = 2(\pi - \theta)$.

Now, it remains to apply the inclusion-exclusion formula. This formula starts with $\sum_{i=1}^{q} \mu_2(F_i, B)$, which is $2q\pi - 2\sum_i \theta_i$. Then comes the sum of the measures associated with the intersection of two faces: this intersection is an edge for the q pairs of adjacent faces, and just $\{v\}$ for the $\binom{q}{2} - q$ remaining ones. This yields the contribution $2q\pi - 4\binom{q}{2}\pi$. We finally need to sum all terms for intersections of three or more sets, which is always equal to $\{v\}$. This is

$$4\pi \sum_{k\geq 3} (-1)^{k-1} \binom{q}{k} = 4\pi \left(1 - q + \binom{q}{2}\right),$$

where we used the fact that

$$\sum_{k\geq 0}(-1)^{k-1}\binom{q}{k} = (1-1)^q = 0.$$

Summing all the terms, we obtain $\mu_2(M, B) = 4\pi - 2\sum_{i=1}^{q}\theta_i$ so that

$$\mu_2(\Omega, B) = 2\pi - \sum_{i=1}^{q}\theta_i.$$

We have therefore obtained the curvature measures associated to an oriented triangulated surface [71]. For any set $B \in \mathbb{R}^3$, they are:

- The mean curvature measure:

$$\mu_1(M, B) = \sum_{e\in E}\varepsilon_e\beta_N(e)\text{length}(B\cap e),$$

  where β_N is the angle between the normals to the faces at e in absolute value and $\varepsilon_e = 1$ if the edge is convex and -1 otherwise.
- The Gauss curvature measure:

$$\mu_2(M, B) = \sum_{v\in V\cap B}\mu_2(M, v),$$

  with

$$\mu_2(M, v_i) = 2\pi - \sum_{k\in\mathcal{F}_i}\theta_v(F_k).$$

Using these expressions, we can make approximations of the curvature at a given vertex by letting

$$\mu_1(M, B) \simeq |B|H_i \text{ and } \mu_2(M, B) \simeq |B|K_i$$

for a vertex v_i in the triangulation. Taking $B = A_i$, as defined in the previous section (see Eq. 4.1), we obtain the same approximation of the Gauss curvature as the one obtained from the discretization of the Gauss–Bonnet theorem. The formulas for the mean curvature differ, however.

4.2.4 Discrete Gradient and Laplace–Beltrami Operators

We conclude this section on triangulated surfaces with a computation of the discrete equivalent of the gradient and Laplacian on surfaces.

Let S be a triangulated surface, $\mathcal{V} = \{v_1, \ldots, v_M\}$ and $\mathcal{F} = \{F_1, \ldots, F_K\}$ denoting, respectively, the sets of vertices and faces of S. To simplify the discussion, we will assume that the surface has no boundary, i.e., each edge belongs to exactly two faces.

A function ψ defined on S assigns a value $\psi(v_i)$ to each vertex, and the gradient of ψ will be defined as a vector indexed over faces. To compute it, we first focus on a face, $F = F_k$ for some $k \in \{1, \ldots, K\}$ that we will drop from the notation until further notice. Let (v_1, v_2, v_3) be the vertices of F (ordered consistently with the orientation), and let $e_1 = v_3 - v_2$, $e_2 = v_1 - v_3$ and $e_3 = v_2 - v_1$. Let $c = (v_1 + v_2 + v_3)/3$ be the center of the face.

We define the gradient of ψ on F, denoted $\nabla_S \psi(F)$, as the gradient of the linear interpolation of ψ on F, i.e.,

$$\nabla_S \psi(F) = \nabla_F \hat{\psi}_F,$$

where $\hat{\psi}_F(a_1 v_1 + a_2 v_2 + a_3 v_3) = a_1 \psi(v_1) + a_2 \psi(v_2) + a_3 \psi(v_3)$ for $a_1 + a_2 + a_3 = 1$ and ∇_F is the gradient on the face F considered as a regular surface. From a computational viewpoint, $u = \nabla_S \psi(F)$ is such that $u = \alpha_1 e_1 + \alpha_2 e_2$ and $u^T(v_k - v_l) = \psi(v_k) - \psi(v_l)$ $(k, l = 1, 2, 3)$, which gives

$$\begin{aligned} \psi(v_3) - \psi(v_2) &= (\alpha_1 e_1 + \alpha_2 e_2)^T e_1, \\ \psi(v_1) - \psi(v_3) &= (\alpha_1 e_1 + \alpha_2 e_2)^T e_2. \end{aligned}$$

Let ψ_F be the column vector $[\psi(v_1), \psi(v_2), \psi(v_3)]^T$, M the 2 by 3 matrix

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

and G_F the matrix

$$G_F = \begin{pmatrix} |e_1|^2 & e_1^T e_2 \\ e_1^T e_2 & |e_2|^2 \end{pmatrix}.$$

With this notation, the previous system is $M\psi_F = G_F \alpha$. We therefore have

$$u = [e_1, e_2]\alpha = [e_1, e_2] G_F^{-1} M \psi_F.$$

We first notice that $\det G_F = |e_1|^2 |e_2|^2 - (e_1^T e_2)^2 = (|e_1||e_2| \sin\theta_3)^2$, where θ_3 is the angle at v_3. It is therefore equal to $4a(F)^2$, where $a(F)$ is the area of F. Moreover, we can write:

$$\begin{aligned} G_F^{-1} M \psi_F &= \det(G_F)^{-1} \begin{pmatrix} |e_2|^2 & -e_1^T e_2 \\ -e_1^T e_2 & |e_1|^2 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \psi_F \\ &= \det(G_F)^{-1} \begin{pmatrix} -e_2^T e_1 & -e_2^T e_2 & -e_2^T e_3 \\ e_1^T e_1 & e_1^T e_2 & e_1^T e_3 \end{pmatrix} \psi_F, \end{aligned} \tag{4.9}$$

in which we have used the identity $e_3 = -e_1 - e_2$. Introducing the vector

$$h_\psi(F) = \psi(v_1) e_1 + \psi(v_2) e_2 + \psi(v_3) e_3$$

and the matrix

$$D_F = e_2 e_1^T - e_1 e_2^T = e_1 e_3^T - e_3 e_1^T = e_3 e_2^T - e_2 e_3,$$

a little computation yields

$$\nabla_S \psi(F) = \frac{D_F}{4a(F)^2} h_\psi(F).$$

We now pass to the computation of the discrete Laplace–Beltrami operator, which we define via the discrete analog of the property

$$\int_S |\nabla_S \psi|^2 d\sigma_S = -\int_S \psi \Delta_S \psi d\sigma_S$$

that characterizes the operator on smooth surfaces without boundary. For triangulated surfaces, we will identify $\Delta_S \psi$ via

$$\sum_{k=1}^{K} |\nabla_S \psi(F_k)|^2 a(F_k) = -\sum_{i=1}^{N} \psi(v_i)(\Delta_S \psi)(v_i)|A_i|,$$

where $|A_i|$ is the area attributed to vertex v_i (using, for example, Voronoï cells).

For a given face F, we can write (using the previous notation): $|\nabla_S \psi(F)|^2 = \alpha^T G_F \alpha = \psi_F^T M^T G_F^{-1} M \psi_F$. Applying M^T to (4.9), we get

$$M^T G_F^{-1} M = \det(G_F)^{-1} \begin{pmatrix} |e_1|^2 & e_1^T e_2 & e_1^T e_3 \\ e_1^T e_2 & |e_2|^2 & e_2^T e_3 \\ e_1^T e_3 & e_2^T e_3 & |e_3|^2 \end{pmatrix}.$$

Let Σ_F denote this last matrix. We can write:

$$\sum_{k=1}^{K} \frac{\psi_{F_k}^T \Sigma_{F_k} \psi_{F_k}}{4a(F_k)} = \frac{1}{4} \sum_{i=1}^{N} \psi(v_i) \sum_{k \in \mathcal{F}_i} (|e_{ik}|^2 \psi(v_i) + e_{ik}^T e'_{ik} \psi(v'_{ik}) + e_{ik}^T e''_{ik} \psi(v''_i k))/a(F_k),$$

where v'_{ik} and v''_{ik} are the other two vertices of F_k, $k \in \mathcal{F}_i$ (in addition to v_i, ordered according to the orientation) and e_{ik}, e'_{ik} and e''_{ik} are, respectively, the edges opposed to v_i, v'_{ik} and v''_{ik} in F_k. This implies that one should define

$$\Delta_S \psi(v_i) = -\frac{1}{4|A_i|} \sum_{k \in \mathcal{F}_i} \left(|e_{ik}|^2 \psi(v_i) + e_{ik}^T e'_{ik} \psi(v'_{ik}) + e_{ik}^T e''_{ik} \psi(v''_{ik})\right)/a(F_k).$$

One can rewrite this discrete Laplacian in terms of angles. Denoting as before by θ'_{ik} and θ''_{ik} the angles at v'_{ik} and v''_{ik}, one has

$$e_{ik}^T e'_{ik} = -\cos\theta''_{ik}|e_{ik}|\,|e'_{ik}| = -2\operatorname{ctn}\theta''_{ik}a(F_k).$$

Similarly, $e_{ik}^T e''_{ik} = -2\operatorname{ctn}\theta'_{ik}a(F_k)$ and, since the sum of the edges is 0,

$$|e_{ik}|^2 = -e_{ik}^T(e'_{ik} + e''_{ik}) = 2(\operatorname{ctn}\theta'_{ik} + \operatorname{ctn}\theta''_{ik})a(F_k).$$

One can therefore write

$$\Delta_S\psi(v_i) = \frac{1}{2|A_i|}\sum_{k\in\mathcal{F}_i}\big(\operatorname{ctn}\theta''_{ik}(\psi(v'_{ik}) - \psi(v_i)) + \operatorname{ctn}\theta'_{ik}(\psi(v''_{ik}) - \psi(v_i))\big),$$

which provides a discrete definition of the Laplace–Beltrami operator on S. This formula is sometimes called the "cotangent formula" [232].

4.3 Consistent Approximation

So far the concepts we have defined for triangulated surfaces have been directly inspired by the corresponding notions in the theory of smooth surfaces. Here, we provide some results evaluating how well a triangulated surface can approximate a smooth one, and whether quantities defined on triangulated surfaces are good estimates of the same quantities computed on the surface that is being approximated.

Because this analysis will be important for further purposes, we focus on the approximation of integrals $\int_\Sigma h(p)\,d\sigma_\Sigma(p)$ over a C^2 regular surface Σ by sums

$$\sum_{F\in\mathcal{F}} h(c_F)a(F),$$

where $\mathcal{F}$ is the set of faces of a triangulated surface S, c_F is the center of mass of face F and $a(F)$ its area.

To handle situations in which Σ is a surface with boundary, we also assume that another C^2 regular surface, Σ', is given, extending Σ so that $\Sigma \cup \partial\Sigma \subset \Sigma'$. If Σ is a closed surface, we can take $\Sigma' = \Sigma$. We let $\varphi : \Sigma' \times (-\rho, \rho) \to \mathbb{R}^3$ be the normal map, so that $\varphi(p,t) = p + tN_{\Sigma'}(p)$, and we assume that ρ is small enough that φ is a diffeomorphism onto its image, denoted by U. For $q \in U$, we let $\xi(q)$ be the closest point to q in Σ', i.e., the unique p such that $\varphi(p,t) = q$ for some $t \in \rho$. We note that

$$\xi(\varphi(p,t)) = p$$

for all p, t, so that $d\xi \circ \varphi\, \partial_p \varphi = \mathrm{Id}$, with $\partial_p \varphi = \mathrm{Id} + t d N_{\Sigma'}$. Fixing p and letting u_1, u_2 denote principal directions at p, with principal curvatures κ_1 and κ_2, we therefore have, for $q = \varphi(p, t)$

$$d\xi(q)u_i = \frac{u_i}{1 + t\kappa_i},\ i = 1, 2.$$

Similarly, $d\xi \circ \varphi\, \partial_t \varphi = 0$, so that $d\xi(q) N_{\Sigma'}(p) = 0$. We assume in the following that ρ is chosen small enough that $1 + t\kappa_i$ is bounded away from zero for $|t| \leq \rho$ and $i = 1, 2$.

Let S be a triangulated surface, with the usual notation $\mathcal{V}, \mathcal{F}, \mathcal{E}$ for the sets of vertices, faces and edges in S. We will assume that $S \subset U$. For such a surface, we define the following constants:

- $\varepsilon_1(S, \Sigma') = \sup_{q \in S} |\xi(q) - p|$, the distance from S to Σ'.
- $\varepsilon_2(S, \Sigma') = 1 - \min_{F \in \mathcal{F}} \min_{q \in F} N_S(F)^T N_{\Sigma'}(\xi(q))$.
- $\delta(S) = \max_{(v_1, v_2) \in \mathcal{E}} |v_1 - v_2|$, the maximum edge size in S.
- $\varepsilon_3(S, \Sigma) = d_H(\xi(S), \Sigma)$, where d_H is the Hausdorff distance

$$d_H(A, A') = \max\left(\sup_{x \in A} \mathrm{dist}(x, A'), \sup_{x \in A'} \mathrm{dist}(x, A) \right).$$

We will also let $\varepsilon(S, \Sigma') = \max(\varepsilon_1, \varepsilon_2, \varepsilon_3, \delta)$, and we will say that a sequence of triangulations $S^{(n)}$ converge to Σ if $\varepsilon(S^{(n)}, \Sigma) \to 0$.

With this notation, we have the following theorem.

Theorem 4.3 *Let h be a continuous function on a compact neighborhood of U. Then*

$$\int_\Sigma h(p)\, d\sigma_\Sigma(p) = \sum_{f \in \mathcal{F}} h(c_f) a(f) + |S| O(\varepsilon). \tag{4.10}$$

Proof Consider $F \in \mathcal{F}$. Let v_1, v_2, v_3 denote the vertices of f and $c = (v_1 + v_2 + v_3)/3$. Let $e_{ij} = v_j - v_i$. We first compute $\int_{\xi(f)} h(p)\, d\sigma_{\Sigma'}(p)$, for which we can use the local chart $\psi : (x, y) \mapsto \xi(v_1 + x e_{12} + y e_{13})$, for $x, y > 0$, $x + y < 1$. Then, letting $\mathcal{T}$ denote the triangle $\{(x, y) : x, y > 0, x + y < 1\}$,

$$\int_{\xi(F)} h(p)\, d\sigma_{\Sigma'}(p) = \int_{\mathcal{T}} h(\psi(x, y)) |\partial_x \psi \times \partial_y \psi|\, dx\, dy.$$

Note that, using the same notation as above for principal directions and curvatures on Σ' at $p = \psi(x, y)$,

$$\partial_x \psi = d\xi \circ \psi\, e_{12} = \frac{e_{12}^T u_1}{1 + t\kappa_1} u_1 + \frac{e_{12}^T u_2}{1 + t\kappa_2} u_2$$
$$\partial_y \psi = d\xi \circ \psi\, e_{13} = \frac{e_{13}^T u_1}{1 + t\kappa_1} u_1 + \frac{e_{13}^T u_2}{1 + t\kappa_2} u_2$$

so that

$$\begin{aligned}\partial_x\psi \times \partial_y\psi &= \frac{e_{12}^T u_1\, e_{13}^T u_2 - e_{12}^T u_2\, e_{13}^T u_1}{(1+t\kappa_1(p))(1+t\kappa_2(p))} N_{\Sigma'}(p) \\ &= \frac{(e_{12}\times e_{13})^T N_{S'}(p)}{(1+t\kappa_1(p))(1+t\kappa_2(p))} N_{\Sigma'}(p) \\ &= 2a(F)\frac{N_F^T N_{S'}(p)}{1+2tH(p)+t^2K(p)} N_{\Sigma'}(p),\end{aligned}$$

where H and K denote the mean and Gauss curvatures. We therefore have

$$\int_{\xi(F)} h(p)\, d\sigma_{\Sigma'}(p) = 2a(F)\int_{\mathcal{T}} \frac{N_f^T N_{\Sigma'}(\psi(x,y))\, h(\psi(x,y))}{1+2tH(\psi(x,y))+t^2K(\psi(x,y))}\, dx\, dy,$$

from which one immediately deduces that

$$\begin{aligned}\left|\int_{\xi(F)} h(p)\, d\sigma_{\Sigma'}(p) - a(F)h(c)\right| &= \left|\int_{\xi(F)} h(p)\, d\sigma_{\Sigma'}(p) - 2a(F)\int_{\mathcal{T}} h(c)\, dx\, dy\right| \\ &\le a(F)\max_{p\in F}\left|h(c) - \frac{N_F^T N_{\Sigma'}(\xi(p))\, h(\xi(p))}{1+2tH(\xi(p))+t^2K(\xi(p))}\right|.\end{aligned}$$

Introduce the modulus of continuity of h

$$\omega_h(\eta) = \max_{x,y\in U, |x-y|<\eta} |h(x)-h(y)|.$$

One has,

$$\begin{aligned}&\left|h(c) - \frac{N_F^T N_{\Sigma'}(\xi(p))\, h(\xi(p))}{1+2tH(\xi(p))+t^2K(\xi(p))}\right| \le \\ &\qquad \omega_h(\delta+\varepsilon_1) + \|h\|_\infty\varepsilon_2 + \|h\|_\infty\varepsilon_1 \max_{|t|\le\varepsilon_1, p\in\Sigma'} \frac{2H(p)+tK(p)}{1+2tH(p)+t^2K(p)},\end{aligned}$$

so that

$$\left|\int_{\xi(F)} h(p)\, d\sigma_{\Sigma'}(p) - a(F)h(c)\right| = a(F)\, O(\varepsilon).$$

Summing over all faces, we obtain the fact that

$$\int_{\xi(S)} h(p)\, d\sigma_{\Sigma'}(p) = \sum_{F\in\mathcal{F}} h(c_F)a(F) + |S|O(\varepsilon).$$

It now suffices to write

$$\int_{\xi(S)} h(p)\, d\sigma_{\Sigma'}(p) = \int_{\Sigma} h(p)\, d\sigma_{\Sigma}(p) + O(\varepsilon_3)$$

to conclude the proof of (4.10). □

Note also that we can replace ε_2 by $\varepsilon_2' = 1 - \min_{F\in\mathcal{F}} N_S(F)^T N_{\Sigma'}(\xi(c_F))$ (or any other point in F), because $\|d\xi\|$ being bounded implies $|N_{\Sigma'}(\xi(q)) - N_{\Sigma'}(c_F)| = o(\delta)$ for $q \in F$.

Almost the same proof can be applied to sums involving normal vectors, yielding, for example

$$\int_{\Sigma} h(p)^T N_{\Sigma}(p)\, d\sigma_{\Sigma}(p) = \sum_{F\in\mathcal{F}} h(c_F)^T N_F a(F) + |S| o(\varepsilon) \tag{4.11}$$

for continuous vector-valued functions h.

Note also that, if $h : \mathbb{R}^3 \to \mathbb{R}$ is a C^1 function, one has

$$\max_F |\nabla_{\Sigma} h(\xi(c_F)) - \nabla_S h(F)| = o(\varepsilon).$$

Indeed, we have defined $\nabla_S h(F) = \nabla_F(\hat{h}_F)$, where $\hat{h}_F$ is a linear interpolation of h on F, but $\nabla_F(\hat{h}_F) = \nabla_F h(c_F) + o(\delta)$ because h is C^1. Moreover, letting ∇h denote the $\mathbb{R}^3$ gradient of h, we have

$$\nabla_F h(c_F) = \nabla h(c_F) - (N_F^T \nabla h(c_F)) N_F$$

and

$$\nabla_{\Sigma} h(\xi(c_F)) = \nabla h(\xi(c_F)) - (N_{\Sigma}(\xi(c_F))^T \nabla h(\xi(c_F))) N_{\Sigma}(\xi(c_F))$$

and these two quantities differ as $O(\varepsilon)$.

One can consider other approximations of geometric quantities and their convergence when triangulated surfaces approximate smooth ones with increasing accuracy. See, for example, [146, 211], in which an equivalence is shown between correct approximation of normals, or metric tensors, of area and of the Laplace–Beltrami operator.

4.4 Isocontours and Isosurfaces

To conclude this chapter, we discuss methods that compute shapes (curves or surfaces) from discrete image data. We will discuss approaches based on energy minimization in Chap. 5. Here, we focus on what is probably the simplest approach, which is to define curves of surfaces implicitly based on interpolation of the image values.

Assuming that the image f is defined on a discrete grid, we will interpolate it as a function $\hat{f} : D \subset \mathbb{R}^d \to \mathbb{R}$ and define a shape as the level set

$$S_\lambda = \{m : f(m) = \lambda\}$$

for a properly chosen threshold, λ. As we know, if the gradient of $\hat{f}$ does not vanish on S_λ, this provides a smooth curve or surface (or a union of such).

The concrete implementation of such an approach presents a few challenges, however. We will start with a discussion of the simpler two-dimensional case, which will help in addressing the computation of isosurfaces, which is more intricate.

4.4.1 Computing Isocontours

We consider here a two-dimensional grid, $\mathcal{G}$, which is formed by points $p(s,t) = (x_s, y_t)$, where $(x_s, s = 1, \ldots, M)$ is a discretization of the horizontal axis and $(y_t, t = 1, \ldots, N)$ a discretization of the vertical axis. We assume that a discretization of a smooth function f is observed, via the collection

$$(f_{st} = f(p(s,t)), s = 1, \ldots, M, t = 1, \ldots, N).$$

The problem is to compute the isocontour $(f = \lambda)$ for a given λ, in the form of a polygon or a union of polygons. Without loss of generality, we can and will assume $\lambda = 0$ in the following discussion.

Since the exact function f is not observed, some interpolation must be done, and we will use bilinear interpolation for this. This means that the true (unobserved) f will be replaced by the interpolation (that we still denote by f, with some abuse of notation) which is defined as follows. Let $C(s,t)$ denote the cell (square) with vertices $p(s+\varepsilon_1, t+\varepsilon_2), \varepsilon_i \in \{0,1\}, i = 1,2$. Then, for $p = x, y \in C(s,t)$, let

$$f(p) = \sum_{\varepsilon_1,\varepsilon_2=0}^{1} \prod_{i=1}^{2} \big(\varepsilon_i r_i(p) + (1-\varepsilon_i)(1-r_i(p))\big) f_{s+\varepsilon_1, t+\varepsilon_2}, \tag{4.12}$$

with $r_1(p) = x - x_s, r_2(p) = y - y_t$.

Obviously, the set $\{f = 0\}$ is the union of its intersections with each cell in the grid, so that we can restrict to these intersections. Within a cell, f is given by (4.12), and the set $\{f = 0\}$ is either empty, or a line segment, or one or two branches of a hyperbola. This is because, introducing the coordinates $\xi = (x - x_s)/(x_{s+1} - x_s)$ and $\eta = (y - y_s)/(y_{s+1} - y_s)$, we can rewrite $f(p)$ in the cell as (up to a positive multiplicative constant):

$$f(p) = f_{--}(1-\xi)(1-\eta) + f_{+-}\xi(1-\eta) + f_{-+}(1-\xi)\eta + f_{++}\xi\eta$$
$$= \rho\Big((\xi + \frac{f_{-+} - f_{--}}{\rho})(\eta + \frac{f_{+-} - f_{--}}{\rho}) - \frac{f_{++}f_{--} - f_{-+}f_{+-}}{\rho^2}\Big)$$

if $\rho := f_{++} - f_{+-} - f_{-+} + f_{--} \neq 0$ and

$$f(p) = (f_{+-} - f_{--})\xi + (f_{-+} - f_{--})\eta + f_{--}$$

if $\rho = 0$. In this formula, f_{++}, f_{+-}, f_{-+} and f_{--} are the values of f at the vertices of the cell.

We will approximate the intersection by line segments intersecting the edges of the cell. There can be 0, 1 or 2 such line segments, and we now discuss when these situations occur. An important observation is that, because the bilinear interpolation is linear when restricted to the edges of the cell, there is at most one intersection of the set $\{f = 0\}$ with each edge, and this is only possible when f takes different signs at each of the edge end-points. When this occurs, the points on the edges at which $f = 0$ can be easily computed by solving a linear equation. They will form the vertices of the polygonal line. The following, the proof of which we skip, can be justified directly from the quadratic expression of f in the cell.

(a) If all f_{++}, f_{+-}, f_{-+} and f_{--} have the same sign: there is no intersection with the edges, and therefore no intersection with the cell.
(b) If three of the values have the same sign, the last one having the opposite sign, there are two vertices in the cell, and one edge connecting them.
(c) If two values have the same sign on one edge and two have the opposite sign on the opposite edge, here also, there are two vertices and one edge.
(d) If the function changes sign on all the edges, there are four vertices and two edges. There are two subcases, letting $\delta = f_{++}f_{--} - f_{-+}f_{+-}$.
    (i) If $\delta > 0$, then one edge links the vertex on $\{\xi = 0\}$ to the one on $\{\eta = 1\}$, and the other the vertex on $\{\eta = 0\}$ to the one on $\{\xi = 1\}$.
    (ii) If $\delta < 0$, then one edge links the vertex on $\{\xi = 0\}$ to the one on $\{\eta = 0\}$, and the other the vertex on $\{\eta = 1\}$ to the one on $\{\xi = 1\}$.

Cases (a), (b) and (c) can be decided based on the signs of f only. Case (d) is called ambiguous because it requires the exact numerical values of f. There are a few additional exceptional cases that are left aside in this discussion. When $f = 0$ at one of the vertices of the cell, this vertex is also in the polygonal line. It connects to other vertices at opposite edges of the cell, unless one of the cell edges that contain it is included in the polygon. There is no ambiguous situation in that case.

Case (d) with $\delta = 0$ is more of a problem, because it corresponds to a situation in which the interpolated surface is the intersection of two lines and therefore has a singular point. One cannot lift this ambiguity, and one of the options (i) and (ii) should be selected. The selection cannot be completely arbitrary because this could create holes in the reconstructed polygons. One possible rule is to take one option (say (i)) when $\rho > 0$ and the other one when $\rho < 0$. The combination of case (d) and

$\delta = \rho = 0$ implies that $f = 0$ at all vertices of the cell which therefore should be in the final polygon, but there is an unsolvable ambiguity as to how they should be connected.

There is another way to handle case (d), disregarding δ, based, as we just discussed, on the sign of ρ, yielding

(d)′ In case (d) above, take solution (i) if $\rho > 0$ and (ii) otherwise.

The resulting algorithm is simpler, because, in case (d), the sign of ρ can be computed directly based on the signs of f on the vertices of the cell. It does not correspond to the bilinear approximation anymore, but this approximation was somewhat arbitrary anyway. It does break the symmetry of the solution, in the sense that, if f is replaced by $-f$, the isocontours computed using (d)′ will differ. This is illustrated in Fig. 4.2

In addition to allowing for the segmentation of specific shapes from images, when the interior of the shape is, say, darker than its exterior, isocontours have been used as basic components of image processing algorithms that are contrast-invariant in the sense of mathematical morphology. A good introduction to this and to the related literature can be found in [53, 54].

Finally, let us note that isocontours can be easily oriented in accordance with our convention for implicit contours, by simply ensuring that grid points with negative values of f lie on the left of each oriented edge.

4.4.2 *Computing Isosurfaces*

We now pass to the case of level sets for functions defined over three dimensions, and describe the construction of triangulated isosurfaces. Although the problem is in principle similar to the two-dimensional case, the solution is notably more complex, mainly because of the large number of ambiguous situations in the determination of the boundary. There is indeed a large literature on the subject, and the reader can refer (for example) to [39] for a recent bibliography.

The three-dimensional generalization of the algorithm that we have presented for isocontouring is called *marching cubes* [178], and progressively builds a triangulation by exploring every grid cell on which the function changes sign. We will use a notation similar to the previous section, and let $\mathcal{G}$ be a regular three-dimensional grid, with grid coordinates $p(s,t,u) = (x_s, y_t, z_u)$ where $s = 1, \ldots, M, t = 1, \ldots, N, u = 1, \ldots, P$. Denote by $f_{stu} = f(p(s,t,u))$ the observed values of f on the grid. Like in two dimensions, we assume that f extends to the continuum with a trilinear interpolation as follows: Let $C(s,t,u)$ denote the cube (cell) with vertices $p(s+\varepsilon_1, t+\varepsilon_2, q+\varepsilon_3), \varepsilon_i \in \{0,1\}, i = 1,2,3$. Then, for $p = x, y, z \in C(s,t,u)$, let

$$f(p) = \sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3=0}^{1} \prod_{i=1}^{3} (\varepsilon_i r_i(p) + (1-\varepsilon_i)(1-r_i(p)))\, f_{s+\varepsilon_1,t+\varepsilon_2,q+\varepsilon_3}$$

with $r_1(p) = x - x_s, r_2(p) = y - y_t, r_3(p) = z - z_u$.

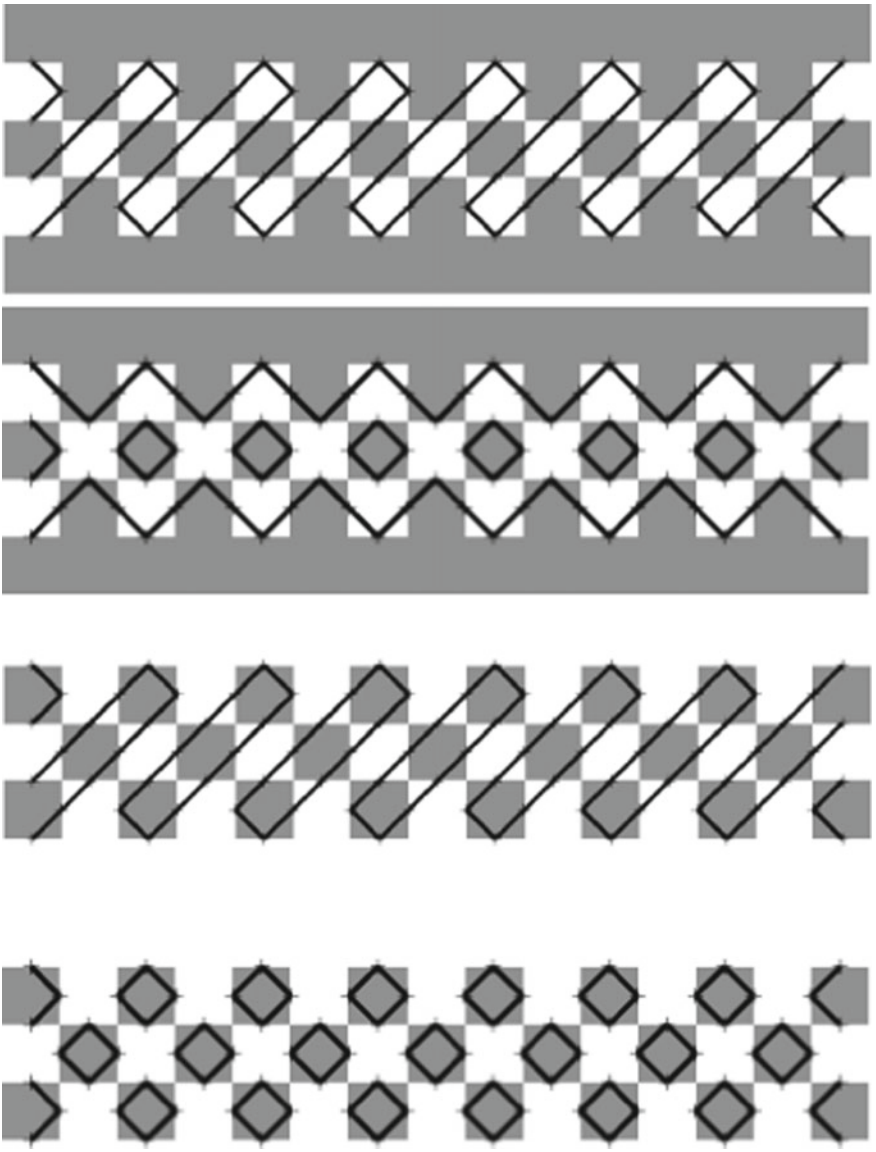

Fig. 4.2 Isocontouring a checkerboard strip using exact bilinear rule (d) (first row) and sign-based rule (d)' (second row). Note that the solutions are different, although both are plausible isocontours for the image. Gray levels are switched in the last two rows, without changing the solution for rule (d) (third row) and significantly altering it for rule (d)' (fourth row), yielding a third plausible solution

The determination of the vertices of the triangulation is similar to the two-dimensional case: the intersections of the level set $f = 0$ and the edges of the cubes $C(s, t, u)$ can be computed by solving a simple linear equation; on a given edge,

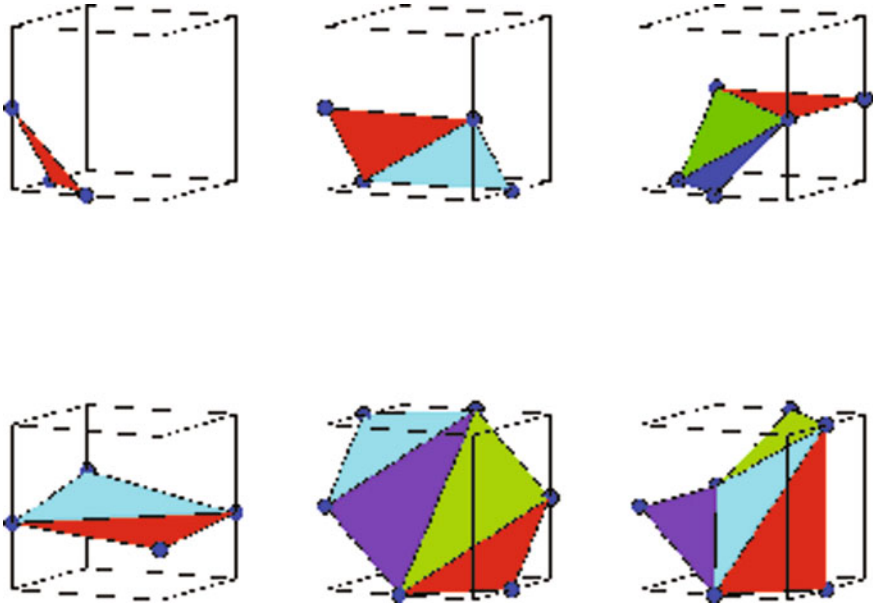

Fig. 4.3 Two-component (non-ambiguous) cases for the marching cubes algorithm

such an intersection exists only if f takes different signs at the end-points, and there can be at most one intersection. The difficulty is how to group these vertices into faces that provide a topologically consistent triangulation.

The main contribution of the marching cubes algorithm is to provide a method in which each cube is considered independently, yielding a reasonably simple implementation. The method works by inspection of the signs of f at the eight vertices of the cube. Like in two dimensions, there are some easy cases. The simplest is when all signs are the same, in which case the triangulation has no node on the cube. Other simple configurations are when the cube vertices of positive sign do not separate the other vertices in two or more regions and vice-versa. In this case, the triangulation has to separate the cube into two parts. There are, up to sign and space symmetry and up to rotation, six such cases, which are provided in Fig. 4.3.

Such triangulations can be efficiently described by labeling the vertices and the edges of the cube, as described in Fig. 4.4. We can describe a sign configuration on the cube by listing the vertices which have a positive sign. We can also describe each triangulation by listing, for each triangle, the three edges it intersects. Figure 4.3 therefore describes the six triangulations

{1} : [(1, 4, 9)]
{1, 2} : [(2, 4, 9), (2, 4, 10)]
{2, 5, 6} : [(1, 2, 9), (2, 8, 9), (2, 8, 6)]
{1, 2, 5, 6} : [(2, 6, 4), (4, 6, 8)]
{2, 3, 4, 7} : [(1, 10, 6), (1, 6, 7), (1, 7, 4), (4, 7, 12)]
{1, 5, 6, 7} : [(1, 10, 11), (1, 11, 8), (8, 11, 7), (4, 1, 8)]

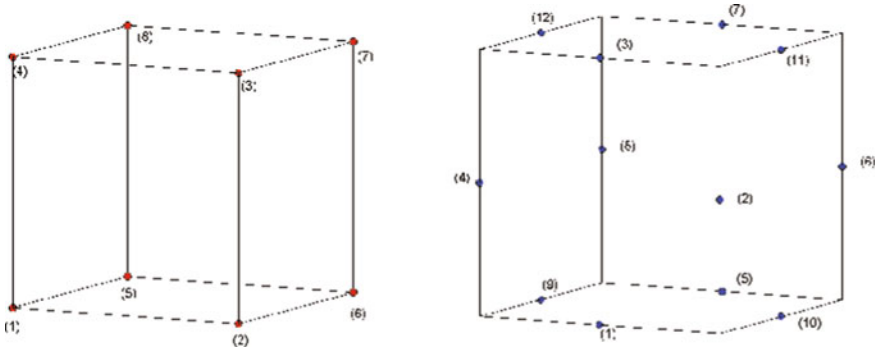

Fig. 4.4 Labels for the vertices (left) and edges (right) of the cube

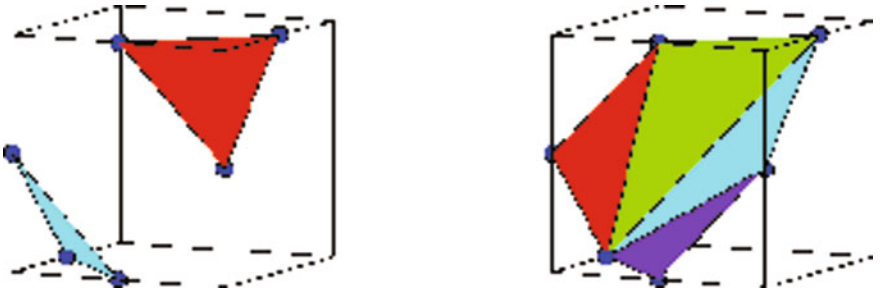

Fig. 4.5 Two triangulations associated to the {3, 8} sign configuration [(2, 3, 11), (1, 4, 9)] and [(4, 9, 3), (3, 9, 11), (9, 11, 2), (1, 2, 9)]

The cases when the signs form more than two connected components on the cube are problematic. They are ambiguous, because the way the surface crosses the cube cannot be decided from the sign pattern alone. One needs to rely on more information (i.e., the actual values of f at the nodes) to decide how to triangulate the surface within the cube, in order to avoid creating topological inconsistencies.

Take, for example, the case in which the cube vertices labeled (1) and (3) have signs distinct from the rest. Then, there are two possible ways (described in Fig. 4.5) in which the surface can cross the cube.

Another kind of ambiguous configuration is when two vertices in two opposite corners are isolated from the rest. Consider, for example, the situation when vertices 1 and 7 are positive while the rest are negative. Then the surface can do two things: either cut out the corners of the cube, or create a tunnel within the cube (see Fig. 4.6).

There have been successive attempts to improve the marching cubes algorithm from its original version ([178], in which the discussion was incomplete) [64, 209, 215, 218, 291] and untying the ambiguous cases. In addition to the two cases described in Figs. 4.5 and 4.6, five other ambiguous sign configurations can be listed, arising from combinations of these two basic cases. A complete description of all possible cases has been provided in [64], together with disambiguation rules. An

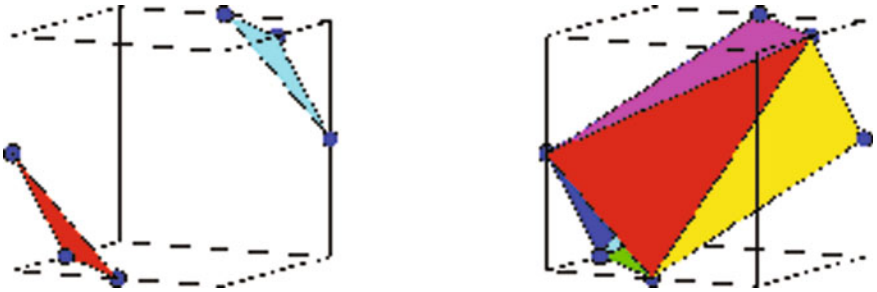

Fig. 4.6 Two triangulations associated to the {1, 7} sign configuration [(1, 4, 9), (7, 6, 11)] and [(1, 4, 11), (1, 11, 6), (1, 9, 6), (9, 6, 7), (9, 4, 7), (4, 11, 7)]

extensive theoretical and numerical analysis of the algorithm has been provided in [217] to which the reader is referred for complementary information, with the listing of all possible topologies within the cube.

If one drops the requirement to provide an accurate triangulation of zero-crossings of the linear interpolation of f within each cube, a reasonably simple option is available [209]. This approach has the disadvantage of breaking the sign-change invariance (which ensures that the computed triangulation should not change if f is

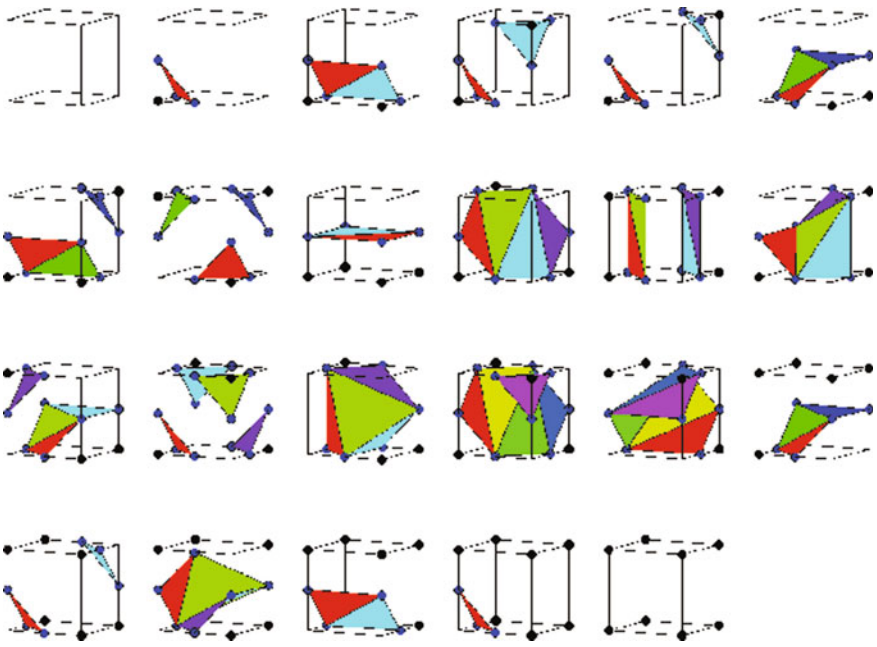

Fig. 4.7 Twenty-three configurations for consistent within-cube triangulation based on vertex signs. Dotted vertices correspond to positive values of the function

replaced by $-f$), but provides a very simple algorithm, still based on the signs of f on the vertices (it can be seen as a generalization of (d)′ in our discussion of the two-dimensional case). This results in 23 different cases (up to rotation invariance), listed in Fig. 4.7. This had to be compared to the 15 cases initially proposed in [178], which was invariant under sign change, but created topological errors.

An alternative to the marching cubes algorithm replaces cubic cells by tetrahedrons before computing the triangulation, which, when properly handled [57], provides a simpler and more stable procedure.

Extracting surfaces as level sets of functions is important even when the original data is not a three-dimensional image from which the region of interest is an isosurface. For example, when the original data is a set of unstructured points that roughly belong to the surface (i.e., they are subject to small errors) some of the commonly used algorithms that reconstruct the surface first reduce to the isosurface problem, trying to infer the signed distance function to the surface, at least in a neighborhood of the observed points. The approach used in [154] is to first approximate the tangent plane to the surface and then build the signed distance function. A similar goal is pursued in [11], using an approach based on computational topology.

Marching cubes (or tetrahedrons) have the drawback of providing a very large number of triangles, sometimes with very acute angles. Simplifying meshes is also the subject of a large literature, but this will not be addressed here (see, for example [100]).

Chapter 5
Evolving Curves and Surfaces

In this chapter, we discuss how curve or surface evolution can be formulated using partial differential equations, and discuss some applications in curve smoothing and image segmentation.

5.1 Curve Evolution

We consider, in this section, curves that depend on time, which will be denoted by t, the curve parameter being denoted by u (or s for arc length). A time-dependent curve is a function of two variables

$$\begin{aligned} m : [0,1] \times [0,\Delta] &\to \mathbb{R}^2 \\ (t,u) &\mapsto m(t,u). \end{aligned}$$

We therefore assume that the domain over which the curve is parametrized (the interval $[0,\Delta]$) is fixed. The curve at time t will be denoted $m_t : u \mapsto m(t,u)$. Its length will be denoted L_t, and the arc length $s_t : [0,1] \to [0,L_t]$. The curvature at a point p will be $\kappa_t(p)$; $T_t(p)$ and $N_t(p)$ will be the unit tangent and normals at p. Differentiation with respect to time will be denoted ∂_t. Differentiation with respect to curve (or, later in this chapter, surface) parameters will use the same convention as in the previous chapters, using dots or ∂ without subscript for univariate functions, and $\partial_1, \partial_2, \ldots$ for multivariate functions.

We consider curves evolving according to differential equations of the kind:

$$\partial_t m_t(u) = A(t,u)T_t(m_t(u)) + B(t,u)N_t(m_t(u)). \tag{5.1}$$

In this equation, $A(t,u)$ and $B(t,u)$ depend on the curve at time t and are scalar functions that depend on the parameter u. Most of the time, they will involve local properties of the curve at u (such as the curvature).

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_5

The decomposition of the evolution into tangent and normal terms is useful, because each of them is associated to different properties of the evolution. The normal term is directly related to the geometric evolution of the curve, as implied by the following lemma:

Lemma 5.1 ([101]) *Assume that m is twice continuously differentiable in space, continuously differentiable in time, regular at all times $t \in [0, \Delta]$ and satisfies the equation*

$$\partial_t m = AT + BN.$$

Then, there exists a time-dependent change of parameter on m, denoted ψ_t, such that $\psi_0(u) = u$ and $\tilde{m}_t(u) := m_t(\psi_t(u))$ is a solution of

$$\partial_t \tilde{m} = \tilde{B} N$$

with $\tilde{B}(t, u) = B(t, \psi_t(u))$.

Proof Let $u \mapsto \psi_t(u)$ be as in the lemma. The evolution of $\tilde{m}_t(u) = m_t(\psi_t(u))$ is

$$\begin{aligned} \partial_t \tilde{m}_t(u) &= \partial_t m_t(\psi_t(u)) + \partial_t \psi_t(u)\, \partial(m_t(\psi_t(u))) \\ &= (A(t, \psi_t(u)) + \partial_t \psi_t(u)\, |\dot{m}_t(\psi_t(u))|)T + B(t, \psi_t(u))N. \end{aligned}$$

We therefore need to show that there exists a ψ such that

$$A(t, \psi_t) + \partial_t \psi_t\, |\dot{m}_t(\psi_t)| = 0.$$

This results from the general theory of ordinary differential equations (cf. Appendix C). For fixed u, let $\xi(t) = \psi_t(u)$. This function must satisfy $\xi(0) = u$ and the equation

$$\partial_t \xi = -A(t, \xi)/|\dot{m}_t(\xi)|.$$

Existence and uniqueness of the solution is ensured by the fact that A and $\dot{m}$ are C^1. For $\dot{m}$, this is true because m is assumed to be C^2 in space, and for A, it suffices to observe that $A(t, \cdot) = \partial_t m_t^T T_t$ and therefore is C^1 with respect to its second variable. □

This lemma implies that the evolution of the curve is essentially captured by the function B, where A only induces changes of parameter.

Very often, the curve variation at time t only depends on the curve at the same time, so that there exist transformations $m \mapsto (\alpha_m, \beta_m)$ with $\alpha_m, \beta_m : [0, \Delta] \to \mathbb{R}$ such that

$$A(t, u) = \alpha_{m_t}(u) \text{ and } B(t, u) = \beta_{m_t}(u).$$

(α and β could also be made to depend on time.)

As usual, we say that these functions are reparametrization-invariant if

$$\beta_{m\circ\psi} = \beta_m \circ \psi$$

for any change of parameter ψ (and similarly for α, although this is of less interest). In view of Lemma 5.1, we see that if β_m is reparametrization-invariant, the evolution

$$\partial_t m = \alpha_m T + \beta_m N$$

can be transformed, after a change of parameter, to

$$\partial_t \tilde{m} = \beta_{\tilde{m}} N.$$

We now discuss the evolution of contours and domain integrals associated to curves evolving according to (5.1). The results are summarized in the following theorem.

Theorem 5.2 *Let $V : \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function, and assume that $m : (t.u) \mapsto m(t,u)$ satisfies (5.1), has continuous partial derivatives $\partial_t \dot{m}_t$, and is such that $u \mapsto m_t(u)$ is a C^2 regular curve for all $t \in [0, t_0]$. Define*

$$F(t) = \int_{m_t} V d\sigma_{m_t} = \int_0^{\Delta} V(m_t(u))|\dot{m}_t(u)|du.$$

Then

$$\partial_t F = [V(m)(\partial_t m)^T T]_0^{\Delta} + \int_{m_t} (\nabla V^T N_t - \kappa_t) N_t^T \partial_t m d\sigma_{m_t} \tag{5.2}$$

(the first term vanishing if the curve is closed). Here κ_t is the curvature of m_t.

Assuming that m_t is simple (and closed) for $t \in [0, t_0]$, and letting Ω_t denote its interior, define

$$G(t) = \int_{\Omega_t} V(x)dx.$$

Then

$$\partial_t G = -\int_{m_t} V N_t^T (\partial_t m_t) d\sigma_{m_t}. \tag{5.3}$$

Finally, let $W : \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 vector field, and

$$H(t) = \int_{m_t} W^T N_t d\sigma_{m_t},$$

where m_t can be open or closed. Then

$$\partial_t H = -[\det(W(m_t), \partial_t m_t)]_{u=0}^{\Delta} + \int_{m_t} \operatorname{div}(W) N_t{}^T (\partial_t m_t) d\sigma_{m_t}. \tag{5.4}$$

Proof First consider the line integral, F. We have $\partial_t |\dot{m}_t|^2 = 2\partial_t \dot{m}_t^T \dot{m}_t$ and, since this is also equal to $2\,|\dot{m}_t|\,\partial_t(|\dot{m}_t|)$, we can write

$$\begin{aligned}
\partial_t F &= \int_0^{\Delta} \nabla V(m_t)^T \partial_t m_t |\dot{m}_t| du + \int_0^{\Delta} V(m_t)\, \partial_t \dot{m}_t^T T_t du \\
&= \int_0^{\Delta} \nabla V(m_t)^T \partial_t m_t |\dot{m}_t| du + [V(m_t)\, \partial_t m_t{}^T T_t]_0^{\Delta} \\
&\quad - \int_0^{\Delta} ((\nabla V(m_t)^T \dot{m}_t)(T_t{}^T \partial_t m_t) + V(m_t)\, \partial_t m_t{}^T \dot{T}_t) du \\
&= [V(m)\, \partial_t m_t{}^T T_t]_0^{\Delta} + \int_0^1 (\nabla V(m_t) - (\nabla V(m_t)^T T_t) T_t)^T \partial_t m_t |\dot{m}_t| du \\
&\quad - \int_0^1 \kappa_t V(m_t)\, (\partial_t m_t^T N_t)\, |\dot{m}_t|\, du \\
&= [\partial_t m_t{}^T T_t]_0^{\Delta} + \int_{m_t} ((\nabla V(m_t)^T N_t) N_t - \kappa_t N_t)^T \partial_t m_t d\sigma_{m_t}.
\end{aligned}$$

This proves (5.2). We now prove (5.3) and (5.4), first observing that the former is a consequence of the latter. Indeed, introduce φ such that $V = \operatorname{div}\varphi$, taking, for example,

$$2\varphi(x_1, x_2) = \left(\int^{x_1} V(x_1', x_2) dx_1', \int^{x_2} V(x_1, x_2') dx_2' \right).$$

Then, from the divergence theorem

$$G(t) = -\int_{m_t} \varphi^T N_t d\sigma_{m_t}$$

so that (5.3) is deduced from (5.4) and the fact that m_t is assumed to be closed.

For (5.4), we can write

$$H(t) = -\int_0^{\Delta} \det(W(m_t), \dot{m}_t) du$$

so that

$$\partial_t H = -\int_0^{\Delta} \det(dW(m_t)\partial_t m_t, \dot{m}_t) du - \int_0^{\Delta} \det(W(m_t), \partial_t \dot{m}_t) du$$

$$= -\int_0^{\Delta} \det(dW(m_t)\partial_t m_t, \dot{m}_t)du - [\det(W(m_t), \partial_t m_t)]_0^{\Delta}$$
$$+\int_0^{\Delta} \det(dW(m_t)\dot{m}_t, \partial_t m_t)du$$

and the conclusion comes from the identity, true for any 2 by 2 matrix A and vectors e_1, e_2,

$$\det(Ae_1, e_2) + \det(e_1, Ae_2) = \text{trace}(A)\det(e_1, e_2)$$

applied to $A = dW(m_t)$, $e_1 = \partial_t m_t$ and $e_2 = \partial_u m_t$. □

Remarks

The reader who has looked ahead at Appendix B may recognize the last statement (the derivative of H) as a special case of Corollary B.30 (in the embedded case, but the immersed case can also be deduced from Theorem B.29). It is interesting to note also that not only (5.3) but also (5.2) are consequences of (5.4) (see the proof of Theorem 5.4).

We also notice that one only needs V to be continuous for (5.3) to hold.

5.1.1 *Grassfires*

As a first example of (5.1), let us consider the simplest case for which $A = 0$ and $B = 1$. This corresponds to

$$\partial_t m_t = N_t, \tag{5.5}$$

which (taking as usual the inward normal) means that m shrinks towards its interior at unit speed. Such evolutions are often called *grassfires* because the evolving curve would look like the boundary of a lawn around which a fire is set at time 0, with the fire propagating inward. They are closely related to medial axes (defined in Chap. 2), because the skeleton is the location at which two separate grassfire fronts meet.

It is quite easy to study (5.5) and prove that solutions exist in small time when starting with a smooth curve, but that singularities are developed in finite time.

First, let us assume that a solution is known, in the form $m(t, u)$ for $t \in [0, t_0]$, $m(0, \cdot)$ being a simple closed curve, and $m(t, \cdot)$ being regular and C^2 for all t in this interval. We first prove that the normals remain unchanged: $N_t(u) = N_0(u)$ for all $u \in [0, \Delta]$. Indeed, because $|N_t| = 1$ at all times, we have $\partial_t N_t^T N_t = 0$. Moreover

$$\partial(\partial_t m) = \dot{N}_t = -|\dot{m}_t|\kappa_t T_t = -\kappa_t \dot{m}_t.$$

This implies

$$0 = \partial_t(\dot{m}_t^T N_t) = \partial_t \dot{m}_t^T N_t + \dot{m}_t^T \partial_t N_t = \dot{m}_t^T \partial_t N_t\,.$$

Therefore, $\partial_t N_t = 0$, since it is perpendicular to both normal and tangent; so the normal is constant over time. Given this, the integration of the evolution equation is straightforward and yields

$$m_t(u) = m_0(u) + t N_0(u) \tag{5.6}$$

and it is easy to show that this provides a solution of (5.5).

In this computation, we have used the assumption that m is smooth (in u) for all $t \in [0, t_0]$. We now check that this fails to be true in finite time, whatever the initial curve is. Indeed, as long as the computation is valid, we have, computing the derivative of (5.6) (assuming that the curve is initially parametrized by arc length)

$$\dot{m}_t(u) = (1 - t\kappa_0(u))T_0(u).$$

In particular, if $\kappa_0(u) > 0$, then, for $t = 1/\kappa_0(u)$, we have $\dot{m}_t = 0$, and the curve is not regular anymore (the previous discussion becomes invalid at this point). Note that there must exist points of positive curvature on the curve, since, for simple positively oriented curves, the integral of the curvature (the rotation index) is 2π. The curvature for small t can be computed using

$$\dot{T}_t(u) = |\dot{m}_t|\, \kappa_t(u) N_t(u).$$

But, since N_t is constant in time, so is T_t, which implies

$$\dot{T}_t(u) = \dot{T}_0(u) = \kappa_0(u) N_0(u).$$

Therefore

$$\kappa_t(u) = \frac{\kappa_0(u)}{1 - t\kappa_0(u)}.$$

This implies that the curvature tends to infinity at the point u_0 of highest curvature in the initial curve, when t tends to $t_0 = 1/\kappa_0(u_0)$.

Even after t_0, we can still define a curve m using (5.6). A very detailed description of what happens immediately after t_0 in the neighborhood of u_0 can be made: this is part of the theory of singularities, and it can be shown that the curve crosses itself, the singularity at u_0 forking into two new singularities, providing a shape called a "swallow tail" (see Fig. 5.1). There are other types of singularity which can be created over time; for example, non-contiguous arcs of the original curve may meet and the region split into two parts (see the second example in Fig. 5.1), creating two new singularities that will evolve.

Returning to the grassfire analogy, however, it makes sense (physically) to require that grass which is already burned does not burn again. So, in the grassfire model, the swallow tail part, and other curve portions after self-intersection should not be included in the evolving curve (Fig. 5.1). An important observation is that both evolutions can be seen as solutions of the original equation (5.5). Its solutions are therefore not uniquely defined once singularities appear.

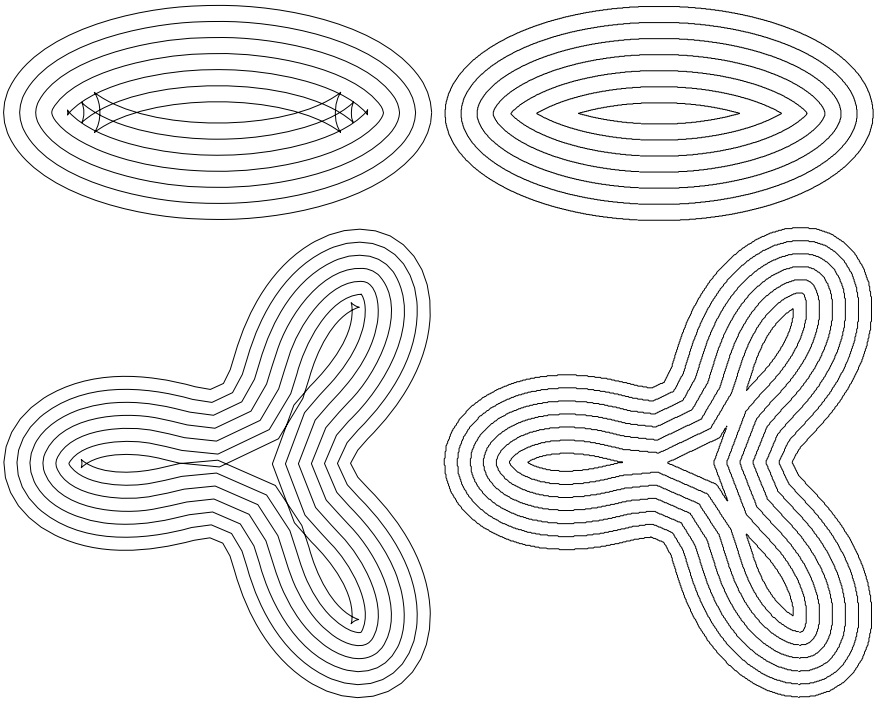

Fig. 5.1 Curve evolution with the grassfire equation. Left: evolution according to (5.6). Right: same evolution after removing the burned-out parts

The location of the points of self-intersection (the first one being at u_0) is interesting since these points belong to the medial axis. (Note that there will be several such points at larger times, each one starting at a local maximum of the curvature.) Tracking them over time provides a feasible algorithm for computing the skeleton [262–264, 274].

5.1.2 Curvature Motion

We now take, in (5.1), $A = 0$ and $B = \kappa_t$, the curvature of m_t. This gives the equation:

$$\partial_t m_t(u) = \kappa_t(u) N_t(u). \tag{5.7}$$

Note that, because κ_t and N_t change signs together when the orientation is changed, this evolution does not depend on the curve's orientation (the previous one did). The following theorem provides a detailed description of how a curve evolves under this equation.

Theorem 5.3 ([118, 133]) *Assume that $u \mapsto m_0(u)$ is a regular, C^2, closed and simple curve in $\mathbb{R}^2$. Then the previous equation has a solution over an interval $[0, t_0]$. The curve remains simple during the evolution. It first becomes convex, then shrinks to a point while its shape becomes circular.*

We will not prove this theorem (the proof is very involved). The interested reader can refer to [66] for a proof and more results on related evolutions. The following simple observations can help, however, in understanding why such a result holds.

The first of these observations is that (5.7) can be interpreted as gradient descent for the function $m \mapsto L(m) = \text{length}(m)$. Indeed, applying Theorem 5.2 with $V = 1$ yields (for any evolving curve $\varepsilon \mapsto \mu(\varepsilon, \cdot)$ with $\mu(0, \cdot) = m$)

$$\partial_\varepsilon L(\mu(\varepsilon, \cdot)) = -\int_{\mu^\varepsilon} \kappa^{\mu^\varepsilon} \partial_\varepsilon \mu^T N^{\mu^\varepsilon} d\sigma_{\mu^\varepsilon}.$$

Taking this at $\varepsilon = 0$, we see that $(-\kappa N)$ is the gradient of the length functional for the metric (cf. Appendix D)

$$\langle h\,,\, \tilde{h} \rangle_m = \int_m h^T \tilde{h}\, d\sigma_m.$$

This implies that the length decreases over time (curvature motion is also called the curve-shortening flow) and we have in fact (if m satisfies (5.7))

$$\partial_t L_t = -\int_0^1 \kappa_t^2 |\dot{m}_t| du = -\int_{m_t} \kappa_t^2 d\sigma_{m_t}.$$

The variation of the enclosed area is also interesting to compute. Letting A_t denote the enclosed area of m_t, we can again apply Theorem 5.2 with $V = 1$ to obtain

$$\partial_t A_t = -\int_{m_t} (N_t^T)(\kappa_t N_t)\, d\sigma_{m_t} = -\int_{m_t} \kappa_t\, d\sigma_{m_t} = -2\pi.$$

The area of the curve therefore decreases at constant speed, 2π. This also shows that the curve disappears in finite time (initial area divided by 2π).

The last interesting quantity is the isoperimetric ratio, which is given by $r_t = A_t / L_t^2$. The isoperimetric inequality (1.19) states that this ratio is always smaller than $1/(4\pi)$, and equal to this value if only if the curve is a circle. We have

$$\begin{aligned} \partial_t r_t &= \frac{1}{L_t^3} \left(L_t \partial_t A_t - 2 A_t \partial_t L_t \right) \\ &= \frac{2A_t}{L_t^3} \left(-\pi \frac{L_t}{A_t} + \int_{m_t} \kappa_t^2 d\sigma_{m_t} \right). \end{aligned}$$

Such a quantity has been analyzed in [117], in which it is shown that

$$\pi \frac{L}{A} \leq \int_m \kappa^2 d\sigma_m$$

as soon as the curve m is convex. When this is true, this implies that the isoperimetric ratio increases during the evolution. Since the ratio is at a maximum for circles, this explains why the curve becomes circular (this does not explain why the curve becomes convex, which is the most difficult part of Theorem 5.3).

This flow can be used as a tool for smoothing curves, since it asymptotically provides a circle. However, this smoothing is combined with an asymptotic reduction to a dot which is a somewhat unwanted behavior. One way to deal with this is to simply let the curve evolve over a time t and rescale it to its original area. The evolution can also be compensated in real time so that the area remains constant: it suffices to use the equation

$$\partial_t m_t = (\kappa_t - 2\pi/L_t) N_t. \tag{5.8}$$

For such curves, the previous computations show that the area satisfies

$$\partial_t A_t = -\int_0^1 (\kappa_t - 2\pi/L_t)\, |\dot{m}_t|\, du = 0$$

and for the length

$$\begin{aligned} \partial_t L_t &= -\int_0^1 \kappa_t (\kappa_t - 2\pi/L_t)\, |\dot{m}_t|\, du \\ &= -\int_0^1 (\kappa_t - 2\pi/L_t)^2\, |\dot{m}_t|\, du. \end{aligned}$$

So the length of the evolving curve decreases, unless κ is constant (equal to $2\pi/L_t$), in which case the considered curve is a circle. For the isoperimetric ratio, we have

$$\begin{aligned} \partial_t r_t &= \frac{1}{L_t^3}(L_t \partial_t A_t - 2A_t \partial_t L_t) \\ &= \frac{2A_t}{L_t^3} \int_{m_t} (\kappa_t - 2\pi/L_t)^2 d\sigma_{m_t}. \end{aligned}$$

This therefore also increases unless the curve is a circle.

5.1.3 *Implicit Representation of the Curvature Motion*

Equation (5.7) may look simple when expressed in terms of geometric quantities, but it is a rather complicated partial differential equation when seen in a fixed parametrization, since it can be rewritten as

$$\partial_t m_t = \frac{\ddot{m}_t}{|\dot{m}_t|^2} - \frac{\ddot{m}_t^T \dot{m}_t}{|\dot{m}_t|^4} \dot{m}_t. \tag{5.9}$$

The direct numerical implementation (by finite differences) of this equation is somewhat unstable (it clearly involves divisions by very small numbers). One obtains a much more stable algorithm if an implicit parametrization is used [227, 255, 256].

For this purpose, assume that the initial curve is the zero level set of a smooth function f^0, so that

$$\mathcal{R}_{m_0} = \left\{ p \in \mathbb{R}^2,\ f^0(p) = 0 \right\}.$$

The principle of an implicit implementation is to make f^0 evolve over time so that its zero level set evolves according to (5.7).

Introduce a time-dependent function $(t, p) \mapsto f_t(p)$, with $f_0 = f^0$. In order for the zero level set of f_t to coincide with the curve m_t defined above, we need $f_t(m_t(u)) = 0$ for all t and all $u \in [0, 1]$, which implies, after differentiation:

$$\partial_t f_t + \nabla f_t(m_t)^T \partial_t m_t = 0.$$

Using $\nabla f_t(m_t) = -|\nabla f_t(m_t)| N_t$, $\partial_t m_t = \kappa_t N_t$ and formula (1.26) for the curvature, this yields the equation

$$\partial_t f_t = |\nabla f_t| \operatorname{div} \frac{\nabla f_t}{|\nabla f_t|}. \tag{5.10}$$

This equation (which is an *anisotropic diffusion*) is very stable and easy to implement (see the next section). Figure 5.2 provides some examples of evolving curves. This

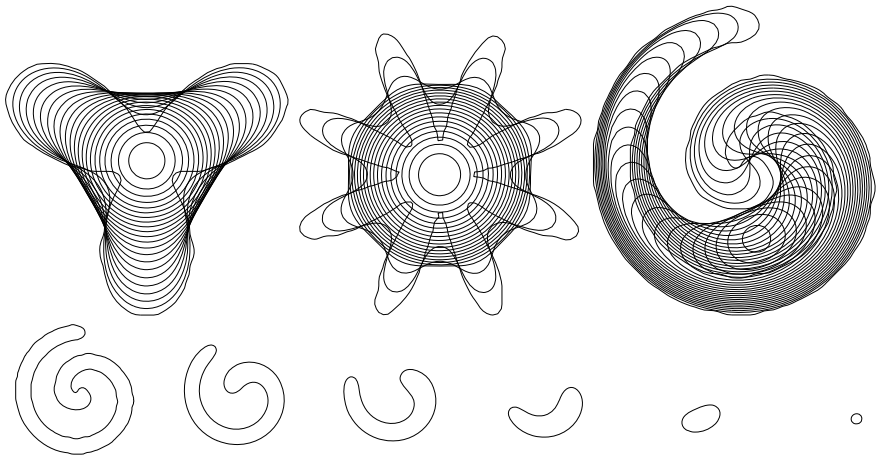

Fig. 5.2 Curves evolving according to the curve-shortening flow. *First row*: three examples of superimposed evolving shapes. *Second row*: details of the evolution of the third shape (spiral) in the first row

equation has also proved itself important in image processing as a way to smooth images while preserving edges (see [9, 10, 103]).

5.1.4 More on the Implementation

We now describe a numerical implementation of Eq. (5.10). It is the simplest one, although not the most efficient (see [255]).

Initializing the Process

The function f^0 must be initialized so that its zero level set coincides with the initial curve $m(0,\cdot)$. A simple choice is to take the signed distance function $f^0(m) = \varepsilon d(m, m(0,\cdot))$ with $\varepsilon = 1$ outside $m(0,\cdot)$ and $\varepsilon = -1$ inside $m(0,\cdot)$. The distance function can be computed efficiently using the algorithm described in Sect. F.4 of Appendix F (Fig. 5.3).

The determination of the points belonging to the interior of the initial curve can be done using standard algorithms in computational geometry [233]. The following algorithm is applicable when the curve is discretized finely enough so that no hole exists in its outline and it does not meet the boundary of the image. Pixels belonging to the curve are assumed to be labeled with 0. Let the initial number of labels be Nlab = 1.

- Label the first column and the first row of the image as 1.
- *First scan*: Scan each row, labeling each pixel like its predecessor in the row, unless this predecessor is a zero. In this case, look at the predecessor in the column: if it has a non-zero label, use this label, otherwise, create a new label, (Nlab + 1), to label the pixel, and add 1 to Nlab.

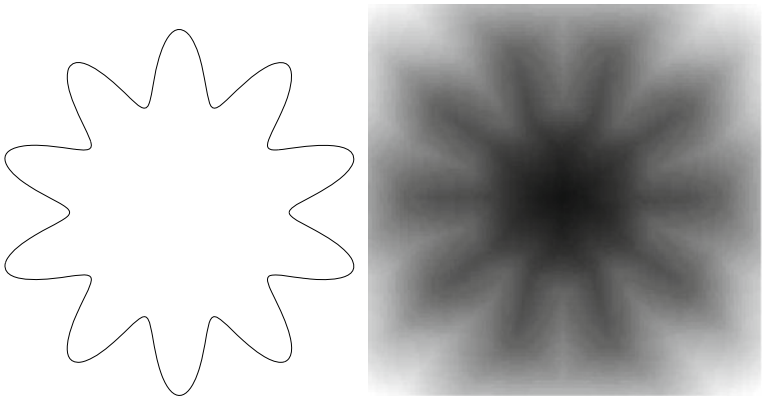

Fig. 5.3 Signed distance map. *Left*: original curve. *Right*: associated signed distance map

- *Second scan*: The previous scan results in over-segmented regions. The last step consists in merging labels by running a backward scan (starting from the last pixel of the last row). Two labels are merged when they are attributed to two neighbor pixels (not separated by a zero).

Finite-Difference Scheme

We now describe how (5.10) can be discretized. Let δt be the time step. The time-discretized evolution algorithm is

$$f^{(n+1)} = f^{(n)} + \delta t \left|\nabla f^{(n)}\right| \operatorname{div}\left(\frac{\nabla f^{(n)}}{\left|\nabla f^{(n)}\right|}\right).$$

δt is the "time" discretization step.

One needs to compute finite-difference approximations of the derivatives, involving a "space" discretization step δx. We first compute an alternative expression for $\operatorname{div}\left(\frac{\nabla f}{|\nabla f|}\right)$. Starting from

$$\partial_1\left(\frac{\partial_1 f}{\sqrt{(\partial_1 f)^2 + (\partial_2 f)^2}}\right) + \partial_2\left(\frac{\partial_2 f}{\sqrt{(\partial_1 f)^2 + (\partial_2 f)^2}}\right),$$

we get

$$\begin{aligned}
&\frac{\partial_1^2 f}{\sqrt{(\partial_1 f)^2 + (\partial_2 f)^2}} - \frac{(\partial_1 f)^2 \partial_1^2 f + \partial_1 f\, \partial_2 f\, \partial_1 \partial_2 f}{((\partial_1 f)^2 + (\partial_2 f)^2)^{3/2}} \\
&\qquad + \frac{\partial_2^2 f}{\sqrt{(\partial_1 f)^2 + (\partial_2) f^2}} - \frac{(\partial_2 f)^2 \partial_2^2 f + \partial_1 f\, \partial_2 f\, \partial_1 \partial_2 f}{((\partial_1 f)^2 + (\partial_2 f)^2)^{3/2}} \\
&\qquad\qquad = \frac{\partial_1^2 f (\partial_2 f)^2 - 2\partial_1 f\, \partial_2 f\, \partial_1 \partial_2 f + \partial_2^2 f (\partial_1 f)^2}{((\partial_1 f)^2 + (\partial_2 f)^2)^{3/2}}.
\end{aligned}$$

To discretize the derivatives, one can use symmetric approximations:

$$\begin{aligned}
\partial_1 f(i,j) &= (f(i+1,j) - f(i-1,j))/2 \\
\partial_2 f(i,j) &= (f(i,j+1) - f(i,j-1))/2 \\
\partial_1^2 f(i,j) &= (f(i+1,j) - 2f(i,j) + f(i-1,j)) \\
\partial_2^2 f(i,j) &= (f(i,j+1) - 2f(i,j) + f(i,j-1)) \\
\partial_1 \partial_2 f(i,j) &= (f(i+1,j+1) - f(i-1,j+1) \\
&\qquad - f(i+1,j-1) + f(i-1,j-1))/4.
\end{aligned}$$

5.2 Surface Evolution

Describing surfaces that evolve over time via a differential equation is not as simple as with curves, because surfaces are, to start with, harder to describe than curves. We can however describe the evolution of *parametrized* surfaces as we did with curves, writing

$$\partial_t m_t(u, v) = \theta(t, u, v) + B(t, u, v)N_t(u, v) \tag{5.11}$$

where $\theta(t, u, v)$ is tangent to m_t at $m_t(u, v)$ (i.e. in the plane generated by $\partial_1 m_t$ and $\partial_2 m_t$ and B is a scalar function).

A slight generalization of this scheme is when another regular, oriented surface, say Σ, is chosen to parametrize the evolving surfaces. We will assume that, for each time t, there exists a function $q \mapsto \mu_t(q)$, defined over Σ such that the evolving surface is $S_t = \mu_t(\Sigma)$ and, for all $q \in S_0$, $d\mu_t(q)$ has rank 2 as a map from $T_q\Sigma$ to $\mathbb{R}^3$ (i.e., μ_t is an immersion).

The set S_t might not be a regular surface (it can have multiple points, for example) so that the tangent planes and normal to S_t at $\mu_t(q)$ may not be defined. However, for any $q \in \Sigma$, one can find a parametrized region in Σ around q (with parametrization $(u, v) = m(u, v)$) which maps into a regular parametrized patch in S_t that has a well-defined tangent plane and normal, that will be respectively denoted $L_{t,q}$ and $N_t(q)$. We can therefore generalize (5.11) by considering surface evolutions of the kind

$$\partial_t \mu_t(q) = \theta(t, q) + B(t, q)N_t(q) \tag{5.12}$$

with $\theta(t, q) \in L_{t,q}$.

If $m^0 : (u, v) \to m^0(u, v)$ is a local chart at q, we get a local parametrization

$$m_t(u, v) = \mu(t, m^0(u, v)) \tag{5.13}$$

on S_t, with

$$\partial_t m_t(u, v) = \theta(t, m^0(u, v)) + B(t, m^0(u, v))N_t(m^0(u, v)).$$

If S_t happens to be an oriented regular surface at all times, then $L_{t,q} = T_{\mu_t(q)}S_t$, and we can think of (5.12) as an equation of the form

$$\partial_t p = \tilde{\theta}(t, p) + \tilde{B}(t, p)N_t(p) \tag{5.14}$$

with $\tilde{\theta}(t, p) \in T_p S_t$.

Similar to curves, the geometric evolution is only captured by B. Let us sketch a justification of this fact. If μ evolves according to (5.12), and $q \mapsto \varphi_t(q)$ is a global change of variable of Σ (a C^1 invertible map from Σ to Σ with a C^1 inverse), then we can consider $\tilde{\mu}_t(q) = \mu_t(\varphi_t(q))$. Then

$$\partial_t \tilde{\mu}_t = d\mu_t \circ \varphi_t \, \partial_t \varphi_t + \theta_t \circ \varphi_t + B_t \circ \varphi_t N_t \circ \varphi_t.$$

So, if

$$\partial_t \varphi_t = -(d\mu_t^{-1} \theta_t) \circ \varphi_t \tag{5.15}$$

then $\tilde{\mu}_t$ follows (5.12) with $\theta = 0$. But (5.15) defines the flow associated to an ordinary differential equation (cf. Appendix C) on Σ and therefore provides a change of parameter.

We have the equivalent of Theorem 5.2:

Theorem 5.4 *Let $V : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function, and assume that a time-dependent regular surface S_t is defined by (5.12) where μ is C^2 as a function of two variables. Define*

$$F(t) = \int_{S_t} V d\sigma_t,$$

where $\sigma_t = \sigma_{S_t}$ is the area measure on S_t. For $p = \mu_t(q) \in S_t$, let $\xi_t(p) = \partial_t \mu_t(q)$. Then

$$\partial_t F = -\int_{\partial S_t} V \, n_t^T \xi_t + \int_{S_t} \left(-2V H_t + \nabla V^T N_t \right) N_t^T \xi_t \, d\sigma_t, \tag{5.16}$$

where H_t is the mean curvature on S_t and n_t is the inward normal to ∂S_t.

Assuming that S_t coincides with the boundary of an open subset $\Omega_t \subset \mathbb{R}^3$, define

$$G(t) = \int_{\Omega_t} V(x) dx.$$

Then

$$\partial_t G = -\int_{S_t} V \, N_t^T \xi_t d\sigma_t. \tag{5.17}$$

Finally, let $W : \mathbb{R}^3 \to \mathbb{R}^3$ be a smooth vector field, and

$$L(t) = \int_{S_t} W^T N_t \, d\sigma_t$$

(where S_t can be open or closed). Then

$$\partial_t L = \int_{\partial S_t} \xi_t^T (\tau_t \times W) \, d\sigma_t + \int_{S_t} \operatorname{div}(W) \, N_t^T \xi_t d\sigma_t, \tag{5.18}$$

where τ_t is the unit tangent to ∂S_t, oriented so that (τ, n, N) forms a direct orthonormal basis.

Proof ("Elementary Proof") We start with a proof which is rather lengthy, but does not require any additional concepts than those that have been introduced so far. We will then provide an alternative shorter proof, based on results from Appendix B.

To analyze the variations of F, take a family $((U_i, m_t^i), i = 1, \dots, n)$ of positively oriented local parametrizations of S_t (the U_i's being independent of time), as given by (5.13), and an adapted partition of unity $(\omega^i, i = 1, \dots, n)$ on Σ. Define, for $t \geq 0$, $\bar{\omega}_t^i(m_t^i(u, v)) = \omega^i(m_0^i(u, v))$, to obtain a partition of unity on S_t and write

$$F(t) = \sum_{i=1}^{N} \int_{U_i} \omega^i(m_0^i(u, v)) V(m_t^i(u, v)) \, |\partial_1 m_t^i \times \partial_2 m_t^i| \, du \, dv. \tag{5.19}$$

We now focus on the variation of one of the terms of the sum; if (U, m_t) is a local patch on S_t, $\omega \circ m_0$ a scalar function on U, define

$$f(t) = \int_U \omega(m_0(u, v)) V(m_t(u, v)) \, |\partial_1 m_t \times \partial_2 m_t| \, du \, dv.$$

Let $h_t = \partial_t m_t$; we have, $\partial_t V(m_t(u, v)) = \nabla V(m_t(u, v))^T h_t$. Also,

$$\partial_t \big(|\partial_1 m_t \times \partial_2 m_t| \big) = \big(\partial_1 h_t \times \partial_2 m_t + \partial_1 m_t \times \partial_2 h_t \big)^T N_t^m,$$

where N_t^m, the normal in the chart, is given by

$$N_t^m(u, v) = \frac{\partial_1 m_t \times \partial_2 m_t}{|\partial_1 m_t \times \partial_2 m_t|}.$$

Rearranging the cross products, we get

$$\partial_t \big(|\partial_1 m_t \times \partial_2 m_t| \big) = (\partial_1 h_t)^T (\partial_2 m_t \times N_t^m) + (\partial_2 h_t)^T (N_t^m \times \partial_1 m_t).$$

We now apply Lemma 3.19, using, as in the lemma,

$$\rho(A) = \operatorname{trace}((\mathrm{Id} - N_t^m (N_t^m)^T) A)$$

for an operator $A : T_p S \to \mathbb{R}^3$. This yields

$$\partial_t \big(|\partial_1 m \times \partial_2 m_t| \big) = \rho(d\xi_t \circ m_t) \, |\partial_1 m_t \times \partial_2 m_t|$$

using the fact that $h_t = \xi_t \circ m_t$ is the local chart representation of ξ_t. We can then write

$$\begin{aligned} \partial_t f &= \int_U \omega \circ m^0 (V \circ m_t \, \rho(d\xi_t \circ m_t) + (\nabla V \circ m_t)^T (\xi_t \circ m_t)) \, |\partial_1 m_t \times \partial_2 m_t| \, du \, dv \\ &= \int_{m_t(U) \cap S_t} \bar{\omega} (V \rho(d\xi_t) + \nabla V^T \xi_t) \, d\sigma_t, \end{aligned}$$

with $\bar{\omega} \circ m = \omega \circ m^0$.

We now can apply this to all the terms in (5.19) to obtain

$$\partial_t F = \int_S (\rho(d\xi_t)V + \nabla V^T \xi_t) d\sigma_t. \tag{5.20}$$

We now eliminate $d\xi_t$ in (5.20), which will require an integration by parts via the divergence theorem. Decompose ξ_t into a tangent and a normal part, namely

$$\xi_t = \xi_t^0 + \eta_t N_t,$$

with $\eta_t = \xi_t^T N_t$ (and therefore $\xi_t^0(p) \in T_p S_t$), so that, by definition of ρ and Eq. (3.24),

$$\rho(d\xi_t) = \mathrm{div}'_{S_t}(\xi_t) = \mathrm{div}_{S_t}\xi_t^0 - 2\eta_t H_t. \tag{5.21}$$

Applying Green's formula, letting n_t be the inward normal to ∂S_t, we get

$$\int_{S_t} V\rho(\xi_t^0) d\sigma_t = -\int_{\partial S_t} V(\xi_t^0)^T n - \int_{S_t} \nabla V^T \xi_t^0 d\sigma_t,$$

because $\mathrm{div}_{S_t}(V\xi_t^0) = V\mathrm{div}_{S_t}(\xi_t^0) + \nabla V^T \xi_t^0$. This in turn implies that

$$\partial_t F = -\int_{\partial S_t} V(\xi_t^0)^T n_t + \int_{S_t} \big(-2V\eta_t H_t - \nabla V^T \xi_t^0 + \nabla V^T \xi_t\big) d\sigma_t,$$

which yields, since $\nabla V^T \xi_t = \nabla V^T \xi_t^0 + \eta_t \nabla V^T N_t$, and using $\eta_t = {\xi_t}^T N_t$, ${\xi_t^0}^T n_t = {\xi_t}^T n_t$:

$$\partial_t F = -\int_{\partial S_t} V\, \xi_t^T n_t + \int_{S_t} \big(-2V H_t + \nabla V^T N_t\big)\,(N_t^T \xi_t) d\sigma_t, \tag{5.22}$$

as needed.

Using the same argument as for curves, one can show that (5.17) is a consequence of (5.18) and of the divergence theorem. We therefore directly prove (5.18). As above, we can decompose the integral over charts using a partition of unity. So we introduce the variation on a single chart and consider, for a scalar function $\omega \circ m_0$,

$$\begin{aligned}\ell(t) &= \int_U \omega \circ m_0(u,v)(W(m_t(u,v))^T(\partial_1 m_t \times \partial_2 m_t))\, du dv \\ &= \int_U \omega \circ m_0(u,v)\det(W(m_t(u,v)), \partial_1 m_t, \partial_2 m_t)\, du dv.\end{aligned}$$

Computing time derivatives, we get, using again $h_t = \partial_t m_t$,

$$\partial_t \ell = \int_U \omega \circ m_0 (dW(m_t)h_t)^T (\partial_1 m_t \times \partial_2 m_t)\, du dv + \int_U \omega \circ m_0 W(m_t)^T (\partial_1 h_t \times \partial_2 m_t + \partial_1 m_t \times \partial_2 h_t)\, du dv.$$

From Lemma 3.19, we know that the normal component of $\zeta_t := \partial_1 h_t \times \partial_2 m_t + \partial_1 m_t \times \partial_2 h_t$ is, writing, with some notation abuse, $\rho(dh_t) = \rho(d\xi_t \circ m_t)$,

$$\zeta_t^T N_t = \rho(dh_t) |\partial_1 m_t \times \partial_2 m_t|.$$

To compute the tangential component, note that

$$\zeta_t^T \partial_1 m_t = (\partial_1 h_t \times \partial_2 m_t)^T \partial_1 m_t = -\partial_1 h_t^T N_t\, |\partial_1 m_t \times \partial_2 m_t|$$

and similarly

$$\zeta_t^T \partial_2 m_t = -\partial_2 h_t^T N_t |\partial_1 m_t \times \partial_2 m_t|.$$

This implies that, for any $\tilde{\zeta} \in T_p S$, we have

$$\zeta^T \tilde{\zeta} = -(dh_t(p)\tilde{\zeta})^T N_t |\partial_1 m_t \times \partial_2 m_t|.$$

So, if we decompose $W = W_t^0 + w_t N_t$ into tangent and a normal components to S_t (taking $w_T = W_t^T N_t$), we get

$$W^T (\partial_1 h_t \times \partial_2 m_t + \partial_1 m_t \times \partial_2 h_t) = (w_t \rho(dh_t) - (dh_t W_t^0)^T N_t)\, |\partial_1 m_t \times \partial_2 m_t|.$$

This gives (taking as above $\bar{\omega}_t(m_t) = \omega \circ m_0$)

$$\partial_t \ell = \int_U \bar{\omega}_t(m_t(u,v))\, (dW(m_t)h_t)^T N_t\, |\partial_1 m_t \times \partial_2 m_t|\, du\, dv + \int_U \bar{\omega}_t(m_t(u,v)) \big(w_t(m_t)\rho(dh_t) - (dh_t W_t^0(m_t))^T N_t\big) |\partial_1 m_t \times \partial_2 m_t|\, du\, dv.$$

This yields (after summing over the partition of the unity)

$$\partial_t L(t) = \int_{S_t} (dW \xi_t + \rho(d\xi_t) W - d\xi_t\, W_t^0)^T N_t d\sigma^t.$$

Using (5.21), we get, letting again $\xi_t = \xi_t^0 + \eta_t N_t$ be the decomposition of ξ_t into tangent and normal components,

$$\partial_t L(t) = -\int_{\partial S_t} w_t \xi_t^T n_t + \int_{S_t} \Big((dW\xi_t)^T N_t - 2\eta_t H_t w_t - (\nabla_{S_t} w_t)^T \xi_t^0 - (d\xi_t W_t^0)^T N_t \Big) d\sigma_t. \tag{5.23}$$

We have (by definition of the gradient)

$$(\nabla_{S_t} w_t)^T \xi_t^0 = dw_t \xi_t^0 = (dW\xi_t^0)^T N_t + (dN_t \xi_t^0)^T W. \tag{5.24}$$

We will also use

$$\begin{aligned} \mathrm{div}_{S_t}(\eta_t W_t^0) &= (\nabla_{S_t}\eta_t)^T W_t^0 + \eta_t \mathrm{div}_{S_t} W_t^0 \\ &= (d\xi_t W_t^0)^T N_t + (dN_t W_t^0)^T \xi_t + \eta_t \mathrm{div}_{S_t} W_t^0 \end{aligned} \tag{5.25}$$

and

$$(dN_t W_t^0)^T \xi_t = (dN_t W_t^0)^T \xi_t^0 = (W_t^0)^T dN_t \xi_t^0 = W^T dN_t \xi_t^0$$

to write

$$\begin{aligned} &(dW\xi_t)^T N_t - 2\eta_t H_t w_t - (\nabla_{S_t} w_t)^T \xi_t^0 - (d\xi_t^0 W_t^0)^T N_t \\ &= \eta_t \big((dWN_t)^T N_t + \mathrm{div}_{S_t} W_t^0 - 2H_t w_t \big) - \mathrm{div}_{S_t}(\eta_t W_t^0) \\ &= \eta_t \big((dWN_t)^T N_t + \rho(dW) \big) - \mathrm{div}_{S_t}(\eta_t W_t^0) \\ &= \eta_t \mathrm{div}(W) - \mathrm{div}_{S_t}(\eta_t W_t^0). \end{aligned}$$

Using this in (5.23) and applying the divergence theorem yields

$$\partial_t L(t) = -\int_{\partial S_t} \big((W^T N_t)(\xi_t^T n_t) - (W^T n_t)(\xi_t^T N_t) \big) d\sigma_{\partial S_t} + \int_{S_t} \mathrm{div}(W)\, \xi_t^T N_t d\sigma_t. \tag{5.26}$$

It now suffices to observe that

$$\xi_t^T(\tau_t \times W) = \det(\tau_t, W, \xi_t) = (W^T n_t)(\xi_t^T N_t) - (W^T N_t)(\xi_t^T n_t)$$

to retrieve (5.18). □

Proof (*Alternative Proof*) Equation (5.26) is exactly the statement of Corollary B.30, which was obtained as a consequence of Stokes's theorem (Theorem B.28), and this yields (5.18) and (5.17), which is a consequence of (5.18) and of the divergence theorem.

We now deduce (5.16) from (5.18) as follows. Fix t_0 and let, for each time t in a neighborhood of t_0, $\hat{N}_t(p)$ be a vector field on $\mathbb{R}^3$ that extends N_t on S_t, i.e., such that $\hat{N}_t(\mu_t(q)) = N_t(\mu_t(q))$ for all $q \in M$. One way to build such a mapping is to introduce the function

$$\varphi_t(q,u) = \mu_t(q) + uN_t(\mu_t(q)).$$

Then φ is C^1 and such that, for any t_0 and $\varepsilon > 0$, there exists an $r_0 > 0$ such that for all $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$, φ_t is a diffeomorphism from $M \times (-r_0, r_0)$ onto an open subset, say Ω_t of $\mathbb{R}^3$ (details are left to the reader). Take any smooth function ρ such that $\rho(0) = 1$ and $\rho(u) = 0$ if $|u| \geq r_0$ and set $\hat{N}_t(\varphi_t(q,u)) = \rho(u)N_t(\mu(s,q))$ on $(t_0 - \varepsilon, t_0 + \varepsilon) \times M \times (-r_0, r_0)$ (which defines $\hat{N}_t$ on Ω_t) and $\hat{N}_t(p) = 0$ if $p \notin \Omega_t$.

Given $\hat{N}_t$, let $W_t(p) = V(p)\hat{N}_t(p)$, so that, for t in a neighborhood of t_0,

$$F(t) = \int_{S_t} W_t^T N_t d\sigma_t.$$

Apply (5.18) (taking into account the fact that W depends on time) to write

$$\partial_t F = \int_{\partial S_t} \xi_t^T(\tau_t \times W_t)d\sigma_t + \int_{S_t} (\partial_t W_t)^T N_t d\sigma_t + \int_{S_t} \operatorname{div}(W_t)\, N_t^T \xi_t d\sigma_t. \quad (5.27)$$

Now $\xi_t^T(\tau_t \times W_t) = V\xi_t^T(\tau_t \times N_t) = -V\xi_t^T n_t$, so that the boundary terms in (5.16) and (5.27) coincide.

We have

$$(\partial_t W_t)^T N_t + \operatorname{div}(W_t)\, N_t^T \xi_t = V\left((\partial_t \hat{N}_t)^T N_t + \operatorname{div}(\hat{N}_t)\, N_t^T \xi_t\right) + (\nabla V^T N_t)(N_t^T \xi_t)$$

so that we need to prove that

$$(\partial_t \hat{N}_t)^T N_t + \operatorname{div}(\hat{N}_t)\, N_t^T \xi_t = -2H_t N_t^T \xi_t,$$

or, since $\operatorname{div}(\hat{N}_t) = -2H_t + N_t^T d\hat{N}_t N_t$, we need to show that $(\partial_t \hat{N}_t)^T N_t + N_t^T d\hat{N}_t N_t = 0$ (all these identities being true on S_t). By assumption, we have $|\hat{N}_t(\mu_t)|^2 = 1$, and taking the derivative in t, we find

$$0 = \hat{N}_t^T \partial_t \hat{N}_t + \hat{N}_t^T d\hat{N}_t \xi_t = N_t^T \partial_t \hat{N}_t + N_t^T d\hat{N}_t \xi_t$$

when these terms are evaluated at $\mu_t(q)$, which is what we needed. □

5.2.1 Mean Curvature Surface Flow

The equivalent of the curve-shortening flow for surfaces is gradient descent for the area functional, the variation of which being given by (5.16) with $V = 1$. The gradient of the area, for the dot product

$$\langle h\,,\,\tilde{h}\rangle_S = \int_S h^T \tilde{h} d\sigma_S$$

therefore is $-2HN$, yielding the gradient descent evolution, called mean curvature flow

$$\partial_t p = 2H(p)N(p). \tag{5.28}$$

This is a regularizing flow for surfaces, and solutions exist in short time. However, singularities may form in the evolving surface, even when starting with relatively simple ones. This could not happen with the curve-shortening flow.

The mean curvature flow in implicit form has exactly the same expression as the curve-shortening flow. Indeed, if the evolving surface is defined (at time t) by $f(t, p) = 0$ (with $\nabla f \neq 0$ on the surface and f differentiable in time), then, using equation (3.26) and the same argument as the one we made with curves, f satisfies the evolution equation

$$\partial_t f = |\nabla f| \mathrm{div}\left(\frac{\nabla f}{|\nabla f|}\right). \tag{5.29}$$

This gives an alternative formulation of the mean curvature flow, that has the advantage of carrying over the solution beyond singularities and allowing for changes of topology. Equation (5.29) is a diffusion that smoothes f tangentially to the surface, since

$$|\nabla f| \,\mathrm{div}\left(\frac{\nabla f}{|\nabla f|}\right) = \Delta f - \langle N\,,\, d^2 f N\rangle.$$

5.3 Gradient Flows

The curvature and mean curvature flows for curves and surfaces are special cases of gradient flows that evolve curves along steepest descent directions relative to a given objective function (the length or the area in these cases), for a given metric (the L^2 metric relative to the arc length or area form).

This approach can be applied to any objective function that has a gradient relative to a chosen dot product; considering, as we did, curves over some interval, or surfaces parametrized over some reference surface, we can consider energies $E(m)$ (or $E(\mu)$) that have a gradient ∇E with respect to some metric, and implement

$$\partial_t m = -\nabla E(m),$$

which can obviously be put into the form (5.1) or (5.12).

As an example beyond mean curvature flows, let us consider plane curves and the energy given by the integral of the squared curvature

$$E(m) = \frac{1}{2}\int_m \kappa^2 d\sigma_m, \tag{5.30}$$

which is often called the bending energy of the curve. To compute the gradient, let us introduce a small perturbation $m(\varepsilon, \cdot)$ of a curve m_0 depending on a parameter ε, such that $m(0, \cdot) = m_0$. We want to compute $\partial_\varepsilon E(m(\varepsilon, \cdot))$.

A first observation that will simplify the computation is that, because E is parametrization-invariant, we can assume that $\partial_\varepsilon m$ is oriented along the normal to $m(\varepsilon, \cdot)$. We can indeed always restrict ourselves to this case using an ε-dependent change of variable that does not affect the value of the energy.

A second simplification will come by introducing the variation with respect to arc length along a curve m, denoted ∂_s^m defined by

$$\partial_s^m f = \frac{\dot{f}}{|\dot{m}|}. \tag{5.31}$$

We will drop the m superscript from the notation and only write ∂_s, although it is important to remember that this operator depends on the curve that supports it.

Assuming that m depends on ε, computing the derivative in ε yields

$$\partial_\varepsilon \partial_s = \partial_s \partial_\varepsilon - \left(\frac{\partial_\varepsilon \dot{m}}{|\dot{m}|}\right)^T T\, \partial_s.$$

If $\partial_\varepsilon m = BN$, this gives

$$\partial_\varepsilon \partial_s = \partial_s \partial_\varepsilon + \kappa B \partial_s. \tag{5.32}$$

Using this identity, we get

$$\partial_\varepsilon \partial_s m = \partial_s(BN) + \kappa BT = (\partial_s B)N$$

and

$$\begin{aligned}\partial_\varepsilon \partial_s^2 m &= \partial_s(\partial_\varepsilon \partial_s m) + \kappa B \partial_s^2 m \\ &= -(\partial_s B)\kappa T + (\partial_s^2 B + \kappa^2 B)N.\end{aligned}$$

Because

$$\partial_\varepsilon \partial_s^2 m = \kappa \partial_\varepsilon N + (\partial_\varepsilon \kappa)N$$

and $\partial_\varepsilon N$ is perpendicular to N, we find

$$\partial_\varepsilon \kappa = \partial_s^2 B + \kappa^2 B \text{ when } \partial_\varepsilon m = BN. \tag{5.33}$$

(We find in passing that the evolution of the curvature with (5.7) is $\partial_t \kappa = \partial_s^2 \kappa + \kappa^3$.)

Combining these results and Theorem 5.2, we can easily compute the variation of E in ε (assuming that m is closed for simplicity), yielding

$$\begin{aligned}\partial_\varepsilon E &= \int_m (\partial_s^2 B + \kappa^2 B)\kappa d\sigma_m - \int_m \kappa^3 B d\sigma_m \\ &= \int_m (\kappa \partial_s^2 B + \kappa^3 B/2) d\sigma_m \\ &= \int_m (\partial_s^2 \kappa + \kappa^3/2) B d\sigma_m,\end{aligned}$$

which provides the gradient of E. Minimizing E by itself leads to the uninteresting solution of a curve blowing up to infinity; indeed, since scaling a curve m by a factor a divides its curvature by a, we have $E(am) = E(m)/a$. However, minimizing E over curves with constant length, or the related problem of minimizing

$$E_\lambda(m) = E(m) + \lambda \,\text{length}(m) \tag{5.34}$$

has been widely studied, in dimension 2 or larger, and curves for which the gradient of E_λ vanishes for some positive λ are called elasticae. We have, using the same notation

$$\partial_\varepsilon E_\lambda = \int_m (\partial_s^2 \kappa + \kappa^3/2 - \lambda\kappa) B \, d\sigma_m \tag{5.35}$$

and the related gradient descent algorithm

$$\partial_t m = -(\partial_s^2 \kappa + \kappa^3/2 - \lambda\kappa) N$$

provides a well-defined and converging evolution [97, 172].

The counterpart of (5.30) for surfaces is the Willmore energy

$$E(S) = \int_S H^2 d\sigma_S, \tag{5.36}$$

and the associated minimizing flow is called the Willmore flow. It is given by

$$\partial_t p = (\Delta_S H + 2H(H^2 - K))N \tag{5.37}$$

(see [302]).

All the flows above were defined as gradient descent for the metric given by the L^2 norm relative to the arc length or the area form. As remarked in Appendix D, changing the metric can induce fundamental changes in the resulting algorithms. For example, as proposed in [272], gradient flows of curves associated to Sobolev metrics

$$\langle h\ ,\ \tilde{h}\rangle_m = \int_m (Ah)^T \tilde{h} d\sigma_m,$$

where A is a differential operator (assumed to be symmetric and positive), can create interesting variants of the original flows. More precisely, they transform an L^2 flow that would initially take the form

$$\partial_t m = -\nabla_{L^2} E(m)$$

into the flow

$$\partial_t m = -A^{-1}(\nabla_{L^2} E(m)),$$

where A^{-1} is the inverse of A and therefore a smoothing operator. This results in an evolution that favors smooth changes.

For closed plane curves, many interesting operators can be represented in terms of Fourier series. For a curve of length L, we can consider the normalized arc length $\tilde{s} = s/L$ defined over $[0, 1]$. Now, a periodic function f defined over $[0, 1]$ can be decomposed (if smooth enough) in terms of its Fourier series

$$f(\tilde{s}) = \sum_{k\in\mathbb{Z}} c_k(f) e^{2i\pi k\tilde{s}}$$

with

$$c_k(f) = \int_0^1 f(u) e^{-2i\pi ku} du.$$

Let $(a_k, k \in \mathbb{Z})$ be a double sequence of positive numbers satisfying $a_{-k} = a_k$. One defines the associated operator by

$$Af(\tilde{s}) = \sum_{k\in\mathbb{Z}} a_k c_k(f) e^{2i\pi k\tilde{s}}.$$

This operator is defined over functions f such that

$$\sum_{k\in\mathbb{Z}} a_k^2 |c_k(f)|^2 < \infty.$$

The inverse operator is then immediately defined by

$$A^{-1} f(\tilde{s}) = \sum_{k\in\mathbb{Z}} \frac{c_k(g)}{a_k} e^{2i\pi k\tilde{s}}.$$

For example, the differential operator $Af = -2\partial_{\tilde{s}}^2 f$ is such that

$$c_k(Af) = 4\pi^2 k^2 c_k(f),$$

as can be easily computed by integration by parts, and is therefore associated to $a_k = 4\pi^2 k^2$. The operator $Af = -\partial_s^2 f + \lambda f$ (with $\lambda > 0$) corresponds to $a_k = 4\pi^2 k^2 + \lambda$, which is positive (implying that A is invertible). Computations such as these, and the fact that Fourier coefficients and Fourier series are discretized by the fast Fourier transform and its inverse, lead to very simple variations of a gradient flow associated to a given objective function (see Sect. 5.4.11).

5.4 Active Contours

5.4.1 Introduction

Active contour methods let curves or surfaces evolve in order to minimize a segmentation energy (i.e., an energy that measures the alignment of the contour or surface with the boundary of a shape present in the image). They provide an important class of curve (and surface) evolution based on specific gradient flows.

We start with a parametrization-dependent formulation (in two dimensions), which is the original method introduced in [165]. We will assume that a function $p \mapsto V(p)$ is defined for $p \in \mathbb{R}^2$. Typically, V is small where a contour is likely to be present in the image (based on some measure of discontinuity within an image, for example). The goal is to ensure that the evolving contour settles along regions where V is small while remaining a smooth closed curve.

To a smooth parametrized curve $m : [0, 1] \to \mathbb{R}^2$, we associate the energy

$$E(m) = \alpha \int_0^1 |\ddot{m}|^2 du + \beta \int_0^1 |\dot{m}|^2 du + \gamma \int_0^1 V(m(u)) du \,. \tag{5.38}$$

Minimizing this energy results in a compromise between smoothness constraints (provided by the first two integrals) and the fact that m aligns with image contours, which comes from the last integral. The minimization is made subject to constraints at the extremities. Typical constraints are:

(i) $m(0)$ and $m(1)$ are fixed, together with $\dot{m}(0)$ and $\dot{m}(1)$; or
(ii) $m(0) = m(1)$, $\dot{m}(0) = \dot{m}(1)$ (closed curves).

We will assume that one of these two conditions is enforced in the following computation.

5.4.2 First Variation and Gradient Descent

Given a curve $m_0(u)$, $u \in [0, 1]$, we evaluate the impact on the cost function E of a small variation $\varepsilon \mapsto m(\varepsilon, \cdot)$. If the extremities ($m(0)$ and $m(1)$) are fixed, we will

have $\partial_\varepsilon m(\varepsilon, 0) = \partial_\varepsilon m(\varepsilon, 1) = 0$. The closedness condition requires $\partial_\varepsilon m(\varepsilon, 0) = \partial_\varepsilon m(1, 0)$. Letting $h = \partial_\varepsilon m$, we have

$$\partial_\varepsilon E = 2\alpha \int_0^1 \ddot{m}^T \ddot{h}\, du + 2\beta \int_0^1 \dot{m}^T \dot{h}\, du + \gamma \int_0^1 \nabla V(m(u))^T h(u)\, du.$$

Assume that m is C^4 and perform two integrations by parts of the first integral and one of the second to obtain:

$$\begin{aligned}\partial_\varepsilon E = {} & -2\alpha[(m^{(3)})^T h]_0^1 + 2\alpha[\ddot{m}^T \dot{h}]_0^1 + 2\beta[\dot{m}^T h]_0^1 \\ & + 2\alpha \int_0^1 (m^{(4)})^T h\, du - 2\beta \int_0^1 \ddot{m}^T h\, du \\ & + \gamma \int_0^1 \nabla E(m)^T h\, du. \end{aligned} \tag{5.39}$$

The boundary terms (first line) disappear for both types of boundary conditions. We can therefore write

$$\partial_\varepsilon E = \varepsilon \int_0^1 \nabla E(m)^T h\, du \tag{5.40}$$

with

$$\nabla E(m) = 2\alpha m^{(4)} - 2\beta \ddot{m} + \gamma \nabla V(m)\,.$$

Note that, in this formula, $\nabla E(m)$ is the variational gradient, therefore a function $u \mapsto \nabla E(m)(u)$, whereas $\nabla V(m)$ is the ordinary gradient (a vector). Using the L^2 metric on $[0, 1]$, we get the following gradient descent evolution:

$$\partial_t m(t, u) = -2\alpha m^{(4)}(t, u) + 2\beta \ddot{m}(t, u) - \gamma \nabla V(m(t, u)). \tag{5.41}$$

5.4.3 Numerical Implementation

The discretization of (5.41) is relatively straightforward. Assume that m is discretized into a sequence of points $(x_1, y_1), \ldots (x_n, y_n)$, stacked into a matrix M with n rows and two columns.

The finite-difference derivatives with respect to the parameter u are linear operations on M. For example, $M' = D_1 M$ with (in the case of closed curves)

$$D_1 = \frac{1}{2\delta u} \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ & & & \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 0 \end{pmatrix},$$

where δu is the discretization step. The second derivative is obtained with a tridiagonal matrix, with $1, -2, 1$ on the diagonal (and periodic assignment of values at the extremities), divided by $(\delta u)^2$. The fourth derivative is pentadiagonal, with values $1, -4, 6, -4, 1$ divided by $(\delta u)^4$. Therefore, the expression

$$-2\alpha m^{(4)} + 2\beta \ddot{m}$$

is discretized in the form $A \cdot M_t$, where A is a pentadiagonal matrix which depends only on α and β.

The function V is rarely known analytically, and most often discretized on a grid ($V = V_{ij}, i, j = 1, \ldots, N$). To compute $\nabla V(m)$, the partial derivatives of V must be estimated at points (x, y) which are not necessarily on the grid. The bilinear interpolation of V at a point $(x, y) \in \mathbb{R}^2$ is

$$\hat{V}(x, y) = (1 - \alpha_x)(1 - \alpha_y)V_{ij} + (1-\alpha_x)\alpha_y V_{ij+1} + \alpha_x(1 - \alpha_y)V_{i+1j} + \alpha_x\alpha_y V_{i+1j+1},$$

where i and j are such that $i\delta x \leq x < (i + 1)\delta x$ and $j\delta x \leq y < (j + 1)\delta x$, and $\alpha_x = (x - i\delta x)/\delta x, \alpha_y = (y - j\delta x)/\delta x$ (with a spatial discretization step δx). This implies that

$$\partial_x \hat{V} = (1 - \alpha_y)(V_{i+1j} - V_{ij}) + \alpha_y(V_{i+1j} - V_{ij+1}),$$

except along lines $x = i\delta x$ at which $\hat{V}$ is not differentiable. One can use central finite difference at such points, i.e.,

$$\partial_1^0 V_{ij} = (V_{i+1j} - V_{i-1j})/(2\delta x).$$

Similar formulas can be used for derivatives with respect to the second variable. This yields a discrete version of the gradient of V, that we will denote by $\nabla V(M)$, which is nonlinear in M.

The direct discretization of (5.41), using a time step δt, yields

$$M_{t+1} = (I + \delta t\, A)M_t + \delta t\, \nabla V(M_t).$$

This algorithm will typically converge to a local minimum of E when δt is small enough.

Its numerical behavior can, however, be improved by a semi-implicit scheme (the previous scheme being called explicit), given by

$$(I - \delta t\, A) M_{t+1} = M_t + \delta t\, \nabla V(M_t). \tag{5.42}$$

This requires the computation of the inverse of $(I - \delta t\, A)$, which can be costly, but this operation has to be done only once, since A does not depend on M.

The discretization needs to be adapted when the distance between two consecutive points in M becomes too large. In this case, the estimation of derivatives becomes unreliable. When this occurs, one should rediscretize the evolving curve. This creates a new matrix M (and also requires one to recompute $(I - \delta t\, A)^{-1}$ if the number of points has changed).

5.4.4 Initialization and Other Driving Techniques

The global minimum of E is not interesting. For closed contours, it is a degenerate curve reduced to a point, as will be shown below. So, the method must start with a reasonable initial curve and let it evolve to the closest local minimum. This is why active contours are often run in an interactive way, the user initializing the process with a coarse curve around the region of interest (using a graphic interface), and the evolution providing the final contour, hopefully a fine outline of the shape. Since, as we will see, the active contour energy generally shrinks curves, the initial contour has to be drawn outside the targeted shape.

The "balloon" technique [70] allows the user to only specify an initial region *within* the shape, which is often easier and can sometimes be done automatically (because the interior of a shape is often more specific than its exterior). The idea is to complete the evolution (5.41) with an outward normal force, yielding the equation

$$\dot{m}_t = -2\alpha m^{(4)} + 2\beta \ddot{m} - \gamma \nabla_m V - pN, \tag{5.43}$$

p being a small positive constant and N the inward normal to the contour. Once the curve stabilizes (generally slightly outside the shape because of the effect of the normal force), the algorithm must be continued with $p = 0$ to provide a correct alignment with the shape boundary. Note that this normal force can also be obtained using a variational approach, adding to the energy an area term

$$-p \int_{\Omega_m} dx,$$

where Ω_m is the interior of the curve m.

Because the "gradient force" $\nabla V(m)$ only carries information near points of high gradient, some improvement can be obtained by first extending it to a vector field v defined on the whole image, and using v instead of $\nabla V(m)$ in (5.43). This results

in the gradient vector flow (GVF) method developed in [304], in which v is built by solving an auxiliary variational problem: minimize

$$\int_\Omega (\text{trace}(dv^T dv) + |\nabla V|\,|\nabla V - v|^2) dm.$$

Thus, $v \simeq \nabla V$ when $|\nabla V|$ is large, and is a smooth interpolation otherwise. Extrapolating the driving force to the whole image in this way allows the algorithm to be more robust to initial data. It also provides improved convergence results, allowing in particular the contour to enter into concave regions in the shape, which is much harder with the original formulation.

5.4.5 Evolution of the Parametrization

The active contour energy is not parametrization-independent (such methods will be discussed in the next section). Its evolution reparametrizes the curve in some optimal way and it is interesting to study how this is associated with the minimization of the active contour energy. In numerical implementations, reparametrization is related to how the density of points changes along the discretized curve.

Let σ_m be the (Euclidean) arc length of a curve m, so that $\dot\sigma_m = |\dot m_u|$, and $L = L_m$ be its length. We let $\psi : [0, L] \to [0, 1]$ be the inverse change of parameter $\psi(\sigma_m(u)) = u$. We also let $\bar m(s) = m(\psi(s))$ be the arc-length parametrization of m. We have $\dot m \circ \psi = \dot{\bar m}/\dot\psi$, which implies in turn

$$\dot\psi\, \ddot m \circ \psi = \frac{\ddot{\bar m}}{\dot\psi} - \frac{\ddot\psi}{\dot\psi^2}\dot{\bar m}\,.$$

Since $\bar m$ is parametrized by arc length, we have

$$|\dot m \circ \psi|^2 = \frac{1}{\dot\psi^2}\,, \quad |\ddot m \circ \psi|^2 = \frac{\kappa^2}{\dot\psi^4} + \frac{\ddot\psi^2}{\dot\psi^6},$$

where κ is the curvature.

We can make the change of variable $u \to \psi(u)$ in $E(m)$, which yields

$$E(m) = \int_0^L \left(\alpha \frac{\ddot\psi^2}{\dot\psi^5} + \alpha \frac{\kappa^2}{\dot\psi^3} + \frac{\beta}{\dot\psi} + \gamma V(\bar m)\dot\psi \right) ds. \tag{5.44}$$

To separate the length-shrinking effect from the rest, we renormalize the geometric and parametric quantities. We let $\tilde\kappa(s) = L\kappa(Ls)$ (which is the curvature of m rescaled by a factor $1/L$) and $\chi(s) = \psi(Ls)$. To lighten the expression, we also let $\tilde V(s) = V(\bar m(Ls))$. These functions are all defined on $[0, 1]$, and a linear change of

variables in the previous energy yields (with $\dot{\chi}(s) = L\dot{\psi}(Ls)$ and $\ddot{\chi}(s) = L^2\ddot{\psi}(Ls)$),

$$E(m) = L^2 \int_0^1 \left(\alpha \frac{\ddot{\chi}^2}{\dot{\chi}^5} + \alpha \frac{\tilde{\kappa}^2}{\dot{\chi}^3} + \frac{\beta}{\dot{\chi}} \right) ds + \gamma \int_0^1 \tilde{V} \dot{\chi}\, ds.$$

The first integral is scale-independent (it only depends on the rescaled curve) and is multiplied by L^2. This therefore shows the length reduction effect of the smoothing part of the snake energy. As L tends to 0, $\tilde{V}$ becomes equal to $V(m(0))$ and the limiting integral has the same value since $\int_0^1 \dot{\chi}_s = 1$. It is therefore minimal when $m(0)$ is a point of lowest value for V, and we see that the global minimum of the energy is a dot at the minimum of V (not a very interesting solution).

Beside the shrinking effect, it is interesting to analyze how the parametrization is optimized for a fixed geometry (i.e., fixed L, $\tilde{m}$ and $\tilde{\kappa}$, which also implies that $\tilde{V}$ is fixed). Recall that $1/\dot{\chi}$ is proportional to ds/du, the "speed" of the parametrized curve. The term $\tilde{\kappa}^2/\dot{\chi}^3$ shows that this speed is penalized when the curvature is high. For a discrete curve, this implies that points have a tendency to accumulate at corners. On the other hand, the term $\dot{\chi}\tilde{V}$ creates more sparsity in regions where V is large.

We now specialize to the case $\alpha = 0$, for which the computation can be pushed further. In this case, we have

$$E(m) = \int_0^L \left(\frac{\beta}{\dot{\psi}} + \gamma V(\bar{m})\dot{\psi} \right) ds\,.$$

We fix the geometry and optimize the parametrization, i.e., we minimize E with respect to ψ subject to the constraints $\int_0^L \dot{\psi} = 1$ and $\dot{\psi} > 0$.

First we can see that E is a convex function of $\dot{\psi}$, minimized over a convex set. This implies, in particular, uniqueness of the minimum if it exists. We disregard the positivity constraint, which, as we will see, will be automatically satisfied by the solution. Using Lagrange multipliers, we obtain the Lagrangian

$$\int_0^L \left(\frac{\beta}{\dot{\psi}} + \gamma V(\bar{m})\dot{\psi} \right) ds - \lambda \int_0^L \dot{\psi} ds\,.$$

A variation of this with respect to $\dot{\psi}$ yields the equation

$$\frac{\beta}{\dot{\psi}^2} = \gamma V(\bar{m}) - \lambda\,.$$

The solution must therefore take the form $\dot{\psi}(s) = \dfrac{\sqrt{\beta}}{\sqrt{\gamma V(\bar{m}(s)) - \lambda}}$, for a suitable value of λ, which must be smaller than $\lambda_* = \gamma \min_s V(\bar{m}(s))$ and such that $\int_0^L \dot{\psi}(s) ds = 1$.

Let us prove that such a λ exists, in the case when m is closed and V and m are smooth. Consider the function

$$f : \lambda \mapsto \int_0^L \frac{\sqrt{\beta}}{\sqrt{\gamma V(\bar{m}(s)) - \lambda}} d\lambda$$

defined on $(-\infty, \lambda_*)$. It is continuous, and tends to 0 as λ tends to $-\infty$. Now, let $\lambda \to \lambda_*$ and take s_0 such that $\lambda_* = \gamma V(\bar{m}(s_0))$. Since λ_* is a minimum and we assume that V and $\bar{m}$ are smooth, the difference $\gamma V(\bar{m}(s)) - \gamma V(\bar{m}(s_0))$ must be an $O((s - s_0)^2)$ or smaller when s is close to s_0. But this implies that the integral diverges to $+\infty$ as $\lambda \to \lambda_*$. Therefore, a value of λ exists with $f(\lambda) = 1$. (If m is an open curve, the above argument is valid with the additional assumption that the minimum of V is not attained at one of the extremities of m.)

From the computation above, we see that the optimal parametrization must satisfy

$$\dot{\psi}^{-2} - (\gamma/\beta)V(\bar{m}) = \text{constant}.$$

Consider now the following class of problems: minimize E under the constraint $\dot{\psi}^{-2} - (\gamma/\beta)V(\bar{m}) = \mu$ for some constant μ. The previous discussion shows that the solution of the original problem is also a solution of this constrained problem for some value of μ (one says that the two classes of problems are equivalent). However, considering the latter problem, we see that it boils down to the purely geometric problem: minimize

$$\int_0^L W(\bar{m}(s))ds$$

with $W(p) = \sqrt{\beta}(\beta\mu + 2\gamma V(p))/\sqrt{\beta\mu + \gamma V(p)}$. This new variational problem fits into the category of geodesic active contours, which are addressed in the next section.

5.4.6 *Parametrization-Invariant Methods*

To obtain a parametrization-invariant formulation of (5.38), it suffices to consider it over arc-length parametrized curves, leading to

$$E(m) = \int_m (\alpha\kappa^2 + \beta + \gamma V(m))d\sigma_m = \int_0^L (\alpha\kappa^2(s) + \beta + \gamma V(m(s)))ds\,. \quad (5.45)$$

The first term is the elastica energy that we defined in (5.34). Combining (5.35) and Theorem 5.2, we see that the gradient flow associated to E is

$$\partial_t m = \big(-\alpha(\ddot{\kappa} + \kappa^3/2) + (\beta + \gamma V)\kappa - \nabla V^T N\big)N. \quad (5.46)$$

Geodesic Active Contours

The case $\alpha = 0$ corresponds to what has been called *geodesic active contours* [56, 250], which correspond to minimizing

$$E(m) = \int_m W(m) d\sigma_m \tag{5.47}$$

(letting $W = \beta + \gamma V$) with the associated evolution

$$\partial_t m = \big(W(m)\kappa - \nabla W(m^T N)\big)N\,. \tag{5.48}$$

This equation can be conveniently implemented using level sets. So, consider a function $(t,\, p) \mapsto F_t(p)$ such that its zero level set at fixed time t is a curve, denoted $u \mapsto m_t(u)$. The equation $F_t(m_t(u)) = 0$ is by definition valid for all t and all u. Computing its derivative with respect to t yields

$$\partial_t F_t(m_t(u)) + \nabla F_t(m_t(u))^T \partial_t m_t(u) = 0\,.$$

Assume, as is the case here, that the evolution of m is prescribed, in the form

$$\partial_t m_t(u) = \beta_t(m_t(u)) N_t(u).$$

This yields, using $N_t = -\nabla F_t / |\nabla F_t|$,

$$\partial_t F_t(m_t(u)) = \beta_t(m_t(u))\, |\nabla F_t(m_t(u))|\,.$$

Assume that β_t, which describes the speed of the normal evolution of γ, can be extended to the whole domain Ω (it is a priori only defined on the evolving curve). Then, we can introduce the global evolution

$$\partial_t F_t(p) = \beta_t(p)\, |\nabla F_t(p)|\,.$$

This equation being valid for all p, it is a fortiori valid on the zero level set of f. Therefore, if this level set did coincide with the initial curve m_0, it would contain at all times the curve m_t that satisfies the evolution equation $\partial_t m_t = \beta N_t$ (implicitly assuming that the evolution is well-defined over the considered time interval).

Returning to geodesic active contours, there is a natural extension for the function β, namely

$$\beta_t(p) = W(p)\, \mathrm{div}\left(\frac{\nabla F_t(p)}{|\nabla F_t(p)|}\right) + \nabla W(p)^T \left(\frac{\nabla F_t(p)}{|\nabla F_t(p)|}\right).$$

This choice yields the partial differential equation:

$$\partial_t F = W\,|\nabla F_t|\,\mathrm{div}\left(\frac{\nabla F_t}{|\nabla F_t|}\right) + \nabla W^T \nabla F_t.$$

The implementation details are similar to those provided in Sect. 5.1.4.

5.4.7 Controlled Curve Evolution

For completeness, we mention another curve evolution method for contour estimation [82], which consists in using a smoothing evolution equation such as the one introduced in Sect. 5.1.3, with an additional factor that controls the speed of the evolution, and essentially stops it in the presence of contours.

The curve evolution equation, in this case, takes the form

$$\partial_t m = V(m)\kappa N, \tag{5.49}$$

where V is now the stopping function. This no longer derives from a variational approach, but we can see that when V is constant, the evolution is similar to (5.10), but that points stop moving when V is close to 0. So V retains the same interpretation as before, as a function which vanishes near regions of high gradient in the image.

The level-set formulation of this equation is

$$\partial_t F = V(t,\cdot)|\nabla F_t|\mathrm{div}\frac{\nabla F_t}{|\nabla F_t|}\,.$$

5.4.8 Geometric Active Surfaces

The transcription of (5.45) to surfaces is

$$E(S) = \int_S (\alpha H^2(p) + \beta) d\sigma_S(p) + \gamma \int_S V(p) d\sigma_S(p), \tag{5.50}$$

with associated gradient flow (using (5.37) and Theorem 5.4)

$$\partial_t p = \big(\alpha(\Delta_S H + 2H(H^2 - K)) + 2(\beta + \gamma V)H - \nabla V^T N\big)N$$

if the surface has no boundary.

The case $\alpha = 0$ simplifies to (letting $W = \beta + \gamma V$)

$$\partial_t p = \big(2WH - \nabla W^T N\big)N$$

for a surface without a boundary and to

$$\partial_t p = \big(2WH - \nabla W^T N\big)N + \delta_{\partial S}(p) W n_S$$

if the surface has a boundary.

For closed surfaces, a level-set evolution can be derived similarly to curves, yielding [250]

$$\partial_t F_t = W|\nabla F_t|\mathrm{div}\left(\frac{\nabla F_t}{|\nabla F_t|}\right) + \nabla W^T \nabla F_t.$$

5.4.9 Designing the V Function

For most applications, boundaries of shapes in images correspond to rapid variation in image intensity, which results in large values of the image gradient. It is therefore natural to use functions V related to the gradient, such as

$$V = -|\nabla I|$$

(recall that active contours align with small V). Designing bounded V's is, however, numerically preferable, a possible choice being given by

$$V = \frac{1}{1 + |\nabla I|/C}$$

for some constant C.

Another option is to rely on an edge detector [52, 186], and start from a binary image indicating edge points. Letting E denote the set of detected edges, one can define V to be the distance map to E

$$V(m) = \inf\{d(m, e), e \in E\}$$

(see Appendix F for quick implementations). Another commonly used approach is to convolve the binary edge image with a Gaussian kernel and let

$$V(m) = \sum_{e \in E}(1 - e^{-|g-m|^2/2\sigma^2}).$$

5.4.10 Inside/Outside Optimization

When information about the image values inside and outside a closed shape is available, it is possible to add area or volume integrals (depending on the dimension) to the geometric formulations. This results in objective functions that, unlike their active-contours counterparts, should be minimized globally to obtain the segmentation.

Letting Ω denote the interior of the shape, the additional term typically takes the form

$$\int_{\Omega} \tilde{V}_{in}(x)dx + \int_{D\setminus\Omega^c} \tilde{V}_{out}(x)dx = \int_{\Omega} (\tilde{V}_{in}(x) - \tilde{V}_{out})dx + \int_{D} \tilde{V}_{out}(x)dx\,. \quad (5.51)$$

Here, $\tilde{V}_{in}$ and $\tilde{V}_{out}$ are defined on some fixed bounded set $D \subset \mathbb{R}^d$, the image domain, and they take large values at points that are unlikely to belong to the interior or the exterior of the shape, respectively.

Letting $\tilde{V} = \tilde{V}_{in} - \tilde{V}_{out}$ and disregarding the last integral, which does not depend on Ω, we end up with an expression taking the form

$$\int_{\Omega} \tilde{V}(x)dx\,.$$

Using Theorems 5.2 and 5.4, adding this term to a segmentation energy simply adds the term $\tilde{V}N$ to the gradient descent algorithms.

A simple example is when the image values are expected to average around a constant, say c_{in} over the interior, and around c_{out} over the exterior of the shape, leading to choosing

$$\tilde{V}_{in}(x) = \lambda_{in}(I(x) - c_{in})^2 \text{ and } \tilde{V}_{out}(x) = \lambda_{out}(I(x) - c_{out})^2.$$

This assumption is made in the Chan–Vese segmentation model [58], which defines (in two dimensions)

$$E(m) = \mu\,\text{length}(m) + \nu\,\text{area}(\Omega_m) + \int_{\Omega_m} \Big(\lambda_{in}(I(x) - c_{in})^2 - \lambda_{out}(I(x) - c_{out})^2\Big)dx, \quad (5.52)$$

where Ω_m is the interior of the shape (with an obvious generalization to surfaces in three dimensions). This energy is a simplified version of the Mumford–Shah functional [213], which is designed for the approximation of an observed image by a piecewise smooth function. (Here the approximation is by a piecewise constant function, and the contour separates the image into exactly two regions.) The associated shape evolution is

$$\partial_t m = (\mu\kappa + \nu + \lambda_{in}(I(x) - c_{in})^2 - \lambda_{out}(I(x) - c_{out})^2)N \quad (5.53)$$

with a level-set formulation

$$\partial_t F_t = |\nabla F_t| \left(\mu\,\text{div}\left(\frac{\nabla F_t}{|\nabla F_t|}\right) + \nu + \lambda_{in}(I(x) - c_{in})^2 - \lambda_{out}(I(x) - c_{out})^2\right). \quad (5.54)$$

Instead of being seen as a function of a curve, energies such as (5.52), which can be put in the form

$$E(m) = \int_m V(p) d\sigma_m + \int_{\Omega_m} \tilde{V}(x) dx, \tag{5.55}$$

can also be considered as functions of the domain Ω, simply setting

$$E(\Omega) = \int_{\partial\Omega} V(p) d\sigma_{\partial\Omega} + \int_{\Omega} \tilde{V}(x) dx. \tag{5.56}$$

The interesting feature of this point of view is that this energy applies to domains that are more general than the interior of a Jordan curve, allowing for multiple connected components, for example, which can be useful when the topology of the targeted set is not known in advance. The associated level-set evolution

$$\partial_t F = |\nabla F| \left(V \operatorname{div} \left(\frac{\nabla F}{|\nabla F|} \right) + \nabla V^T \nabla F + \tilde{V} \right) \tag{5.57}$$

allows for such changes of topology, but it is also interesting to express (5.56) directly in terms of level sets of a function.

To motivate the reformulation of the method, consider a δ-function approximation of (5.56), as described Sect. 3.10. Using the notation of that section,

$$E(\Omega) \simeq \int_{\mathbb{R}^d} V(x) |\nabla (H_\varepsilon \circ F)(x)| dl + \int_{\mathbb{R}^d} (1 - H_\varepsilon \circ F(x)) \tilde{V}(x) dx \tag{5.58}$$

when $\Omega = \{F \leq 0\}$ (V must be continuous for this approximation to be valid). Let $E_\varepsilon(F)$ denote the right-hand side in (5.58). Introduce the function $u = 1 - H_\varepsilon \circ F$ to reparametrize the problem so that

$$E_\varepsilon(F) = \tilde{E}(u) = \int_{\mathbb{R}^d} V(x) |\nabla u(x)| dx + \int_{\mathbb{R}^d} u(x) \tilde{V}(x) dx. \tag{5.59}$$

We now focus on the minimization of $E(u)$ over all functions u taking values in $[0, 1]$. Let $\Omega(t)$ denote the level set $u \geq t$, for $t \in [0, 1]$. We can write

$$\begin{aligned} \int_{\mathbb{R}^d} u(x) \tilde{V}(x) dx &= \int_{\mathbb{R}^d} \int_0^1 \mathbf{1}_{t \leq u(x)} V(x) \, dt \, dx \\ &= \int_0^1 \int_{\Omega_t} V(x) \, dx \, dt. \end{aligned}$$

We also have

$$\int_{\mathbb{R}^d} V(x) |\nabla u(x)| dx \text{“=”} \int_0^1 \int_{\partial\Omega_t} V(p) \, d\sigma_{\partial\Omega_t} \, dt.$$

The latter expression is an application of the co-area formula. We have put the equal sign between quotes because a rigorous application of this formula requires more advanced mathematical concepts than those that are assumed in this discussion. For example, the boundary of Ω_t is not necessarily a smooth curve, but one can still define a generalization of the inner integral on the left-hand side for which the identity is valid. Disregarding these issues, we therefore have

$$\begin{aligned}\tilde{E}(u) \text{“=”} & \int_0^1 \left(\int_{\partial\Omega_t} V(p)\, d\sigma_{\partial\Omega_t} + \int_{\Omega_t} V(x)\, dx \right) dt \\ &= \int_0^1 E(\Omega_t)\, dt.\end{aligned}$$

Now, if Ω^* is a minimizer of E, then $E(\Omega^*) \leq \tilde{E}(u)$ for all u, and, $\tilde{E}(u^*) = E(\Omega^*)$ for $u^* = 1 - \mathbf{1}_{\Omega^*}$. Conversely, if $0 \leq u \leq 1$ minimizes $\tilde{E}$, then Ω_t is a minimizer of E for almost all $t \in (0, 1)$ [286]. Passing from E to $\tilde{E}$ is actually a significant simplification of the original problem, because $\tilde{E}$ is a convex function of u (minimized subject to convex constraints), allowing one to leverage the large collection of optimization methods available for convex functions.

Returning to the Chan–Vese energy (5.52), (5.59) becomes

$$\begin{aligned}\tilde{E}(u) = & \int_{\mathbb{R}^d} |\nabla u(x)|\, dx + \\ & \int_{\mathbb{R}^d} u(x) \left(\nu + \lambda_{in}(I(x) - c_{in})^2 - \lambda_{out}(I(x) - c_{out})^2 \right) dx.\end{aligned} \tag{5.60}$$

The numbers c_{in} and c_{out} that characterize the image values inside and outside the shape are not always known in advance. In such a situation, it is natural to also minimize (5.52) with respect to c_{in} and c_{out}. For this to be possible, one needs to limit the size of the integration domain for the outside integral to a fixed region, D. When u is fixed, their optimal values are easily computed and are given by

$$c_{in} = \frac{\int_D u(x) I(x) dx}{\int_D u(x) dx} \text{ and } c_{out} = \frac{\int_D (1 - u(x)) I(x) dx}{\int_D (1 - u(x)) dx}.$$

Minimization in F, c_{in} and c_{out} can be implemented by alternating the computation of a minimizer of (5.59) (or a few steps of the associated minimization algorithm) with a periodic updating of c_{in} and c_{out} according to these expressions. One can even include the optimization of these constants within a complete convex optimization problem, as detailed in [46].

5.4.11 Sobolev Active Contours

We now follow-up on our discussion at the end of Sect. 5.3, and describe the interesting variants of the active contour algorithms introduced in [272]. Using the notation

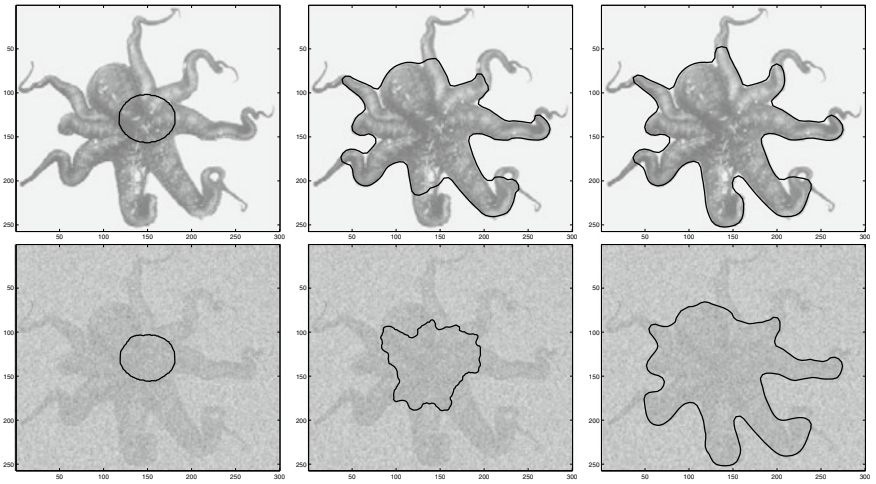

Fig. 5.4 Comparison between geometric and Sobolev active contours. In each row, the left image is the initial contour for both algorithms, the center one is the result of geometric active contours once the evolution stabilizes, and the right one is for Sobolev active contours. The first row presents a clean image and the second a noisy one. In both cases, geometric active contours get stuck at an unsatisfactory local minimum (especially in the noisy case), while Sobolev active contours reach a much more accurate solution

of Sect. 5.3, Sobolev active contours modify an evolution such as

$$\partial_t m = (\kappa V - V^T N + \tilde{V})N$$

(which corresponds to the minimization of (5.55)) and transform it into

$$\partial_t m = A^{-1}\big((\kappa V - V^T N + \tilde{V})N\big), \tag{5.61}$$

where A is a differential operator, or more generally an operator $f \mapsto Af$ such that the kth Fourier coefficient of Af is equal to $a_{|k|}c_k(f)$, where a_k is a positive sequence of numbers that tend to infinity (e.g., $a_k = 4\pi^2k^2 + \lambda$) and $c_k(f)$ is the kth Fourier coefficient of f. When implementing (5.61), high-order Fourier coefficients of $(\kappa V - V^T N + \tilde{V})N$ are divided by large numbers, which results in attenuating high-frequency motion and focusing on low frequency variations, yielding global changes in the shape. The resulting algorithm makes the evolution less sensitive to noise (which typically induces high-frequency motion), and also accelerates convergence because large moves are easier to make. The counterpart is that small details in the shape boundary are harder to acquire, and may require running the algorithm over longer periods. Examples illustrating the robustness of Sobolev active contours are provided in Fig. 5.4.

Chapter 6
Deformable Templates

Deformable templates represent shapes as transformations of a given prototype, or template. One of the advantages of this approach is that the template needs to be specified only once, for a whole family of shapes. If the template is well chosen, describing the transformation leading to a shape results in a simpler representation, typically involving a small number of parameters. The conciseness of the description is important for detection or tracking algorithms in which the shape is a variable, since it reduces the number of degrees of freedom. Small-dimensional representations are also more easily amenable to probabilistic modeling, leading, as we will see, to interesting statistical shape models.

The methods that we describe provide a parametrized family of shapes, $(m(\theta), \theta \in \Theta)$, where Θ is a parameter set. Most of the time, Θ will be some subset of $\mathbb{R}^d$ but it can also be infinite-dimensional. We will always assume, as a convention, that $0 \in \Theta$ and that $m(0)$ represents the template.

To simplify the presentation, we will restrict to curves, therefore assuming that $m(\theta)$ is a parametrized curve $u \mapsto m(u, \theta)$ defined over a fixed interval $[a, b]$. Other situations can easily be transposed from this one. For example, one commonly uses configurations of *labeled* points, or landmarks, with $m(\theta) = (m_1(\theta), \ldots, m_N(\theta))$ as a finite-dimensional descriptor of a shape. Transposition from curves to surfaces is also easy.

6.1 Linear Representations

We start with a description of linear methods, in which

$$m(\theta) = m(0) + \sum_{k=1}^{n} \theta_k u_k,$$

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_6

where u_k is a displacement applied to $m(0)$: for example, if $m(0)$ is a closed curve, u_k is defined on $[a, b]$, taking values in $\mathbb{R}^d$ with $u_k(a) = u_k(b)$. If $m(0)$ is a configuration of points, u_k is a list of two-dimensional vectors.

The issue in this context is obviously how to choose the u_k's. We will provide two examples, the first one based on a deterministic approach, and the second relying on statistical learning.

6.1.1 Energetic Representation

The framework developed in this section characterizes an object using a "small-deformation" model partially inspired by elasticity or mechanics. The object is described, not by its aspect, but by how it deforms. Our presentation is inspired by that developed in [230] for face recognition. It includes the principal warps described in [41] as a particular case, and provides an interesting way of decomposing shape variations in a basis that depends on the geometry of the considered shape.

For such a shape m, we will consider small variations, represented by transformations $h \mapsto F(m, h)$. For example, one can take $F(m, h) = m + h$ when this makes sense. We assume that the small variations, h, belong to a Hilbert space H (see Appendix A), with dot product $\langle \cdot , \cdot \rangle_m$, possibly depending on m.

Associate to h some deformation energy, denoted $E(h)$. Attribute to a time-dependent variation, $t \mapsto h(t)$, the total energy:

$$J(h) = \frac{1}{2} \int \|\partial_t h(t)\|_m^2 dt + \int E(h(t)) dt.$$

Inspired by the Hamilton principle, we consider shape trajectories that are extremals of the Lagrangian $\|\partial_t h\|_m^2/2 - E(h)$, therefore characterized by

$$\partial_t^2 h + \nabla E(h(t)) = 0,$$

where ∇E is the Hilbert gradient, defined by

$$\partial_\varepsilon E(h + \varepsilon w)_{|_{\varepsilon=0}} = \left\langle \nabla E(h) \, , \, w \right\rangle_m.$$

We make the assumption that this gradient exists. In fact, because we only analyze small variations, we will assume that a second derivative exists at $h = 0$, i.e., we assume that, for some symmetric operator Σ_m,

$$\nabla E(h) = \Sigma_m h + o(\|h\|_m).$$

Typically, we will have $E \geq 0$ with $E(0) = 0$, which ensures that Σ_m is a non-negative operator. The linearized equation for h now becomes

$$\partial_t^2 h + \Sigma_m h = 0. \tag{6.1}$$

This equation has a simple solution when Σ_m is diagonalizable. Making this assumption (which is always true in finite dimensions), letting $(f_1, f_2, \ldots)$ be the eigenvectors and $(\lambda_1, \lambda_2, \ldots)$ the corresponding eigenvalues (in decreasing order), solutions of (6.1) take the form

$$h(t) = \sum_{k \geq 1} \alpha^{(k)}(t) f_k$$

with $\partial_t^2 \alpha^{(k)} + \lambda_k \alpha^{(k)} = 0$, so that $\alpha^{(k)}$ oscillates with frequency $\omega_k = 1/\sqrt{\lambda_k}$.

The $\omega_k' s$ form what was called a *modal representation* in [253]. These vibration modes can be used to describe and compare shapes (so that similar shapes should have similar vibration modes). It is also possible to use this model for a template-based representation: let m be a template, with a modal decomposition as before, and represent small variations as

$$(\alpha_1, \ldots, \alpha_N) \to \tilde{m} = F\left(m, \sum_{k=1}^{N} \alpha_k f_k\right),$$

which has a linearized deformation energy given by $\sum_k \lambda_k \alpha_k^2$.

Consider a first example of such a construction using plane curves. Let $m(\cdot) = m(0, \cdot)$ be the prototype and Ω_m its interior. Assume that m is parametrized by arc length. A deformation of m can be represented as a vector field $s \mapsto h(s)N(s)$, where h is a scalar function and N the unit normal to m. The deformed template is $s \mapsto m(s) + h(s)N(s)$. A simple choice for E is

$$E(h) = \frac{1}{2} \int_0^L \partial_s h^2 ds,$$

for which $\Sigma_m h = -\partial_s^2 h$ and Eq. (6.1) is

$$\partial_t^2 h = \partial_s^2 h,$$

which is the classical wave equation in one dimension. Since this equation does not depend on the prototype, m, it is not really interesting for our purposes, and we need to consider energies that depend on geometric properties of m. The next simplest choice is probably

$$E(h) = \frac{1}{2} \int_0^L \rho_m(s) \partial_s h^2 ds,$$

where ρ_m is some function defined along m, for example $\rho_m = 1 + \kappa_m^2$ (where κ_m is the curvature along m). In this case, we get $\Sigma_m h = -2\partial_s(\rho_m \partial_s h)$. The vibration modes are the eigenvectors of this inhomogeneous diffusion operator along the curve.

One can obviously consider many variations of this framework. Consider, for example, discrete shapes, represented by a finite collection of landmarks, so that a shape is now a finite collection $m = (x_1, \ldots, x_N)$ with each $x_i \in \mathbb{R}^2$. Given displacements $h = (h_1, \ldots, h_N)$, define $h^{(m)}(x)$, for $x \in \mathbb{R}^2$ by

$$h^{(m)}(x) = \sum_{i=1}^{N} g(|x - x_i|^2)\alpha_i$$

with $g(t) = e^{-t/2\sigma^2}$, where $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^2$ are chosen so that $h(x_i) = h_i$, $i = 1, \ldots, N$. Then, we can define

$$\begin{aligned} E_m(h) &= \int_{\mathbb{R}^2} |h^{(m)}(x)|^2 dx \\ &= \sum_{i,j=1}^{N} \alpha_i^T \alpha_j \int_{\mathbb{R}^2} g(|x_i - x|^2) g(|x - x_j|^2) dx \\ &= \sum_{i,j=1}^{N} c_{ij}(m) \alpha_i^T \alpha_j \end{aligned}$$

with

$$c_{ij} = \int_{\mathbb{R}^2} e^{-\frac{|x_i - x|^2}{2\sigma^2} - \frac{|x_j - x|^2}{2\sigma^2}} dx = \pi\sigma^2 e^{-\frac{|x_i - x_j|^2}{4\sigma^2}}.$$

Finally, notice that, from the constraints, $\alpha = S(m)^{-1}h$ with $s_{ij}(m) = g(|x_i - x_j|^2)$, we have

$$E_m(h) = \mathbf{1}_d^T h^T S(m)^{-1} C(m) S(m)^{-1} h \mathbf{1}_d,$$

where, in this expression, h is organized in an N by d matrix and $\mathbf{1}_d$ is the d-dimensional vector with all coordinates equal to 1. The modal decomposition will, in this case, be provided by eigenvalues and eigenvectors of $S(m)^{-1}C(m)S(m)^{-1}$.

The principal warp representation [41] is very similar to this one, and corresponds to

$$E_m(h) = \mathbf{1}_d^T h^T S(m)^{-1} h \mathbf{1}_d. \tag{6.2}$$

It is also associated to some energy computed as a function of $h^{(m)}$, as will be clear to the reader after the description of reproducing kernel Hilbert spaces in Chap. 8.

One can also define

$$E_m(h) = \int_{\mathbb{R}^2} \text{trace}((dh^{(m)})^T dh^{(m)}) dx$$

or some other function of $(dh^{(m)})^T dh^{(m)}$, which corresponds to elastic energies. Closed-form computation can still be done as a function of $h_1, \ldots, h_N$, and provides a representation similar to the one introduced in [230, 253].

6.2 Probabilistic Decompositions

6.2.1 Deformation Axes

One can build another kind of modal decomposition, based on a training set of shapes, using principal component analysis (PCA).

We will work with parametrized curves. The following discussion can be applied, however, with any of the shape representations described in Chap. 1, or, as considered in [73], with finite collections of points (landmarks) placed along (or within) the shape.

Assume that a training set is given containing N shapes that we will consider as versions of the same object or class of objects. We shall denote its elements by $m^{(k)}(\cdot)$, $k = 1, \ldots, N$, and assume they are all defined on the same interval, I. The average is given by

$$\bar{m}(u) = \frac{1}{N} \sum_{k=1}^{N} m^{(k)}(u).$$

A PCA (cf. Appendix E) applied to $m^{(k)}$, $k = 1, \ldots, N$, with the L^2 inner product provides a finite-dimensional approximation called the *active shape representation*

$$m^{(k)}(u) = \bar{m}(u) + \sum_{i=1}^{p} \alpha_{ki} e^{(i)}(u), \tag{6.3}$$

where the principal directions $e^{(1)}, \ldots, e^{(p)}$ provide deformation modes along which the shape has the most variations.

This provides a new, small-dimensional curve representation, in terms of variations of the template $\bar{m}$. One can use it, for example, to detect shapes in an image, which requires the estimation of p parameters, plus three parameters (in two dimensions) describing the shape position in the image (rotation and translation).

One must be aware, when using this method, of the limits of the validity of the PCA approach, which is a linear method. It is not always "meaningful" to compute linear combinations of deformation vectors, even though, once the data is represented by an array of numbers, such a computation is always possible and easy. The important issue, however, is whether one can safely go back, that is, whether one can associate a valid shape (which can be interpreted as an object of the same category as the initial dataset) to any such linear combination. The answer, in general, is yes, provided the coefficients in the decomposition are not too large. Large coefficients, however, lead to large distortions, singularities, and do not model interesting shapes. Because of this, PCA-based decompositions should be considered as first-order linear approximations of more complex, nonlinear, variations. Plane curves, for example, can certainly be considered as elements of some functional space, on which linear combinations are valid, but their result does not always lead to satisfactory shapes. To take an example,

assume that the training set only contains triangles. The PCA decomposition includes no mechanism ensuring that the shapes remain triangular after decomposition on a few principal components. Most often, the representation will be very poor, as far as shape interpretation is concerned.

In fact, shape decomposition must always be, in one way or another, coupled with some feature alignment on the dataset. In [73], this is implicit, since the approach is based on landmarks that have been properly selected by hand. To deal with general curves, it is important to preprocess the parametrizations to ensure that they are consistent, in the sense that points with the same parameter have similar geometric properties. The curves cannot, in particular, all be assumed to be arc-length parametrized. One way to proceed is to assume that the parametrization is arc length for only one curve, say $m^{(0)}$. For the other curves, say $m^{(k)}$, $k = 1, \ldots, N$, we want to make a change of parametrization, $\varphi^{(k)}$, such that $m^{(k)}(\varphi^{(k)}(s)) = m_0(s) + \delta^{(k)}(s)$ with $\delta^{(k)}$ as small as possible. Methods to achieve such simultaneous parametrizations implement curve registration algorithms. They will be presented later in this book.

In addition to aligning the parametrization, it is important to also ensure that the geometries are aligned, with respect to linear transformations (such as rotations, translations, scaling). All these operations have the effect of representing all the shapes in the same "coordinate system", within which linear methods will be more likely to perform well.

Finally, we notice that this framework can be used to generate stochastic models of shapes. We can use the expression

$$m(u) = \bar{m}(u) + \sum_{i=1}^{p} \alpha_i e^{(i)}(u)$$

and generate random curves m by using randomly generated α_i's. Based on the statistical interpretation of PCA, the α_i's are uncorrelated, and their respective variances are the eigenvalues λ_i^2 that correspond to the eigenvector $e^{(i)}$. Simple models generate the α_i's as independent Gaussian variables with variance λ_i^2, or uniformly distributed on $[-\sqrt{3}\lambda_i, \sqrt{3}\lambda_i]$.

6.3 Stochastic Deformation Models

6.3.1 Generalities

The previous approaches analyzed variations directly in the shape representation. We now discuss a point of view which first models deformations as a generic process, before applying them to the template.

We consider here the (numerically important) situation in which the deformed curves are polygons. Restricting ourselves to this finitely generated family will

simplify the mathematical formulation of the theory. The template will therefore be represented as a list of contiguous line segments, and we will model a deformation as a process that can act on each line segment separately. The whole approach is a special case of Grenander's theory of deformable templates, and we refer to [14, 134–136, 138] for more references and information. The general principles of deformable templates assume that an "object" can be built by assembling elementary components (called generators), with specified composition rules. In the case we consider here, generators are line segments and composition rules imply that exactly two segments are joined at their extremities. One then introduces a set of transformations (via a suitable group action) that modify the generators, under the constraints of maintaining the composition rules. In our example, the transformation group will consist of collections of planar similitudes.

6.3.2 *Representation and Deformations of Planar Polygonal Shapes*

The formulas being much simpler when expressed using complex notation, we identify a point $p = (x, y)$ in the plane with the complex number $x + iy$, that we also denote by p. A polygonal line can either be defined by the ordered list of its vertices, say $s_0, \ldots, s_N \in \mathbb{C}$ or, equivalently, by one vertex s_0 and the sequence of vectors $v_k = s_{k+1} - s_k$, $k = 0, \ldots, N-1$. The latter representation has the advantage that the sequence $(v_0, \ldots, v_{N-1})$ is a translation-invariant representation of the polygon. A polygonal line modulo translations will therefore be denoted $\pi = (v_0, \ldots, v_{N-1})$. The polygonal line is a polygon if it is closed, i.e., if and only if $v_0 + \cdots + v_{N-1} = 0$. A polygonal line with origin s_0 will be denoted (s_0, π).

A polygonal line can be deformed by a sequence of rotations and scalings applied separately to each edge v_k. In $\mathbb{C}$, such a transformation is just a complex multiplication. Therefore, a deformation is associated with an N-tuple of non-vanishing complex numbers $z = (z_0, \ldots, z_{N-1})$, the action of z on π being

$$z \cdot \pi = (z_0 v_0, \ldots, z_{N-1} v_{N-1}) \,. \tag{6.4}$$

This defines a group action (cf. Sect. B.5) of $G = (\mathbb{C} \setminus \{0\})^N$ on the set of polygonal lines with N vertices.

In this group, some transformations play a particular role. Introduce the set

$$\Delta = \{z \in G, z = z(1, \ldots, 1), z \in \mathbb{C}\}$$

(the diagonal in G). An element in Δ provides a single similitude applied simultaneously to all edges, i.e., Δ represents the actions of similitudes on polygons. Similarly, the set

$$\Delta_0 = \{z \in G, z = z(1, \ldots, 1), z \in \mathbb{C}, |z| = 1\}$$

represents the action of rotations.

A polygonal line modulo similitudes (resp. rotations) can be represented as an orbit $\Delta \cdot \pi$ (resp. $\Delta_0 \cdot \pi$). We can define the quotient groups G/Δ and G/Δ_0, namely the sets of orbits $\Delta \cdot z$ (resp. $\Delta_0 \cdot z$) for $z \in G$ (they have a group structure because G is commutative). One obtains a well posed action of, say, G/Δ on polygonal lines modulo similitudes, by defining

$$(\Delta \cdot \mathbf{z}) \cdot (\Delta \cdot \pi) = \Delta \cdot (z \cdot \pi).$$

Given a polygon, π, we define $F(\pi)$ as the set of group elements z in G that transform π into another polygon, namely

$$F(\pi) = \{z \in G, z_0 v_0 + \cdots + z_{N-1} v_{N-1} = 0\}. \tag{6.5}$$

Note that $F(\pi)$ is not a subgroup of G.

We can use this representation to provide a stochastic model for polygonal lines. It suffices for this to choose a template π and a random variable ζ on G and to take $\zeta \cdot \pi$ to obtain a random polygonal line. Because we are interested in shapes, however, we will restrict ourselves to closed lines. Therefore, given $\pi = (v_0, \ldots, v_{N-1})$, we will assume that ζ takes values in $F(\pi)$.

We now build simple probability distributions on G and $F(\pi)$ for a fixed π. Consider the function:

$$E(z) = (\alpha/2) \sum_{k=0}^{N-1} |z_k - 1|^2 + (\beta/2) \sum_{k=0}^{N-1} |z_k - z_{k-1}|^2.$$

The first term is large when z is far from the identity, and the second one penalizes strong variations between consecutive z_i's. Here and in the following, we let $z_{-1} = z_{N-1}$.

We want to choose a probability distribution on G which is small when E is large. A natural choice would be to take the measure with density proportional to $\exp(-E(z))/\prod_{k=1}^{N-1} |z_k|$ with respect to the Lebesgue measure on $\mathbb{C}^{N-1}$. This is the "Gibbs measure", with energy E, relative to the Haar measure, $\prod_{k=1}^{N-1} dz_k/|z_k|$ which is the uniform measure on G. Such a choice is of interest in that it gives a very small probability to small values of $|z_k|$, which is consistent with the fact that the z_k's are non-vanishing on G. Unfortunately, this model leads to intractable computations, and we will rely on the simpler, but less accurate, model with density, f, proportional to $\exp(-E(z))$. This choice will greatly simplify the simulation algorithms, and in particular, the handling of the closedness constraint.

With $\pi = (v_0, \ldots, v_{N-1})$, this constraint is expressed by $\sum_k v_k z_k = 0$, and we will use the conditional density for f given this identity. This conditional distribution can be computed by using a discrete Fourier transform. Define

$$u_l = \hat{z}_l = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} z_k e^{-2i\pi \frac{kl}{N}} .$$

One can easily prove that E can be written

$$E(z) = \alpha |u_0 - \sqrt{N}|^2 + \sum_{l=1}^{N-1} \left(\alpha + 2\beta \left(1 - \cos \frac{2\pi l}{N} \right) \right) |u_l|^2 ,$$

and that the constraint becomes

$$\sum_{l=0}^{N-1} \hat{v}_l u_l = 0$$

with $\hat{v}_l = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} v_k e^{-2i\pi \frac{kl}{N}}$. Notice that, because π is closed, we have $\hat{v}_0 = 0$.

Let $w_0 = \sqrt{\alpha}(u_0 - \sqrt{N})$, and, for $l \geq 1$, $w_l = \sqrt{\alpha + 2\beta(1 - \cos \frac{2\pi l}{N})} u_l$, so that

$$E(z) = \sum_{l=0}^{N-1} |w_l|^2.$$

Without the constraint, the previous computation implies that the real and imaginary parts of $w_0, \ldots, w_{N-1}$ are mutually independent standard Gaussian variables: they therefore can be easily simulated, and the value of $z_0, \ldots z_{N-1}$ directly computed after an inverse Fourier transform. Conditioning on closedness only slightly complicates the procedure. Replacing u_l by its expression as a function of $w_1, \ldots, w_{N-1}$, and using $\hat{v}_0 = 0$, the constraint can be written in the form

$$\sum_{l=0}^{N-1} c_l w_l = 0$$

with $c_0 = 0$ ($c_p = \sqrt{\alpha + 2\beta(1 - \cos \frac{2\pi l}{N})}$). The following standard lemma from the theory of Gaussian variables solves our problem.

Lemma 6.1 *Let $\boldsymbol{w}$ be a standard Gaussian vector in $\mathbb{R}^{2N}$, and let V be a vector subspace of $\mathbb{R}^{2N}$. Let Π_V be the orthogonal projection on V. Then, the random variable $\Pi_V(\boldsymbol{w})$ follows the conditional distribution given that $\boldsymbol{w} \in V$.*

Assume that $\boldsymbol{c} = (c_0, \ldots, c_{N-1})$ has been normalized so that $\sum |c_i|^2 = 1$. To sample closed random polygonal lines, it suffices to sample a standard Gaussian $\boldsymbol{w}^*$ in $\mathbb{C}^N$, and set

$$\boldsymbol{w} = \boldsymbol{w}^* - \left(\sum_{l=0}^{N-1} c_l w_l^* \right) \boldsymbol{c}.$$

Some examples of random shapes simulated with this process are provided in Fig. 6.1.

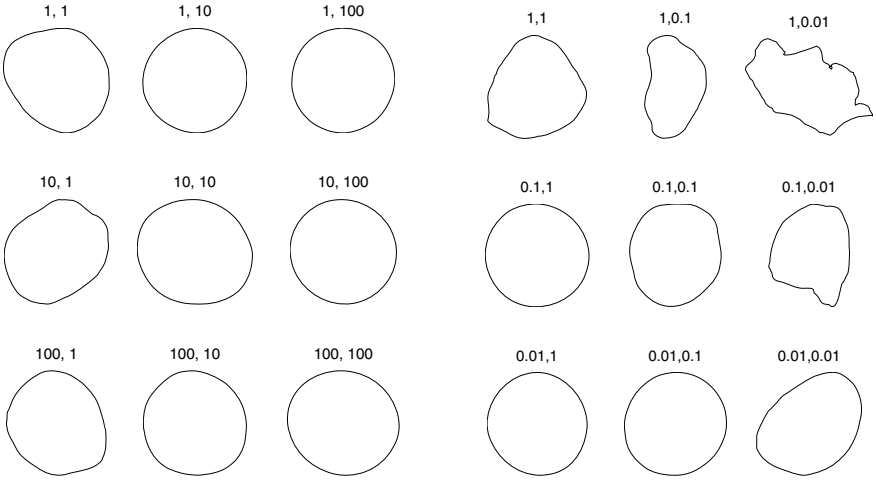

Fig. 6.1 Random deformations of a circle (with different values for α and β)

6.4 Segmentation with Deformable Templates

Using deformable templates in shape segmentation algorithms incorporates much stronger constraints than with active contours, which only implement the fact that shapes can be assumed to be smooth. If one knows the kind of shapes that are to be detected, one obviously gains in robustness and accuracy by using a segmentation method that looks for small variations of an average shape in this category.

Detection algorithms can be associated with the models provided in Sects. 6.2 and 6.3. Let us start with Sect. 6.2 with a representation that takes the form, denoting $\alpha = (\alpha_1, \dots \alpha_{p_0})$:

$$m^\alpha = \bar{m} + \sum_{i=1}^{p_0} \alpha_i K^{(i)},$$

for some template $\bar{m}$ and vector fields $K^{(i)}$. This information also comes with the variance of α_i, denoted λ_i^2.

The "pose" of the shape within the image is also unknown. It is associated with a Euclidean or affine transformation g applied to m^α. The problem is then to find g and α such that gm^α is close to regions of low deformation energy within the image.

One can use a variational approach for this purpose. As described in Sect. 5.4.9, one starts with the definition of a potential V which is small when evaluated as a point close to contours. One can then define

$$E(g, \alpha) = \sum_{i=1}^{n} \frac{\alpha_i^2}{\lambda_i^2} + \beta \int_I V(gm^\alpha(u))du.$$

The derivatives of E are

$$\partial_{\alpha_i} E = 2\frac{\alpha_i}{\lambda_i^2} + \beta \int_I \nabla V(gm^\alpha(u))^T (gK^i(u))\, du$$

and

$$\partial_g E = \int_I m^\alpha(u) \nabla V(gm^\alpha(u))^T du$$

(it is a matrix). A similar computation can be made for variants of the definition of the cost function. One can, for example add a penalty (such as $|\log\det(g)|$) to penalize shapes that are too small or too large. One can also replace the quadratic term in α_i by boundedness constraints, such as $|\alpha_i| < \sqrt{3}\lambda_i$.

If scaling to very small curves is penalized, it is plausible that, in contrast to the case of active contours, the global minimum of E provides an acceptable solution. However, from a practical point of view, minimizing E is a difficult problem, with many local minima. It is therefore still necessary to start the algorithm with a good guess of the initial curve.

Consider now the representation of Sect. 6.3. We will use the same notation as in this section, a shape being modeled by a polygon π with N edges denoted $(v_0, \ldots, v_{N-1})$. A deformation is represented by N complex numbers $z = (z_0, \ldots, z_{N-1})$, with the action

$$z \cdot \pi = (z_0 v_0, \ldots, z_{N-1} v_{N-1}).$$

We have denoted by Δ (resp. Δ_0) the set of z's for which all z_i's coincide (resp. coincide and have modulus 1); these subgroups of $G = (\mathbb{C} \setminus \{0\})^N$ correspond to plane similitudes (resp. plane rotations).

We denote by $[z]$ and $[z]_0$ the classes of z modulo Δ and Δ_0. Similarly, when π is a polygon, we denote by $[\pi]$ and $[\pi]_0$ the classes of π modulo Δ and Δ_0. For example,

$$[z] = \{c \cdot z, c \in \Delta\}$$

and

$$[\pi] = \{z \cdot \pi, z \in \Delta\}.$$

A template should be considered as a polygon modulo Δ or Δ_0 (depending on whether scale invariance is required), whereas a shape embedded in an image should be a polygon with an origin. Let $\overline{\pi}$ denote the template, although we should use the notation $[\overline{\pi}]$ or $[\overline{\pi}]_0$. Introduce a function $V(\cdot)$, defined over the image, that is large for points that are far from image contours. The quantity that can be minimized is

$$Q(z, s_0) = E([z]) + \int_{m = s_0 + z \cdot \overline{\pi}} V d\sigma_m,$$

with $s_0 \in \mathbb{C}$ and $z \in G$. The deformation energy E is a function defined on G/Δ (or equivalently a function defined on G, invariant under similitude transformations), that measures the difference between z and a similitude. For example, with $z = (z_0, \dots, z_{N-1})$, and $z_k = r_k e^{i\theta_k}$, one can take

$$E([z]) = \sum_{k=1}^{N} (\log r_k - \log r_{k-1})^2 + \sum_{k=1}^{N} \arg(e^{i\theta_k - i\theta_{k-1}})^2 .$$

Here, we have defined $\arg z$, for $z \neq 0$, as the unique $\theta \in]-\pi, \pi]$ such that $z = re^{i\theta}$ with $r > 0$. We also use the convention $r_N = r_0$, $\theta_N = \theta_0$ for the last term of the sum (assuming we are dealing with closed curves).

If scale invariance is relaxed, a simpler choice is

$$E([z]_0) = \sum_{k=1}^{N} |z_k - z_{k-1}|^2 .$$

Notice that for closed curves, it must be ensured that $z \cdot \overline{\pi}$ remains closed, which induces the additional constraint, taking $\pi = (v_0, \dots, v_{N-1})$:

$$\sum_{k=0}^{N-1} z_k v_k = 0 .$$

It is interesting to compute the continuum limit of this energy. Still using complex numbers, we consider a C^1 template curve $m : I \to \mathbb{C}$, where $I = [0, L]$ is an interval, with arc-length parametrization. For a given N, we consider the polygon $\pi = (v_0, \dots, v_{N-1})$, with

$$v_k = m\left(\frac{kL}{N}\right) - m\left(\frac{(k-1)L}{N}\right) \simeq \frac{L}{N} \partial_s m\left(\frac{(k-1)L}{N}\right) .$$

A deformation, represented by $z = (z_0, \dots, z_{N-1})$ will also be assumed to come from a continuous curve ζ defined on $[0, 1]$ with $z_k = \zeta(k/N)$. The continuum equivalent of $\pi \mapsto \pi \cdot z$ can then be written as a transformation of derivatives:

$$\partial_s m(s) \mapsto \zeta(s/L) \partial_s m(s),$$

which leads us to define an action of non-vanishing complex-valued curves ζ on closed curves by

$$(\zeta \cdot m)(s) = \int_0^s \zeta(u/L) \dot{m}(u) du .$$

In the rotation-invariant case, the energy of the action should be given by the limit of

$$\sum_{k=1}^{N-1} |z_k - z_{k-1}|^2.$$

Using the fact that $z_k - z_{k-1} \simeq \dot{\zeta}((k-1)/N)/N$, we have the continuum equivalent

$$N \sum_{k=1}^{N-1} |z_k - z_{k-1}|^2 \to \int_0^1 |\dot{\zeta}(s)|^2 du\,.$$

This is the H^1 norm of the deformation generator along the curve.

Chapter 7
Ordinary Differential Equations and Groups of Diffeomorphisms

7.1 Introduction

This chapter introduces some spaces of diffeomorphisms, and describes how ordinary differential equations provide a convenient way of generating deformations. The discussion will rely on several results in Appendix C that the reader may want consult at this point.

Let Ω be an open subset of $\mathbb{R}^d$. A "deformation" can be represented as a function φ which assigns to each point $x \in \Omega$ a displaced position $y = \varphi(x) \in \Omega$. There are two undesired behaviors that we would like to avoid:

- The deformation should not create holes: every point $y \in \Omega$ should be the image of some point $x \in \Omega$, i.e., φ should be onto.
- Folds should also be prohibited: two distinct points x and x' in Ω should not target the same point $y \in \Omega$, i.e., φ must be one-to-one.

Thus deformations must be bijections of Ω. In addition, we require some smoothness for φ. The next definition recalls some previously introduced terminology:

Definition 7.1 A homeomorphism of Ω is a continuous bijection $\varphi : \Omega \to \Omega$ such that its inverse, φ^{-1}, is continuous.

A diffeomorphism of Ω is a continuously differentiable homeomorphism $\varphi : \Omega \to \Omega$ such that φ^{-1} is continuously differentiable.

If φ is a diffeomorphism, the chain rule applied to the identity $\varphi \circ \varphi^{-1} = \mathrm{id}$ implies that its derivative, $d\varphi$, is such that $d\varphi(x)$ is invertible for all $x \in \Omega$. Conversely, using the inverse mapping theorem, one shows that a continuously differentiable homeomorphism with an invertible derivative is a diffeomorphism. From now on, most of the deformations we shall consider will be diffeomorphisms of some open set $\Omega \subset \mathbb{R}^d$. If φ and φ' are diffeomorphisms, then $\varphi \circ \varphi'$ is a diffeomorphism, and so is φ^{-1} by definition. Diffeomorphisms of Ω form a group with respect to composition of functions, denoted $\mathrm{Diff}(\Omega)$.

Throughout the rest of this book, we will write formulas that combine differentiation, inversion and composition of diffeomorphisms or functions. We will always

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_7

assume (unless specified by parentheses) that differentiation is performed first, followed by composition and inversion. For example,

$$df \circ \varphi = (df) \circ \varphi \neq d(f \circ \varphi).$$

Similarly,

$$d\varphi^{-1} = (d\varphi)^{-1} \neq d(\varphi^{-1}).$$

We will be concerned with specific subgroups of $\mathrm{Diff}(\Omega)$ associated with additional smoothness requirements. Using the notation of Appendix A, we define, for a multivariate, vector-valued function f, $\|f\|_{p,\infty}$ as the maximum of the supremum norms of the partial derivatives of order less than or equal to p of the components of f.

Definition 7.2 Let $p \geq 1$. We define $\mathrm{Diff}^{p,\infty}(\Omega)$ as the set of diffeomorphisms of Ω, φ, such that

$$\max(\|\varphi - \mathrm{id}\|_{p,\infty}, \|\varphi^{-1} - \mathrm{id}\|_{p,\infty}) < \infty.$$

We will write $\mathrm{Diff}^{p,\infty}$ instead of $\mathrm{Diff}^{p,\infty}(\Omega)$ whenever Ω is clear from the context. We show that $\mathrm{Diff}^{p,\infty}$ is a subgroup of $\mathrm{Diff}(\Omega)$. The facts that $\mathrm{id} \in \mathrm{Diff}^{p,\infty}$ and that $\varphi \in \mathrm{Diff}^{p,\infty} \Rightarrow \varphi^{-1} \in \mathrm{Diff}^{p,\infty}$ are obvious. Stability for composition will be clear from Lemma 7.3, which requires some notation. For $k \in \{1, \ldots, d\}$, denote by $\partial_k f$ the partial derivative of f with respect to the kth coordinate. For any p-tuple $J = (k_1, \ldots, k_p)$, let

$$\partial_J f = \partial_{k_1} \ldots \partial_{k_p} f. \tag{7.1}$$

(Notice that indices can be repeated in J and that the operator does not depend on how the elements of J have been ordered.) We say that a q-tuple I is a subtuple of J, and write $I \subset J$, if $I = (k_{i_1}, \ldots, k_{i_q})$ with $1 \leq i_1 < \cdots < i_q \leq p$, and we define the set

$$\partial_{(J)} f = \{\partial_I f, I \subset J\}. \tag{7.2}$$

We first recall the product rule (or Leibnitz formula). Given any two C^p scalar-valued functions u and v defined over $\mathbb{R}^d$, one has

$$\partial_J(uv) = \sum_{I \subset J} \partial_I u \; \partial_{J \setminus I} v\,. \tag{7.3}$$

This can be shown by induction, the proof being left to the reader. Here, $J \setminus I$ denotes the indices in J that are not listed in I.

Letting $|I| = q$ if I is a q-tuple, we have the following lemma, in which we let $\mathscr{D}_p$ denote the set of all partitions of $\{1, \ldots, p\}$.

Lemma 7.3 *Let $g : \Omega \to \Omega$ and $f : \Omega \to \mathbb{R}$ be C^p functions on Ω, and $J = (k_1, \ldots, k_p) \subset \{1, \ldots, d\}^p$. Then $f \circ g$ is C^p and, for $h_1, \ldots, h_p \in \mathbb{R}^d$,*

$$d^p(f\circ g)(h_1,\dots,h_p)=\sum_{\{J_1,\dots,J_k\}\in\mathscr{D}_p} d^k f\circ g(d^{|J_1|}g(h_{J_1}),\dots,d^{|J_k|}g(h_{J_k})) \quad (7.4)$$

with h_J representing the family $(h_j,\ j\in J)$.

Proof The result is true for $p=1$, with $d(f\circ g)h=df\circ g\,dg\,h$, with $\mathscr{D}_1=\{\{1\}\}$. For $p=2$, we have $d^2(f\circ g)(h_1,h_2)=d^2f\circ g(dg\,h_1,dg\,h_2)+df\circ g\,d^2g(h_1,h_2)$, which is consistent with $\mathscr{D}_2$ having two elements: $\{\{1\},\{2\}\}$ and $\{\{1,2\}\}$. We now prove the result by induction, assuming that it holds for all $k\le p$, with $p\ge 1$ and proving that it holds for $p+1$. Using the induction hypothesis, we have

$$\begin{aligned}
&d^{p+1}(f\circ g)(h_1,\dots,h_p,h_{p+1})\\
&=\sum_{\{J_1,\dots,J_k\}\in\mathscr{D}_p} d\left(d^k f\circ g(d^{|J_1|}g(h_{J_1}),\dots,d^{|J_k|}g(h_{J_k}))\right)h_{p+1}\\
&=\sum_{\{J_1,\dots,J_k\}\in\mathscr{D}_p} d^{k+1}f\circ g(d^{|J_1|}g(h_{J_1}),\dots,d^{|J_k|}g(h_{J_k}),dg\,h_{p+1})\\
&+\sum_{\{J_1,\dots,J_k\}\in\mathscr{D}_p}\sum_{j=1}^{k} d^k f\circ g\Big(d^{|J_1|}g(h_{J_1}),\dots,d^{|J_{j-1}|}g(h_{J_{j-1}}),d^{|J_j|+1}g(h_{J_j},h_{p+1}),\\
&\qquad\qquad d^{|J_{j+1}|}g(h_{J_{j+1}}),\dots,d^{|J_k|}g(h_{J_k})\Big)
\end{aligned}$$

It now suffices to observe that one can enumerate all partitions of $\{1,\dots,p+1\}$ without repetition by taking partitions in $\mathscr{D}_p$ and either adding $\{p+1\}$ as an additional subset, or replacing one of its subset by its union with $\{p+1\}$. □

From this lemma directly follows:

Proposition 7.4 *Let $g:\Omega\to\Omega$ and $f:\Omega\to\mathbb{R}$ be C^p functions on Ω. Then*

$$\|f\circ g\|_{p,\infty}\le C(p,d)\|f\|_{p,\infty}\sum_{\substack{q_1,\dots,q_j:\\ q_1+\cdots+q_j=p-j}}\|dg\|_{q_1,\infty}\cdots\|dg\|_{q_j,\infty} \quad (7.5)$$

for some fixed constant $C(p,d)$ depending on p and on the dimension.

This proposition (or Lemma 7.3) can be used in turn to prove the following inequalities involving the composition map:

$$\|f\circ g-\tilde f\circ g\|_{p,\infty}\le C(\|dg\|_{p-1,\infty})\|f-\tilde f\|_{p,\infty} \quad (7.6)$$

and

$$\|f\circ g-f\circ\tilde g\|_{p,\infty}\le C(\|dg\|_{p-1,\infty},\|d\tilde g\|_{p-1,\infty})\|f\|_{p+1,\infty}\|g-\tilde g\|_{p,\infty}, \quad (7.7)$$

where C is in both cases a polynomial function of its variables. The second inequality can be obtained, for example, by writing

$$f \circ g - f \circ \tilde{g} = \int_0^1 df \circ \big((1-\varepsilon)g + \varepsilon\tilde{g}\big)(g - \tilde{g})\, d\varepsilon$$

and applying the previous estimates and the Leibnitz formula to the integrand. If $f \in C_0^p(\Omega)$, Inequality (7.7) can be improved by introducing moduli of continuity:

$$\mu(d^q f, \delta) = \sup_{x,y\in\mathbb{R}, |x-y|<\delta} |d^q f(y) - d^q f(x)|.$$

Then, letting $\mu^{(p)}(f, \delta) = \max_{q\le p} \mu(d^q f, \delta)$ and applying Lemma 7.3, we get

$$\|f \circ g - f \circ \tilde{g}\|_{p,\infty} \le$$
$$C(\|dg\|_{p-1,\infty}, \|d\tilde{g}\|_{p-1,\infty})(\mu^{(p)}(f, \|g - \tilde{g}\|_\infty) + \|f\|_{p,\infty}\|g - \tilde{g}\|_{p,\infty}). \quad (7.8)$$

Obviously, if $f \in C_0^{p+1}(\Omega)$, then $\mu^{(p)}(f, \|g - \tilde{g}\|_\infty) \le \|f\|_{p+1,\infty}\|g - \tilde{g}\|_\infty$, so that (7.8) implies (7.7).

Based on these remarks, the following corollary of Lemma 7.3 holds.

Corollary 7.5 $\mathrm{Diff}^{p,\infty}(\Omega)$ *is a subgroup of* $\mathrm{Diff}(\Omega)$.

In the following, we will most of the time work with diffeomorphisms that tend to the identity at infinity, defining

$$\mathrm{Diff}_0^{p,\infty}(\Omega) = \left\{\varphi \in \mathrm{Diff}^{p,\infty}(\Omega) : \varphi - \mathrm{id} \in C_0^p(\Omega, \mathbb{R}^d)\right\}.$$

Here $C_0^p(\Omega, \mathbb{R}^d)$ is the space of functions $v : \Omega \to \mathbb{R}^d$ whose partial derivatives up to order p can be made arbitrarily small outside of large enough compact subsets of Ω (see Definition A.14). It will be convenient to identify $C_0^p(\Omega, \mathbb{R}^d)$ with a subset of $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$:

$$C_0^p(\Omega, \mathbb{R}^d) = \left\{f \in C_0^p(\mathbb{R}^d, \mathbb{R}^d) : f(x) = 0 \text{ if } x \notin \Omega\right\}$$

and similarly

$$\mathrm{Diff}_0^{p,\infty}(\Omega) = \left\{\varphi \in \mathrm{Diff}_0^{p,\infty}(\mathbb{R}^d) : \varphi(x) = x \text{ if } x \notin \Omega\right\}.$$

Theorem 7.6 $\mathrm{Diff}_0^{p,\infty}(\Omega)$ *is a subgroup of* $\mathrm{Diff}^{p,\infty}(\Omega)$ *and the composition map* $(\varphi, \psi) \mapsto \varphi \circ \psi$ *is continuous in both variables over* $\mathrm{Diff}_0^{p,\infty}(\Omega) \times \mathrm{Diff}_0^{p,\infty}(\Omega)$.

Moreover, for a given $\psi \in \mathrm{Diff}_0^{p,\infty}(\Omega)$, *the right composition map* $\mathscr{C}_\varphi^R : \psi \mapsto \psi \circ \varphi$ *is infinitely differentiable from* $\mathrm{Diff}_0^{k,\infty}(\Omega)$ *to itself for any* $k \le p$, *with derivative* $d\mathscr{C}_\varphi^R(\psi)\eta = \eta \circ \varphi$ *and vanishing higher derivatives, and the left composition map* $\mathscr{C}_\varphi^L : \psi \mapsto \varphi \circ \psi$ *is differentiable from* $\mathrm{Diff}_0^{k,\infty}$ *to itself if* $k \le p - 1$, *with derivative* $d\mathscr{C}_\varphi^L(\psi)\eta = (d\varphi \circ \psi)\,\eta$.

Proof The proof that $\varphi \in \mathrm{Diff}_0^{p,\infty}$ implies that $\varphi^{-1} \in \mathrm{Diff}_0^{p,\infty}$ will be given later as a consequence of Lemma 7.7. If $\varphi, \psi \in \mathrm{Diff}_0^{p,\infty}(\Omega)$, then, writing $\varphi \circ \psi - \mathrm{id} = (\varphi - \mathrm{id}) \circ \psi + \psi - \mathrm{id}$, we see that we need to prove that $(\varphi - \mathrm{id}) \circ \psi \in C_0^p(\Omega, \mathbb{R}^d)$ to ensure that $\varphi \circ \psi \in \mathrm{Diff}^{p,\infty}(\Omega)$. But the latter fact is a direct consequence of Lemma 7.3.

To prove the continuity of the composition map, fix $\varphi_0, \psi_0 \in \mathrm{Diff}^{p,\infty}(\Omega)$ and write

$$\varphi \circ \psi - \varphi_0 \circ \psi_0 = (\varphi - \varphi_0) \circ \psi + (\varphi_0 \circ \psi - \varphi_0 \circ \psi_0).$$

Assume that $\max(\|\varphi - \varphi_0\|_{p,\infty}, \|\psi - \psi_0\|_{p,\infty}) < \delta$. Then Lemma 7.3 implies that

$$\|(\varphi - \varphi_0) \circ \psi\|_{p,\infty} \leq C\|\varphi - \varphi_0\|_{p,\infty} \|\psi\|_{p,\infty} \leq C\delta(1+\delta)\|\psi_0\|_{p,\infty}$$

for some constant C. The same lemma implies that (for $q \leq p$)

$$\begin{aligned} &d^q(\varphi_0 \circ \psi - \varphi_0 \circ \psi_0)(h_1, \ldots, h_q) \\ &= \sum_{\{J_1,\ldots,J_k\} \in \mathscr{D}_q} (d^k\varphi_0 \circ \psi - d^k\varphi_0 \circ \psi_0)(d^{|J_1|}\psi(h_{J_1}), \ldots, d^{|J_k|}\psi(h_{J_k})) \\ &+ \sum_{\{J_1,\ldots,J_k\} \in \mathscr{D}_q} (d^k\varphi_0 \circ \psi_0(d^{|J_1|}\psi(h_{J_1}), \ldots, d^{|J_k|}\psi(h_{J_k})) \\ &\qquad\qquad - d^k\varphi_0 \circ \psi_0(d^{|J_1|}\psi_0(h_{J_1}), \ldots, d^{|J_k|}\psi_0(h_{J_k})). \end{aligned}$$

This implies that

$$\begin{aligned} &\|\varphi \circ \psi - \varphi_0 \circ \psi_0\|_{p,\infty} \\ &\qquad \leq C(1+\delta)^q \max_k \|d^k\varphi_0 \circ \psi - d^k\varphi_0 \circ \psi_0\|_\infty \|\psi_0\|_{p,\infty}^q + C\delta\|\varphi_0\|_{p,\infty}. \end{aligned}$$

Because $d^k\varphi_0$ is uniformly continuous on Ω (by Proposition A.15), $\|(d^k\varphi_0) \circ \psi - (d^k\varphi_0) \circ \psi_0\|_\infty$ can be made arbitrarily small by taking δ small enough. This shows that $\|\varphi \circ \psi - \varphi_0 \circ \psi_0\|_{p,\infty}$ can be made arbitrarily small too and proves the continuity of the composition map.

The differentiability of the right composition map is a consequence of its continuity, since it is linear. For the left composition, one can write

$$\varphi \circ (\psi + \eta) - \varphi \circ \psi - d\varphi \circ \varphi\,\eta = \int_0^1 (d\varphi \circ (\varphi + t\eta) - d\varphi \circ \varphi)\eta, dt$$

and the previous argument showing the continuity of the left composition can now be applied to the integrand to prove that $\|d\varphi \circ (\varphi + t\eta) - d\varphi \circ \varphi\|_{p-1,\infty}$ tends to 0 with $\|\eta\|_{p,\infty}$, which, combined with the product rule, implies that the $(p-1,\infty)$-norm of the right-hand side is an $o(\|\eta\|_{p-1,\infty})$. □

We now review properties of the inversion map $Inv : \varphi \mapsto \varphi^{-1}$. We first relate the derivatives of φ^{-1} to those of φ under the assumption that the latter is a C^p diffeomorphism of Ω. In this discussion, we will fix φ and define $\psi = \varphi^{-1}$. Recall that $d\psi\, h = d\varphi^{-1} \circ \psi\, h$ for all $h \in \mathbb{R}^d$. Differentiating again and applying the chain rule, we have

$$d^2\psi(h_1, h_2) = -d\varphi^{-1} \circ \psi\, d^2\varphi(d\varphi^{-1} \circ \psi\, h_1, d\varphi^{-1} \circ \psi\, h_2) \circ \psi$$

or

$$d^2\psi(h_1, h_2) \circ \varphi = -d\varphi^{-1} d^2\varphi(d\varphi^{-1} h_1, d\varphi^{-1} h_2).$$

Things rapidly get more complicated with higher derivatives. We let the reader check that

$$\begin{aligned} d^3\psi(h_1, h_2, h_3) \circ \varphi = &-d\varphi^{-1} d^3\varphi(d\varphi^{-1} h_1, d\varphi^{-1} h_2, d\varphi^{-1} h_3) \\ &+ d\varphi^{-1} d^2\varphi(d\varphi^{-1} d^2\varphi(d\varphi^{-1} h_1, d\varphi^{-1} h_2), d\varphi^{-1} h_3) \\ &+ d\varphi^{-1} d^2\varphi(d\varphi^{-1} d^2\varphi(d\varphi^{-1} h_1, d\varphi^{-1} h_3), d\varphi^{-1} h_2) \\ &+ d\varphi^{-1} d^2\varphi(d\varphi^{-1} d^2\varphi(d\varphi^{-1} h_2, d\varphi^{-1} h_3), d\varphi^{-1} h_1). \end{aligned}$$

To state a general formula, we fix $h_1, \ldots, h_p \in \mathbb{R}^d$ and we introduce the following subsets of mappings from Ω to the sets of multilinear forms from $\mathbb{R}^d \to \mathbb{R}^d$.

Let $\mathcal{P}_k$ denote the family of all subsets of $V_k = \{1, \ldots, k\}$. A recursive partition of $\{1, \ldots, k\}$ can be represented as a tree T whose vertices are subsets of $(1, \ldots, k)$ with the additional conditions: (i) The root of T is the whole set V_k; (ii) The children of each node form a partition of the subset associated to this node; (iii) Each non-terminal node has at least two children; (iv) Terminal nodes are singletons. Let $\mathcal{T}_k$ be the set of recursive partitions of V_k. To each $T \in \mathcal{T}_k$, we recursively define the function U_T by

$$U_T(h_1, \ldots, h_k) = d\varphi^{-1} d^k\varphi(h_1, \ldots, h_k)$$

if all children of the root of T are singletons and

$$U_T(h_1, \ldots, h_k) = d\varphi^{-1} d^m\varphi(U_{T_1}(h_{V_1}), \ldots, U_{T_m}(h_{V_m})),$$

where m is the number of children of V_k (the root node), with associated subsets V_1, ..., V_m, and $T_1, \ldots, T_m$ are the subtrees of T rooted at $V_1, \ldots, V_m$. With this notation, we have

Lemma 7.7

$$d^p\psi(d\varphi h_1, \ldots, d\varphi h_p) \circ \varphi = (-1)^p \sum_{T \in \mathcal{T}_p} (-1)^{|T|} U_T(h_1, \ldots, h_p), \tag{7.9}$$

where $|T|$ is the number of nodes in T.

Proof We prove this by induction. For a tree with one node, we have $d\psi \circ \varphi h = d\varphi^{-1} h$ so that $d\psi \circ \varphi d\varphi h = d\varphi^{-1} d\varphi h$. For $p = 2$, the only tree in $\mathscr{T}_2$ has $\{1, 2\}$ as root and $\{1\}$, $\{2\}$ as terminal nodes, and

$$d^2\psi(d\varphi h_1, d\varphi h_2) \circ \varphi = -d\varphi^{-1} d^2\varphi(h_1, h_2) = (-1)^{2+|T|} U_T(h_1, h_2).$$

Let us now prove the formula by induction, assuming that the formula is true for all $k \in \{2, \ldots, p\}$ and showing that it remains true for $p + 1$. Taking the derivative of the left-hand side of (7.9), we find

$$\begin{aligned} d(d^p\psi(d\varphi h_1, \ldots, d\varphi h_p) \circ \varphi) h_{p+1} &= d^{p+1}(\psi)(d\varphi h_1, \ldots, d\varphi h_{p+1}) \circ \varphi \\ &+ \sum_{j=1}^{p} d^p\psi(d\varphi h_1, \ldots, d\varphi j_{j-1}, d^2\varphi(h_j, h_{p+1}), d\varphi h_{j+1}, \ldots, d\varphi h_p). \end{aligned}$$

Let $T_{p+1}^{(1)}$ be the set of trees obtained from trees $T \in \mathscr{T}_p$ by choosing a terminal node $\{j\}$ in T, replacing it by $\{j, p + 1\}$, adding the two singleton children, and adding $p + 1$ as a new element of all the ancestors of $\{j\}$. Similarly, let $\mathscr{T}_{p+1}^{(2)}$ be the set of trees obtained by adding $\{p + 1\}$ as a child to some non-terminal node, and propagating upward. Finally let $\mathscr{T}_{p+1}^{(3)}$ denote the trees T' obtained from $T \in \mathscr{T}_p$ by placing $\{p + 1\}$ as a sibling of the root of T (which is V_p), both becoming children of the root of T' (which is V_{p+1}). These operations add two nodes to the original trees for groups (1) and (3) (so that $(-1)^{|T|}$ remains unchanged), and only one node for group (2), changing the sign of $(-1)^{|T|}$. Notice that $\mathscr{T}_{p+1}^{(1)}$, $\mathscr{T}_{p+1}^{(2)}$, $\mathscr{T}_{p+1}^{(3)}$ form a partition of $\mathscr{T}_{p+1}$.

We can already see that

$$\begin{aligned} \sum_{j=1}^{p} d\varphi\, d^p\psi(d\varphi h_1, \ldots, d^2\varphi(h_j, h_{p+1}), \ldots, d\varphi h_p) \\ = (-1)^p \sum_{T \in \mathscr{T}_{p+1}^{(1)}} (-1)^{|T|} U_T(h_1, \ldots, h_p, h_{p+1}). \end{aligned}$$

Now consider the derivatives in the right-hand side of (7.9). From the recursive definition of U_T, we find that

$$\begin{aligned} dU_T(h_1, \ldots, h_p, h_{p+1}) &= -d\varphi^{-1} d^2\varphi(U_T(h_1, \ldots, h_p), h_{p+1}) \\ &+ d\varphi^{-1} d^{m+1}\varphi(U_{T_1}(h_{V_1}), \ldots, U_{T_m}(h_{V_m}), h_{p+1}) \\ &+ \sum_{j=1}^{m} d\varphi^{-1} d^m\varphi(U_{T_1}(h_{V_1}), \ldots, dU_{T_j}(h_{V_j}, h_{p+1}), \ldots, U_{T_m}(h_{V_m}) \\ &= -(-1)^p \sum_{T \in \mathscr{T}_{p+1}^{(3)}} (-1)^{|T|} U_T(h_1, \ldots, h_{p+1}) \\ &+ (-1)^p \sum_{T \in \mathscr{T}_{p+1}^{(2)}} (-1)^{|T|-1} U_T(h_1, \ldots, h_{p+1}). \end{aligned}$$

Putting all the terms together, we retrieve

$$d^{p+1}(\psi)(d\varphi\, h_1, \dots, d\varphi\, h_{p+1}) \circ \varphi = (-1)^{p+1} \sum_{T \in \mathscr{T}_{p+1}} (-1)^{|T|} U_T(h_1, \dots, h_{p+1}),$$

which proves the lemma. □

An immediate consequence of this lemma is that, for some continuous function C, one has $\|\varphi^{-1}\|_{p,\infty} \le C(\|\varphi\|_{p,\infty})\|d\varphi^{-1}\|_\infty^{p+1}$. Using this, we can obtain the fact that $\mathrm{Diff}_0^{p,\infty}(\Omega)$ is invariant by function inversion, completing the proof of Theorem 7.6. Here is another consequence.

Proposition 7.8 *If $\varphi \in \mathrm{Diff}_0^{p,\infty}(\Omega)$, the mapping $\mathcal{I}_\varphi : \eta \mapsto (\varphi + \eta)^{-1}$ is differentiable at $\eta = 0$ when considered as a mapping from $C_0^p(\Omega, \mathbb{R}^d)$ to $C^{p-1}(\Omega, \mathbb{R}^d)$, with derivative $d\mathcal{I}_\varphi(0)\eta = -d(\varphi^{-1})\eta \circ \varphi^{-1} = -(d\varphi^{-1}\eta) \circ \varphi^{-1}$.*

Proof We first consider the case $\varphi = \mathrm{id}$. Then $d(\mathrm{id} + \eta) = \mathrm{Id} + d\eta$ and $d^k(\mathrm{id} + \eta) = d^k\eta$ for $k \ge 2$. We have

$$(\mathrm{id} + \eta)^{-1} - \mathrm{id} + \eta = (\mathrm{id} + \eta)^{-1} - (\mathrm{id} + \eta) \circ (\mathrm{id} + \eta)^{-1} + \eta = \eta - \eta \circ (\mathrm{id} + \eta)^{-1}$$

so that

$$\begin{aligned} \|(\mathrm{id} + \eta)^{-1} - \mathrm{id} + \eta\|_\infty &= \|\eta - \eta \circ (\mathrm{id} + \eta)^{-1}\|_\infty \\ &= \|\eta \circ (\mathrm{id} + \eta) - \eta\|_\infty \\ &\le \|d\eta\|_\infty \|\eta\|_\infty = O(\|\eta\|_{1,\infty}^2). \end{aligned}$$

Similarly, with $d((\mathrm{id} + \eta)^{-1}) = -(\mathrm{Id} + d\eta)^{-1} \circ (\mathrm{id} + \eta)^{-1} = \mathrm{Id} - (d\eta(\mathrm{Id} + d\eta)^{-1}) \circ (\mathrm{id} + \eta)^{-1}$,

$$\begin{aligned} \|d((\mathrm{id} + \eta)^{-1}) - \mathrm{Id} + d\eta\|_\infty &= \|d\eta \circ (\mathrm{id} + \eta) - d\eta(\mathrm{Id} + d\eta)^{-1}\|_\infty \\ &= \|\eta \circ (\mathrm{id} + \eta) - \eta\|_\infty \le \|d\eta\|_\infty \|\eta\|_\infty = O(\|\eta\|_{2,\infty}^2). \end{aligned}$$

Using again the fact that $(\mathrm{Id} + d\eta)^{-1} - \mathrm{id} = -d\eta(\mathrm{Id} + d\eta)^{-1}$, we can check that, for $T \in \mathscr{T}_k$, $k \ge 2$,

$$U_T((\mathrm{Id} + d\eta)^{-1}h_1, \dots, (\mathrm{Id} + d\eta)^{-1}h_k) = o(\|\eta\|_{k,\infty}^2),$$

unless T is such that all children of the root are singletons, for which

$$U_T((\mathrm{Id} + d\eta)^{-1}h_1, \dots, (\mathrm{Id} + d\eta)^{-1}h_k) = -d^k\eta(h_1, \dots, h_k) + O(\|\eta\|_{k,\infty}^2).$$

It follows that

$$\|d^k((\mathrm{id} + \eta)^{-1})(h_1, \dots, h_k) \circ (\mathrm{id} + \eta) + d^k\eta(h_1, \dots, h_k)\|_\infty = O(\|\eta\|_{k,\infty}^2).$$

The left-hand side is also equal to

$$\|d^k((\mathrm{id}+\eta)^{-1})(h_1,\dots,h_k)+d^k\eta(h_1,\dots,h_k)\circ(\mathrm{id}+\eta)^{-1}\|_\infty$$

and

$$\|d^k\eta(h_1,\dots,h_k)\circ(\mathrm{id}+\eta)^{-1}-d^k\eta(h_1,\dots,h_k)\|_\infty\le\|d^{k+1}\eta\|_\infty\|\eta\|_\infty.$$

This implies that

$$\|d^k((\mathrm{id}+\eta)^{-1})(h_1,\dots,h_k)+d^k\eta(h_1,\dots,h_k)\|_\infty=O(\|\eta\|^2_{k+1,\infty}),$$

proving that $d\mathcal{I}_{\mathrm{id}}(0)\eta=-\eta$.

To prove the result for $\varphi\neq\mathrm{id}$, we note that $(\varphi+\eta)^{-1}=\varphi^{-1}\circ(\mathrm{id}+\eta\circ\varphi^{-1})^{-1}$. The conclusion then results from the chain rule applied to

$$\eta\mapsto\mathrm{id}+\eta\circ\varphi^{-1}\mapsto(\mathrm{id}+\eta\circ\varphi^{-1})^{-1}\mapsto\varphi^{-1}\circ(\mathrm{id}+\eta\circ\varphi^{-1})^{-1}$$

respectively mapping $C^p_0(\Omega,\mathbb{R}^d)$ to $\mathrm{Diff}^{p,\infty}_0$ to $\mathrm{Diff}^{p-1,\infty}_0$ to $\mathrm{Diff}^{p-1,\infty}_0$. The differentiability of each map results from Theorem 7.16, the first part of the proof and the fact that $\varphi^{-1}\in\mathrm{Diff}^{p,\infty}_0(\Omega)$. □

Diffeomorphisms act on various structures supported by Ω. Consider, for example, a function $I:\Omega\to\mathbb{R}$ (an image) and a diffeomorphism φ of Ω. The associated deformation creates a new image I' on Ω by letting $I'(y)$ be the value of I at the position x which has been moved to y, i.e., $I'(y)=I(\varphi^{-1}(y))$ or $I'=I\circ\varphi^{-1}$. We will be specifically interested in the inverse problem of estimating the best diffeomorphism from the output of its action. For example, the image matching problem consists in finding an algorithm which, given two functions I and I' on Ω, is able to recover a plausible diffeomorphism φ such that $I'=I\circ\varphi^{-1}$.

To be able to develop these algorithms, we will need a computational construction of diffeomorphisms (this is not provided by Definition 7.2). In order to motivate this general construction, we start with a direct, but limited, way of building diffeomorphisms, by small perturbations of the identity.

Proposition 7.9 *Let $u\in C^1(\Omega,\mathbb{R}^d)$, and assume that,*

(i) $u(x)$ and $du(x)$ tend to 0 as x tends to infinity.
(ii) There exists a constant C_0 such that $|u(x)|<C_0\,\mathrm{dist}(x,\Omega^c)$ for all $x\in\Omega$.

Then, for small enough ε, $\varphi:x\mapsto x+\varepsilon u(x)$ is a diffeomorphism of Ω.

(Of course, if Ω has no boundary, one can take C_0 arbitrarily small.)

Proof The function φ is obviously continuously differentiable. Since du is continuous and tends to 0 at infinity, it is bounded and there exists a constant C such

that $\left|u(x) - u(x')\right| \leq C\left|x - x'\right|$. We will take $\varepsilon < 1/(C + C_0)$. This implies that $x + \varepsilon u(x)$ takes Ω into itself. Indeed, if $y \in \Omega^c$, then

$$|x + \varepsilon u(x) - y| \geq (1 - C_0\varepsilon)\mathrm{dist}(x, \Omega^c) > 0.$$

We first show that φ is one-to-one. If $\varphi(x) = \varphi(x')$ we have

$$\left|x - x'\right| = \varepsilon\left|u(x) - u(x')\right| \leq C\varepsilon\left|x - x'\right| < \frac{C}{C + C_0}\left|x - x'\right|$$

which implies $x = x'$ as needed.

We now show that φ is onto. Take $y \in \Omega$ and $\rho = \mathrm{dist}(y, \Omega^c)$. If $|\eta| < \rho$, then $u(y + \eta)$ is well defined and satisfies

$$\varepsilon|u(y + \eta)| \leq \varepsilon|u(y)| + \varepsilon C\eta \leq \varepsilon(C_0 + C)\rho < \rho.$$

(The conclusion obviously also holds when $\rho = +\infty$.) Therefore, $\psi_y(\eta) := -\varepsilon u(y + \eta)$ maps $B(0, \rho)$ into itself. For $\eta, \eta' \in B(0, \rho)$, we have

$$\left|\psi_y(\eta) - \psi_y(\eta')\right| \leq \varepsilon C\left|\eta - \eta'\right|.$$

Since $\varepsilon C < 1$, ψ_y is contractive, and the fixed-point theorem (Theorem C.1) implies that there exists an $\eta \in B(0, \rho)$ such that $\psi_y(\eta) = \eta$. But in this case,

$$\varphi(y + \eta) = y + \eta + \varepsilon u(y + \eta) = y + \eta - \psi_y(\eta) = y,$$

so that $y \in \varphi(\Omega)$ and φ is onto.

It remains to prove that φ^{-1} is continuous. Assume that $\varphi(x) = y$ and $\varphi(x') = y'$. Then, from $x - x' = y - y' - \varepsilon(u(x) - u(x'))$ we get

$$|x - x'| \leq |y - y'| + C\varepsilon|x - x'|,$$

so that

$$|x - x'| < \frac{1}{1 - C\varepsilon}|y - y'|,$$

which proves the continuity of φ^{-1}. □

We therefore know how to build small deformations. Of course, we cannot be satisfied with this, since they correspond to a rather limited class of diffeomorphisms. However, we can use them to generate large deformations, because diffeomorphisms can be combined using composition.

Thus, let $\varepsilon_0 > 0$ and $u_1, \ldots, u_n, \ldots$ be vector fields on Ω which are such that, for $\varepsilon < \varepsilon_0$, $\mathrm{id} + \varepsilon u_i$ is a diffeomorphism of Ω. Consider

$$\varphi_n = (\mathrm{id} + \varepsilon u_n) \circ \cdots \circ (\mathrm{id} + \varepsilon u_1).$$

We have

$$\varphi_{n+1} = (\mathrm{id} + \varepsilon u_n) \circ \varphi_n = \varphi_n + \varepsilon u_n \circ \varphi_n,$$

which can also be written as $(\varphi_{n+1} - \varphi_n)/\varepsilon = u_n \circ \varphi_n$. Fixing $x \in \Omega$ and letting $x_0 = x$, $x_n = \varphi_n(x)$, we have the relation $(x_{n+1} - x_n)/\varepsilon = u_n(x_n)$. This can be viewed as a discretization of a differential equation, of the form (introducing a continuous time variable t):

$$\partial_t x(t) = u(t, x(t)).$$

This motivates the rest of this chapter, which will be devoted to building diffeomorphisms as flows associated to ordinary differential equations (ODEs).

7.2 Flows and Groups of Diffeomorphisms

7.2.1 Definitions

We let $\Omega \subset \mathbb{R}^d$ be open and denote as before by $C_0^p(\Omega, \mathbb{R}^d)$ the Banach space of continuously differentiable vector fields v on Ω that tend to 0 at infinity, according to Definition A.14. Elements $v \in C_0^p(\Omega, \mathbb{R}^d)$ can be considered as defined on $\mathbb{R}^d$ by setting $v(x) = 0$ if $x \notin \Omega$, and therefore may also be considered as elements of $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$.

We define the set $\mathcal{X}^{p,1}(T, \Omega) = L^1([0, T], C_0^p(\Omega, \mathbb{R}^d))$ of absolutely integrable functions from $[0, T]$ to $C_0^p(\Omega, \mathbb{R}^d)$. An element of $\mathcal{X}^{p,1}(T, \Omega)$ is a time-dependent vector field, $(v(t, \cdot), t \in [0, 1])$ such that, for each t, $v(t) := v(t, \cdot) \in C_0^p(\Omega, \mathbb{R}^d)$ and

$$\|v\|_{\mathcal{X}^{p,1},T} := \int_0^T \|v(t)\|_{p,\infty}\, dt < \infty. \tag{7.10}$$

$\mathcal{X}^{p,1}$ is a Banach space provided one identifies time-dependent vector fields v and v' such that $v(t, \cdot) = v'(t, \cdot)$ for almost all $t \in [0, T]$.

7.2.2 Variation in the Initial Condition

The results developed in Appendix C for ODEs in Banach spaces clearly apply to this context with $\mathbb{B} = \mathbb{R}^d$ and noting that $C_0^p(\Omega, \mathbb{R}^d) \subset C_{(0)}^p(\Omega, \mathbb{R}^d)$ (see the notation in Appendix C). In particular, for $v \in \mathcal{X}^{1,1}(T, \Omega)$, the ordinary differential equation $\partial_t y = v(t, y)$ has a unique solution over $[0, T]$ given any initial condition $y(s) = x$, and the associated flow, φ_{st}^v, defined by

$$\partial_t \varphi^v_{st} = v(t) \circ \varphi^v_{st}, \tag{7.11}$$

and $\varphi^v_{ss} = \mathrm{id}$, is a diffeomorphism of Ω, with derivative described in Theorems C.15 and C.18. We repeat these results in the following theorem.

Theorem 7.10 *If $p \geq 1$ and $\int_0^T \|v(t)\|_{p,\infty} dt < \infty$, then, for all $s, t \in [0, T]$, φ^v_{st} is p times differentiable and for all $q \leq p$*

$$\partial_t d^q \varphi^v_{st} = d^q (v(t) \circ \varphi^v_{st}) \tag{7.12}$$

with initial condition $d^q \varphi^v_{ss} = \mathrm{Id}_d$ for $q = 1$ and $d^q \varphi^v_{ss} = 0$ for $q > 1$. In particular, $d\varphi^v_{st}$ satisfies

$$\partial_t W(t) = dv(t, \varphi^v_{st}(x)) W(t). \tag{7.13}$$

Moreover, there exist constants C, C' (independent of v) such that

$$\sup_{s \in [0,1]} \left\| \varphi^v_{st} - \mathrm{id} \right\|_{p,\infty} \leq C \exp\left(C' \int_0^1 \|v(t)\|_{p,\infty} dt \right). \tag{7.14}$$

The following theorem also controls the behavior of φ_{st} at infinity.

Theorem 7.11 *If $p \geq 1$ and $\int_0^t \|v(t)\|_{p,\infty} dt < \infty$, then $\varphi^v_{st} \in \mathrm{Diff}^{p,\infty}_0(\Omega)$.*

Proof We only need to complete Theorem 7.10 by proving that $\varphi(x) - x$ and its derivatives vanish at infinity. We have

$$\varphi^v_{st}(x) - x = \int_s^t v(u, \varphi_{su}(x))\, du,$$

which implies that $|\varphi^v_{st}(x) - x| \leq (t - s)\|v\|_\infty$. Note that $\varphi^v_{st}(x) - x = 0$ for $x \notin \Omega$, so that there is nothing to prove for these x's.

Let

$$\mathcal{K}_n = \left\{ x \in \Omega : |x| \leq n \text{ and } \mathrm{dist}(x, \Omega^c) \geq \frac{1}{n} \right\}$$

so that $(\mathcal{K}_n, n \geq 1)$ is an increasing sequence of compact subsets of Ω and any compact subset of Ω is included in some $\mathcal{K}_n$ for n large enough. Assume (to reach a contradiction) that there exists an $\varepsilon > 0$ such that, for all n, there exists $x_n \in \Omega \setminus \mathcal{K}_n$ such that $|\varphi^v_{st}(x_n) - x_n| > \varepsilon$. Because $\varphi^v_{su} - \mathrm{id}$ is bounded, $\varphi^v_{su}(x_n)$ also tends to infinity, for all $u \in [s, t]$, so that $|v(u, \varphi^v_{su}(x_n))| \to 0$. It suffices to apply the dominated convergence theorem to conclude that

$$\int_0^T |v(u, \varphi_{su}(x_n))|\, du \to 0,$$

which provides our contradiction. We have therefore shown that, for all $\varepsilon > 0$, one can take n large enough so that $|\varphi^v_{st}(x) - x| < \varepsilon$ for all $x \in \Omega \setminus \mathcal{K}_n$, so that $\varphi_{st} - \mathrm{id} \in$

$C_0^0(\Omega, \mathbb{R}^d)$. This argument can be extended to derivatives using Lemma 7.3 to prove the proposition. The details are left to the reader □

7.2.3 Variations with Respect to the Vector Field

The previous equation can be analyzed in a more abstract form, by considering the mapping $\mathscr{V}^v : C_0^p(\Omega, \mathbb{R}^d) \to C_0^p(\Omega, \mathbb{R}^d)$ defined by $\mathscr{V}^v(t,u)(x) = v(t, x + u(x))$ for $v \in \mathcal{X}^{p+1}(T, \Omega)$, or $\mathscr{V}^v(t,u) = v(t) \circ (\mathrm{id} + u)$. Letting $\mathbb{B} = C_0^p(\Omega, \mathbb{R}^d)$ the assumption that $v \in \mathcal{X}^{p+1}(T, \Omega)$ implies that $\mathscr{V}^v \in L^1([0,T], C_b^1(\mathbb{B}, \mathbb{B}))$. This statement relies on the observation that if we let (for given t and u)

$$\begin{aligned} R(t,h) &= \mathscr{V}^v(t, u+h) - \mathscr{V}^v(t,u) - dv(t) \circ (\mathrm{id} + u)h \\ &= v(t) \circ (\mathrm{id} + u + h) - v(t) \circ (\mathrm{id} + u) - dv(t) \circ (\mathrm{id} + u)h \end{aligned}$$

then Lemma 7.3 and the product rule imply that

$$\begin{aligned} \|R(t,h)\|_{p,\infty} &= \int_0^1 (dv(t) \circ (\mathrm{id} + u + \varepsilon h) - dv(t) \circ (\mathrm{id} + u))\, h\, d\varepsilon \\ &\le \mu^{(p)}(dv(t), \|h\|_{p,\infty}) \|h\|_{p,\infty} \\ &= o(\|h\|_{p,\infty}), \end{aligned}$$

so that $\mathscr{V}^v(t)$ is C^1 with derivative $d\mathscr{V}^v(t,u) = dv(t) \circ (\mathrm{id} + u)$, which is continuous in u with respect to the (p,∞)-norm if $v(t) \in C_0^{p+1}(\Omega, \mathbb{R}^d)$. Moreover,

$$\int_0^T \|d\mathscr{V}^v(t)\|_\infty dt \le \int_0^T \|dv(t)\|_{p,\infty} dt \le \|v\|_{\mathcal{X}^{p+1,1},T},$$

proving that $\mathscr{V}^v \in L^1([0,T], C_b^1(\mathbb{B}, \mathbb{B}))$. This shows that $\partial_t u = \mathscr{V}^v(t,u)$ has a unique solution with $u_s^v(t) \in C_0^p(\Omega, \mathbb{R}^d)$, $u^v(s) = 0$ and $\varphi_{st}^v = \mathrm{id} + u_s^v(t) \in \mathrm{Diff}_0^{p,\infty}(\Omega)$. We therefore retrieve the previous result (but under a stronger assumption on v).

We now apply Theorem C.17 in this framework to study the differentiability of the flow with respect to v. The (linear) mapping $v \mapsto \mathscr{V}^v$ is C^1 from $\mathcal{X}^{p+1}(T, \Omega)$ to $L^1([0,T], C_b^1(\mathbb{B}, \mathbb{B}))$ with $\partial_v \mathscr{V}^v h = \mathscr{V}^h$, so that the theorem's assumptions are satisfied. We therefore have the following result (see Eq. (C.13) after Theorem C.17).

Theorem 7.12 *The mapping $v \mapsto \varphi_{st}^v$ is differentiable from $\mathcal{X}^{p+1,1}(T, \Omega)$ to $\mathrm{Diff}_0^{p,\infty}(\Omega)$ and*

$$\partial_v \varphi_{st}^v h = \int_s^t (d\varphi_{ut}^v h(u)) \circ \varphi_{su}^v du. \tag{7.15}$$

Note that, even though Theorem 7.11 states that $\varphi^v_{st} \in \mathrm{Diff}^{p+1,\infty}_0$, it is with respect to the (p,∞) norm that this mapping is differentiable. Combined with Theorem 7.10 this result implies that, for $v, \tilde{v} \in \mathcal{X}^{p+1,1}$, we have

$$\begin{aligned}\left\|\varphi^v_{st} - \varphi^{\tilde{v}}_{st}\right\|_{p,\infty} &= \int_0^1 \int_s^t d\varphi^{(1-\varepsilon)v+\varepsilon\tilde{v}}_{su}(\tilde{v}-v)\,du\,d\varepsilon \\ &\le \left(1 + C\exp\left(C'\max\left(\|v\|_{\mathcal{X}^{p+1,1}}, \|\tilde{v}\|_{\mathcal{X}^{p+1,1}}\right)\right)\right)\|\tilde{v}-\tilde{v}'\|_{\mathcal{X}^{p+1,1}}.\end{aligned} \tag{7.16}$$

7.2.4 Weak Continuity

The previous theorem implies the continuity of $v \mapsto \varphi^v_{st}$ in the norm topology, when this mapping is considered from $\mathcal{X}^{p+1,\infty}$ to $\mathrm{Diff}^{p,\infty}_0$. We now show that this mapping is also weakly continuous, or, more precisely, that if v^n is any bounded weakly converging sequence in $\mathcal{X}^{p+1,\infty}$, with limit v, then $\varphi^{v^n}_{st}$ and its first p derivatives converge to φ^v_{st} uniformly over compact sets. Write, for a d-tuple J and $x \in \Omega$,

$$\partial_J\varphi^v_{st}(x) - \partial_J\varphi^{v^n}_{st}(x) = \int_s^t \left(\left(\partial_J(v(u)\circ\varphi^v_{su}(x)) - \partial_J(v^n(u)\circ\varphi^{v^n}_{su}(x))\right)\right.\,du.$$

We have

$$\left|\partial_J\varphi^v_{st}(x) - \partial_J\varphi^{v^n}_{st}(x)\right| \le \left|\int_s^t \left(\partial_J(v(u)\circ\varphi^v_{su}(x)) - \partial_J(v^n(u)\circ\varphi^v_{su}(x))\right)du\right| \tag{7.17}$$

$$+ \int_s^t \left|\partial_J(v^n(u)\circ\varphi^v_{su}(x)) - \partial_J(v^n(u)\circ\varphi^{v^n}_{su}(x))\right|du. \tag{7.18}$$

Applying Lemma 7.3, we see that the second integral is less than

$$C\int_s^t \|v^n(u)\|_{p+1,\infty}\left(\max_{|J|\le p}|\partial_J\varphi^v_{su}(x) - \partial_J\varphi^{v^n}_{su}(x)|\right)du$$

for some constant C that depends on $\max_u \|\varphi^v_{su}\|_{p,\infty}$ and $\max_u \|\varphi^{v^n}_{su}\|_{p,\infty}$, both quantities being controlled by $\|v\|_{p,\infty}$ and $\sup_n \|v^n\|_{p,\infty}$. We are therefore in position to apply Theorem C.11 (Gronwall's lemma) to (7.17) with

$$u(\tau, c) = \max_{|J|\le p}\left|\partial_J\varphi^v_{s\,s+\tau} - \partial_J\varphi^{v^n}_{s+\tau}\right|,$$

$\alpha(\tau) = \|v^n(s+\tau)\|_{p+1,\infty}$ and

$$c^n(\tau, x) = \left| \int_s^{s+\tau} \left(\partial_J (v(\theta) \circ \varphi_{s\theta}^v) - \partial_J (v^n(\theta) \circ \varphi_{s\theta}^v) \right) d\theta \right|. \tag{7.19}$$

This yields the inequality

$$\max_{|J|\leq p} \left\| \partial_J \varphi_{st}^v(x) - \partial_J \varphi_{st}^{v^n}(x) \right\| \leq c^n(t-s) + C \int_s^t c^n(\theta - s) \left\| v^n(\theta) \right\|_{p+1,\infty} \exp\left(\int_\theta^t \left\| v^n(\theta') \right\|_{p+1,\infty} d\theta' \right) d\theta. \tag{7.20}$$

The weak convergence of v^n to v implies that

$$\int_s^{s+\tau} \partial_J v^n(\theta, y(\theta)) d\theta \to \int \partial_J v(\theta, y(\theta)) d\theta$$

for any measurable function y, since the right-hand side is a continuous linear form of $v \in \mathcal{X}^{p+1,\infty}$. Using this and Lemma 7.3, we see that $c^n(\tau, x) \to 0$ for all τ. Adding to this the fact that $c^n(\tau, x)$ is bounded uniformly in τ (again as a consequence of Lemma 7.3), it suffices to apply the dominated convergence theorem to obtain the fact that the left-hand side of (7.20) goes to 0 as $n \to \infty$.

Theorem 7.10 and Eq. (C.5) imply that there exists a constant C such that, for all $n > 0$, and $x, y \in \Omega$

$$\left| \partial_J (\varphi_{st}^{v^n}(x) - \varphi_{st}^{v^n}(y)) \right| \leq C |x - y|.$$

This implies that the family $\partial_J \varphi_{st}^{v^n}$ is equicontinuous and a similar argument shows that it is bounded. Letting Q be any compact subset of Ω, Ascoli's theorem [306] implies that $(\partial_J \varphi_{st}^{v^n}, n \geq 0)$ is relatively compact with respect to the uniform convergence on Q. But the limit of any subsequence that converges uniformly must be $\partial_J \varphi_{st}^v$ since it is already the pointwise limit of the whole sequence. This implies that the uniform limit exists and is equal to $\partial_J \varphi_{st}^v$. Thus, we have just proved the following theorem:

Theorem 7.13 ([126]) *If $v \in \mathcal{X}^{p+1,1}(T, \Omega)$ and v^n is a bounded sequence in $\mathcal{X}^{p+1,1}(T, \Omega)$ which weakly converges to v, then, for all $s, t \in [0, T]$, for every compact subset $Q \subset \overline{\Omega}$ and every tuple J such that $|J| \leq p$,*

$$\lim_{n\to\infty} \max_{x\in Q} \left| \partial_J (\varphi_{st}^{v^n}(x) - \varphi_{st}^v(x)) \right| = 0.$$

We will say that φ^{v^n} converges to φ^v in the (p, ∞)-compact topology. This topology is metrizable, with possible metric

$$d(\varphi, \psi) = \sum_{n=0}^{\infty} 2^{-n} \max_{x\in Q_n} \max_{|J|\leq p} \frac{|\partial_J \varphi(x) - \partial_J \psi(x)|}{1 + |\partial_J \varphi(x) - \partial_J \psi(x)|},$$

where Q_n is an increasing sequence of compact sets such that any compact set $Q \in \Omega$ belongs to Q_n for large enough n.

7.2.5 Admissible Banach Spaces

The previous results are true a fortiori for vector fields v belonging to Banach or Hilbert spaces that are continuously embedded in $C_0^1(\Omega, \mathbb{R}^d)$. We formalize this with the following definitions.

Definition 7.14 A Banach space $V \subset C_0^1(\Omega, \mathbb{R}^d)$ is admissible if it is (canonically) embedded in $C_0^1(\Omega, \mathbb{R}^d)$, i.e., there exists a constant C such that, for all $v \in V$,

$$\|v\|_V \geq C \, \|v\|_{1,\infty} . \tag{7.21}$$

If V is admissible, we denote by $\mathcal{X}_V^1 = L^1([0, 1], V)$ the set of time-dependent vector fields $(v(t), t \in [0, 1])$ such that, for each t, $v(t) \in V$ and

$$\|v\|_{\mathcal{X}_V^1} := \int_0^1 \|v(t)\|_V \, dt < \infty.$$

If the interval $[0, 1]$ is replaced by $[0, T]$, we will use the notation $\mathcal{X}_V^1(T)$ and $\|v\|_{\mathcal{X}_V^1, T}$.

Definition 7.15 If $V \subset C_0^1(\Omega, \mathbb{R}^d)$ is admissible, denote by

$$\mathrm{Diff}_V = \left\{ \varphi_{01}^v, v \in \mathcal{X}_V^1 \right\}$$

the set of diffeomorphisms provided by flows associated to elements $v \in \mathcal{X}_V^1$ at time 1.

Theorem 7.16 Diff_V *is a subgroup of* $\mathrm{Diff}_0^{1,\infty}(\Omega)$.

Proof The inclusion is obvious from Theorem 7.11, so we focus on the subgroup property. The identity function belongs to Diff_V: it corresponds, for example, to φ_{01}^v when $v = 0$. If $\psi = \varphi_{01}^v$ and $\psi' = \varphi_{01}^{v'}$, with $v, v' \in \mathcal{X}_V^1$, then $\psi' \circ \psi = \varphi_{01}^w$ with $w(t) = v(2t)$ for $t \in [0, 1/2]$ and $w(t) = v'(2t - 1)$ for $t \in (1/2, 1]$ (the details are left to the reader) and w belongs to $\mathcal{X}_V^1$. Similarly, if $\psi = \varphi_{01}^v$, then $\psi^{-1} = \psi_{01}^w$ with $w(t) = -v(1 - t)$. Indeed, we have

$$\varphi_{0,1-t}^w(y) = y - \int_0^{1-t} v(1 - s) \circ \varphi_{0s}^w(y) ds = y + \int_1^t v(s) \circ \varphi_{0,1-s}^w ds,$$

which implies (by the uniqueness theorem) that $\varphi_{0,1-t}^w(y) = \varphi_{1t}^v(y)$ and in particular $\varphi_{01}^w = \varphi_{10}^v$. This proves that Diff_V is a subgroup of $\mathrm{Diff}_0^{1,\infty}$. □

Thus, by selecting a certain Banach space V, we can in turn specify a group of diffeomorphisms. In particular, elements in Diff_V inherit the smoothness properties of elements of V. Theorems 7.10 and 7.11 indeed imply that Diff_V is a subgroup of $\mathrm{Diff}_0^{p,\infty}$ as soon as V is embedded in $C_0^p(\Omega, \mathbb{R}^d)$.

We can interpret this framework in terms of a control system, in which vector fields in V control diffeomorphisms through the flow equation $\partial_t \varphi(t) = v(t) \circ \varphi(t)$. The set Diff_V can then be seen as the group of *attainable diffeomorphisms*, starting from the identity and using control trajectories with finite cost $\|v\|_{\mathcal{X}_V^1}$.

7.2.6 A Distance on G_V

Let V be an admissible Banach space. For ψ and ψ' in Diff_V, we let

$$d_V(\psi, \psi') = \inf_{v \in \mathcal{X}_V^1(\Omega)} \left\{ \|v\|_{\mathcal{X}_V^1}, \psi' = \varphi_{01}^v \circ \psi \right\}. \tag{7.22}$$

We have the following theorem:

Theorem 7.17 (Trouvé) *The function d_V is a distance on* Diff_V*, and* (Diff_V, d_V) *is a complete metric space.*

Recall that d_V is a distance if it is symmetric, satisfies the triangle inequality $d_V(\psi, \psi') \leq d_V(\psi, \psi'') + d_V(\psi'', \psi')$ and is such that $d_V(\psi, \psi') = 0$ if and only if $\psi = \psi'$.

Proof Note that the set over which the infimum is computed is not empty: if $\psi, \psi' \in \mathrm{Diff}_V$, then $\psi' \circ \psi^{-1} \in \mathrm{Diff}_V$ (because Diff_V is a group) and therefore can be written in the form φ_{01}^v for some $v \in \mathcal{X}_V^1$.

Let us start with the symmetry: fix $\varepsilon > 0$ and v such that $\|v\|_{\mathcal{X}_V^1} \leq d(\psi, \psi') + \varepsilon$ and $\psi' = \varphi_{01}^v \circ \psi$. This implies that $\psi = \varphi_{10}^v \circ \psi'$, but we know (from the proof of Theorem 7.16) that $\varphi_{10}^v = \varphi_{01}^w$ with $w(t) = -v(1-t)$. Because $\|w\|_{\mathcal{X}_V^1} = \|v\|_{\mathcal{X}_V^1}$, we have, from the definition of d_V:

$$d_V(\psi', \psi) \leq \|w\|_{\mathcal{X}_V^1} \leq d_V(\psi, \psi') + \varepsilon$$

and since this is true for every ε, we have $d_V(\psi', \psi) \leq d_V(\psi, \psi')$. Inverting the roles of ψ and ψ' yields $d_V(\psi', \psi) = d_V(\psi, \psi')$.

For the triangle inequality, let v and v' be such that $\|v\|_{\mathcal{X}_V^1} \leq d(\psi, \psi'') + \varepsilon$, $\|v'\|_{\mathcal{X}_V^1} \leq d(\psi'', \psi') + \varepsilon$, $\psi'' = \varphi_{01}^v \circ \psi$ and $\psi' = \varphi_{01}^{v'} \circ \psi''$. We thus have $\psi' = \varphi_{01}^{v'} \circ \varphi_{01}^v \circ \psi$ and we know, still from the proof of Theorem 7.16, that $\varphi_{01}^v \circ \varphi_{01}^{v'} = \varphi_{01}^w$ with $w(t) = v'(2t)$ for $t \in [0, 1/2]$ and $w(t) = v(2t-1)$ for $t \in (1/2, 1]$. But, in this case, $\|w\|_{\mathcal{X}_V^1} = \|v\|_{\mathcal{X}_V^1} + \|v'\|_{\mathcal{X}_V^1}$ so that

$$d(\psi, \psi') \leq \|w\|_{\mathcal{X}_V^1} \leq d(\psi, \psi'') + d(\psi'', \psi') + 2\varepsilon,$$

which implies the triangle inequality, since this is true for every $\varepsilon > 0$.

We obviously have $d(\psi, \psi) = 0$ since $\varphi_{01}^0 = \mathrm{id}$. Assume that $d(\psi, \psi') = 0$. This implies that there exists a sequence v_n such that $\|v_n\|_{\mathcal{X}_V^1} \to 0$ and $\psi' \circ \psi^{-1} = \varphi_{01}^{v_n}$. The continuity of $v \mapsto \varphi_{st}^v$ implies that $\varphi_{01}^{v_n} \to \varphi_{01}^0 = \mathrm{id}$ so that $\psi = \psi'$.

Let us now check that we do indeed have a complete metric space. Let ψ^n be a Cauchy sequence for d_V, so that, for any $\varepsilon > 0$, there exists an n_0 such that, for any $n \geq n_0$, $d_V(\varphi^n, \psi^{n_0}) \leq \varepsilon$. Taking recursively $\varepsilon = 2^{-n}$, it is possible to extract a subsequence ψ^{n_k} of ψ^n such that

$$\sum_{k=0}^{\infty} d_V(\psi^{n_k}, \psi^{n_{k+1}}) < \infty.$$

Since a Cauchy sequence converges whenever one of its subsequences does, it is sufficient to show that ψ^{n_k} has a limit.

From the definition of d_V, there exists, for every $k \geq 0$, an element v^k in $\mathcal{X}_V^1$ such that $\psi^{n_{k+1}} = \varphi_{01}^{v^k} \circ \psi^{n_k}$ and

$$\left\| v^k \right\|_{\mathcal{X}_V^1} \leq d_V(\psi^{n_k}, \psi^{n_{k+1}}) + 2^{-k-1}.$$

Let us define a time-dependent vector field v by $v(t) = 2v^0(2t)$ for $t \in [0, 1/2[$, $v(t) = 4v^1(4t - 2)$ for $t \in [1/2, 3/4[$, and so on: to define the general term, introduce the dyadic sequence of times $t_0 = 0$ and $t_{k+1} = t_k + 2^{-k-1}$ and let

$$v(t) = 2^{k+1} v^k(2^{k+1}(t - t_k))$$

for $t \in [t_k, t_{k+1}[$. Since t_k tends to 1 as $t \to \infty$, this defines $v(t)$ on $[0, 1)$, and we fix $v(1) = 0$. We have

$$\begin{aligned} \|v(t)\|_{\mathcal{X}_V^1} &= \sum_{k=0}^{\infty} 2^{k+1} \int_{t_k}^{t_{k+1}} \left\| v^k(2^{k+1}(t - t_k)) \right\|_V dt \\ &= \sum_{k=0}^{\infty} \int_0^1 \left\| v^k(t) \right\|_V dt \\ &\leq 1 + \sum_{k=0}^{\infty} d_V(\psi^{n_k}, \psi^{n_{k+1}}), \end{aligned}$$

so that $v \in \mathcal{X}_V^1$. Now, consider the associated flow φ_{0t}^v: it is obtained by first integrating $2v^0(2t)$ between $[0, 1/2)$, which yields $\varphi_{0,1/2}^v = \varphi_{01}^{v^0}$. Iterating this, we have

$$\varphi_{0t_{k+1}}^v = \varphi_{01}^{v^k} \circ \cdots \circ \varphi_{01}^{v^0},$$

so that

$$\psi^{n_{k+1}} = \varphi^v_{0t_{k+1}} \circ \psi^{n_0}.$$

Let $\psi^\infty = \varphi^v_{01} \circ \psi^{n_0}$. We also have $\psi^\infty = \varphi^v_{t_k 1} \circ \psi^{n_k}$. Because $\varphi^v_{t_k 1} = \varphi^{w^k}_{01}$ with $w^k(t) = v((t - t_k)/(1 - t_k))/(1 - t_k)$, and

$$\|w\|_{\mathcal{X}^1_V} = \int_{t_k}^1 \|v(t)\|_V \, dt,$$

we obtain the fact that $d_V(\psi^{n_k}, \psi^\infty) \to 0$, which completes the proof of Theorem 7.17. □

7.2.7 *Properties of the Distance*

We first introduce the set of square integrable (in time) time-dependent vector fields:

Definition 7.18 Let V be an admissible Banach space of vector fields $v : \Omega \to \mathbb{R}^d$. We define $\mathcal{X}^2_V$ as the set of time-dependent vector fields $v = (v(t), t \in [0, 1])$ such that, for each t, $v_t \in V$ and

$$\int_0^1 \|v(t)\|^2_V \, dt < \infty.$$

We state without proof the important result (in which one identifies time-dependent vector fields that coincide for almost all t):

Proposition 7.19 *$\mathcal{X}^2_V$ is a Banach space with norm*

$$\|v\|_{\mathcal{X}^2_V} = \left(\int_0^1 \|v_t\|^2_V \, dt \right)^{1/2}.$$

Moreover, if V is a Hilbert space, then $\mathcal{X}^2_V$ is also a Hilbert space with

$$\langle v \, , \, w \rangle_{\mathcal{X}^2_V} = \int_0^1 \langle v_t \, , \, w_t \rangle_V dt.$$

Because $\left(\int_0^1 \|v(t)\|_V \, dt \right)^2 \leq \int_0^1 \|v(t)\|^2_V \, dt$, we have $\mathcal{X}^2_V \subset \mathcal{X}^1_V$ and if $v \in \mathcal{X}^2_V$, $\|v\|_{\mathcal{X}^1_V} \leq \|v\|_{\mathcal{X}^2_V}$. The computation of d_V can be reduced to a minimization over $\mathcal{X}^2_V$ by the following theorem.

Theorem 7.20 *If V is admissible and $\psi, \psi' \in G_V$, we have*

$$d_V(\psi, \psi') = \inf_{v \in \mathcal{X}^2_V} \left\{ \|v\|_{\mathcal{X}^2_V} \, , \, \psi' = \varphi^v_{01} \circ \psi \right\}. \tag{7.23}$$

Proof Let $\delta_V(\psi, \psi')$ be given by the right-hand side of (7.23) and d_V be given by (7.22). Because d_V is the infimum over a larger set than δ_V, and minimizes a quantity which is always smaller, we have $d_V(\psi, \psi') \leq \delta_V(\psi, \psi')$ and we just need to prove the reverse inequality. For this, consider $v \in \mathcal{X}_V^1$ such that $\psi' = \varphi_{01}^v \circ \psi$. It suffices to prove that, for any $\varepsilon > 0$, there exists a $w \in \mathcal{X}_V^2$ such that $\psi' = \varphi_{01}^w \circ \psi$ and $\|w\|_{\mathcal{X}_V^2} \leq \|v\|_{\mathcal{X}_V^1} + \varepsilon$. Let α be an absolutely continuous increasing function from $[0, 1]$ onto $[0, 1]$, which implies $\alpha(0) = 0$ and $\alpha(1) = 1$, and

$$\alpha(t) = \int_0^t \dot{\alpha}(s)\, dt$$

for all $t \in [0, 1]$. Applying the change of variable formula ([246], sect. 26), we have

$$\begin{aligned} \varphi_{0\alpha(t)}^v(x) &= x + \int_0^{\alpha(t)} v(u, \varphi_{0u}(x))\, du \\ &= x + \int_0^t \dot{\alpha}(s)\, v(\alpha(s), \varphi_{0\alpha(s)}(x))\, ds, \end{aligned}$$

so that the flow generated by $w = \dot{\alpha}(s)v(\alpha(s))$ is $\varphi_{0\alpha(t)}^v$, which coincides with φ_{01}^v at $t = 1$. We have

$$\|w\|_{\mathcal{X}_V^1} = \int_0^1 \dot{\alpha}(s)\, \|v(\alpha(s))\|_V\, ds = \int_0^1 \|v(t)\|_V\, dt = \|v\|_{\mathcal{X}_V^1},$$

so that this time change does not affect the minimization in (7.22). However, we have, denoting by $\beta(t)$ the inverse of $\alpha(t)$,

$$\|w\|_{\mathcal{X}_V^2}^2 = \int_0^1 \dot{\alpha}(s)^2\, \|v(\alpha(s))\|_V^2\, ds = \int_0^1 \dot{\alpha} \circ \beta(s)\, \|v(s)\|_V^2\, ds,$$

so that this transformation can be used to reduce $\|v\|_{\mathcal{X}_V^2}$. Define, for some $\eta > 0$, $\dot{\alpha} \circ \beta(t) = c/(\eta + \|v(t)\|_V)$, which yields $\dot{\beta}(t) = (\eta + \|v(t)\|_V)/c$ and $c = \eta + \|v\|_{\mathcal{X}_V^1}$. This gives

$$\|w\|_{\mathcal{X}_V^2}^2 = c \int_0^1 \frac{\|v(t)\|_V^2}{\eta + \|v(t)\|_V} dt \leq c\, \|v\|_{\mathcal{X}_V^1} = (\|v\|_{\mathcal{X}_V^1} + \eta)\, \|v\|_{\mathcal{X}_V^1}.$$

By choosing η small enough, we can always arrange that $\|w\|_{\mathcal{X}_V^2} \leq \|v\|_{\mathcal{X}_V^1} + \varepsilon$, which is what we wanted to prove. □

A consequence of this result is the following fact.

Corollary 7.21 *If the infimum in (7.23) is attained at some $v \in \mathcal{X}_V^2$, then $t \mapsto \|v(t)\|_V$ is constant.*

Proof Indeed, let v achieve the minimum in (7.23): we have

$$d_V(\psi, \psi') = \|v\|_{\mathcal{X}_V^2} \geq \|v\|_{\mathcal{X}_V^1},$$

but $\|v\|_{\mathcal{X}_V^1} \geq d_V(\psi, \psi')$ by definition. Thus, we must have $\|v\|_{\mathcal{X}_V^2} = \|v\|_{\mathcal{X}_V^1}$. This corresponds to the equality case in Schwartz's inequality, which can only be achieved by (almost everywhere) constant functions. □

Corollary 7.21 is usefully completed by the following theorem:

Theorem 7.22 *If V is Hilbert and admissible, and $\psi, \psi' \in \mathrm{Diff}_V$, there exists a $v \in \mathcal{X}_V^2$ such that*

$$d_V(\psi, \psi') = \|v\|_{\mathcal{X}_V^2}$$

and $\psi' = \varphi_{01}^v \circ \psi$.

Proof By Proposition 7.19, $\mathcal{X}_V^2$ is a Hilbert space. Let us fix a minimizing sequence for $d_V(\psi, \psi')$, i.e., a sequence $v^n \in \mathcal{X}_V^2$ such that $\|v^n\|_{\mathcal{X}_V^2} \to d_V(\psi, \psi')$ and $\psi' = \varphi_{01}^{v^n} \circ \psi$. This implies that $(\|v^n\|_{\mathcal{X}_V^2})$ is bounded and by Theorem A.20, one can extract a subsequence of v^n (which we still denote by v^n) that weakly converges to some $v \in \mathcal{X}_V^2$, such that

$$\|v\|_{\mathcal{X}_V^2} \leq \liminf \left\|v^n\right\|_{\mathcal{X}_V^2} = d_V(\psi, \psi').$$

We now apply Theorem 7.13 (using the fact that weak convergence in $\mathcal{X}_V^2$ implies weak convergence in the bigger space $\mathcal{X}^{1,1}(\Omega)$) and obtain the fact that φ^{v^n} converges to φ^v so that $\psi' = \varphi_{01}^v \circ \psi$ remains true: this proves Theorem 7.22. □

Remark 7.23 This theorem is true for spaces V that are more general than Hilbert spaces, such as reflexive Banach spaces, for which $L^2([0, 1], V) = L^2([0, 1], V^*)^*$, which allows one to apply the same compactness argument (thanks to the Banach–Alaoglu theorem). We refer to [83] for more details on L^p spaces of vector-valued functions. We will, however, primarily work with admissible Hilbert spaces in the rest of the book.

Chapter 8
Building Admissible Spaces

In the previous chapter we defined a family of admissible spaces V that induce groups of diffeomorphisms using flows associated to ordinary differential equations. We now show how such spaces can be explicitly constructed, focusing on Hilbert spaces. This construction is fundamental, because it is intimately related to computational methods. We will in particular introduce the notion of *reproducing kernels* associated to an admissible space, which will provide our main computational tool. We introduce this in the next section.

8.1 Reproducing Kernel Hilbert Spaces

8.1.1 The Scalar Case

Although we build diffeomorphisms from Hilbert spaces of vector fields, it will be easier to introduce reproducing kernel Hilbert spaces for scalar-valued functions, which has its own interest anyway [15, 16, 20, 96, 299].

Let $\Omega \subset \mathbb{R}^d$. Consider a Hilbert space V included in $L^2(\Omega, \mathbb{R})$. We assume that elements of V are smooth enough, and require the inclusion and the canonical embedding of V in $C^0(\Omega, \mathbb{R})$. For example, it suffices (from Morrey's theorem, see Theorem A.16) that $V \subset H^m(\Omega, \mathbb{R})$ with $m > d/2$. (Here the inclusion is assumed to be continuous, and H^m is the Sobolev space of functions with square integrable derivatives up to order m.) This assumption implies that there exists a constant C such that, for all $v \in V$,

$$\|v\|_\infty \leq C \|v\|_V .$$

We make another assumption on V.

Assumption 8.1 We assume that a relation of the kind $\sum_{i=1}^N \alpha_i v(x_i) = 0$ cannot be true for every $v \in V$ unless $\alpha_1 = \cdots = \alpha_N = 0$, $(x_1, \ldots, x_N)$ being an arbitrary family of distinct points in Ω.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_8

This is true, for example, if V contains functions supported on arbitrary compact sets.

Each x in Ω specifies a linear form δ_x defined by $(\delta_x \mid v) = v(x)$ for $x \in V$. We have

$$|(\delta_x \mid v)| \leq \|v\|_\infty \leq C \|v\|_V$$

so that $\delta_x \in V^*$. Let $\mathbb{K}$ denote the inverse duality operator of V, associated with Riesz's theorem (Theorem A.12), so that, for every $v \in V$,

$$v(x) = (\delta_x \mid v) = \langle \mathbb{K}\delta_x \,,\, v \rangle_V. \tag{8.1}$$

Since it belongs to V, $\mathbb{K}\delta_x$ is a continuous function $y \mapsto (\mathbb{K}\delta_x)(y)$. This defines a function of two variables, denoted $K : \Omega \times \Omega \to \mathbb{R}$, by $K(y, x) = (\mathbb{K}\delta_x)(y)$, i.e., $K(\cdot, x) = \mathbb{K}\delta_x$.

This function K has several interesting properties. First, applying Eq. (8.1) to $v = K(\cdot, y)$ yields

$$K(x, y) = \langle K(\cdot, x) \,,\, K(\cdot, y) \rangle_V.$$

Since the last term is symmetric, we have $K(x, y) = K(y, x)$, and because of the obtained identity, K is called the *reproducing kernel* of V.

A second property is the fact that K is positive definite, in the sense that, for any family $x_1, \ldots, x_N \in V$ and any sequence $\alpha_1, \ldots, \alpha_N$ in $\mathbb{R}$, the double sum

$$\sum_{i,j=1}^{N} \alpha_i \alpha_j K(x_i, x_j)$$

is non-negative, and vanishes if and only if all α_i equal 0. Indeed, by the reproducing property, this sum may be written $\left\| \sum_{i=1}^N \alpha_i K(\cdot, x_i) \right\|_V^2$ and this is non-negative. If it vanishes, then $\sum_{i=1}^N \alpha_i K(\cdot, x_i) = 0$, which implies, by Eq. (8.1), that, for every $v \in V$, one has $\sum_{i=1}^N \alpha_i v(x_i) = 0$, and our assumption on V implies that $\alpha_1 = \cdots = \alpha_N = 0$.

Scalar Spline Interpolation

As a first (and important) example of application of kernels, we discuss the following interpolation problem [299].

($\mathcal{S}_V$) *Fix a family of distinct points $x_1, \ldots, x_N$ in Ω. Find a function $v \in V$ of minimal norm satisfying the constraints $v(x_i) = \lambda_i$, where $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ are prescribed values.*

To solve this problem, define V_0 to be the set of v's for which the constraints vanish:

$$V_0 = \{v \in V : v(x_i) = 0, i = 1, \ldots, N\}.$$

Using the kernel K, we may write

$$V_0 = \left\{ v \in V : \left\langle K(., x_i)\,,\ v \right\rangle_V = 0, i = 1, \ldots, N \right\}$$

so that

$$V_0 = \text{span}\,\{K(., x_1), \ldots, K(., x_N)\}^{\perp},$$

the orthogonal being taken with respect to the V inner product. We have the following first result.

Lemma 8.1 *If there exists a solution $\hat{v}$ of problem $\mathcal{S}_V$, then $\hat{v} \in V_0^{\perp} = \text{span}\,\{K(\cdot, x_1), \ldots, K(\cdot, x_N)\}$. Moreover, if $\hat{v} \in V_0^{\perp}$ is a solution of $\mathcal{S}_{V_0^{\perp}}$, then it is a solution of $\mathcal{S}_V$.*

Proof Let $\hat{v}$ be a solution of $\mathcal{S}_V$, and let v^* be its orthogonal projection on $\text{span}\,\{K(\cdot, x_1), \ldots, K(\cdot, x_N)\}$. From the properties of orthogonal projections, we have $\hat{v} - v^* \in V_0$, which implies, by the definition of V_0, that $\hat{v}(x_i) = v^*(x_i)$ for $i = 1, \ldots, N$. But, since $\|v^*\|_V \leq \left\|\hat{v}\right\|_V$ (by the variational characterization of the projection), and $\left\|\hat{v}\right\|_V \leq \|v^*\|_V$ by assumption, both norms are equal, which is only possible when $\hat{v} = v^*$. Therefore, $\hat{v} \in V_0^{\perp}$ and the proof of the first assertion is complete.

Now, if $\hat{v}$ is a solution of $\mathcal{S}_{V_0^{\perp}}$ and v is any function in V which satisfies the constraints, then $v - \hat{v} \in V_0$ and $\|v\|_V^2 = \left\|\hat{v}\right\|_V^2 + \left\|v - \hat{v}\right\|_V^2 \geq \left\|\hat{v}\right\|_V^2$, which shows that $\hat{v}$ is a solution of $\mathcal{S}_V$. □

This lemma allows us to restrict the search for a solution of $\mathcal{S}_V$ to the set of linear combinations of $K(\cdot, x_1), \ldots, K(\cdot, x_N)$, which places us in a convenient finite-dimensional situation. We look for $\hat{v}$ in the form

$$\hat{v}(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i)$$

and we introduce the $N \times N$ matrix S with coefficients $s_{ij} = K(x_i, x_j)$. The whole problem may now be reformulated as a function of the vector $\alpha = (\alpha_1, \ldots, \alpha_N)^T$ (a column vector) and of the matrix S. Indeed, by the reproducing property of K, we have

$$\left\|\hat{v}\right\|_V^2 = \sum_{i=1}^{N} \alpha_i \alpha_j K(x_i, x_j) = \alpha^T S \alpha \tag{8.2}$$

and each constraint may be written as $\lambda_i = v(x_i) = \sum_{j=1}^{N} \alpha_j K(x_i, x_j)$, so that, letting $\lambda = (\lambda_1, \ldots, \lambda_N)^T$, the whole system of constraints may be expressed as $S\alpha = \lambda$.

Our hypotheses imply that S is invertible; indeed, if $S\alpha = 0$, then $\alpha^T S \alpha = 0$ which, by Eq. (8.2) and the positive definiteness of K, is only possible when $\alpha = 0$ (we assume that the x_i's are distinct). Therefore, there is only one $\hat{v}$ in $V_0^{\perp}$ which satisfies the constraints, and it corresponds to $\alpha = S^{-1}\lambda$. These results are summarized in the next theorem.

Theorem 8.2 *Problem $\mathcal{S}_V$ has a unique solution in V, given by*

$$\hat{v}(x) = \sum_{i=1}^{N} K(x, x_i)\alpha_i$$

with

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} = \begin{pmatrix} K(x_1, x_1) & \dots & K(x_1, x_N) \\ \vdots & \vdots & \vdots \\ K(x_N, x_1) & \dots & K(x_N, x_N) \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}.$$

Another important variant of the same problem comes when the "hard" constraints $v(x_i) = \lambda_i$ are replaced by "soft" constraints, in the form of a penalty function added to the minimized norm. This may be expressed as the minimization of a function of the form

$$E(v) = \|v\|_V^2 + C \sum_{i=1}^{N} \varphi(|v(x_i) - y_i|)$$

for some increasing, convex function on $[0, +\infty)$ and $C > 0$. Since the second term of E does not depend on the projection of v on V_0, Lemma 8.1 remains valid, again reducing the problem to finding v of the form

$$v(x) = \sum_{i=1}^{N} K(x, x_i)\alpha_i$$

for which

$$E(v) = \sum_{i,j=1}^{N} \alpha_i \alpha_j K(x_i, x_j) + C \sum_{i=1}^{N} \varphi\left(\left| \sum_{j=1}^{N} K(x_i, x_j)\alpha_j - \lambda_i \right| \right).$$

Assume, to simplify, that φ is differentiable and $\dot{\varphi}(0) = 0$. We have, letting $\psi(x) = \operatorname{sign}(x)\dot{\varphi}(x)$,

$$\partial_{\alpha_j} E = 2 \sum_{i=1}^{N} \alpha_i K(x_i, x_j) + C \sum_{i=1}^{N} K(x_i, x_j)\psi\left(\left| \sum_{l=1}^{N} K(x_i, x_l)\alpha_l - \lambda_i \right| \right).$$

Assuming, still, that the x_i are distinct, we can apply S^{-1} to the system $\partial_{\alpha_j} E = 0$, $j = 1, \dots, N$, which characterizes the minimum, yielding

$$2\alpha_i + C\psi\left(\left| \sum_{j=1}^{N} K(x_i, x_j)\alpha_j - \lambda_i \right| \right) = 0. \tag{8.3}$$

One can use convex optimization methods to minimize E [40, 221]. The particular case of $\varphi(x) = x^2$ is much simpler to solve, since in this case $\psi(x) = 2x$ and Eq. (8.3) becomes

$$2\alpha_i + 2C\left(\sum_{j=1}^{N} K(x_i, x_j)\alpha_j - \lambda_i\right) = 0.$$

The solution of this equation is $\alpha = (S + \mathrm{Id}/C)^{-1}\lambda$, yielding a result very similar to Theorem 8.2:

Theorem 8.3 *The unique minimum over V of*

$$\|v\|_V^2 + C\sum_{i=1}^{N} |v(x_i) - \lambda_i|^2$$

is attained at

$$\hat{v}(x) = \sum_{i=1}^{N} K(x, x_i)\alpha_i$$

with

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} = (S + \mathrm{Id}/C)^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}$$

and

$$S = \begin{pmatrix} K(x_1, x_1) & \dots & K(x_1, x_N) \\ \vdots & \vdots & \vdots \\ K(x_N, x_1) & \dots & K(x_N, x_N) \end{pmatrix}.$$

To conclude this section, we prove that V is the Hilbert space generated by the kernel functions. Recall that the Hilbert space generated by a family of vectors is the closure of the vector space formed by all finite linear combinations of these vectors.

Theorem 8.4 *If $V \subset L^2(\Omega, \mathbb{R})$ is an RKHS with kernel K, then*

$$V = \mathrm{Hilb}\{K(\cdot, x), x \in \Omega\}.$$

Proof Let $v \in V$ and

$$W = \mathrm{Hilb}\{K(\cdot, x), x \in \Omega\}.$$

We have $W \subset V$ since the latter is a Hilbert space containing all $K(\cdot, x)$. Let w be the orthogonal projection of v on W. Then, for all $x \in \Omega$, $w(x) = \langle w\,,\ K(\cdot, x)\rangle_V = \langle v\,,\ K(\cdot, x)\rangle_V = v(x)$ by definition of the kernel and of the orthogonal projection. This proves that $v = w \in W$ and therefore $W = V$. □

8.1.2 The Vector Case

In the previous section, elements of V were functions from $\Omega \subset \mathbb{R}^d$ to $\mathbb{R}$. When working with deformations, which is our goal here, functions of interest describe displacements of points in Ω and therefore must be vector-valued. This leads us to address the problem of spline approximation for vector fields in Ω, which, as will be seen, is handled quite similarly to the scalar case.

So, in this section, V is a Hilbert space, canonically embedded in $L^2(\Omega, \mathbb{R}^k)$ and in $C^0(\Omega, \mathbb{R}^k)$. We will mostly be interested in the case $k = d$, i.e., V contains vector fields on Ω. In this case, the assumption is obviously true if V is admissible. Fixing $x \in \Omega$, the evaluation function $v \mapsto v(x)$ is a continuous linear map from V to $\mathbb{R}^k$. This implies that, for any $a \in \mathbb{R}^k$, the function $v \mapsto a^T v(x)$ is a continuous linear functional on V. We will denote this linear form by $a\delta_x$, so that

$$(a\delta_x \mid v) = a^T v(x). \tag{8.4}$$

Let, as before, $\mathbb{K}$ denote the inverse duality operator of V, so that, for any $v \in V$

$$\left\langle \mathbb{K}(a\delta_x)\,,\, v \right\rangle_V = a^T v(x). \tag{8.5}$$

The map $a \mapsto \mathbb{K}(a\delta_x)$ is linear from $\mathbb{R}^k$ to V. (This is because $a \mapsto a^T v(x)$ is linear and because of the uniqueness of the Riesz representation.) Therefore, for $y \in \Omega$, the map $a \mapsto \mathbb{K}(a\delta_x)(y)$ is linear from $\mathbb{R}^k$ to $\mathbb{R}^k$. This implies that there exists a function taking values in the set of k by k matrices, that we will denote by $(y, x) \mapsto K(y, x)$, such that, for $a \in \mathbb{R}^k$, $x, y \in \Omega$, $\mathbb{K}(a\delta_x)(y) = K(y, x)a$.

The kernel K here being matrix-valued, the reproducing property is

$$\left\langle K(\cdot, x)a\,,\, K(\cdot, y)b \right\rangle_V = a^T K(x, y)b.$$

From the symmetry of the first term, we obtain the fact that, for all $a, b \in \mathbb{R}^k$, $a^T K(x, y)b = b^T K(y, x)a$, which implies that $K(y, x) = K(x, y)^T$.

To ensure the positivity of K, we make an assumption similar to the scalar case:

Assumption 8.2 If $x_1, \ldots, x_N \in \Omega$ and $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^k$ are such that, for all $v \in V$, $\alpha_1^T v(x_1) + \cdots + \alpha_N^T v(x_N) = 0$, then $\alpha_1 = \cdots = \alpha_N = 0$.

Under this assumption, it is easy to prove that, for all $\alpha_1, \ldots, \alpha_N \in \mathbb{R}^k$,

$$\sum_{i,j=1}^{N} \alpha_i^T K(x_i, x_j)\alpha_j \geq 0,$$

with equality if and only if all α_i vanish.

The generalization of Theorem 8.4 is straightforward.

Theorem 8.5 *If $V \subset L^2(\Omega, \mathbb{R}^k)$ is an RKHS with kernel K, then*

$$V = \text{Hilb}\left\{K(\cdot, x)a, x \in \Omega, a \in \mathbb{R}^k\right\}.$$

Vector Spline Interpolation

The interpolation problem in the vector case is

($\mathcal{S}_V$) *Given $x_1, \ldots, x_N$ in Ω, $\lambda_1, \ldots, \lambda_N$ in $\mathbb{R}^k$, find v in V, with minimum norm, such that $v(x_i) = \lambda_i$.*

As before, we let

$$V_0 = \{v \in V : v(x_i) = 0, i = 1, \ldots, N\}.$$

Then, Lemma 8.1 remains valid (we omit the proof, which duplicates the scalar case):

Lemma 8.6 *If there exists a solution $\hat{v}$ of problem $\mathcal{S}_V$, then $\hat{v} \in V_0^\perp$. Moreover, if $\hat{v} \in V_0^\perp$ is a solution of $\mathcal{S}_{V_0^\perp}$, then it is a solution of $\mathcal{S}_V$.*

The characterization of $V_0^\perp$ is similar to the scalar case:

Lemma 8.7

$$V_0^\perp = \left\{v = \sum_{i=1}^N K(., x_i)\alpha_i, \alpha_1, \ldots, \alpha_N \in \mathbb{R}^k\right\}.$$

Proof It is clear that $w \in V_0$ if and only if, for any $\alpha_1, \ldots, \alpha_N$, one has

$$\sum_{i=1}^N \alpha_i^T\ w(x_i) = 0.$$

Thus $w \in V_0$ if and only if $\langle v\ ,\ w\rangle_V = 0$ for all v of the form $v = \sum_{i=1}^N K(\cdot, x_i)\alpha_i$. Thus

$$V_0 = \left\{v = \sum_{i=1}^N K(\cdot, x_i)\alpha_i, \alpha_1, \ldots, \alpha_N \in \mathbb{R}^k\right\}^\perp$$

and since $\left\{v = \sum_{i=1}^N K(\cdot, x_i)\alpha_i, \alpha_1, \ldots, \alpha_N \in \mathbb{R}^k\right\}$ is finite-dimensional, hence closed, one has

$$V_0^\perp = \left\{v = \sum_{i=1}^N K(\cdot, x_i)\alpha_i, \alpha_1, \ldots, \alpha_N \in \mathbb{R}^k\right\}.$$

□

When $v = \sum_{j=1}^{N} K(\cdot, x_j)\alpha_j \in V_0^{\perp}$, the constraint $v(x_i) = \lambda_i$ yields

$$\sum_{j=1}^{N} K(x_i, x_j)\alpha_j = \lambda_i.$$

Since we also have, in this case,

$$\|v\|_V^2 = \sum_{i,j=1}^{N} \alpha_i^T \; K(x_i, x_j)\alpha_j,$$

the whole problem can be rewritten quite concisely with matrices, introducing the notation

$$S = S(x_1, \ldots, x_N) = \begin{pmatrix} K(x_1, x_1) & \ldots & K(x_1, x_N) \\ \vdots & \vdots & \vdots \\ K(x_N, x_1) & \ldots & K(x_N, x_N) \end{pmatrix}, \tag{8.6}$$

which is now a block matrix of size $Nk \times Nk$,

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix}, \quad \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix},$$

each α_i, λ_i being k-dimensional column vectors. The whole set of constraints now becomes $S\alpha = \lambda$ and $\|v\|_V^2 = \; \alpha^T \; S\alpha$. Thus, replacing numbers by blocks, the problem has exactly the same structure as in the scalar case, and we can repeat the results we have obtained.

Theorem 8.8 (Interpolating splines) *Problem ($\mathcal{S}_V$) has a unique solution in V, given by*

$$\hat{v}(x) = \sum_{i=1}^{N} K(x, x_i)\alpha_i$$

with

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} = \begin{pmatrix} K(x_1, x_1) & \ldots & K(x_1, x_N) \\ \vdots & \vdots & \vdots \\ K(x_N, x_1) & \vdots & K(x_N, x_N) \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}.$$

Theorem 8.9 (Smoothing splines) *The minimum over V of*

$$\|v\|_V^2 + C \sum_{i=1}^{N} |v(x_i) - \lambda_i|^2$$

is attained at

$$\hat{v}(x) = \sum_{i=1}^{N} K(x, x_i)\alpha_i$$

with

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} = (S + \mathrm{Id}/C)^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}$$

and $S = S(x_1, \dots, x_N)$ *given by Eq.* (8.6).

8.1.3 Derivatives of the Kernel

If the RKHS V is continuously embedded in $C^p(\Omega, \mathbb{R}^k)$, then, for any $y \in \Omega$, the function $x \mapsto K(x, y)$ is in $C^p(\Omega, \mathcal{M}_k(\mathbb{R}))$ (because $K(\cdot, y)b \in V$ for all $b \in \mathbb{R}^k$) and so is $y \mapsto K(x, y)$ by symmetry. This proves that, for any $r \leq b$, both $\partial_1^r K$ and $\partial_2^r K$ are well defined. We now go further and prove that all derivatives $\partial_1^q \partial_2^r K$ exist for $q, r \leq p$.

By assumption, the linear form $b_1 \delta_y^{(r)}(a_1, \dots, a_r) : v \mapsto b_1^T d^r v(y)(a_1, \dots, a_r)$ is continuous on V for $r \leq p$, for any $y \in \Omega$, $b_1 \in \mathbb{R}^k$ and $a_1, \dots, a_r \in \mathbb{R}^d$. (Recall that $d^r v(y)$ is a multilinear map from $(\mathbb{R}^d)^r$ to $\mathbb{R}^k$). Taking $v = K(\cdot, x)b_2$ we can write

$$\begin{aligned} b_1^T \partial_1^r (K(y, x)b_2)(a_1, \dots, a_r) &= \left\langle K(\cdot, x)b_2 \,,\, \mathbb{K}(b_1 \delta_y^{(r)}(a_1, \dots, a_r)) \right\rangle_V \\ &= b_2^T \mathbb{K}(b_1 \delta_y^{(r)}(a_1, \dots, a_r))(x). \end{aligned}$$

This allows us to identify $\mathbb{K}(b_1 \delta_y^{(r)}(a_1, \dots, a_r))$, noticing that

$$\begin{aligned} b_1^T \partial_1^r (K(y, x)b_2)(a_1, \dots, a_r) &= \partial_1^r (b_1^T K(y, x)b_2)(a_1, \dots, a_r) \\ &= b_2^T \partial_2^r (K(x, y)b_1)(a_1, \dots, a_r), \end{aligned}$$

which yields

$$\mathbb{K}(b_1 \delta_y^{(r)}(a_1, \dots, a_r))(x) = \partial_2^r (K(x, y)b_1)(a_1, \dots, a_r),$$

proving that the right-hand side belongs to V and is therefore p times differentiable in x. We summarize this in the following proposition.

Proposition 8.10 *If V is a Hilbert space continuously embedded in $C^p(\Omega, \mathbb{R}^k)$, then its reproducing kernel K is such that $\partial_1^q \partial_2^r K$ exists for all $q, r \leq p$. More precisely, for all $b \in \mathbb{R}^k$ and $a_1, \dots, a_r \in \mathbb{R}^d$, the vector field*

$$\partial_2^r(K(\cdot, y)b)(a_1, \ldots, a_r) : x \mapsto \partial_2^r(K(x, y)b)(a_1, \ldots, a_r)$$

belongs to V and satisfies

$$d^r(b^T v)(y)(a_1, \ldots, a_r) = \left\langle \partial_2^r(K(\cdot, y)b)(a_1, \ldots, a_r)\,,\ v \right\rangle_V \tag{8.7}$$

for all $v \in V$.

Applying (8.7) to $v = \partial_2^{r'}(K(\cdot, y')b')(a'_1, \ldots, a'_{r'})$ we get the identity

$$\begin{aligned}\left\langle \partial_2^r(K(\cdot, y)b)(a_1, \ldots, a_r)\,,\ \partial_2^{r'}(K(\cdot, y')b')(a'_1, \ldots, a'_{r'}) \right\rangle_V \\ = \partial_1^r \partial_2^{r'}(b^T K(y, y')b')(a_1, \ldots, a_r, a'_1, \ldots, a'_{r'})\end{aligned} \tag{8.8}$$

for $r, r' \leq p$.

In the following, it will be convenient to use the notation

$$\partial_2^r K(\cdot, y)(b, a_1, \ldots, a_r) = \partial_2^r(K(\cdot, y)b)(a_1, \ldots, a_r). \tag{8.9}$$

8.2 Building V from Operators

One way to define a Hilbert space V of smooth functions or vector fields is to use inner products associated to operators. For example, consider the spaces $H^m(\mathbb{R}^d, \mathbb{R})$, that can be equipped with the norm

$$\|u\|_{m,2}^2 = \sum_{|\alpha| \leq m} \|\partial_\alpha u\|_2^2\,,$$

where $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index, $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and $\partial_\alpha = \partial_{x_1}^{\alpha_1} \ldots \partial_{x_d}^{\alpha_d}$. If u is a smooth (C^∞) and compactly supported function, the partial derivatives can be integrated by parts to write

$$\|u\|_{m,2}^2 = \langle Au\,,\ u \rangle_2$$

with

$$Au = \sum_{|\alpha| \leq m} (-1)^{-|\alpha|} \partial_{2\alpha} u.$$

By construction, one has, for any pair of smooth compactly supported functions,

$$\langle Au\,,\ v \rangle_2 = \langle u\,,\ v \rangle_{m,2} = \langle u\,,\ Av \rangle_2,$$

which expresses the fact that A is a symmetric operator. From this identity, it appears that A coincides with the duality operator $\mathbb{L}$ on $H^m(\Omega, \mathbb{R})$ (modulo the identification of the function Au with the linear form

$$Au\,dx : v \mapsto \int (Au)\, v dx,$$

i.e., modulo the identification of L^2 with its dual).

Obviously, one can build an infinity of operators that provide a Hilbert structure equivalent to the H^m inner product, i.e., such that

$$c_0 \langle u\ ,\ v \rangle_{m,2} \leq (Au \mid v)_2 \leq c_1 \langle u\ ,\ v \rangle_{m,2}$$

for some positive constants c_0 and c_1. Any operator of the form

$$Au = \sum_{|\alpha| \leq m} (-1)^{-|\alpha|} \rho_\alpha \partial_{2\alpha} u$$

with positive ρ_α's will work. Actually, most of these coefficients can vanish and still provide an equivalent norm. This can be shown by considering the operator in the Fourier domain.

Indeed, by Proposition A.22, the Fourier transform of Au is $\mathcal{F}(Au)(\xi) = P_A(\xi)\hat{u}(\xi)$ with

$$P_A(\xi) = \sum_{|\alpha| \leq m} \rho_\alpha (2\pi)^{|\alpha|} \xi^{2\alpha}$$

and $\xi^{2\alpha} = \xi_1^{2\alpha_1} \ldots \xi_d^{2\alpha_d}$.

Using the isometry property of the Fourier transform, one has

$$\langle Au\ ,\ u \rangle_2 = \langle P_A \hat{u}\ ,\ \hat{u} \rangle_2$$

and we see that, as soon as there exist two constants $0 < c_0 < c_1$ such that

$$c_0 \leq \frac{P_A(\xi)}{\sum_{|\alpha| \leq m} (2\pi)^{|\alpha|} \xi^{2\alpha}} \leq c_1,$$

the norm associated to A will be equivalent to the H^m norm. It is not hard to show, in addition, that this happens if and only if the constant coefficient ρ_0 and the leading coefficients ρ_α with $\max_j \alpha_j = m$ are all positive. For example, the polynomial

$$P_A(\xi) = 1 + a(2\pi)^m \sum_{j=1}^{d} \xi_j^{2m}$$

corresponds to the operator

$$Au = u + (-1)^m a \sum_{j=1}^{d} \partial_{x_j}^{2m} u$$

with norm

$$\|u\|_A^2 = \|u\|_2^2 + a \sum_{j=1}^{d} \|\partial_{x_j}^m u\|^2,$$

which is equivalent to the H^m norm. Another equivalent norm is provided by taking

$$P_A(\xi) = \left(1 + 2\pi a|\xi|^2\right)^m, \tag{8.10}$$

which corresponds to the operator $A = (\mathrm{Id} - a\Delta)^m$, where Δ is the Laplacian. This operator has important invariance properties.

More generally, we will say that V is a Hilbert space associated to an operator A if A can be defined on a domain $D(A) \subset V$, with $D(A)$ dense in V and

$$\langle Au\ ,\ v\rangle_2 = \langle u\ ,\ v\rangle_V$$

for $u, v \in D(A)$. In such a case, A coincides with the restriction of the duality operator of V to $D(A)$.

It is also interesting to consider the situation in which one starts from the operator A and its domain $D(A)$ to build the associated Hilbert space. This is called the Friedrichs extension of an operator [315]. Since the construction is not restricted to subspaces of $L^2(\Omega, \mathbb{R})$, we make the following presentation with an arbitrary Hilbert space H.

To start, we need a subspace D, included in H and dense in this space, and an operator (i.e., a linear functional), $L : D \to H$. Our typical application will be with $D = C_c^\infty(\Omega, \mathbb{R}^d)$ (the set of C^∞ functions with compact support in Ω) and $H = L^2(\Omega, \mathbb{R}^d)$. In such a case, L may be chosen as a differential operator of any degree, since derivatives of C^∞ functions with compact support obviously belong to L^2. However, L will be assumed to satisfy an additional monotonicity constraint:

Assumption 8.3 The operator L is assumed to be symmetric and strongly monotonic on D, which means that there exists a constant $c > 0$ such that, for all $u \in D$,

$$\langle u\ ,\ Lu\rangle_H \geq c\langle u\ ,\ u\rangle_H \tag{8.11}$$

and for all $u, v \in D$

$$\langle u\ ,\ Lv\rangle_H = \langle Lu\ ,\ v\rangle_H. \tag{8.12}$$

An example of strongly monotonic operator on $C_c^\infty(\Omega, \mathbb{R})$ is given by $Lu = -\Delta u + \lambda u$, where Δ is the Laplacian: $\Delta u = \sum_{i=1}^k \partial_{x_i}^2 u$. Indeed, in this case, and when u has compact support, an integration by parts yields

$$-\int_\Omega \Delta u(x)u(x)dx = \sum_{i=1}^{k} \int_\Omega \left(\partial_{x_i} u\right)^2 dx \geq 0,$$

so that $\langle u\ ,\ Lu\rangle_H \geq \lambda \langle u\ ,\ u\rangle_H$.

Returning to the general case, the operator L induces an inner product on D, defined by

$$\langle u\ ,\ v\rangle_L = \langle u\ ,\ Lv\rangle_H.$$

Assumption 8.3 ensures the symmetry of this product and its positive definiteness. But D is not complete with respect to $\|\cdot\|_L$, so we need to enlarge it (and simultaneously extend L) to obtain a Hilbert space. Note that there always exists an extension of a pre-Hilbertian structure (such as the one we have on D) to an abstract Hilbert space. The following result states that this extension is actually a subspace embedded in H.

Theorem 8.11 (Friedrichs extension) *The inner product*

$$\langle \cdot\ ,\ \cdot\rangle_L : D \times D \to \mathbb{R}$$

can be extended to an inner product

$$\langle \cdot\ ,\ \cdot\rangle_V : V \times V \to \mathbb{R}$$

where V is a dense subspace of H with respect to $\|\cdot\|_H$, and such that D is a dense subspace of V with respect to $\|\cdot\|_V$. The operator L can also be extended to the duality operator of V, $\mathbb{L} : V \to V^*$. The extensions have the properties that:

- *$(V, \|\cdot\|_V)$ is continuously embedded in $(H, \|\cdot\|_H)$.*
- *If $u, v \in D$, $\langle u\ ,\ v\rangle_V = \langle Lu\ ,\ v\rangle_H = (\mathbb{L}u \mid v)$.*
- *V is a Hilbert space with respect to $\langle \cdot\ ,\ \cdot\rangle_V$.*

The fact that $\mathbb{L}$ is an extension of L comes modulo the identification $H = H^*$. Indeed, we have $V \subset H = H^* \subset V^*$ (by the "duality paradox"), so that L, defined on $D \subset V$, can be seen as an operator with values in H^*.

Definition 8.12 The restriction of the operator $\mathbb{L}$ defined in Theorem 8.11 to the space

$$V_{\mathbb{L}} = \left\{u \in V : \mathbb{L}u \in H^* = H\right\}$$

is called the Friedrichs extension of L.

(Notice that, because V is densely embedded in H, H^* is in turn densely embedded in V^*, by Theorem A.13.) In the following, we will stop using a distinct notation for L and $\mathbb{L}$, and use the same letter ($\mathbb{L}$) for the Friedrichs extension.

We will not prove Theorem 8.11, but the interested reader may refer to [315]. The Friedrichs extension has other interesting properties:

Theorem 8.13 *The operator $\mathbb{L} : V_{\mathbb{L}} \to H$ is bijective and self-adjoint, i.e., $\langle \mathbb{L}u \, , \, v \rangle_H = \langle u \, , \, \mathbb{L}v \rangle_H$ for all $u, v \in V_{\mathbb{L}}$.*

Its inverse, $\mathbb{K} = \mathbb{L}^{-1} : H \to H$ is continuous and self-adjoint.

If the embedding $V \subset H$ is compact, then $\mathbb{K} : H \to H$ is a compact operator.

Note that if V is any Hilbert space continuously embedded and dense in H, then, letting $\mathbb{L}$ denote the duality operator $\mathbb{L} : V \to V^*$, we have by definition $(\mathbb{L}u \mid v) = \langle u \, , \, v \rangle_V$ for $u, v \in V$. One can define $V_{\mathbb{L}}$ as the set of $u \in V$ such that $\mathbb{L}u \in H$, i.e., such that $(\mathbb{L}u \mid v) = \langle u \, , \, v \rangle_V \leq C \|v\|_H$ for some constant C and all $v \in V$. Then the statement of Theorem 8.13 is also true.

In the following, we will mostly be interested in embeddings stronger than the L^2 embedding implied by the monotony assumption. It is important that such embeddings are conserved by the extension whenever they are true in the initial space D. This is stated in the next proposition.

Proposition 8.14 *Let D, V and H be as in Theorem 8.11, and B be a Banach space such that $D \subset B \subset H$ and B is canonically embedded in H (there exists a $c_1 > 0$ such that $\|u\|_B \geq c_1 \|u\|_H$). Assume that there exists a constant c_2 such that, for all $u \in D$, $\sqrt{\langle \mathbb{L}u \, , \, u \rangle_H} \geq c_2 \|u\|_B$. Then $V \subset B$ and $\|u\|_V \geq c_2 \|u\|_B$ for all $u \in V$.*

In particular, if B is compactly embedded in H, then $\mathbb{K} = \mathbb{L}^{-1} : H \to H$ is a compact operator.

Proof Let $u \in V$. Since D is dense in V, there exists a sequence $u_n \in D$ such that $\|u_n - u\|_V \to 0$. Thus u_n is a Cauchy sequence in V and, by our assumption, it is also a Cauchy sequence in B, so that there exists a $u' \in B$ such that $\|u_n - u'\|_B$ tends to 0. But since V and B are both embedded in H, we have $\|u_n - u\|_H \to 0$ and $\|u_n - u'\|_H \to 0$, which implies that $u = u'$. Thus u belongs to B, and since $\|u_n\|_V$ and $\|u_n\|_B$ respectively converge to $\|u\|_V$ and $\|u\|_B$, passing to the limit in the inequality $\|u_n\|_V \geq c_2 \|u_n\|_B$ completes the proof of Proposition 8.14. □

When Ω is bounded, one can use the fact that $C^0(\Omega, \mathbb{R}^k)$ is compactly embedded in $L^2(\Omega, \mathbb{R}^k)$ to derive the following consequence of Theorem 8.13. This theorem indeed implies that $\mathbb{K}$ is a compact, self-adjoint operator. Such operators have the important property of admitting an orthonormal sequence of eigenvectors: more precisely, there exists a decreasing sequence, (ρ_n), of positive numbers, which is either finite or tends to 0, and an orthonormal sequence φ_n in $L^2(\Omega, \mathbb{R}^k)$, such that, for $u \in L^2(\Omega, \mathbb{R}^k)$,

$$\mathbb{K}u = \sum_{n=1}^{\infty} \rho_n \langle u \, , \, \varphi_n \rangle_{L^2} \varphi_n.$$

This directly characterizes $V_{\mathbb{L}}$ as the set

$$V_{\mathbb{L}} = \left\{ u \in L^2(\Omega, \mathbb{R}^k) : \sum_{n=1}^{\infty} \frac{\langle u \, , \, \varphi_n \rangle_{L^2}^2}{\rho_n^2} < \infty \right\}$$

and for $u \in V_{\mathbb{L}}$, we have

$$\mathbb{L}u = \sum_{n=1}^{\infty} \rho_n^{-1} \langle u\,,\, \varphi_n \rangle_{L^2} \varphi_n,$$

so that, for $u, v \in V_{\mathbb{L}}$

$$\langle u\,,\, v \rangle_V = \langle \mathbb{L}u\,,\, v \rangle_{L^2} = \sum_{n=1}^{\infty} \rho_n^{-1} \langle u\,,\, \varphi_n \rangle_{L^2} \langle u\,,\, \varphi_n \rangle_{L^2}.$$

This indicates that V should be given by

$$V = \left\{ u \in L^2(\Omega, \mathbb{R}^k) : \sum_{n=1}^{\infty} \rho_n^{-1} \langle u\,,\, \varphi_n \rangle_{L^2}^2 < \infty \right\}.$$

This is indeed the case, because $V_{\mathbb{L}}$ is dense in this set: if $u \in V$, then $u_N = \sum_{n=1}^{N} \langle u\,,\, \varphi_n \rangle \varphi_n$ belongs to $V_{\mathbb{L}}$ and $\|u_N - u\|_V \to 0$. We summarize what we have just obtained in the following theorem.

Theorem 8.15 *Assume that Ω is bounded, $D = C_c^{\infty}(\Omega, \mathbb{R}^k)$, $H = L^2(\Omega, \mathbb{R}^k)$ and $\mathbb{L} : D \to H$ is symmetric and satisfies*

$$\langle \mathbb{L}u\,,\, u \rangle_{L^2} \geq c \, \|u\|_{\infty}^2$$

for some constant $c > 0$. Then the space V associated to $\mathbb{L}$ via Theorem 8.11 is continuously embedded in $C^0(\Omega, \mathbb{R}^k)$ and there exists an orthonormal basis, (φ_n), in $L^2(\Omega, \mathbb{R}^k)$ and a decreasing sequence of positive numbers, (ρ_n), which tends to 0 such that

$$V = \left\{ u \in L^2(\Omega, \mathbb{R}^k) : \sum_{n=1}^{\infty} \rho_n^{-1} \langle u\,,\, \varphi_n \rangle_{L^2}^2 < \infty \right\}.$$

Moreover,

$$\mathbb{L}u = \sum_{n=1}^{\infty} \rho_n^{-1} \langle u\,,\, \varphi_n \rangle_{L^2} \varphi_n$$

whenever

$$\sum_{n=1}^{\infty} \left(\frac{\langle u\,,\, \varphi_n \rangle_{L^2}}{\rho_n} \right)^2 < \infty.$$

8.3 Invariance of the Inner Product

We now discuss admissible Hilbert spaces that share the important property of being invariant under Euclidean transformations. This property is especially important in shape recognition. This analysis will also lead to explicit constructions of operators and kernels. We consider here the situation $\Omega = \mathbb{R}^d$ and $k = d$.

8.3.1 Invariance: The Operator Side

To motivate the invariance requirements, let us return to the interpolation problem for vector fields. Let $x_1, \ldots, x_N \in \Omega$ and vectors $v_1, \ldots, v_N$ be given, and h be the optimal interpolation, i.e., the function in V with minimal norm satisfying $h(x_i) = v_i$ for $i = 1, \ldots, N$. Given a rotation R and a vector $b \in \mathbb{R}^d$, this problem can be modified in (at least) three ways, leading to three possible invariance conditions.

I1. Point transformation: if each x_i is replaced by $Rx_i + b$, then the new optimal function $\tilde{h}$ should be such that $\tilde{h}(Rx + b) = h(x)$ and $\|\tilde{h}\|_V = \|h\|_V$.
I2. Vector transformation: if each v_i is replaced by Rv_i, then $\tilde{h}$ should be such that $\tilde{h}(x) = Rh(x)$ and $\|\tilde{h}\|_V = \|h\|_V$.
I3. Point and vector transformation: if each x_i is replaced by $Rx_i + b$ and each v_i by Rv_i, then $\tilde{h}(Rx + b) = Rh(x)$ and $\|\tilde{h}\|_V = \|h\|_V$.

Obviously, cases I1 to I3 implicitly require that V is closed under the transformation $h(\cdot) \mapsto h(R^{-1}((\cdot) - b))$, $h(\cdot) \mapsto Rh(\cdot)$ and $h(\cdot) \mapsto Rh(R^{-1}((\cdot) - b))$ respectively. All three invariance conditions extend the transformation applied to the input of the problem to the whole space. For our purposes, the most relevant condition is I3 because it corresponds to the invariance we need when considering velocities (or displacements): if $v = \partial_t x$, then $Rv = \partial_t(Rx + b)$, which leads to the transformations in case I3. The invariance in I1 and I2 and their consequences provides some interesting insights into the construction of kernels, which is why we consider them here too.

Notice that vector transformations do not involve translations. There are two reasons for this: in the cases we are interested in, v_i is either considered as a difference between two points in $\mathbb{R}^d$ (x_i and its target), or as a velocity ($\partial_t x_i$), and both are not affected by translations. Moreover, the spaces V we are considering typically contain functions that vanish at infinity, so that V is not closed under the transformation $h \mapsto h + b$.

In all three cases, we have a transformation that we will denote by $h \mapsto (R, b) \star h$ (the case I1, I2 or I3 associated with this notation will always be clear from the context). Notice that the transformation $h \to (R, b) \star h$ is, in all cases, an action of the Euclidean group (translations and rotations) on functions defined on Ω. We will also denote by $(R, b) \star (x, v)$ the associated transformation on points and vectors (with $x = (x_1, \ldots, x_N)$, $v = (v_1, \ldots, v_N)$, N being arbitrary).

In the following, we will assume that V is a Hilbert space continuously embedded and dense in $L^2(\mathbb{R}^d, \mathbb{R}^d)$. We start with the following proposition.

Proposition 8.16 *Assume that V is an RKHS such that, for any pair (x, v), the optimal interpolation associated to $(R, b) \star (x, b)$ is $(R, b) \star h$, where h is the optimal interpolation associated to (x, b), and that $\|h\|_V = \|(R, b) \star h\|_V$. Then $h \mapsto (R, b) \star h$ is an isometry from V onto itself, i.e., for all $h \in V$, $(R, b) \star h \in V$ and $\|(R, b) \star h\|_V = \|h\|_V$.*

Proof Let $h \in V$. For $n > 0$, let x^n form a grid of step length 2^{-n} discretizing the cube $[-2^n, 2^n]^d$. Let $v_i^n = h(x_i^n)$ and let h^n be the optimal interpolant associated with (x^n, v^n). Let $\tilde{h}^n = (R, b) \star h^n$, which belongs to V by assumption.

Let V^n be the set of functions in V that satisfy the constraints (x^n, v^n). Then $h^n = \pi_{V^n}(h)$, the orthonormal projection of h on V^n. Because $h^m \in V^n$ for all $m \geq n$, we have

$$\|h^m\|_V^2 = \|h^m - h^n\|_V^2 + \|h^n\|_V^2.$$

Also, because $h \in V^n$ for all n, the same identity is true with h in place of h^m, so that $\|h^n\|_V$ is a bounded increasing sequence, therefore converging to a limit. Moreover, we have

$$\|h^m - h^n\|_V^2 = \left| \|h^m\|_V^2 - \|h^n\|_V^2 \right|,$$

which proves that h^m is Cauchy, and therefore has a limit h'. Because V is admissible, we know that

$$|h^n(x) - h^n(y)| \leq C \|h\|_V \, |x - y|$$

for some constant C (and we have used the fact that $\|h^n\|_V \leq \|h\|_V$). Because h^n coincides with h on the sets x^n, this implies that $h^n(x) \to h(x)$ for all n, therefore implying that $h = h'$. Using the same argument, $\tilde{h}^n$ is Cauchy in V and its limit in V must also be its pointwise limit, which is $(R, b) \star h$ (which therefore must belong to V). It then suffices to take the limit of the identity $\|h^n\|_V = \|(R, b) \star h^n\|_V$, which is true for all n by assumption, to show that the transformation is an isometry. □

We now study the consequences of the invariance conditions. Introducing the operator $\mathbb{L}$ such that $(\mathbb{L}u \mid v) = \langle u, v \rangle_V$, the isometry property implies that

$$(\mathbb{L}((R, b) \star u) \mid (R, b) \star v) = (\mathbb{L}u \mid v)$$

for all $u, v \in V$, or, denoting by $(R, b) \star \mathbb{L}$ the operator such that

$$\Big(((R, b) \star \mathbb{L})u \,\Big|\, v \Big) = \Big(\mathbb{L}((R, b) \star u) \,\Big|\, (R, b) \star v \Big),$$

that $(R, b) \star \mathbb{L} = \mathbb{L}$.

We let, as above, $V_{\mathbb{L}}$ denote the space of $v \in V$ such that $\mathbb{L}v \in L^2(\mathbb{R}^d, \mathbb{R}^d)$. We first note that the $\star$ action maps $V_{\mathbb{L}}$ onto itself. Indeed, if $u \in V_{\mathbb{L}}$, then, for all $v \in V$

$$(\mathbb{L}((R,b)\star u)\mid v)=\left(\mathbb{L}u\;\middle|\;(R,b)^{-1}\star v\right)\le\|\mathbb{L}u\|_2\,\|(R,b)^{-1}\star v\|_2=\|\mathbb{L}u\|_2\|v\|_2$$

(using a change of variable in the last identity). This implies that $\mathbb{L}((R,b)\star u)\in L^2(\mathbb{R}^d,\mathbb{R}^d)$, so that $(R,b)\star u\in V_{\mathbb{L}}$.

Let us start with case I2, which does not involve translation and is simpler to analyze. In this case, the requirement is that $v\in V\Rightarrow Rv\in V$ for all rotations R, and that $\|Rv\|_V=\|v\|_V$. Let $(e_1,\dots,e_d)$ denote the canonical basis of $\mathbb{R}^d$ and let

$$W=\left\{v^Te_1:v\in V\right\}$$

be the set of scalar-valued functions provided by the first coordinate of elements of v. Because of rotation invariance, we clearly have $W=\left\{v^Tu:v\in V\right\}$ for any $u\in\mathbb{R}^d$, $u\neq 0$. We have the following lemma.

Lemma 8.17 *Under invariance condition I2, and if $d\ge 3$, one has*

$$V=\{v=w_1e_1+\cdots+w_de_d:w_1,\dots,w_d\in W\}.$$

Moreover, with the above decomposition, we have

$$\|v\|_V^2=\|w_1e_1\|_V^2+\cdots+\|w_de_d\|_V^2. \tag{8.13}$$

The result is true for $d=2$ if the invariance condition is extended to include all $R\in O_d(\mathbb{R})$ (rotations and symmetries).

Proof Let $\tilde V=We_1+\cdots+We_d$. If $v\in V$, we have $v=(v^Te_1)e_1+\cdots+(v^Te_d)e_d$ and we have seen that $v^Te_i\in W$ for all i. This shows that $V\subset\tilde V$.

To prove the converse, it suffices to show that $we_1\in V$ for all $w\in W$, because rotation invariance immediately implies that $we_i\in V$ for all $i=1,\dots,d$. So take $w\in W$ and $v\in V$ such that $v^Te_1=w$. Let $v_i=v^Te_i$ (so that $v_1=w$).

Let $S=\{-1,1\}^{d-1}$ and S^+ be the set of $\epsilon\in S$ such that $\epsilon_1\cdots\epsilon_{d-1}=1$. For every $\epsilon\in S^+$, the linear transformation that maps $(e_1,\dots,e_d)$ to $(e_1,\epsilon_1e_2,\dots,\epsilon_{d-1}e_d)$ is a rotation, which implies that

$$v^\epsilon:=v_1e_1+\epsilon_1v_2e_2+\cdots+\epsilon_{d-1}v_de_d\in V$$

and so is

$$\hat v:=\sum_{\epsilon\in S^+}v^\epsilon=|S^+|v_1e_1+\sum_{\epsilon\in S^+}(\epsilon_1v_2e_2+\cdots+\epsilon_{d-1}v_de_d).$$

The number of $\epsilon\in S^+$ such that $\epsilon_i=1$ is the number of subsets with even cardinality in $\{2,\dots,d-1\}$, namely

$$a_+=\sum_{0\le k\le(d-2)/2}\binom{d-2}{2k}.$$

Similarly, the number of $\epsilon \in S^+$ such that $\epsilon_i = -1$ is

$$a_- = \sum_{0 \leq k \leq (d-3)/2} \binom{d-2}{2k+1}.$$

When $d > 2$, these numbers are equal, because

$$a_+ - a_- = \sum_{k=1}^{d-2} (-1)^k \binom{d-2}{k} = 0.$$

Notice also that $|S^+| = a_+ + a_- = 2^{d-2}$. This implies that $\hat{v} = 2^{d-2} v_1 e_1$. Since $\hat{v}$ belongs to V, so does $v_1 e_1$, which completes our proof that $\tilde{V} = V$ for the case $d > 2$.

If $d = 2$ one has $S = \{-1, 1\}$, $S^+ = \{1\}$ and we cannot reach the same conclusion. If invariance by symmetry is added, however, then we can sum over S instead of S^+ in the definition of $\hat{v}$, which gives $\hat{v} = 2 v_1 e_1$ and proves the result in this case.

To prove the statement about $\|v\|_V^2$, it suffices to show that $\langle w_i e_i \, , \, w_j e_j \rangle_V = 0$ for $i \neq j$. But for $d > 2$, there exists a rotation that maps e_i to itself and e_j to $-e_j$ and the invariance of the inner product implies $\langle w_i e_i \, , \, w_j e_j \rangle_V = -\langle w_i e_i \, , \, w_j e_j \rangle_V$, which must therefore vanish. For $d = 2$, one can take the symmetry $(e_1, e_2) \mapsto (e_1, -e_2)$. □

The space W is a Hilbert space of scalar functions, with norm $\|w\|_W = \|w e_1\|_V$. As such, it also has a duality operator, $\mathbb{L}_W$, such that $\left(\mathbb{L}_W w \;\middle|\; w'\right) = \langle w \, , \, w' \rangle_W$. Lemma 8.17 and Eq. (8.13) imply that the operator $\mathbb{L} = \mathbb{L}_V$ associated with V operates coordinate-wise as $\mathbb{L} v = (\mathbb{L}_W v_1, \ldots, \mathbb{L}_W v_d)$. It is easy to check that this property (or the conclusion of Lemma 8.17) implies that I2 is true.

We now pass to I1 and I3, and focus first on translation invariance, which, in both cases (taking $R = \mathrm{Id}$), requires that, for all $h \in V$ and $v \in \mathbb{R}^d$, the vector field $\tilde{h} : x \mapsto h(x - b)$ belongs to V and has the same norm as h. To handle this problem, we first introduce the Fourier transform of the operator $\mathbb{L}$.

Following Sect. A.10, denote the Fourier transform of a square integrable vector field u either by $\hat{u}$ or $\mathcal{F}(u)$, and recall that $\mathcal{F}$ is an isometry of $L^2(\mathbb{R}^d, \mathbb{C}^d)$ (L^2 functions defined on $\mathbb{R}^d$ taking values $\mathbb{C}^d$). For this discussion, we denote $L^2(\mathbb{R}^d, \mathbb{R}^d)$ by H. We have, by assumption, $V \subset H = L^2(\mathbb{R}^d, \mathbb{R}^d) \subset L^2(\mathbb{R}^d, \mathbb{C}^d)$ with V dense in H. For L^2 functions u and v with values in $\mathbb{C}^d$, we will define

$$\langle\!\langle u \, , \, v \rangle\!\rangle_2 = \int_{\mathbb{R}^d} u^T \bar{v} \, dx$$

(so that $\langle\!\langle v \, , \, u \rangle\!\rangle_2 = \overline{\langle\!\langle u \, , \, v \rangle\!\rangle_2}$). Recall that the isometric property of the Fourier transform states that

$$\langle\!\langle \hat{u} \, , \, \hat{v} \rangle\!\rangle_2 = \langle\!\langle u \, , \, v \rangle\!\rangle_2$$

for all square integrable functions u and v.

Let $\hat{V} = \mathcal{F}(V)$ and, for $\varphi \in \hat{V}$, let $\|\varphi\|_{\hat{V}} = \|\mathcal{F}^{-1}\varphi\|_V$. Then $\hat{V}$ is a Hilbert space isometric to V via the Fourier transform, dense in $\hat{H} = \mathcal{F}(H)$, where $\hat{H}$ is the space of functions $\varphi \in L^2(\mathbb{R}^d, \mathbb{C})$ such that $\varphi(-\xi) = \bar{\varphi}(\xi)$ almost surely. Denote by $\hat{\mathbb{L}}$ its duality operator and let

$$V_{\hat{\mathbb{L}}} = \left\{ \varphi \in \hat{V} : \hat{\mathbb{L}}\varphi \in \hat{H} \right\}.$$

Then, it is easy to show (and left to the reader) that $V_{\hat{\mathbb{L}}} = \hat{V}_{\mathbb{L}} = \mathcal{F}(V_{\mathbb{L}})$ and $\hat{\mathbb{L}}\varphi = \mathcal{F}\left(\mathbb{L}\mathcal{F}^{-1}(\varphi)\right)$ for $\varphi \in V_{\hat{\mathbb{L}}}$. In addition, the following lemma holds.

Lemma 8.18 *Assume that translations act as isometries on V and that for all $\xi \in \mathbb{R}^d$, there exists a family of continuous functions $\psi_1, \ldots, \psi_d \in \hat{V}_{\mathbb{L}}$ such that $\psi_1(\xi), \ldots, \psi_d(\xi)$ are linearly independent.*

Then $\hat{\mathbb{L}}$ has a matrix multiplier: there exists a function $f : \mathbb{R}^d \to \mathcal{M}_d(\mathbb{C})$ such that for all $\xi \in \mathbb{R}^d$, $f(\xi)$ is a positive definite Hermitian matrix and for all $\varphi \in \hat{V}_{\mathbb{L}}$ and $\xi \in \mathbb{R}^d$, $(\hat{\mathbb{L}}\varphi)(\xi) = f(\xi)\varphi(\xi)$. Moreover, there exists a constant $c > 0$ such that

$$\xi^T f(\xi)\xi \geq c\,|\xi|^2$$

for almost all $\xi \in \mathbb{R}^d$

Conversely, if $\hat{\mathbb{L}}$ has a matrix multiplier, it is translation-invariant.

Proof The Fourier transform of $(\mathrm{Id}, b) \star u$ is $\gamma_b \hat{u}$ where $\gamma_b : \xi \mapsto e^{-2\imath\pi\xi^T b}$. Since $V_{\mathbb{L}}$ in closed under the action of translations, $\hat{V}_{\mathbb{L}}$ is also closed under the transformations $\varphi \mapsto \gamma_b\varphi$ and

$$\left(\hat{\mathbb{L}}(\gamma_b\varphi) \;\middle|\; \gamma_b\psi\right) = \left(\hat{\mathbb{L}}\varphi \;\middle|\; \psi\right)$$

for all $\varphi, \psi \in \hat{V}_{\mathbb{L}}$ and $b \in \mathbb{R}^d$. Consider the space M of functions m such that (i) for all $\varphi \in \hat{V}_{\mathbb{L}}$, both $m\varphi$ and $\bar{m}\varphi$ are in $\hat{V}_{\mathbb{L}}$ and (ii) for all $\varphi, \psi \in \hat{V}_{\mathbb{L}}$,

$$\left(\hat{\mathbb{L}}(m\varphi) \;\middle|\; \psi\right) = \left(\hat{\mathbb{L}}\varphi \;\middle|\; \bar{m}\psi\right)$$

holds. (Note that the left-hand side is also equal to $\overline{\left(\hat{\mathbb{L}}\psi \;\middle|\; m\varphi\right)}$.) This space contains all γ_b's (because $\bar{\gamma}_b = \gamma_b^{-1} = \gamma_{-b}$), and therefore also their linear combinations. Moreover, since

$$\left(\hat{\mathbb{L}}\varphi \;\middle|\; \bar{m}\psi\right) \leq \|\hat{\mathbb{L}}\varphi\|_2 \|\psi\|_2 \|m\|_\infty,$$

the map $m \mapsto \left(\hat{\mathbb{L}}\varphi \;\middle|\; \bar{m}\psi\right)$ is continuous with respect to the supremum norm, so that the considered space contains all limits in the supremum norm of linear combinations of γ_b's with real coefficients, which is the space of all bounded continuous functions from $\mathbb{R}^d$ to $\mathbb{C}$ that satisfy $m(-\xi) = \bar{m}(\xi)$. This statement can be extended to all

bounded measurable functions on $\mathbb{R}^d$: indeed, if m is such a function, Lusin's theorem [249] implies that, for every $\varepsilon > 0$, there exists a continuous bounded function m_ε such that $\|m_\varepsilon\|_\infty \le \|m\|_\infty$ and the set $N_\varepsilon = \{\xi : m(\xi) \neq m_\varepsilon(\xi)\}$ has measure less than ε. Replacing m_ε by $x \mapsto (m_\varepsilon(x) + \bar{m}_\varepsilon(-x))/2$ if needed, we can also assume that $m_\varepsilon(-x) = \bar{m}_\varepsilon(x)$. Then the identities

$$\overline{\left(\hat{\mathbb{L}}\psi \;\middle|\; m_\varepsilon\varphi\right)} = \left(\hat{\mathbb{L}}\varphi \;\middle|\; \bar{m}_\varepsilon\psi\right)$$

can be taken to the limit $\varepsilon \to 0$ yielding the same identity for m, based on the fact that

$$\left(\hat{\mathbb{L}}\varphi \;\middle|\; (\bar{m} - \bar{m}_\varepsilon)\psi\right) = \int_{N_\varepsilon} (\bar{m} - \bar{m}_\varepsilon)(\hat{\mathbb{L}}\varphi)^T \bar{\psi}\, d\xi \to 0$$

using the dominated convergence theorem.

Take now $\xi \in \mathbb{R}^d$ and a family of continuous functions $\psi_1, \dots, \psi_d \in \hat{V}_\mathbb{L}$ such that $(\psi_1(\xi), \dots, \psi_d(\xi))$ is linearly independent. By continuity, linear independence will remain true in a neighborhood of ξ, say, for $|\eta - \xi| < \delta$, and for such η (and small enough δ) the conjugate basis $\chi_1(\eta), \dots, \chi_d(\eta)$ such that

$$\sum_{i=1}^{d} \psi_i(\eta)\chi_i(\eta)^T = \mathrm{Id}_{\mathbb{R}^d}$$

is well defined and continuous. Because $\bar{\psi}_i(-\xi) = \psi_i(\xi)$ for all i, $\chi_1(\eta), \dots, \chi_d(\eta)$ are also well defined for $|\xi + \eta| < \delta$, with $\bar{\chi}_i(-\eta) = \chi_i(\eta)$. Finally, let ω be a continuous, bounded function such that $\omega(\eta) = 1$ if $|\xi - \eta| < \delta/2$ or $|\xi + \eta| < \delta/2$ and $\omega(\eta) = 0$ if $|\xi - \eta| > \delta$ and $|\xi + \eta| > \delta$. Since we can replace ω by $\eta \mapsto (\omega(\eta) + \omega(-\eta))/2$, we can also assume that ω is even. Define, for $i = 1, \dots, d$, $\tilde{\chi}_i(\eta) = \omega(\eta)\chi_i(\eta)$ if $|\eta - \xi| < \delta$ or $|\eta + \xi| < \delta$ and $\tilde{\chi}_i(\eta) = 0$ otherwise.

If $\varphi \in \hat{V}_\mathbb{L}$ is bounded, one has $\omega\hat{\mathbb{L}}\varphi = \hat{\mathbb{L}}(\omega\varphi)$ because $\omega \in M$ and

$$\hat{\mathbb{L}}(\omega\varphi) = \sum_{i=1}^{d} \hat{\mathbb{L}}(\tilde{\chi}_i^T \varphi\, \psi_i) = \sum_{i=1}^{d} \tilde{\chi}_i^T \varphi \hat{\mathbb{L}}(\psi_i)$$

because $\tilde{\chi}_i^T \varphi \in M$. So letting

$$f = \sum_{i=1}^{d} \hat{\mathbb{L}}(\psi_i)\chi_i^T,$$

we find $\hat{\mathbb{L}}\varphi = f\varphi$ in an open neighborhood of ξ. Since the same construction can be made near every ξ, one can piece together a function f defined on $\mathbb{R}^d$ such that $\hat{\mathbb{L}}\varphi = f\varphi$ for all bounded $\varphi \in \hat{V}_\mathbb{L}$. The extension from bounded to arbitrary φ can be done by truncation: if $\rho_N = \mathbf{1}_{|\varphi|>N}$, then $\hat{\mathbb{L}}(\rho_N\varphi) = f(\rho_N\varphi) = \rho_N f\varphi$ because

$\rho_N\varphi \in M$. In addition, $\rho_N \in M$, so that $\mathbb{L}(\rho_N\varphi) = \rho_N\mathbb{L}\varphi$. So $\rho_N\mathbb{L}\varphi = \rho_N f\varphi$, which implies that $\mathbb{L}\varphi = f\varphi$ by letting N go to infinity.

Since $\mathbb{L}$ is symmetric and real, we have

$$\int_{\mathbb{R}^d} (\hat{\mathbb{L}}\varphi)^T \bar{\psi}\, d\xi = \int_{\mathbb{R}^d} (\hat{\mathbb{L}}\psi)^T \bar{\varphi}\, d\xi \in \mathbb{R}$$

for all $\varphi, \psi \in \hat{V}_L$. This implies

$$\int_{\mathbb{R}^d} \bar{\psi}^T f\varphi\, d\xi = \int_{\mathbb{R}^d} \bar{\varphi}^T f\psi\, d\xi = \int_{\mathbb{R}^d} \bar{\psi}^T \bar{f}^T \varphi\, d\xi.$$

So, replacing f by $(f + \bar{f}^T)/2$ if necessary, we can assume that f is Hermitian, with moreover

$$\int_{\mathbb{R}^d} \bar{\psi}^T f\psi\, d\xi \geq c\|\psi\|_2^2$$

for all $\psi \in \hat{V}_{\mathbb{L}}$ and a fixed $c > 0$. This implies that, for all $\psi \in \hat{V}_{\mathbb{L}}$, one has $\bar{\psi}^T f\psi \geq c|\psi|^2$ a.e. Indeed, letting A be any set over which $\bar{\psi}^T f\psi < c|\psi|^2$, one has $\mathbf{1}_A\psi \in \hat{V}_{\mathbb{L}}$, which implies

$$\int_A \bar{\psi}^T f\psi\, d\xi \geq c\int_A |\psi|^2\, d\xi,$$

which is only possible if A has measure 0. Our hypothesis on the existence of linearly independent continuous functions in $\hat{V}_{\mathbb{L}}$ at any ξ implies that $f(\xi) \geq c\,\mathrm{Id}_{\mathbb{R}^d}$ for almost all ξ.

The last statement of the lemma is obvious from $f(\xi)(\gamma_b(\xi)\varphi(\xi)) = \gamma_b(\xi) f(\xi)\varphi(\xi)$. □

We now assume that the conclusion of the lemma holds, and that $\hat{\mathbb{L}}$ is given by a Hermitian matrix multiplier $f \geq c\,\mathrm{Id}_{\mathbb{R}^d}$, which, as just seen, ensures translation invariance. This assumption will allow us to focus only on rotation from now on. As we will see, an interesting special case is when the entries of f are polynomials in ξ, which corresponds, from standard properties of the Fourier transform, to $\mathbb{L}$ being a multi-dimensional differential operator.

To consider rotation invariance, we will write, for short, $R \star h = (R, 0) \star h$ for a rotation matrix R. We compute $\mathcal{F}(R \star h)$ as a function of $\hat{h}$. We have, in case I1:

$$\begin{aligned}
\mathcal{F}(R \star h)(\xi) &= \int_{\mathbb{R}^d} e^{-2\iota\pi\xi^T x}\, R \star h(x)\, dx \\
&= \int_{\mathbb{R}^d} e^{-2\iota\pi\xi^T x}\, h(R^{-1}x)\, dx \\
&= \int_{\mathbb{R}^d} e^{-2\iota\pi\xi^T (Ry)}\, h(y)\, dy \\
&= \hat{h} \circ R^{-1}(\xi),
\end{aligned}$$

where we have used the fact that $\det R = 1$ in the change of variables. In case I3, the same computation gives $\mathcal{F}(R \star h) = R\hat{h} \circ R^{-1}$.

We can therefore write, in case I1:

$$\begin{aligned}(\mathbb{L}(R \star h) \mid R \star h) &= \left(f\hat{h} \circ R^{-1} \,\middle|\, \hat{h} \circ R^{-1}\right) \\ &= \left((f \circ R)\hat{h} \,\middle|\, \hat{h}\right)\end{aligned}$$

and in case I3

$$\begin{aligned}(\mathbb{L}(R \star h) \mid R \star h) &= \left(fR\hat{h} \circ R^{-1} \,\middle|\, R\hat{h} \circ R^{-1}\right) \\ &= \left(R^T(f \circ R)R\hat{h} \,\middle|\, \hat{h}\right),\end{aligned}$$

in which we have used the facts that $\det R = 1$, $R^T = R^{-1}$. So the invariance conditions boil down, in case I1, to $f(R\xi) = f(\xi)$ and in case I3, to $R^T f(R\xi)R = f(\xi)$, for all ξ and all rotation matrices R.

We now investigate the consequences of these conditions. In case I1, using the fact that one can always find a rotation that maps ξ to $|\xi|e_1$, one sees that f must only depend on $|\xi|$ (f is radial). Note that we assume that $\mathbb{L}$ transforms real vector fields into real ones. The Fourier transform u of a real-valued function is characterized by the property $\overline{u}(-x) = u(x)$. In order that $\overline{u}(-x) = u(x) \Rightarrow \overline{(fu)}(-x) = fu(x)$, we need $\bar{f}(-x) = f(x)$. If f is radial, then $f(-x) = f(x)$, which implies that f is a real matrix.

Case I3 requires a little more work and is summarized in the next lemma.

Lemma 8.19 *Assume that a function f defined on $\mathbb{R}^d$ and taking values in the set of Hermitian matrices satisfies*

$$f(R\xi) = Rf(\xi)R^T$$

for all $\xi \in \mathbb{R}^d$ and all rotation matrices R if $d \geq 3$, and for all orthogonal matrices R if $d = 2$. Then there exist two real-valued functions λ and μ, defined over $[0, +\infty)$ such that $\mu(0) = 0$, and

$$f(\xi) = \lambda(|\xi|)\mathrm{Id}_{\mathbb{R}^d} + \mu(|\xi|)\frac{\xi\xi^T}{|\xi|^2}. \tag{8.14}$$

Proof We will use the following linear algebra result. Let M be a $d \times d$ Hermitian matrix satisfying $R^T MR = M$ for all rotation matrices M. Then $M = \lambda\mathrm{Id}$ for some real number λ. Indeed, let a be an eigenvector of M, satisfying $Ma = \lambda a$ and $|a| = 1$ (with λ necessarily real because M is Hermitian). Then, for every rotation matrix R, $MRa = RMa = \lambda Ra$, so that Ra is also an eigenvector of M. Choosing

rotations $R_1 = \mathrm{Id}_{\mathbb{R}^d}, R_2, \ldots, R_d$ such that $R_1 a, \ldots, R_d a$ is a basis of $\mathbb{C}^d$, we obtain $M = \lambda \mathrm{Id}_{\mathbb{R}^d}$.[1]

Taking $\xi = 0$, the identity $R^T f(0) R = f(0)$ for all R therefore implies that $f(0) = \lambda_0 \mathrm{Id}_{\mathbb{R}^d}$ for some $\lambda_0 \in \mathbb{R}$. Let us now fix $\xi_0 \neq 0$. For any rotation R that leaves ξ_0 invariant, we have

$$f(\xi_0) = R^T f(R\xi_0) R = R^T f(\xi_0) R$$

This yields the fact that, for any R such that $R\xi_0 = \xi_0$, we have $R^T f(\xi_0) R = f(\xi_0)$.

We now separate the cases $d = 2$ and $d > 2$, starting with the latter, and show that in this case one has $f(\xi_0)\xi_0 = \alpha\xi_0$ for some $\alpha \in \mathbb{R}$. Indeed, assume that $f(\xi_0)\xi_0 = \alpha\xi_0 + \beta v$, with $|v| = 1$, $v \perp \xi_0$ and $\beta \neq 0$. Since $d \geq 3$, one can find a third unit vector, w, such that $(\xi_0/|\xi_0|, v, w)$ is an orthonormal family, and one can form a rotation R such that $R\xi_0 = \xi_0$, $Rv = w$ and $Rw = -v$ (leaving invariant any vector perpendicular to ξ_0, v and w). For this rotation, we have $R^T f(\xi_0) R\xi_0 = \alpha\xi_0 + \beta w$ and this can be equal to $\alpha\xi_0 + \beta v$ if and only if $\beta = 0$.

So, ξ_0 is an eigenvector of $f(\xi_0)$, which implies that $f(\xi_0)$ maps the space $\xi_0^\perp$ into itself (because it is Hermitian). But the fact that $R^T f(\xi_0) R = f(\xi_0)$ for any rotation R of $\xi_0^\perp$ implies that $f(\xi_0)$ restricted to $\xi_0^\perp$ is a homothety, i.e., that there exists a $\lambda(\xi_0)$ such that $f(\xi_0)\eta = \lambda(\xi_0)\eta$ if $\eta \perp \xi_0$.

The orthogonal projection of a vector $a \in \mathbb{C}^d$ on the line generated by ξ_0 is

$$P_{\xi_0} a = \frac{\xi_0^T a}{|\xi_0|^2} \xi_0 = \frac{\xi_0 \xi_0^T}{|\xi_0|^2} a$$

so that, for $a \in \mathbb{C}^d$, and letting $f(\xi_0)\xi_0 = \alpha(\xi_0)\xi_0$

$$f(\xi_0) = \alpha(\xi_0) \frac{\xi_0 \xi_0^T}{|\xi_0|^2} + \lambda(\xi_0) \left(\mathrm{Id} - \frac{\xi_0 \xi_0^T}{|\xi_0|^2} \right) = \mu(\xi_0) \frac{\xi_0 \xi_0^T}{|\xi_0|^2} + \lambda(\xi_0) \mathrm{Id}$$

(taking $\mu = \alpha - \lambda$). Now, for any rotation R

$$\begin{aligned} \mu(R\xi_0) \frac{R\xi_0 (R\xi_0)^T}{|\xi_0|^2} + \lambda(R\xi_0)\mathrm{Id} = f(R\xi_0) &= R f(\xi_0) R^T \\ &= \mu(\xi_0) \frac{R\xi_0 (R\xi_0)^T}{|\xi_0|^2} + \lambda(\xi_0)\mathrm{Id} \end{aligned}$$

so that $\mu(R\xi_0) = \mu(\xi_0)$ and $\lambda(R\xi_0) = \lambda(\xi_0)$. This implies that λ and μ must be radial functions and proves (8.14).

Finally, since we know that $f(0) = \lambda \mathrm{Id}$, we must have $\mu(0) = 0$.

Let us now consider the case $d = 2$. Let $\tilde{\xi}_0$ be obtained from ξ_0 by a rotation of $\pi/2$. Consider the rotation $R_0^T = (\xi, \tilde{\xi})/|\xi|$. Then, R_0 maps ξ_0 to $e_1 = (1, 0)^T$ and

[1] This standard result is in fact true for any matrix M (not only Hermitian) in dimension $d > 2$, and also in dimension 2 if one adds symmetries to rotations.

letting $g(t) = f(te_1)$, we obtain

$$f(\xi) = (\xi, \tilde{\xi})\frac{g(|\xi|)}{|\xi|^2}(\xi, \tilde{\xi})^T$$

and this matrix clearly satisfies $f(R\xi) = Rf(\xi)R^T$. If g is diagonal, then letting $\tilde{\mu}$ and $\tilde{\lambda}$ denote its entries, we find

$$f(\xi_0) = \tilde{\mu}(|\xi_0|)\frac{\xi_0\xi_0^T}{|\xi_0|^2} + \tilde{\lambda}(|\xi_0|)\frac{\tilde{\xi}_0\tilde{\xi}_0^T}{|\xi_0|^2} = (\tilde{\mu}(|\xi_0|) - \tilde{\lambda}(|\xi_0|))\frac{\xi_0\xi_0^T}{|\xi_0|^2} + \tilde{\lambda}(|\xi_0|)|\mathrm{Id},$$

which takes the same form as the one obtained with $d \geq 3$. The off-diagonal coefficients of g can be non-zero, however, unless one assumes the additional constraint that the norm is also invariant under symmetry. In this case, we can also use $-\tilde{\xi}_0/|\xi_0|$ to complete ξ_0 and get

$$f(\xi) = (\xi_0, -\tilde{\xi}_0)\frac{g(|\xi_0|)}{|\xi_0|^2}(\xi_0, -\tilde{\xi}_0)^T$$

which, combined with the previous constraint, is only possible when g is diagonal. □

To ensure that $f(\xi) \geq c\,\mathrm{Id}$ for all ξ, we need its eigenvalues (which are $\lambda(|\xi|)$ and $\lambda(|\xi|) + \mu(|\xi|)$) to be larger than c. We obtain the condition that $\lambda \geq c + \max(-\mu, 0)$. We summarize all this in the following theorem.

Theorem 8.20 *Let $\mathbb{L}$ be such that $\hat{\mathbb{L}}$ is a multiplier with $\hat{\mathbb{L}}\varphi = f\,\varphi$ and $f \geq c\,\mathrm{Id}$. Then $\mathbb{L}$ is rotation-invariant (rotation/symmetry-invariant for $d = 2$) if and only if f takes the form*

$$f(\xi) = \mu(|\xi|)\frac{\xi\xi^T}{|\xi|^2} + \lambda(|\xi|)\mathrm{Id}$$

for some functions λ and μ such that $\mu(0) = 0$ and $\lambda \geq c + \max(-\mu, 0)$.

Diagonal Operators

Consider the case $\mu \equiv 0$ in Theorem 8.20. Let $\mathbb{L}_0$ be the scalar operator such that $\hat{\mathbb{L}}_0\varphi = \lambda\varphi$. Then $\mathbb{L}h$ is obtained by applying $\mathbb{L}_0$ to each coordinate of h, i.e., it takes the form

$$\mathbb{L}h = (\mathbb{L}_0 h_1, \ldots, \mathbb{L}_0 h_d).$$

Note that diagonal operators in this form are those that satisfy both I1 and I3.

These simple, diagonal, operators are those that are most commonly used in practice. If one wants, in addition, $\mathbb{L}_0$ to be a differential operator, then $\lambda(|\xi|)$ must be a polynomial in the coefficients of ξ, which is only possible if this function takes the form

$$\lambda(|\xi|) = \sum_{q=0}^{p} \lambda_q |\xi|^{2q}.$$

To prove this statement, just notice that if $\lambda(|\xi|) = P(\xi_1, \ldots, \xi_d)$, where P is a polynomial, then $\lambda(t) = P(t, 0, \ldots, 0)$ is a univariate polynomial and the fact that $P(-t, 0, \ldots, 0) = P(t, 0, \ldots, 0)$ implies that coefficients of odd degree vanish.

Applying the inverse Fourier transform, the corresponding operator is $\mathbb{L}_0 u = \sum_{q=0}^{p} \lambda_q (-1)^q \Delta^q u$. For example, if $\lambda(|\xi|) = (\alpha + |\xi|^2)^2$, we get $\mathbb{L}_0 = (\alpha \mathrm{Id}_d - \Delta)^2$.

General Operators

Let us consider an example of a valid differential operator with $\mu \neq 0$, taking $\mu(|\xi|) = \alpha|\xi|^2$ and λ as above. This yields the operator

$$\mathbb{L}_{kl} u = -\alpha \partial_k \partial_l u + \delta_{kl} \sum_{q=0}^{p} (-1)^q \lambda_q \Delta^q u$$

so that

$$\mathbb{L} h = -\alpha \nabla(\mathrm{div} h) + \sum_{q=0}^{p} (-1)^q \lambda_q \Delta^q h.$$

Similarly, taking $\mu(\xi) = \alpha|\xi|^2 \lambda(\xi)$ yields the operator

$$\mathbb{L} h = \sum_{q=0}^{p} (-1)^q \lambda_q \Delta^q (h - \nabla(\mathrm{div} h)).$$

8.3.2 *Invariance: The Kernel Side*

Let us focus on case I3 in this discussion, and assume that $\hat{\mathbb{L}}$ is associated with a multiplier $f \geq c\,\mathrm{Id}$. Then the kernel operator $\mathbb{K} = \mathbb{L}^{-1}$ is such that $\hat{\mathbb{K}}$ is associated with the multiplier $g = f^{-1}$ in the Fourier domain. Note that we need $\mathbb{K}(a^T \delta_x)$ to be well-defined. The kth component of $a^T \delta_x$ being $a_k \delta_x$, we can write $\mathcal{F}\left((a^T \delta_x)\right)(\xi) = a \exp(-i \xi^T x)$ so that

$$\mathcal{F}\left(\mathbb{K}(a^T \delta_x)\right)(\xi) = g(\xi) a \exp(-i \xi^T x).$$

A sufficient condition for this to have a continuous inverse Fourier transform is that $|g|$ is integrable (taking any matrix norm). Assuming this, we have

$$(\mathbb{K}(a^T \delta_x)(y)) = K(y, x) a = \tilde{g}(y - x) a,$$

where $\tilde{g}$ is the inverse Fourier transform of g.

When $f(\xi) = \lambda(|\xi|)\mathrm{Id} + \mu(|\xi|)\xi\xi^T/|\xi|^2$, we can check that

$$f^{-1}(\xi) = \frac{1}{\lambda(|\xi|)}\left(\mathrm{Id} - \frac{\xi\xi^T}{|\xi|^2}\right) + \frac{1}{\lambda(|\xi|)+\mu(|\xi|)}\left(\frac{\xi\xi^T}{|\xi|^2}\right).$$

Let $\mathbb{A} : L^2(\mathbb{R}^d, \mathbb{R}^d) \to L^2(\mathbb{R}^d, \mathbb{R}^d)$ denote the operator such that $\hat{\mathbb{A}}$ is associated with the matrix multiplier $\xi\xi^T/|\xi|^2$. Let also κ_1 and κ_2 be the inverse Fourier transforms of λ^{-1} and $(\lambda+\mu)^{-1}$. Then

$$\mathbb{K}u = \kappa_1 * (u - \mathbb{A}u) + \kappa_2 * \mathbb{A}u = \kappa_1 u - \mathbb{A}((\kappa_1 - \kappa_2) * u),$$

where $*$ denotes convolution (we used the fact that operators that are multiples of the identity in the Fourier domain commute with any operator associated to a matrix multiplier, such as $\mathbb{A}$). For smooth enough u, the vector field $v = \mathbb{A}u$ can be computed by solving $\Delta v = -\nabla \mathrm{div} u$, which is also equivalent to

$$v = \nabla \mathrm{div}(G * u),$$

where G is the Green's function of the Laplacian operator defined by (see [102])

$$G(x) = \begin{cases} -\dfrac{1}{2\pi}\log|x| & (n=2) \\ \dfrac{c_n}{|x|^{n-2}} & (n \geq 3) \end{cases} \tag{8.15}$$

with $c_n = \Gamma(1+n/2)/(\pi^{n/2} n(n-1))$.

Note that the inverse Fourier transform of a radial function (i.e., a function $\rho(\xi)$ that only depends on $|\xi|$) is also radial. To see this, we can write

$$\begin{aligned} \tilde{\rho}(x) &= \int_{\mathbb{R}^d} \rho(|\xi|) e^{i\xi^T x} d\xi \\ &= \int_0^\infty \rho(t) t^{d-1} \int_{S^{d-1}} e^{i\eta^T t x} ds(\eta) \\ &= \int_0^\infty \rho(t) t^{d-1} \int_{S^{d-1}} e^{i\eta_1 t|x|} ds(\eta), \end{aligned}$$

where S^{d-1} is the unit sphere in $\mathbb{R}^d$; the first change of variable was $\xi = t\eta$ and the last identity comes from the fact that the integral is invariant under rotation of x, so that we could take x parallel to the first coordinate axis. The last integral can in turn be expressed in terms of the Bessel function $J_{d/2}$ [223], yielding an expression which will not be detailed (or used) here. This implies that κ_1 and κ_2 above are radial kernels.

We can therefore describe the kernels associated with spaces V that satisfy I3 as those taking the form

$$K = (\kappa_1 - \mathbb{A}\kappa_1) + \mathbb{A}\kappa_2,$$

where κ_1 and κ_2 are radial kernels with bounded Fourier transform (recall the constraints $\lambda \geq c$ and $\lambda + \mu \geq c$). As remarked in [195] (in which the theory of translation and rotation-invariant kernels is discussed extensively), this expression decomposes K into divergence-free and curl-free parts (in 3D). Indeed, because $\mathbb{A}$ is a gradient, one has $\operatorname{curl}\mathbb{A} = 0$, so that $\mathbb{A}\kappa_2$ is curl-free and $\operatorname{curl} K = \operatorname{curl} \kappa_1$. Moreover, one has

$$\operatorname{div}(\mathbb{A}u) = \Delta \operatorname{div}(G * u) = \operatorname{div} \Delta(G * u) = \operatorname{div} u,$$

so that $\kappa_1 - \mathbb{A}\kappa_1$ is divergence-free and $\operatorname{div} K = \operatorname{div} \kappa_2$.

Examples of Radial Kernels

The following proposition provides a nice dimension-independent characterization of scalar radial kernels [251].

Proposition 8.21 *The scalar kernel $K(x, y) = \gamma(|x - y|)$ is positive definite for all dimensions d if and only there exists a positive measure μ on $(0, +\infty)$ such that*

$$\gamma(t) = \int_0^{+\infty} e^{-t^2 u} d\mu(u).$$

This includes, in particular, the case of all functions of the form

$$\gamma(t) = \int_0^{+\infty} e^{-t^2 u} f(u) du \tag{8.16}$$

for a positive function f. Note that if a function γ provides a positive kernel in dimension d, it trivially provides a positive kernel in dimension $d' \leq d$, but not necessarily for $d' > d$ unless it takes the form given in Proposition 8.21.

Translation-invariant kernels (not necessarily radial) of the kind $K(x, y) = \Gamma(x - y)$ can be characterized in a similar way, by Bochner's theorem [246].

Proposition 8.22 *The kernel $\chi(x, y) = \Gamma(x - y)$ is positive definite if and only if there exists a positive, symmetric measure μ on $\mathbb{R}^d$ such that*

$$\Gamma(x) = \int_{\mathbb{R}^d} e^{-2\imath\pi\, x^T u} d\mu(u).$$

Letting μ be a Dirac measure ($\mu = \delta_{\sigma^{-2}}$) in Eq. (8.16) yields $\gamma(t) = e^{-\frac{t^2}{\sigma^2}}$. The associated kernel

$$K(x, y) = e^{-\frac{|x-y|^2}{\sigma^2}} \mathrm{Id}_d$$

is the *Gaussian kernel* on $\mathbb{R}^d$ and is one of the most commonly used for spline smoothing.

We can also use Proposition 8.21 with $f(u) = e^{-u}$. This provides the Cauchy kernel

$$K(x, y) = \frac{1}{1 + |x - y|^2} \mathrm{Id}.$$

Other choices can be made: if f is the indicator function of the interval $[0, 1]$, then

$$\gamma(t) = \int_0^1 e^{-t^2 u} du = \frac{1 - e^{-t^2}}{t^2}$$

with the corresponding kernel. If one takes $f(u) = \exp(-u^2/2)$, then

$$\begin{aligned} \gamma(t) &= \int_0^\infty e^{-u^2/2 - ut^2} du \\ &= e^{t^4/2} \int_0^\infty e^{-(u+t^2)^2/2} du \\ &= \frac{\sqrt{\pi}}{2} e^{t^4/2} \mathrm{erfc}(t^2), \end{aligned}$$

where $\mathrm{erfc}(q)$ is the probability that a standard Gaussian distribution is larger than q in absolute value,

$$\mathrm{erfc}(q) = \frac{2}{\sqrt{\pi}} \int_q^\infty e^{-u^2/2} du.$$

(This function is widely tabulated and directly available in most computational softwares.)

We now consider scalar kernels that correspond to differential operators that are polynomial in the Laplacian. Using the inverse Fourier form, they correspond to kernels given by $K(x, y) = \Gamma(x - y)$ with

$$\Gamma(z) = \int_{\mathbb{R}^d} \frac{e^{iz^T \xi}}{P(|\xi|^2)} d\xi$$

for some polynomial P such that $P(t) > 0$ for $t \geq 0$. Of particular interest is the case $P(t) = (1 + t)^k$ for some positive integer k, which corresponds to the operator $(\mathrm{Id} - \Delta)^k$, because the associated kernel can be explicitly computed, at least in odd dimensions, which we now assume. Note that $1/P(|\xi|^2)$ must be integrable, which in this particular case means $k \geq (d + 1)/2$.

To compute Γ, we can assume (by rotation invariance) that z is on the positive side of the first coordinate axis, i.e., $z = (|z|, 0, \ldots, 0)$. Write $\xi = (t, \eta)$, with $t \in \mathbb{R}$ and $\eta \in \mathbb{R}^{d-1}$ so that

$$\Gamma(z) = \int_{-\infty}^{+\infty} e^{it|z|} \int_{\mathbb{R}^{d-1}} (1 + t^2 + |\eta|^2)^{-k} d\eta dt.$$

Making the change of variable $\eta = \sqrt{1+t^2}\zeta$ (so that $d\eta = (1+t^2)^{(d-1)/2}d\zeta$) expresses $\Gamma(z)$ as the product of two integrals,

$$\Gamma(z) = \int_{-\infty}^{+\infty} \frac{e^{it|z|}}{(1+t^2)^{k-(d-1)/2}} dt \int_{\mathbb{R}^{d-1}} (1+|\zeta|^2)^{-k} d\zeta.$$

The second integral is a constant $c_d(k)$, which can be explicitly computed, but we will not need the exact value here; for most applications, it suffices to know the expression of the kernel up to a multiplicative constant (and translation-invariant kernels are often normalized so that $K(x,x) = 1$). The first integral is related to the Basset function (or modified Bessel function of the third kind, see [223]), which we will denote by $\boldsymbol{K}_\nu(t)$, via the formula

$$\int_{-\infty}^{+\infty} \frac{e^{it|z|}}{(1+t^2)^{k-(d-1)/2}} dt \propto t^{k-d}\boldsymbol{K}_{k-d}(t).$$

It is known analytically for odd values of d, and can be evaluated numerically when d is even. In the odd case, using for example the method of residues [249], one shows the following lemma, which we give without proof.

Lemma 8.23 *If $c \geq 1$, we have*

$$\int_{-\infty}^{+\infty} \frac{e^{it|z|}}{(1+t^2)^{c+1}} dt = \frac{\pi e^{-|z|}}{4^c c!} \sum_{l=0}^{c} a(c,l)|z|^l \tag{8.17}$$

with $a(c,l) = 2^l(2c-l)\dots(c+1-l)/l!$.

Ignoring the constants, this yields the kernel (letting $c = k - (d+1)/2$, and normalizing K_c so that $K_c(x,x) = 1$)

$$K_c(x,y) = e^{-|x-y|} \sum_{l=0}^{c} b(c,l)|x-y|^l,$$

with $b(c,l) = a(c.l)/a(c,0)$. For $c = 0$, $K_0(x,y) = \exp(-|x-y|)$ is called the Laplacian or Abel's kernel. From Lemma 8.23, we get

$$\begin{aligned}
K_1(x,y) &= (1+|x-y|)\exp(-|x-y|)\\
K_2(x,y) &= (1+|x-y|+|x-y|^2/3)\exp(-|x-y|)\\
K_3(x,y) &= (1+|x-y|+2|x-y|^2/15+|x-y|^3/15)\exp(-|x-y|)\\
K_4(x,y) &= (1+|x-y|+3|x-y|^2/7+2|x-y|^3/21+|x-y|^4/105)\\
&\quad \exp(-|x-y|).
\end{aligned}$$

The resulting family are often called Matérn kernels. Note that K_1 is differentiable (with respect to each variable), and K_2 is twice differentiable. More generally K_c

has c derivatives, still with respect to each variable. As always, the kernel can be scaled, replacing x and y by x/σ and y/σ. For a given c, this corresponds (up to a multiplicative constant) to the operator

$$\mathbb{L}_c = (\mathrm{Id} - \sigma\Delta)^{c+(d+1)/2}.$$

The previous collection of kernels can be extended by the following series of combination rules. It is indeed obvious that the addition of two kernels is a kernel, as is the multiplication of a kernel by a positive number. A kernel can also be scaled by a positive factor: $K(x, y) \to K(x/a, y/a)$. The composition of two kernels is also a kernel, i.e.,

$$(K_1 * K_2)(x, y) = \int K_1(x, z)K_2(z, y)dz.$$

Also, the direct multiplication of two kernels is a kernel (i.e., $(K_1K_2)(x, y) := K_1(x, y)K_2(x, y)$). So, for example, in dimensions 1 or 3, the kernel defined by $K(x, y) = \gamma(|x - y|^2)$ with

$$\gamma(t) = (1 + \sqrt{t})e^{-\sqrt{t}-t/2}$$

is a valid kernel, since it is the direct multiplication of the Gaussian kernel and a Matérn kernel.

8.4 Mercer's Theorem

The interest of the discussion above lies in the fact that it makes it possible to define the Hilbert space V from a positive definite kernel. We gave a description of kernels using Fourier transforms, but another way to achieve this (in particular when Ω is bounded) is by using Mercer's theorem [246], which we cite without proof.

Theorem 8.24 *Let $K : \Omega \times \Omega \to \mathbb{R}$ be a continuous, positive definite kernel, such that*

$$\int_{\Omega\times\Omega} K(x, y)^2 dxdy < \infty.$$

Then, there exists an orthonormal sequence of functions in $L^2(\Omega, \mathbb{R})$, $\varphi_1, \varphi_2, \ldots$ and a decreasing sequence (ρ_n) which tends to 0 as n tends to ∞ such that

$$K(x, y) = \sum_{n=1}^{\infty} \rho_n \varphi_n(x)\varphi_n(y).$$

Note that a kernel satisfying the conditions of Mercer's theorem cannot be translation-invariant over $\Omega = \mathbb{R}^d$ because $\int K(x, y)^2 dy$ would be independent of x (and therefore not integrable) in that case.

Let the conditions of Mercer's theorem be true, and define the Hilbert space V by

$$V = \left\{ v \in L^2(\Omega, \mathbb{R}) : \sum_{n=1}^{\infty} \rho_n^{-1} \langle v\,,\ \varphi_n \rangle_{L^2}^2 < \infty \right\}.$$

Define, for $v, w \in V$:

$$\langle v\,,\ w \rangle_V = \sum_{n=1}^{\infty} \rho_n^{-1} \langle v\,,\ \varphi_n \rangle_{L_2} \langle w\,,\ \varphi_n \rangle_{L_2}.$$

Notice that, for $v \in V$, the series

$$v(x) = \sum_{n=1}^{\infty} \langle v\,,\ \varphi_n \rangle_{L^2} \varphi_n(x)$$

is pointwise convergent, since

$$\left(\sum_{n=p}^{m} \langle v\,,\ \varphi_n \rangle_{L^2} \varphi_n(x) \right)^2 \leq \sum_{n=p}^{m} \rho_n^{-1} \langle v\,,\ \varphi_n \rangle_{L^2}^2 \sum_{n=p}^{m} \rho_n (\varphi_n(x))^2$$

and both terms in the upper bound can be made arbitrarily small (recall that $\sum_n \rho_n \varphi_n(x)^2 = K(x, x) < \infty$). Similarly,

$$v(x) - v(y) = \sum_{n=1}^{\infty} \langle v\,,\ \varphi_n \rangle_{L^2} (\varphi_n(x) - \varphi_n(y))$$

so that

$$\begin{aligned} (v(x) - v(y))^2 &\leq \sum_{n=1}^{\infty} \rho_n^{-1} \langle v\,,\ \varphi_n \rangle_{L^2}^2 \sum_{n=1}^{\infty} \rho_n (\varphi_n(x) - \varphi_n(y))^2 \\ &= \|v\|_V^2 \left(K(x, x) - 2K(x, y) + K(y, y) \right) \end{aligned}$$

and v is continuous. Then,

$$\langle \varphi_m\,,\ K(.\,, x) \rangle_{L^2} = \sum_{n=1}^{\infty} \rho_n \varphi_n(x) \langle \varphi_m\,,\ \varphi_n \rangle_{L^2} = \rho_m \varphi_m(x)$$

so that

$$\sum_{n=1}^{\infty} \rho_n^{-1} \langle \varphi_n\,,\ K(\cdot, x) \rangle_{L^2}^2 = \sum_{n=1}^{\infty} \rho_n \varphi_n(x)^2 = K(x, x) < \infty,$$

which implies $K(\cdot, x) \in V$ and a similar computation shows that

$$\big\langle K(\cdot, x)\,,\ K(\cdot, y)\big\rangle_V = K(x, y),$$

so that K is reproducing.

Finally, if $v \in V$,

$$\begin{aligned}\big\langle v\,,\ K(\cdot, x)\big\rangle_V &= \sum_{n=1}^{\infty} \rho_n^{-1}\big\langle v\,,\ \varphi_n\big\rangle_{L^2}\big\langle \varphi_n\,,\ K(\cdot, c)\big\rangle_{L^2} \\ &= \sum_{n=1}^{\infty} \big\langle v\,,\ \varphi_n\big\rangle_{L^2}\varphi_n(x) = v(x),\end{aligned}$$

so that $K(\cdot, x)$ is the Riesz representation of the evaluation functional on V.

8.5 Thin-Plate Interpolation

Thin-plate theory corresponds to the situation in which the operator $\mathbb{L}$ is some power of the Laplacian. As a tool for shape analysis, it was originally introduced in [42]. Consider the following bilinear form:

$$\big\langle f\,,\ g\big\rangle_{\mathbb{L}} = \int_{\mathbb{R}^2} \Delta f \Delta g dx = \int_{\mathbb{R}^2} f \Delta^2 g dx, \tag{8.18}$$

which corresponds to the operator $\mathbb{L} = \Delta^2$.

We need to define it on a somewhat unusual Hilbert space. We consider the Beppo Levi space $\mathcal{H}_1$ of all functions in $\mathbb{R}^d$ with square integrable second derivatives, which have a bounded gradient at infinity. In this space, $\|f\|_{\mathbb{L}} = 0$ is equivalent to the fact that f is affine, i.e., $f(x) = a^T x + b$, for some $a \in \mathbb{R}^d$ and $b \in \mathbb{R}$. The Hilbert space we consider is the space $\mathcal{H}$ of equivalent classes of functions modulo the addition of affine functions, namely

$$[f] = \big\{g : g(x) = f(x) + a^T x + b, a \in \mathbb{R}^d, b \in \mathbb{R}\big\}$$

for $f \in \mathcal{H}_1$. Obviously, the norm associated to (8.18) is constant over the set $[f]$, and $\|f\|_{\mathbb{L}} = 0$ if and only if $[f] = [0]$. One can then define

$$\big\langle [f]\,,\ [g]\big\rangle_{\mathcal{H}} = \big\langle f\,,\ g\big\rangle_{\mathbb{L}}$$

without ambiguity.

This space also has a kernel, although the analysis has to be different from what we have done previously, since the evaluation functional is not defined on

$\mathcal{H}$ (where functions are only known up to the addition of an affine term). However, the following is true [192]. Let $U(r) = (1/8\pi)r^2 \log r$ if the dimension, d, is 2, and $U(r) = (1/16\pi)r^3$ for $d = 3$. Then, for all $u \in \mathcal{H}_1$, there exist $a_u \in \mathbb{R}^d$ and $b_u \in \mathbb{R}$ such that

$$u(x) = \int_{\mathbb{R}^d} U(|x - y|)\Delta^2 u(y)dy + a_u^T x + b_u.$$

We will denote by $U(\cdot, x)$ the function $y \mapsto U(|x - y|)$.

The spline interpolation problem must also be addressed in a different way in this context. Fix, as in Sect. 8.1.1, landmarks $x_1, \ldots, x_N$ and scalar constraints $(c_1, \ldots, c_N)$. Again, the constraint $h(x_i) = c_i$ has no meaning in $\mathcal{H}$, but the constraint $\langle U(\cdot - x_i)\,,\, h\rangle_{\mathcal{H}} = c_i$ does, and means that there exist a_h and b_h such that

$$h(x_i) + a_h^T x_i + b_h = c_i,$$

i.e., h satisfies the constraints up to an affine term. The corresponding interpolation problem is: minimize $\|h\|_{\mathbb{L}}$ under the constraint that there exist a, b such that $h(x_i) + a^T x_i + b = c_i$, $i = 1, \ldots, N$.

Define, as before, $S_{ij} = U(|x_i - x_j|)$. The function h that is optimal under these constraints must therefore take the form

$$h(x) = \sum_{i=1}^{N} \alpha_i U(|x - x_i|) + a^T x + b.$$

Replacing h by its expression in terms of the α's, a and b yields the finite-dimensional problem: minimize

$$\alpha^T S\alpha$$

under the constraint $S\alpha + Q\gamma = c$ with $\gamma = (a_1, \ldots, a_d, b)^T$ (with size $(d + 1) \times 1$) and Q, with size $N \times (d + 1)$ given by (letting $x_i = (x_i^1, \ldots, x_i^d)$):

$$Q = \begin{pmatrix} x_1^1 & \cdots & x_1^d & 1 \\ \vdots & \vdots & \vdots & \\ x_N^1 & \cdots & x_N^d & 1 \end{pmatrix}.$$

The optimal (α, γ) can be computed by identifying the gradient to 0. One obtains

$$\hat{\gamma} = \left(Q^T S^{-1} Q\right)^{-1} Q^T S^{-1} c$$

and $\hat{\alpha} = S^{-1}(c - Q\hat{\gamma})$.

The inexact matching problem simply consists in minimizing

$$\|h\|_{\mathbb{L}}^2 + \lambda \sum_{i=1}^{N} (h(x_i) + a^T x_i + b - c_i)^2$$

with respect to h *and* a, b. It can be reduced to minimizing

$$\alpha^T S\alpha + \lambda(S\alpha + a^T x_i + b - c)^T (S\alpha + a^T x_i + b - c)\,,$$

and the solution is provided by the same formulas, simply replacing S by $S_\lambda = S + (1/\lambda)\mathrm{Id}$.

When the function h and the c_i take d-dimensional values (e.g., correspond to displacements), the above computation has to be applied to each coordinate, which simply corresponds to using a diagonal operator, each component equal to Δ^2, in the definition of the dot product. This is equivalent to using the diagonal scalar kernel associated to U.

8.6 Asymptotically Affine Kernels

We return to vector fields, and discuss how affine components can be combined with any kernel, beyond the thin-plate spline approach of the previous section. We assume here that $\Omega = \mathbb{R}^d$. We would like to consider spaces V that contain vector fields with an affine behavior at infinity. Note that the spaces V that we have considered so far, either by completion of C^∞ compactly supported functions or using kernels defined by Fourier transforms, only contain functions that vanish at infinity. We recall the definition of a function that vanishes at infinity:

Definition 8.25 A function $f : \Omega \to \mathbb{R}^d$ is said to vanish at infinity if and only if, for all $\varepsilon > 0$, there exists an $A > 0$ such that $|f(x)| < \varepsilon$ whenever $x \in \Omega$ and $|x| > A$.

Here, we let V be a Hilbert space of vector fields that vanish at infinity and define

$$V_{aff} = \left\{ w : \exists w_0 \in V,\, A \in \mathcal{M}_d(\mathbb{R}) \text{ and } b \in \mathbb{R}^d \text{ with } w(x) = w_0(x) + Ax + b \right\}.$$

We have the following important fact:

Proposition 8.26 *If V is a Hilbert space of continuous vector fields that vanish at infinity, then the decomposition $w(x) = w_0(x) + Ax + b$ for $w \in V_{aff}$ is unique.*

Proof Using differences, it suffices to prove this for $w = 0$, and so, if w_0, A and b are such that, for all x, $w_0(x) + Ax + b = 0$, then, for any fixed $x \neq 0$ and $t > 0$, $w_0(tx) + tAx + b = 0$ so that $Ax = -(w_0(tx) + b)/t$. Letting t tend to infinity, we get $Ax = 0$ for all x, so that $A = 0$. Now, for all x, we get $b = -w_0(x)$, which implies $b = 0$ (since w_0 vanishes at infinity), and therefore also $w_0 = 0$. □

So we can speak of the affine part (A_w, b_w) of an element $w \in V_{aff}$. Letting also $q_w(x) = w(x) - A_w x - b_w$, we can extend the inner product in V to define

$$\left\langle w\,,\, \tilde{w} \right\rangle_{V_{aff}} = \left\langle q_w\,,\, q_{\tilde{w}} \right\rangle_V + \left\langle A_w\,,\, A_{\tilde{w}} \right\rangle + \left\langle b_w\,,\, b_{\tilde{w}} \right\rangle,$$

where $\langle A\ ,\ \tilde{A}\rangle$ is some inner product between matrices (e.g., $\mathrm{trace}(A^T\tilde{A})$) and $\langle b\ ,\ \tilde{b}\rangle$ some inner product between vectors (e.g., $b^T\tilde{b}$). With this product V_{aff} is obviously a Hilbert space. We want to compute its kernel, K_{aff}, as a function of the kernel K of V (assuming that V is reproducing). Given $x \in \mathbb{R}^d$ and $a \in \mathbb{R}^d$, we need to express $a^T w(x)$ in the form $\langle K_{aff}(\cdot, x)a\ ,\ w\rangle_{V_{aff}}$. Using the decomposition, we have

$$\begin{aligned} a^T w(x) &= a^T q_w(x) + a^T A_w x + a^T b_w \\ &= \langle K(\cdot, x)a\ ,\ q_w\rangle_V + \langle (ax^T)^\sharp\ ,\ A_w\rangle + \langle a^\sharp\ ,\ b_w\rangle \end{aligned}$$

where, for a matrix M, we define $M^\sharp$ by the identity $\langle M^\sharp\ ,\ A\rangle = \mathrm{trace}(M^T A)$ for all A and for a vector z, $z^\sharp$ is such that $\langle z^\sharp\ ,\ b\rangle = z^T b$ for all b. From the definition of the extended inner product, we can define and compute

$$K_{aff}(y, x)a = K(y, x)a + (ax^T)^\sharp y + a^\sharp.$$

In particular, when $\langle A\ ,\ \tilde{A}\rangle = \lambda\,\mathrm{trace}(A^T\tilde{A})$ and $\langle b\ ,\ \tilde{b}\rangle = \mu b^T\tilde{b}$, we get

$$K_{aff}(y, x)a = K(y, x)a + ax^T y + a = (K(x, y) + (x^T y/\lambda + 1/\mu)\mathrm{Id}_d)a.$$

This provides an immediate extension of spline interpolation of vector fields which includes affine transformations, by just replacing K by K_{aff}. For example, exact interpolation with constraints $v(x_i) = c_i$ is obtained by letting

$$v(x) = \sum_{i=1}^{N} K_{aff}(x_i, x_j)\alpha_j,$$

the vectors α_j being obtained by solving the system

$$\sum_{j=1}^{N}(K(x_i, x_j)\alpha_j + (x_i^T x_j/\lambda + 1/\mu)\alpha_j) = c_i,\ \text{for } i = 1, \ldots, N. \tag{8.19}$$

The case $\lambda = \mu \to 0$ is particularly interesting, since this corresponds to relaxing the penalty on the affine displacement, and we obtain in this way an affine invariance similar to thin plates. More precisely, we have:

Proposition 8.27 *Let v^λ, A^λ and b^λ be the solutions of (8.19) when $\mu = \lambda$. Then v^λ converges to v^*, the unique solution of the affine invariant interpolation problem: minimize $\|v\|_V^2$ under the constraint that there exist A and b with $v(x_i) = c_i - Ax_i - b$ for $i = 1, \ldots, N$.*

Proof Indeed, v^λ, A^λ and b^λ are the unique solutions of the problem: minimize, with respect to v, A, b,

$$\|v\|_V^2 + \lambda(\|A\|^2 + \|b\|^2)$$

under the constraints $v(x_i) = c_i - Ax_i - b$ for $i = 1, \ldots, N$. Because $\|v^\lambda\|_V$ is always smaller that the norm of the optimal v for the problem without affine component, we know that it is bounded. Thus, to prove that v^λ converges to v^*, it suffices to prove that any weakly convergent subsequence of v^λ has v^* as a limit. So take such a sequence, v^{λ_n}, and let v^0 be its weak limit. Because $v \mapsto v(x)$ is a continuous linear form, we have $v^{\lambda_n}(x_i) \to v^0(x_i)$ for all i.

Consider the linear transformation $F : \mathbb{R}^{d^2+d} \to \mathbb{R}^{Nd}$ that maps (A, b) to $(Ax_i + b, i = 1, \ldots, N)$. Because $A^\lambda x_i + b^\lambda + v^\lambda(x_i) = c_i$, we know that $(c_i - v^\lambda(x_i), i = 1, \ldots N) \in \text{Range}(F)$, so that, passing to the limit in λ_n, we get the fact that $(c_i - v^0(x_i), i = 1, \ldots N) \in \text{Range}(F)$, which means that v^0 satisfies the condition of the limit problem. Moreover, for any λ, we have

$$\|v^*\|_V^2 + \lambda(\|A^*\|^2 + \|b^*\|^2) \geq \|v^\lambda\|_V^2 + \lambda(\|A^\lambda\|^2 + \|b^\lambda\|^2) \geq \|v^\lambda\|_V^2.$$

Since the liminf of the last term is larger than $\|v^0\|_V^2$, and the limit of the left-hand side is $\|v^*\|_V^2$, we can conclude that $v^0 = v^*$ since v^* is the unique solution of the limit problem. □

Chapter 9
Deformable Objects and Matching Functionals

9.1 General Principles

In the previous two chapters, we introduced and studied basic tools related to deformations and their mathematical representation using diffeomorphisms. In this chapter, we start investigating relations between deformations and the objects they affect, which we will call deformable objects, and discuss the variations of matching functionals, which are cost functions that measure the quality of the registration between two deformable objects.

Let Ω be an open subset of $\mathbb{R}^d$ and G a group of diffeomorphisms on Ω. Consider a set $\mathcal{I}$ of structures of interest, on which G has an action: for every I in $\mathcal{I}$ and every $\varphi \in G$, the result of the action of φ on I is denoted $\varphi \cdot I$ and is a new element of $\mathcal{I}$. This requires (see Sect. B.5) that $\mathrm{id} \cdot I = I$ and $\varphi \cdot (\psi \cdot I) = (\varphi \circ \psi) \cdot I$. Elements of $\mathcal{I}$ will be referred to as *deformable objects*.

A matching functional is based on a function $D : \mathcal{I} \times \mathcal{I} \to [0, +\infty)$ such that $D(I, I')$ measures the discrepancy between the two objects I and I', and is defined over G by

$$E_{I,I'}(\varphi) = D(\varphi \cdot I, I'). \tag{9.1}$$

So $E_{I,I'}(\varphi)$ measures the difference between the *target* object I' and the deformed one $\varphi \cdot I$. Because it is mapped onto the target by the deformation, the object I will often be referred to as the *template* (and $\varphi \cdot I$ as the deformed template).

Even if our discussion of matching principles and algorithms is rather extensive, and occupies a large portion of this book, the size of the literature, and our choice of privileging methods that implement diffeomorphic matching prevents us from providing an exhaustive account of the registration methods that have been proposed over the last few decades. The interested reader can refer to a few starting points in order to complement the presentation that is made here, including [12, 13, 22, 27, 28, 41, 42, 111, 125, 240, 244, 275], and textbooks such as [132, 139, 208, 214].

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_9

9.2 Differentiation with Respect to Diffeomorphisms

We will review, starting with the next section, a series of matching functionals that are adapted to different types of deformable objects (landmarks, images, curves, etc.). We will also compute the derivative of each of them with respect to the diffeomorphism φ.

We also introduce a special form of differential which is adapted to variational problems over diffeomorphisms. This shape, or Eulerian differential, as we will call it, is a standard tool in shape optimization [80], and we will interpret it later on as a gradient for a specific Riemannian metric over diffeomorphisms.

Recall that we have defined $\mathrm{Diff}^{p,\infty} = \mathrm{Diff}^{p,\infty}(\Omega)$ to be the set of diffeomorphisms ψ such that

$$\max(\|\psi - \mathrm{id}\|_{p,\infty}, \|\psi^{-1} - \mathrm{id}\|_{p,\infty}) < \infty.$$

We have also defined $\mathrm{Diff}_0^{p,\infty}$ as the subgroup of $\mathrm{Diff}^{p,\infty}$ whose elements converge to the identity at infinity.

Definition 9.1 A function $\varphi \mapsto U(\varphi)$ is (p,∞)-compliant if it is defined for all φ in $\mathrm{Diff}_0^{p,\infty}$.

A (p,∞)-compliant U is locally (p,∞)-Lipschitz if, for all $\varphi \in \mathrm{Diff}_0^{p,\infty}$, there exist positive numbers $\varepsilon(\varphi)$ and $C(\varphi)$ such that

$$|U(\psi) - U(\tilde{\psi})| \leq C(\varphi)\|\psi - \tilde{\psi}\|_{p,\infty}$$

whenever ψ and $\tilde{\psi}$ are diffeomorphisms such that

$$\max(\|\psi - \varphi\|_{p,\infty}, \|\tilde{\psi} - \varphi\|_{p,\infty}) < \varepsilon(\varphi).$$

Note that a (p,∞)-compliant (resp. locally Lipschitz) U is (q,∞)-compliant (resp. locally Lipschitz) for any q larger than p.

Because $\mathrm{Diff}_0^{p,\infty}$ is an open subset of $\mathrm{id} + C_0^p(\Omega, \mathbb{R}^d)$, both Gâteaux and Fréchet derivatives are well defined for functions defined on this set (see Sect. C.1). In the following, whenever we speak of a derivative (without a qualifier), this will always mean in the strong (Fréchet) sense. A function U is C^1 on $\mathrm{Diff}_0^{p,\infty}$ if and only if U is Fréchet differentiable and $dU(\psi)$ is continuous in ψ, which is equivalent (by Proposition C.5) to U being Gâteaux differentiable and $dU(\psi)$ continuous in ψ. Note also that U being C^1 implies that U is (p,∞)-Lipschitz.

Using the group structure of $\mathrm{Diff}_0^{p,\infty}$, we can define another type of differential using the infinitesimal action of vector fields. If V is an admissible vector space and $v \in V$, we will denote by φ_{0t}^v the flow associated to the equation

$$\partial_t y = v(y).$$

Note that this is the same notation as the flow associated to a differential equation $\partial_t y = v(t, y)$, where v is now a time-dependent vector field. This is not a conflict of notation if one agrees to identify vector fields, v, in V and the associated constant time-dependent vector field defined by $\tilde{v}(t, \cdot) = v$ for all t.

Definition 9.2 Let V be an admissible Hilbert space continuously embedded in $C_0^p(\Omega, \mathbb{R}^d)$ (so that $\mathrm{Diff}_V \subset \mathrm{Diff}_0^{p,\infty}$). We say that a (p, ∞)-compliant function U over diffeomorphisms has an Eulerian differential in V at ψ if there exists a linear form $\bar{\partial}U(\psi) \in V^*$ such that, for all $v \in V$,

$$\big(\bar{\partial}U(\psi) \;\big|\; v\big) = \partial_\varepsilon U(\varphi_{0\varepsilon}^v \circ \psi)_{|\varepsilon=0}. \tag{9.2}$$

If the Eulerian differential exists, the V-Eulerian gradient of U at ψ, denoted $\overline{\nabla}^V U(\varphi) \in V$, is defined by

$$\big\langle \overline{\nabla}^V U(\varphi)\,,\; v\big\rangle_V = \big(\bar{\partial}U(\varphi) \;\big|\; v\big). \tag{9.3}$$

In this case, $\overline{\nabla}^V U(\varphi) = \mathbb{K}\bar{\partial}U(\varphi)$, where $\mathbb{K}$ is the kernel operator of V.

The following proposition indicates when Eq. (9.2) remains valid with time-dependent vector fields v.

Proposition 9.3 *Let V be an admissible Hilbert space continuously embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$. Let V and U satisfy the hypotheses of Definition 9.2. If U is (p, ∞)-locally Lipschitz and has a V-Eulerian differential at ψ and if $v(t, \cdot)$ is a time-dependent vector field such that*

$$\lim_{\varepsilon \to \infty} \frac{1}{\varepsilon} \int_0^\varepsilon \|v(t, \cdot) - v(0, \cdot)\|_V dt = 0, \tag{9.4}$$

then

$$\big(\bar{\partial}U(\psi) \;\big|\; v(0, \cdot)\big) = \partial_\varepsilon U(\varphi_{0\varepsilon}^v \circ \psi)_{|\varepsilon=0}. \tag{9.5}$$

Proof Letting $v_0 = v(0, \cdot)$, we need to prove that

$$\frac{1}{\varepsilon}(U(\varphi_{0\varepsilon}^v \circ \psi) - U(\varphi_{0\varepsilon}^{v_0} \circ \psi)) \to 0$$

as $\varepsilon \to 0$. From Proposition 7.4, we know that if ψ, φ, $\tilde{\varphi}$ are in $\mathrm{Diff}_0^{p,\infty}$, there exists a constant $C_p(\psi)$ such that

$$\|\varphi \circ \psi - \tilde{\varphi} \circ \psi\|_{p,\infty} \leq C_p(\psi)\|\varphi - \tilde{\varphi}\|_{p,\infty}.$$

Now, since U is Lipschitz, we have, for small enough ε,

$$\begin{aligned}|U(\varphi^v_{0\varepsilon} \circ \psi) - U(\varphi^{v_0}_{0\varepsilon} \circ \psi)| &\leq C(\psi)\|\varphi^v_{0\varepsilon} \circ \psi - \tilde{\varphi}^v_{0\varepsilon} \circ \psi\|_{p,\infty} \\ &\leq C(\psi)C_p(\psi)\|\varphi^v_{0\varepsilon} - \varphi^{v_0}_{0\varepsilon}\|_{p,\infty} \\ &\leq C(\psi)C_p(\psi)\tilde{C}(v_0)\int_0^\varepsilon \|v(t,\cdot) - v_0\|_V dt,\end{aligned}$$

where $\tilde{C}(v_0)$ depends on $\|v_0\|_{p+1,\infty}$ and can be derived from Eq. (7.16), Noting that $\|v\|_{\mathcal{X}^{p+1,1},\varepsilon} \leq \varepsilon(C'' + \|v_0\|_{p+1,\infty})$ for small enough ε, this proves the proposition. □

Note also that, if U is C^1, then the chain rule implies that

$$\left(\bar{\partial}U(\varphi) \mid v\right) = (dU(\varphi) \mid v \circ \varphi). \tag{9.6}$$

To the Eulerian gradient of U, we associate a "gradient descent" process (that we will formally interpret as a Riemannian gradient descent for a suitable metric in Sect. 11.4.3) which generates a time-dependent element of G by setting

$$\partial_t \varphi(t,x) = -\overline{\nabla}^V U(\varphi(t))(\varphi(t,x)). \tag{9.7}$$

As long as $\int_0^t \left\|\overline{\nabla}^V U(\varphi(s))\right\|_V ds$ is finite, this generates a time-dependent element of Diff_V. This therefore provides an evolution within the group of diffeomorphisms, an important property. Assuming that Proposition 9.3 applies at time t (e.g., if U is C^1), we can write

$$\partial_t U(\varphi(t)) = \left\langle \overline{\nabla}^V U(\varphi(t)),\, \partial_t \varphi \right\rangle_V = -\left\|\overline{\nabla}^V U(\varphi(t))\right\|_V^2,$$

so that $U(\varphi(t))$ decreases with time.

9.3 Relation with Matching Functionals

As pointed out in the introduction, matching functionals take the form

$$U(\varphi) = U_I(\varphi) = Z(\varphi \cdot I), \tag{9.8}$$

where I is a fixed deformable object for some function Z (e.g., $Z(I) = D(I, I_1)$ for a fixed I_1). Using the group action property, we have

$$U_I(\psi) = U_{\varphi \cdot I}(\psi \circ \varphi^{-1}).$$

Using this property and the fact that the mapping $\psi \circ \varphi^{-1}$ is smooth (infinitely differentiable) from $\mathrm{Diff}_0^{p,\infty}$ onto itself, we find that if U_I is Gâteau or Fréchet

differentiable at $\psi = \mathrm{id}$ for any $I \in \mathcal{I}$, then it is differentiable at all $\psi \in \mathrm{Diff}_0^{p,\infty}$, and

$$(dU_I(\varphi) \mid h) = \big(dU_{\varphi \cdot I}(\mathrm{id}) \,\big|\, h \circ \varphi^{-1}\big).$$

A similar statement holds for the Eulerian differential, with

$$\big(\bar{\partial} U_I(\psi) \,\big|\, v\big) = \big(\bar{\partial} U_{\psi \cdot I}(\mathrm{id}) \,\big|\, v\big).$$

Notice that when U is differentiable at $\psi = \mathrm{id}$, then $\bar{\partial} U(\mathrm{id}) = dU(\mathrm{id})$. Finally, if we assume that $\mathcal{I}$ is itself a Banach space (or an open subset of a Banach space), that Z is differentiable and that the action $R_I : \varphi \mapsto \varphi \cdot I$ is also differentiable, we have, using the chain rule

$$(dU_I(\varphi) \mid v) = (dZ(\varphi \cdot I) \mid dR_I(\varphi)v)$$

or $dU_I(\varphi) = dR_I(\varphi)^* dZ(\varphi \cdot I)$. At $\varphi = \mathrm{id}$, $dR_I(\mathrm{id})v$ is the *infinitesimal action* of v on I, which we denote by $v \cdot I$, so that

$$(dU_I(\mathrm{id}) \mid v) = (dZ(I) \mid v \cdot I)$$

and

$$\big(\bar{\partial} U_I(\psi) \,\big|\, v\big) = (dZ(\psi \cdot I) \mid v \cdot (\psi \cdot I)).$$

We now present a series of matching problems, involving different types of deformable objects. In each case, we will introduce adapted matching functionals and compute their differentials. As just remarked, derivatives with respect to the diffeomorphisms can all be derived from that of the function Z, on which we will, whenever possible, focus the computations.

9.4 Labeled Point Matching

The simplest way to represent a visual structure is with configurations of labeled points, or *landmarks* attached to the structure. Anatomical shapes or images are typical examples of structures on which landmarks can be easily defined; this includes specific locations in faces (corners of the eyes, tip of the nose, etc.), fingertips for hands, apex of the heart, etc. Many man-made objects, like cars or other vehicles, can be landmarked too. Finally, landmarks can represent the centers of simple objects, like cells in biological images.

In the labeled point-matching problem, objects are ordered collections of N points $x_1, \ldots, x_N \in \Omega$, where N is fixed. Diffeomorphisms act on such objects by:

$$\varphi \cdot (x_1, \ldots, x_N) = (\varphi(x_1), \ldots, \varphi(x_N)). \tag{9.9}$$

The landmark-matching problem is not to find correspondences between two objects, say $I = (x_1, \dots, x_N)$ and $I' = (x'_1, \dots, x'_N)$, since we know that x_i and x'_i are homologous, but to extrapolate these correspondences to the rest of the space.

Here we can take $\mathcal{I} = (\mathbb{R}^d)^N$, or, if one restricts to distinct landmarks, the open subset

$$\mathcal{I} = \left\{(x_1, \dots, x_N) \in (\mathbb{R}^d)^N, x_i \neq x_j \text{ if } i \neq j\right\}.$$

For $I = (x_1, \dots, x_N)$, the action $\varphi \mapsto \varphi \cdot I$ is C^1 on $\mathrm{Diff}_0^{p,\infty}$ for any $p \geq 0$ (it is the restriction of a linear map, with $|\varphi \cdot I|^2 \leq \sqrt{N}\|\varphi\|_\infty$). The simplest matching functional that we can consider for this purpose is associated with

$$Z(I) = |I - I'|^2 = \sum_{k=1}^{N} |x_k - x'_k|^2$$

with $(dZ(I) \mid h) = 2(I - I')^T h$ (considering I, I' and h as dN-dimensional column vectors). We have

$$U_I(\varphi) = E_{I,I'}(\varphi) = |I' - \varphi \cdot I|^2 = \sum_{i=1}^{N} \left|x'_i - \varphi(x_i)\right|^2, \tag{9.10}$$

$$(dU_I(\varphi) \mid v) = 2\sum_{i=1}^{N} (\varphi(x_i) - x'_i)^T v(x_i). \tag{9.11}$$

This can be written as

$$dU_I(\varphi) = 2\sum_{i=1}^{N} (\varphi(x_i) - x'_i)\delta_{x_i}.$$

From (9.6), we have

$$\left(\bar{\partial} U_I(\varphi) \,\middle|\, h\right) = 2\sum_{i=1}^{N} (\varphi(x_i) - x'_i)^T h \circ \varphi(x_i) \tag{9.12}$$

or

$$\bar{\partial} U_I(\varphi) = 2\sum_{i=1}^{N} (\varphi(x_i) - x'_i)\delta_{\varphi(x_i)},$$

and (9.3) gives

$$\overline{\nabla}^V U_I(\varphi) = 2\sum_{i=1}^{N} K(\cdot, \varphi(x_i))(\varphi(x_i) - x'_i). \tag{9.13}$$

The gradient descent algorithm (9.7) takes a very simple form:

$$\partial_t \varphi(t,x) = -2\sum_{i=1}^{N} K(\varphi(t,x),\varphi(t,x_i))(\varphi(t,x_i) - x_i'). \tag{9.14}$$

This system can be solved in two steps: let $y_i(t) = \varphi(t, x_i)$. Applying (9.14) at $x = x_j$ yields

$$\partial_t y_j = -2\sum_{i=1}^{N} K(y_j, y_i)(y_i - x_i').$$

This is a differential system in $y_1, \ldots, y_N$. The first step is to solve it with initial conditions $y_j(0) = x_j$. Once this is done, the extrapolated value of $\varphi(t,x)$ for a general x is the solution of the differential equation

$$\partial_t y = -2\sum_{i=1}^{N} K(y, y_i)(y_i - x_i')$$

initialized at $y(0) = x$. Figure 9.1 gives an example obtained by running this procedure, providing an illustration of the impact of the choice of the kernel for the solution. The last panel in Fig. 9.1 also shows the limitations of this algorithm, in the sense that it is trying to move the points in the direction of their targets at each step, while a more indirect path can sometimes be found generating less distortion (these results should be compared to Fig. 10.1 in Chapter 10).

9.5 Image Matching

Images, or more generally multivariate functions, are also important and widely used instances of deformable objects. They correspond to functions I defined on Ω with values in $\mathbb{R}$. Diffeomorphisms act on them by:

$$(\varphi \cdot I)(x) = I(\varphi^{-1}(x))$$

for $x \in \Omega$. Fixing two such functions I and I', the simplest matching functional which can be considered is the squared L^2 norm of the difference $Z(I) = \|I - I'\|_2^2$, yielding

$$U_I(\varphi) = E_{I,I'}(\varphi) = \int_\Omega \left|I \circ \varphi^{-1}(x) - I'(x)\right|^2 dx. \tag{9.15}$$

We will need the derivative of the mapping Inv : $\varphi \mapsto \varphi^{-1}$. Considering Inv as a mapping from $\mathrm{Diff}_0^{p+1,\infty}$ to $\mathrm{Diff}_0^{p,\infty}$, it is given by (see Proposition 7.8)

$$d\mathrm{Inv}(\varphi)h = -\left(d\varphi \circ \varphi^{-1}\right)^{-1} h \circ \varphi^{-1} = -d(\varphi^{-1})h \circ \varphi^{-1}. \tag{9.16}$$

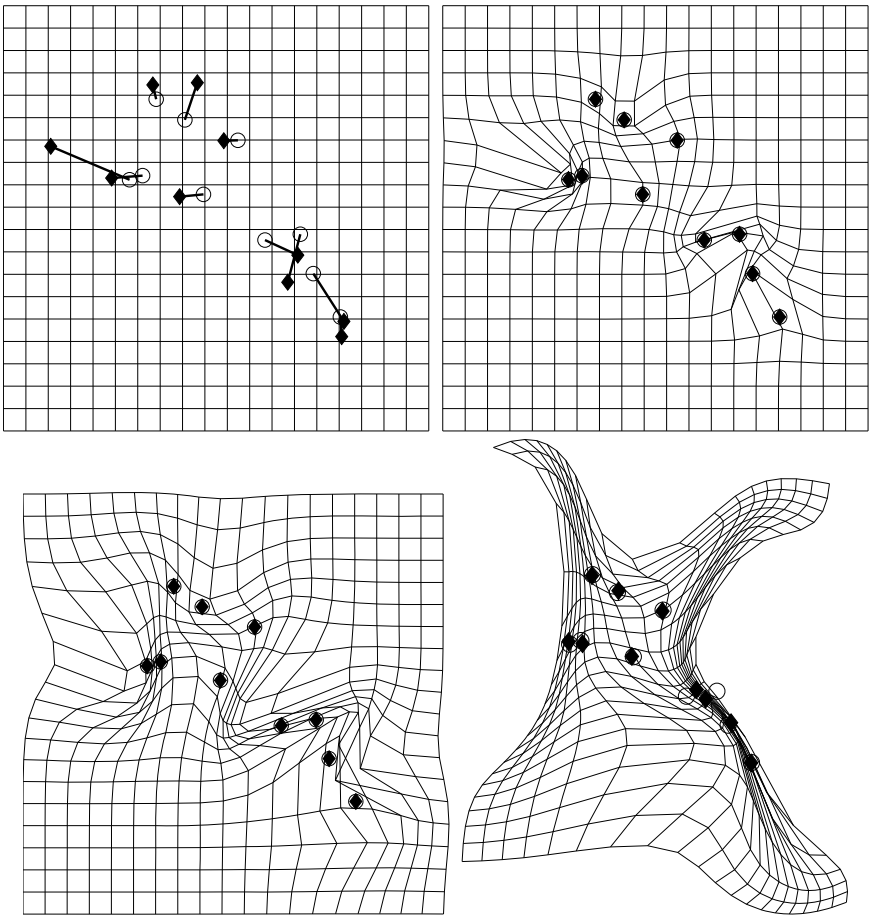

Fig. 9.1 Greedy landmark matching. Implementation of the gradient descent algorithm in (9.14), starting with $\varphi = \mathrm{id}$, for the correspondences depicted in the upper-left image (diamonds moving to circles). The following three images provide the result after numerical convergence for Gaussian kernels $K(x, y) = \exp(-|x - y|^2/2\sigma^2)\mathrm{Id}$ with $\sigma = 1, 2, 4$ in grid units. Larger σ induce increasing smoothness in the final solution, and deformations affecting a larger part of the space. As seen in the figure for $\sigma = 4$, the evolution can result in huge deformations

Similarly Inv is C^k from $\mathrm{Diff}_0^{p+k,\infty}$ to Diff_0^p.

We now compute the derivative of U_I under the assumption that I' is square integrable, I is compactly supported (on some set Q_I) and continuously differentiable. One can relax the differentiability assumption on I (considering, for example, piecewise smooth images), but the analysis is much more difficult, and we refer the reader to results in [293–295] for more details. Define

$$\tilde{U}_I(\varphi) = \int_\Omega \left| I \circ \varphi(x) - I'(x) \right|^2 dx$$

so that $U_I = \tilde{U}_I \circ \mathrm{Inv}$. Fixing $\varphi \in \mathrm{Diff}_0^{p+1}$, $h \in C_0^{p+1}(\Omega, \mathbb{R}^d)$ and letting $\varphi_\varepsilon = \varphi + \varepsilon h$ for $|\varepsilon| \leq 1$, we have

$$\partial_\varepsilon \tilde{U}_I(\varphi_\varepsilon) = 2 \int_\Omega (I \circ \varphi_\varepsilon(x) - I'(x)) \nabla I \circ \varphi_\varepsilon(x)^T h(x)\, dx.$$

The integrand in the right-hand side is dominated by the integrable upper bound $(\|I\|_\infty + |I'(x)|)\|\nabla I\|_\infty \|h\|_\infty \mathbf{1}_{\tilde{Q}}(x)$, where $\tilde{Q}$ is any compact set that contains $\varphi_\varepsilon^{-1}(Q_I)$ for $|\varepsilon| \leq 1$, which justifies the derivation. Taking $\varepsilon = 0$, we obtain the directional derivative of $\tilde{U}_I$, which is

$$\left(d\tilde{U}_I(\varphi) \,\middle|\, h\right) = 2 \int_\Omega (I \circ \varphi(x) - I'(x)) \nabla I \circ \varphi(x)^T h(x)\, dx.$$

Our hypotheses on I imply that $d\tilde{U}_I(\varphi)$ is continuous in $\varphi \in \mathrm{Diff}_0^{p,\infty}$ for any $p \geq 0$. Fixing φ and assuming $\|\varphi' - \varphi\|_\infty \leq 1$, we have

$$\begin{aligned}
&\left|\left(d\tilde{U}_I(\varphi) - d\tilde{U}_I(\varphi') \,\middle|\, h\right)\right| \\
&\quad \leq 2 \int_\Omega (I \circ \varphi(x) - I \circ \varphi'(x)) \nabla I \circ \varphi'(x)^T h(x)\, dx \\
&\qquad + 2 \int_\Omega (I \circ \varphi(x) - I'(x)) (\nabla I \circ \varphi(x) - \nabla I \circ \varphi'(x))^T h(x)\, dx \\
&\quad \leq 2|\tilde{Q}|\, \|I \circ \varphi - I \circ \varphi'\|_\infty \|\nabla I\|_\infty \|h\|_\infty \\
&\qquad + 2\sqrt{|\tilde{Q}|}\, \|I \circ \varphi - I'\|_2 \|\nabla I \circ \varphi - \nabla I \circ \varphi'\|_\infty \|h\|_\infty,
\end{aligned}$$

where $\tilde{Q}$ is a compact set containing all $(\varphi')^{-1}(Q_I)$ for $\|\varphi' - \varphi\|_\infty \leq 1$. The facts that I and ∇I are uniformly continuous on Q_I imply that $\|I \circ \varphi - I \circ \varphi'\|_\infty$ and $\|\nabla I \circ \varphi - \nabla I \circ \varphi'\|_\infty$ tend to 0 as $\|\varphi - \varphi'\|_\infty$ tends to 0. (We have denoted by $|\tilde{Q}|$ the Lebesgue measure of the set $\tilde{Q}$.)

As a composition of C^1 functions, we find that U_I is C^1 on $\mathrm{Diff}_0^{p+1,\infty}$ for any $p \geq 0$, with (applying the chain rule)

$$\begin{aligned}
&(dU_I(\varphi) \mid h) = \\
&\qquad -2 \int_\Omega (I \circ \varphi^{-1}(x) - I'(x)) \nabla I \circ \varphi^{-1}(x)^T d(\varphi^{-1}) h \circ \varphi^{-1}(x)\, dx. \qquad (9.17)
\end{aligned}$$

Notice that $\nabla (I \circ \varphi^{-1})^T = (\nabla I^T \circ \varphi^{-1}) d(\varphi^{-1})$, so that

$$(dU_I(\varphi) \mid h) = -2 \int_\Omega (I \circ \varphi^{-1}(x) - I'(x)) \nabla (I \circ \varphi^{-1})(x)^T h \circ \varphi^{-1}(x)\, dx$$

and we retrieve the formula

$$(dU_I(\varphi) \mid h) = \left(dU_{\varphi \cdot I}(\mathrm{id}) \,\middle|\, h \circ \varphi^{-1}\right).$$

The Eulerian derivative is given by

$$\begin{aligned}\big(\bar\partial U_I(\varphi) \;\big|\; v\big) &= (dU_I(\mathrm{id}) \mid v\circ\varphi) \\ &= -2\int_\Omega (I\circ\varphi^{-1}(x) - I'(x))\nabla(I\circ\varphi^{-1})(x)^T v(x)\,dx.\end{aligned}$$

We introduce a notation that will be used throughout this chapter and that generalizes the one given for point measures in Eq. (8.4). If μ is a measure on Ω and $z:\Omega\to\mathbb{R}^d$ a μ-measurable function, the vector measure $(z\mu)$ is the linear form over vector fields on Ω defined by

$$(z\mu \mid h) = \int_\Omega z^T h d\mu. \tag{9.18}$$

With this notation we can write

$$\bar\partial U_I(\varphi) = -2\big((I\circ\varphi^{-1} - I')\nabla(I\circ\varphi^{-1})\big)dx. \tag{9.19}$$

Notice also that, making a change of variable in (9.17), we have

$$dU_I(\varphi) = -2\big(\det(d\varphi)\,(I - I'\circ\varphi)\,d\varphi^{-T}\nabla I\big)dx. \tag{9.20}$$

To compute the Eulerian gradient of U_I, we need to apply the kernel operator, $\mathbb{K}$, to $\bar\partial U_I(\varphi)$, which requires the following lemma.

Lemma 9.4 *If V is a reproducing kernel Hilbert space (RKHS) of vector fields on Ω with kernel operator $\mathbb{K}$ and kernel K, μ is a measure on Ω and z a μ-measurable function from Ω to $\mathbb{R}^d$, then, for all $x\in\Omega$,*

$$\mathbb{K}(z\mu)(x) = \int_\Omega K(x,y)z(y)d\mu(y).$$

Proof From the definition of the kernel, we have, for any $a\in\mathbb{R}^d$:

$$\begin{aligned}a^T\mathbb{K}(z\mu)(x) &= (a\delta_x \mid \mathbb{K}(z\mu)) \\ &= (z\mu \mid \mathbb{K}(a\delta_x)) \\ &= (z\mu \mid K(.,x)a) \\ &= \int_\Omega z^T(y)K(y,x)a d\mu(y) \\ &= a^T\int_\Omega K(x,y)z(y)d\mu(y),\end{aligned}$$

which proves Lemma 9.4. □

The expression of the Eulerian gradient of U_I is now given by Lemma 9.4:

$$\overline{\nabla}^V U_I(\varphi) = -2 \int_\Omega (I \circ \varphi^{-1}(y) - I'(y)) K(., y) \nabla (I \circ \varphi^{-1})(y) dy. \tag{9.21}$$

This provides the following "greedy" image-matching algorithm [67, 278].

Algorithm 9.5 Greedy image matching Start with $\varphi(0) = \mathrm{id}$ and solve the evolution equation

$$\partial_t \varphi(t, y) = 2 \int_\Omega (J(t, x) - I'(x)) K(\varphi(t, y), x) \nabla J(t, x) dx \tag{9.22}$$

with $I(t, \cdot) = I \circ (\varphi(t))^{-1}$.

This algorithm can also be written uniquely in terms of the evolving image, J, using $\partial_t J \circ \varphi + (J \circ \varphi)^T \partial_t \varphi = 0$. This yields

$$\partial_t J(t, y) = -2 \int_\Omega K(y, x)(J(t, x) - I'(x)) \nabla J(t, x)^T \nabla I(t, y) dx.$$

In contrast to what we did in the landmark case, this algorithm should not be run indefinitely (or until numerical convergence). The fundamental difference is that, in the landmark case, there is an infinity of solutions to the diffeomorphic interpolation problem, and the greedy algorithm would generally run until it finds one of them and then stabilize. In the case of images, it is perfectly possible (and even typical) that there is no solution to the matching problem, i.e., no diffeomorphism φ such that $I \circ \varphi^{-1} = I'$. In that case, Algorithm 9.5 will run indefinitely, creating huge deformations while trying to solve an impossible problem.

To decide when the evolution should be stopped, an interesting suggestion has been made in [278]. Define

$$v(t, x) = 2 \int_\Omega (J(t, x) - I'(x)) K(y, x) \nabla J(t, x) dx$$

so that (9.22) reduces to $\partial_t \varphi = v(t) \circ \varphi$. As we know from Chap. 7, the smoothness of φ at time t can be controlled by

$$\int_0^t \|v(s)\|_V^2 ds,$$

the norm being explicitly given by

$$\|v(s)\|_V^2 = 2 \int_{\Omega \times \Omega} K(y, x)(J(s, x) - I'(x))(J(s, y) - I'(y)) \nabla J(s, x)^T \nabla J(s, y) dx dy.$$

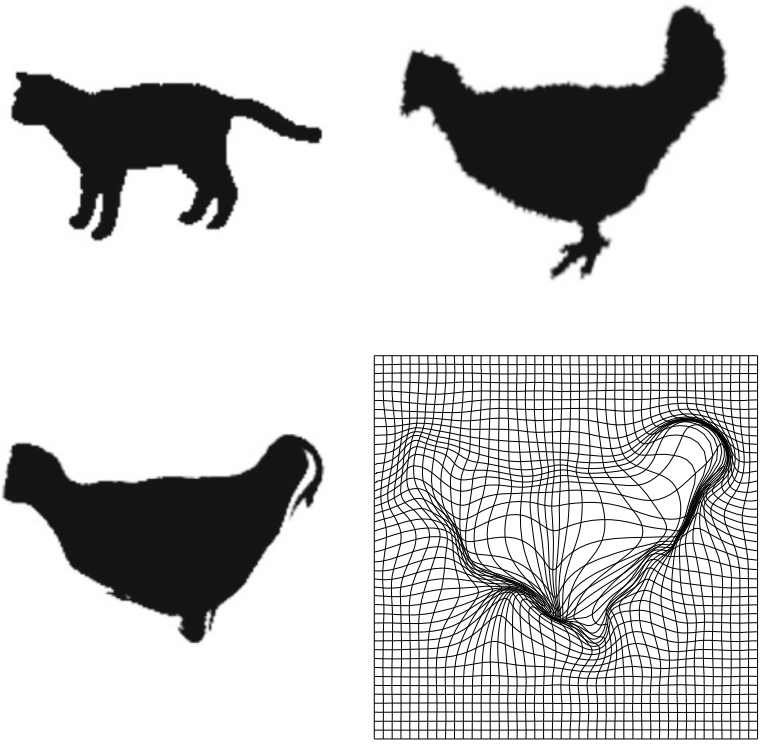

Fig. 9.2 Greedy image matching. Output of Algorithm 9.5 when estimating a deformation of the first image to match the second one. The third image is the obtained deformation of the first one and the last provides the deformation applied to a grid

Define, for some parameter λ,

$$E(t) = \frac{1}{t}\int_0^t \|v(s)\|_V^2 ds + \lambda \int_\Omega (J(t, y) - I'(y))^2 dy.$$

Then, the stopping time proposed in [278] for Algorithm 9.5 is the first t at which $E(t)$ stops decreasing. Some experimental results using this algorithm and stopping rule are provided in Fig. 9.2.

There are many other possible choices for a matching criterion, least squares being, as we wrote, the simplest one. Among other possibilities, comparison criteria involving histograms provide an interesting option, because they allow for contrast-invariant comparisons.

Given a pair of images, I, I', associate to each $x \in \Omega$ and image values λ and λ' the local histogram $H_x(\lambda, \lambda')$, which counts the frequency of simultaneous occurrence of values λ in I and λ' in I' at the same location in a small window around x. One computationally feasible way to define it is to use the kernel estimator

$$H_{I,I'}(x,\lambda,\lambda') = \int_\Omega f(|I(y)-\lambda|) f(|I'(y)-\lambda'|) g(x,y) dy$$

in which f is a positive function such that $\int_{\mathbb{R}} f(t)dt = 1$ and f vanishes when t is far from 0, and $g \geq 0$ is such that for all x, $\int_\Omega g(x,y)dy = 1$ and $g(x,y)$ vanishes when y is far from x.

For each x, $H_{I,I'}(x,\cdot,\cdot)$ is a bi-dimensional probability function, and there exist several ways of measuring the degree of dependence between its components. The simplest one, which is probably sufficient for most applications, is the correlation ratio, given by

$$C_{I,I'}(x) = 1 - \frac{\int_{\mathbb{R}^2} \lambda\lambda' H_{I,I'}(x,\lambda,\lambda')d\lambda d\lambda'}{\sqrt{\int_{\mathbb{R}^2} \lambda^2 H_{I,I'}(x,\lambda,\lambda')d\lambda d\lambda' \int_{\mathbb{R}^2} (\lambda')^2 H_{I,I'}(x,\lambda,\lambda')d\lambda d\lambda'}}.$$

It is then possible to define the matching function by

$$U_I(\varphi) = \int_\Omega C_{I\circ\varphi^{-1},I'}(x)dx.$$

The differential of U_I with respect to φ can be obtained after a lengthy (but elementary) computation. Some details can be found in [145]. A slightly simpler option is to use criteria based on the global histogram, which is defined by

$$H_{I,I'}(\lambda,\lambda') = \int_\Omega f(|I(y)-\lambda|) f(|I'(y)-\lambda'|) dy,$$

and the matching criterion is simply $U_I(\varphi) = C_{I\circ\varphi^{-1},I'}$ or, as introduced in [185, 298], the mutual information computed from the joint histogram.

9.6 Measure Matching

The running assumption in Sect. 9.4 was that the point sets $(x_1,\ldots,x_N)$ were labeled, so that, when comparing two of them, the correspondences were known and the problem was to extrapolate them to the whole space.

In some applications, correspondences are not given and need to be inferred as part of the matching problem. One way to handle this is to include them as new unknowns (in addition to the unknown diffeomorphism), add extra terms to the energy that measures the quality of correspondences, and minimize the whole thing. Such an approach is taken, for example, in [240, 241].

Another point of view is to start with a representation of the point set that does not depend on how the points are ordered. A natural mathematical representation of a subset of $\mathbb{R}^d$ is by the uniform measure on this set, at least when this is well-defined.

For a very general class of sets, this corresponds to the Hausdorff measure for the appropriate dimension [107], which, for finite sets, simply provides the sum of Dirac measures at each point, i.e., $x = (x_1, \dots, x_N)$ is represented by

$$\mu_x = \sum_{i=1}^{N} \delta_{x_i}.$$

For us, this raises the issue of comparing measures using diffeomorphisms, which will be referred to as the measure-matching problem.

In line with all other matching problems we are considering in this chapter, specifying the measure-matching problem requires, first, defining the action of diffeomorphisms on the considered objects, and second, using a good comparison criterion between two objects.

Let us start with the action of diffeomorphisms. The only fact we need here concerning measures is that they are linear forms acting on functions on $\mathbb{R}^d$ via

$$(\mu \mid f) = \int_{\mathbb{R}^d} f d\mu.$$

In particular, if μ_x is as above, then

$$(\mu_x \mid f) = \sum_{i=1}^{N} f(x_i).$$

If φ is a diffeomorphism of Ω and μ a measure, we define a new measure $\varphi \cdot \mu$ by

$$(\varphi \cdot \mu \mid f) = (\mu \mid f \circ \varphi).$$

It is straightforward to check that this provides a group action. If $\mu = \mu_x$, we have

$$(\varphi \cdot \mu_x \mid f) = \sum_{i=1}^{N} f \circ \varphi(x_i) = \big(\mu_{\varphi(x)} \;\big|\; f\big),$$

so that the transformation of the measure associated to a point set x is the measure associated to the transformed point set, which is reasonable.

When μ has a density with respect to Lebesgue measure, say $\mu = z dx$, this action can be translated to a resulting transformation over densities as follows.

Proposition 9.6 *If $\mu = z\,dx$, where z is a positive, Lebesgue integrable function on $\Omega \subset \mathbb{R}^d$, and φ is a diffeomorphism of Ω, then*

$$\varphi \cdot \mu = \det(d(\varphi^{-1}))\, z \circ \varphi^{-1}\, dx. \tag{9.23}$$

The proposition is an immediate consequence of the definition of $\varphi \cdot \mu$ and of the change of variable formula (details are left to the reader). Note that the action of diffeomorphisms does not change the total mass of a positive measure, that is $(\varphi \cdot \mu)(\Omega) = \mu(\Omega)$ if φ is a diffeomorphism of Ω.

Now that we have defined the action, we need to choose a function $D(\mu, \mu')$ to compare two measures μ and μ'. Many such functions exist, especially when measures are normalized to have a unit mass, since this allows for the use of many comparison criteria defined in probability or information theory (such as the Kullback–Leibler divergence [75]). A very general example is the Wasserstein distance [238, 301], which is associated to a positive, symmetric, cost function $\rho : \Omega \times \Omega \to [0, +\infty)$ and defined by

$$d_\rho(\mu, \mu') = \inf_\nu \int_{\Omega^2} \rho(x, y)\nu(dx, dy), \tag{9.24}$$

where the minimization is over all ν with the first marginal given by μ, and the second one by μ'. If μ and μ' are uniform measures on discrete point sets, i.e.,

$$\mu = \frac{1}{N}\sum_{k=1}^{N} \delta_{x_k}, \ \mu' = \frac{1}{M}\sum_{k=1}^{M} \delta_{x'_k},$$

then computing the Wasserstein distance reduces to minimizing

$$\sum_{k=1}^{N}\sum_{l=1}^{M} \rho(x_k, x'_l)\nu(x_k, x'_l)$$

subject to the constraints

$$\sum_{l=1}^{M} \nu(x_k, x'_l) = 1/N \text{ and } \sum_{k=1}^{N} \nu(x_k, x'_l) = 1/M.$$

This *linear assignment* problem is solved by finite-dimensional linear programming. If this is combined with diffeomorphic interpolation, i.e., if one tries to compute a diffeomorphism φ minimizing $d_\rho(\varphi \cdot x, x')$, this results in a formulation that mixes discrete and continuous optimization problems, similar to the methods introduced in [240]. The Wasserstein distance is also closely related to the mass transport problem, which can also be used to estimate diffeomorphisms, and will be discussed in the next chapter. For the moment, we focus on matching functionals associated with measures, and start with the case in which the compared measures are differentiable with respect to Lebesgue measure, i.e., with the problem of matching densities.

9.6.1 Matching Densities

Since densities are scalar-valued functions, we can use standard norms to design matching functionals for them. As an example, we can take the simplest case of the L^2 norm, as we did with images. The difference with the image case is that the action is different, and has the interesting feature of involving the derivative of the diffeomorphism, via the Jacobian determinant.

So, let us consider the action $\varphi \star \zeta$ given by

$$\varphi \star \zeta = \det(d(\varphi^{-1}))\, z \circ \varphi^{-1}$$

and use the matching functional

$$U_\zeta(\varphi) = E_{\zeta,\zeta'}(\varphi) = \int_\Omega (\varphi \star \zeta - \zeta')^2 dx.$$

Since we will need it for the differentiation of the Jacobian, we recall the following standard result on the derivative of the determinant.

Proposition 9.7 *Let $F(A) = \det(A)$ be defined over $\mathcal{M}_n(\mathbb{R})$, the space of all n by n matrices. Then, for any $A, H \in \mathcal{M}_n(\mathbb{R})$,*

$$dF(A)H = \operatorname{trace}(\operatorname{Adj}(A)H), \tag{9.25}$$

where $\operatorname{Adj}(A)$ is the adjugate matrix of A, i.e., the matrix with (i, j) entry given by the determinant of A with the jth row and ith column removed, multiplied by $(-1)^{i+j}$. (When A is invertible $\operatorname{Adj}(A) = \det(A)\, A^{-1}$.)

For $A = \mathrm{Id}$, we have

$$dF(\mathrm{Id})H = \operatorname{trace}(H). \tag{9.26}$$

Proof To prove this proposition, start with $A = \mathrm{Id}$ and use the facts that, if δ_{ij} is the matrix with 1 as the (i, j) entry and 0 everywhere else, then $\det(\mathrm{Id} + \varepsilon\delta_{ij}) = 1 + \varepsilon$ if $i = j$ and 1 otherwise, which directly gives (9.26). Then, prove the result for an invertible A using

$$\det(A + \varepsilon H) = \det(A) \det(\mathrm{Id} + \varepsilon A^{-1} H)$$

and the fact that, when A is invertible, $\det(A)A^{-1} = \operatorname{Adj}(A)$. This also implies the result for a general (not necessarily invertible) A because the determinant is a polynomial in the entries of a matrix, and so are its partial derivatives, and the coefficients of these polynomials are fully determined by the values taken on the dense set of invertible matrices. □

We have

$$U_\zeta(\varphi) = \int_\Omega \left(\det(d(\varphi^{-1}))\, \zeta \circ \varphi^{-1} - \zeta'\right)^2 dx.$$

Under the assumptions that ζ is C^1 and compactly supported and that ζ' is square integrable, one can prove that $E_{\zeta,\zeta'}$ is C^1 when defined over $\mathrm{Diff}_0^{p+2,\infty}$ with $p \geq 0$ (the details are left to the reader). To compute the derivative at any given φ, it will be convenient to use the trick described at the end of Sect. 9.2, starting with the computation of the differential at the identity and deducing from it the differential at any φ by replacing ζ by $\varphi \cdot \zeta$.

If $\varphi(\varepsilon, \cdot)$ is a diffeomorphism that depends on a parameter ε, such that $\varphi(0, \cdot) = \mathrm{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h \in C_0^{p+2}(\Omega, \mathbb{R}^d)$, then, at $\varepsilon = 0$, $\partial_\varepsilon \zeta \circ \varphi(\varepsilon, \cdot)^{-1} = -\nabla \zeta^T h$ and $\partial_\varepsilon \det(d(\varphi(\varepsilon, \cdot)^{-1})) = -\mathrm{trace}(dh) = -\mathrm{div}\, h$. This implies that

$$\partial_\varepsilon \left(\zeta \circ \varphi(\varepsilon, \cdot)^{-1} \det(d(\varphi(\varepsilon, \cdot))^{-1}) \right) = -\nabla \zeta^T h - \zeta \,\mathrm{div}\, h = -\mathrm{div}(\zeta h)$$

at $\varepsilon = 0$ and

$$\partial_\varepsilon U_\zeta(\varphi_\varepsilon)_{|_{\varepsilon=0}} = -2 \int_\Omega (\zeta - \zeta') \mathrm{div}(\zeta h)\, dx.$$

So this gives

$$\left(dU_\zeta(\mathrm{id}) \,\middle|\, h \right) = -2 \int_\Omega (\zeta - \zeta') \mathrm{div}(\zeta h)\, dx$$

and

$$\left(dU_\zeta(\varphi) \,\middle|\, h \right) = -2 \int_\Omega (\varphi \star \zeta - \zeta') (\mathrm{div}((\varphi \star \zeta) h)\, dx. \tag{9.27}$$

We can use the divergence theorem to obtain an alternative expression (using the fact that h vanishes on $\partial\Omega$ or at infinity), yielding

$$\left(dU_\zeta(\varphi) \,\middle|\, h \right) = 2 \int_\Omega \nabla(\varphi \star \zeta - \zeta')(\varphi \star \zeta) h dx \tag{9.28}$$

or

$$dU_\zeta(\varphi) = 2(\varphi \star \zeta) \nabla(\varphi \star \zeta - \zeta') dx. \tag{9.29}$$

One can appreciate the symmetry of this expression compared with the one obtained with images in (9.19).

9.6.2 Dual RKHS Norms on Measures

One of the limitations of functional norms, such as the L^2 norm, is that they do not apply to singular objects such as the Dirac measures that motivated our study of the measure-matching problem. It is certainly possible to smooth out singular objects and transform them into densities that can be compared using the previous matching functional. For example, given a density function ρ (a Gaussian, for example) and a point set $(x_1, \ldots, x_N)$, one can compute a density

$$\zeta_x(y) = \sum_{k=1}^{N} \rho\Big(\frac{y - x_k}{\sigma}\Big), \tag{9.30}$$

where σ is a positive scale parameter (this is a standard kernel density estimator). One can then compare two point sets, say x and x', by comparing the associated ζ_x and $\zeta_{x'}$ using the previous method.

The representation in (9.30) is somewhat imperfect, in the sense that, for the natural actions we have defined, we have in general $\varphi \star \zeta_x \neq \zeta_{\varphi \cdot x}$: the density associated to a deformed point set is not the deformed density. If the goal is to compare two point sets, it makes more sense to use $\zeta_{\varphi \cdot x}$ instead of $\varphi \cdot \zeta_x$ as a density resulting from the deformation, and to rather use the cost function

$$U_x(\varphi) = E_{x,x'}(\varphi) = \int_{\mathbb{R}^d} (\zeta_{\varphi \cdot x} - \zeta_{x'})^2 dy, \tag{9.31}$$

which can be written, if $x = (x_1, \ldots, x_N)$ and $x' = (x'_1, \ldots, x'_M)$, and introducing the function

$$\xi(z, z') = \int_{\mathbb{R}^d} \rho\Big(\frac{y - z}{\sigma}\Big)\rho\Big(\frac{y - z'}{\sigma}\Big) dy, \tag{9.32}$$

as

$$\begin{aligned} U_x(\varphi) = \sum_{k,l=1}^{N} \xi(\varphi(x_k), \varphi(x_l)) \\ - 2\sum_{k=1}^{N}\sum_{l=1}^{M} \xi(\varphi(x_k), x'_l) + \sum_{k,l=1}^{M} \xi(x'_k, x'_l). \end{aligned} \tag{9.33}$$

Before computing the variations of this energy, we make the preliminary remark that the obtained expression is a particular case of what comes from a representation of measures as linear forms over RKHSs of scalar functions. Indeed, since measures are linear forms on functions, we can evaluate their dual norm, given by

$$\|\mu\| = \sup\{(\mu \mid f) : \|f\| = 1\}. \tag{9.34}$$

Following [128], assume that the function norm in (9.34) is that of an RKHS. More precisely, let W be an RKHS of real-valued functions, so that we have an operator $\mathbb{K}_W : W^* \to W$ with $\mathbb{K}_W \delta_x := \xi(\cdot, x)$ and with the identity $(\mu \mid f) = \big\langle \mathbb{K}_W \mu\,,\ f \big\rangle_W$ for $\mu \in W^*$, $f \in W$. With this choice, (9.34) becomes

$$\begin{aligned} \|\mu\|_{W^*} &= \sup\{(\mu \mid f) : \|f\|_W = 1\} \\ &= \sup\big\{\big\langle \mathbb{K}_W \mu\,,\ f\big\rangle_W : \|f\|_W = 1\big\} \\ &= \|\mathbb{K}_W \mu\|_W. \end{aligned}$$

This implies that

$$\|\mu\|_{W^*}^2 = \langle \mathbb{K}_W \mu\,,\ \mathbb{K}_W \mu \rangle_W = (\mu \mid \mathbb{K}_W \mu).$$

If μ is a measure, this expression is very simple and is given by

$$\|\mu\|_{W^*}^2 = \int \xi(x, y) d\mu(x) d\mu(y).$$

This is because $\mathbb{K}_W \mu(x) = (\delta_x \mid \mathbb{K}_W \mu) = (\mu \mid \mathbb{K}_W \delta_x) = \int \xi(y, x) d\mu(y)$. So we can take

$$U_\mu(\varphi) = E_{\mu,\mu'}(\varphi) = \|\varphi \cdot \mu - \mu'\|_{W^*}^2. \tag{9.35}$$

Expanding the norm, we get

$$\begin{aligned} U_\mu(\varphi) &= \langle \varphi \cdot \mu\,,\ \varphi \cdot \mu \rangle_{W^*} - 2\langle \varphi \cdot \mu\,,\ \mu' \rangle_{W^*} + \langle \mu'\,,\ \mu' \rangle_{W^*} \\ &= (\varphi \cdot \mu \mid \xi(\varphi \cdot \mu)) - 2\big(\varphi \cdot \mu \,\big|\, \xi\mu'\big) + \big(\mu' \,\big|\, \xi\mu'\big) \\ &= \int \xi(\varphi(x), \varphi(y)) d\mu(x) d\mu(y) - 2 \int \xi(\varphi(x), y) d\mu(x) d\mu'(y) \\ &\quad + \int \xi(x, y) d\mu'(x) d\mu'(u). \end{aligned}$$

We retrieve (9.33) when μ and μ' are sums of Dirac measures and ξ is chosen as in (9.32), but the RKHS formulation is more general.

Assume that μ is bounded and that ξ is continuously differentiable and bounded, with bounded derivatives. Then (leaving the proof to the reader) U_I is C^1 on $\mathrm{Diff}_0^{p,\infty}$ for any $p \geq 0$ with derivative

$$\begin{aligned} \big(\partial U_\mu(\varphi) \,\big|\, h\big) &= 2 \int \nabla_1 \xi(\varphi(x), \varphi(y))^T h(x) d\mu(x) d\mu(y) \\ &\quad -2 \int \nabla_1 \xi(\varphi(x), z)^T h(x) d\mu(x) d\mu'(z). \end{aligned}$$

In particular,

$$dU_\mu(\mathrm{id}) = \bar{\partial} U_\mu(\mathrm{id}) = 2 \left(\int \nabla_1 \xi(\cdot, y) d\mu(y) - \int \nabla_1 \xi(\cdot, z) d\mu'(z) \right) \mu. \tag{9.36}$$

To obtain the Eulerian differential at a generic φ, it suffices to replace μ by $\varphi \cdot \mu$, which yields:

Proposition 9.8 *The Eulerian derivative and gradient of* (9.35) *are*

$$\bar{\partial} U_\mu(\varphi) = 2 \left(\int \nabla_1 \xi(\cdot, \varphi(y)) d\mu(y) - \int \nabla_1 \xi(\cdot, z) d\mu'(z) \right) (\varphi \cdot \mu) \tag{9.37}$$

and

$$\overline{\nabla}^V U_\mu(\varphi)(\cdot) = 2\int K(\cdot,\varphi(x)) \left(\int \nabla_1\xi(\varphi(x),\varphi(y))d\mu(y) - \nabla_1\xi(\varphi(x),z)d\mu'(z)\right) d\mu(x). \quad (9.38)$$

The derivative of the expression in (9.33) can be directly deduced from this expression. This leads to the following unlabeled point-matching evolution for point sets $x = (x_1, \ldots, x_N)$ and $x' = (x'_1, \ldots, x'_M)$:

$$\partial_t\varphi(z) = -2\sum_{i=1}^{N} K(\varphi(z),\varphi(x_i)) \left(\sum_{j=1}^{N} \nabla_1\xi(\varphi(x_i),\varphi(x_j)) - \sum_{h=1}^{M} \nabla_1\xi(\varphi(x_i),x'_h)\right). \quad (9.39)$$

As discussed in the case of labeled point sets, this equation may be solved in two stages: letting $z_i(t) = \varphi(t, x_i)$, first solve the system

$$\partial_t z_q = -2\sum_{i=1}^{N} K(z_q, z_i) \left(\sum_{j=1}^{N} \nabla_1\xi(z_i, z_j) - \sum_{h=1}^{M} \nabla_1\xi(z_i, x'_h)\right).$$

Once this is done, the trajectory of an arbitrary point $z(t) = \varphi_t(z_0)$ is

$$\partial_t z = -2\sum_{i=1}^{N} K(z, z_i) \left(\sum_{j=1}^{N} \nabla_1\xi(z_i, z_j) - \sum_{h=1}^{M} \nabla_1\xi(z_i, x'_h)\right).$$

9.7 Matching Curves and Surfaces

Curves in two dimensions and surfaces in three dimensions are probably the most natural representations of shapes, and their comparison using matching functionals is a fundamental issue. In this section, we discuss a series of representations that can be seen as extensions of measure-matching methods. (This is not the unique way to compare such objects, and we will see a few more methods in the following chapters, especially for curves.)

Note that we are looking here for correspondences between points in the curves and surfaces that derive from global diffeomorphisms of the ambient space. The curve- (or surface-) matching problems are often studied in the literature as attempts to find diffeomorphic correspondences between points along the curve (or surface)

only. Even if such restricted diffeomorphisms can generally be extended to diffeomorphisms of the whole space, the two approaches generally lead to very different algorithms. The search for correspondences within the structures is often implemented as a search for correspondences between parametrizations. This is easier for curves (looking, for example, for correspondences of the arc-length parametrizations), than for surfaces, which may not be topologically equivalent in the first place (a sphere cannot be matched to a torus); when matching topologically equivalent surfaces, special parametrizations, like conformal maps [72, 269] can be used. In this framework, once parametrizations are fixed, one can look for diffeomorphisms in parameter space that optimally align some well-chosen, preferably intrinsic, representation. In the case of curves, one can choose the representation $s \mapsto \kappa_\gamma(s)$, where κ_γ is the curvature of a curve γ, with the curve rescaled to have length 1 to fix the interval over which this representation is defined. One can then use image-matching functionals to compare them, i.e., find φ (a diffeomorphism of the unit interval) such that $\varphi \cdot \kappa_\gamma \simeq \kappa_{\gamma'}$.

But, as we wrote, the main focus in this chapter is the definition of matching functionals for deformable objects in $\mathbb{R}^d$, and we now address this problem for curves and surfaces.

9.7.1 Curve Matching with Measures

We can arguably make a parallel between point sets and curves in that labeled point sets correspond to parametrized curves and unlabeled point sets to curves modulo parametrization. In this regard we have a direct generalization of the labeled point-matching functional to parametrized curves (assumed to be defined over the same interval, say [0, 1]), simply given by

$$E_{\gamma,\gamma'}(\varphi) = \int_0^1 |\varphi(\gamma(u)) - \gamma'(u)|^2 du.$$

But being given two consistent parametrizations of the curves (to allow for direct comparisons as done above) almost never happens in practice. Interesting formulations of the curve matching problem should therefore consider curves modulo parametrization, so that the natural analogy is with unlabeled point sets. The counterpart of a uniform measure over a finite set of points is the uniform measure on the curve, defined by, if γ, parametrized over an interval $[a, b]$, is C^1 and regular

$$\left(\mu_\gamma \mid f\right) = \int_\gamma f \, d\sigma_\gamma = \int_a^b f(\gamma(u)) \, |\dot{\gamma}(u)| \, du.$$

This is clearly a parametrization-independent representation. Now, if φ is a diffeomorphism, we have, by definition of the action of diffeomorphisms on measures

$$\big(\varphi\cdot\mu_\gamma \;\big|\; f\big) = \int_\gamma f\circ\varphi\, d\sigma_\gamma = \int_a^b f(\varphi(\gamma(u)))\,|\dot\gamma(u)|\,du.$$

However, we have

$$\big(\mu_{\varphi\cdot\gamma} \;\big|\; f\big) = \int_{\varphi(\gamma)} f\, d\sigma_{\varphi(\gamma)} = \int_a^b f(\varphi(\gamma(u)))\,|d\varphi(\gamma(u))\dot\gamma(u)|\,du.$$

So, in contrast to point sets, for which we had $\varphi\cdot\mu_x = \mu_{\varphi(x)}$, the image of the measure associated to a curve is not the measure associated to the image of a curve. When the initial goal is to compare curves, and not measures, it is more natural to use the second definition, $\mu_{\varphi\cdot\gamma}$, rather than the first one. Using the notation of the previous section, and introducing a target curve γ' defined on $[a', b']$, we can set

$$\begin{aligned}
&E_{\gamma,\gamma'}(\varphi) = \|\mu_{\varphi\cdot\gamma} - \mu_{\gamma'}\|^2_{W^*} \qquad (9.40)\\
&= \big\langle \mu_{\varphi\cdot\gamma}\,,\ \mu_{\varphi\cdot\gamma}\big\rangle_{W^*} - 2\big\langle\mu_{\varphi\cdot\gamma}\,,\ \mu_{\gamma'}\big\rangle_{W^*} + \big\langle\mu_{\gamma'}\,,\ \mu_{\gamma'}\big\rangle_{W^*}\\
&= \big(\mu_{\varphi\cdot\gamma}\;\big|\;\xi(\mu_{\varphi\cdot\gamma})\big) - 2\big(\mu_{\varphi\cdot\gamma}\;\big|\;\xi\mu_{\gamma'}\big) + \big(\mu_{\gamma'}\;\big|\;\xi\mu_{\gamma'}\big)\\
&= \int_a^b\int_a^b \xi(\varphi(\gamma(u)),\varphi(\gamma(v)))\,|d\varphi(\gamma(u))\dot\gamma(u)|\,|d\varphi(\gamma(v))\dot\gamma(v)|\,dudv\\
&- 2\int_a^b\int_{a'}^{b'} \xi(\varphi(\gamma(u)),\gamma'(v))\,|d\varphi(\gamma(u))\dot\gamma(u)|\,|\dot\gamma'(v)|\,dudv\\
&+ \int_{a'}^{b'}\int_{a'}^{b'} \xi(\gamma'(u),\gamma'(v))\,|\dot\gamma'(u)|\,|\dot\gamma'(v)|\,dudv.
\end{aligned}$$

If ξ is C^1, then $E_{\gamma,\gamma'}$ is C^1 on $\mathrm{Diff}_0^{p+1,\infty}$ for any $p\geq 0$. To explicitly compute the derivative, take $\varphi(\varepsilon,\cdot)$ such that $\varphi(0,\cdot) = \mathrm{id}$ and $\partial_\varepsilon\varphi(0,\cdot) = h$, so that

$$\partial_\varepsilon E(\varphi(\varepsilon,\cdot)) = 2\partial_\varepsilon\big\langle\mu_{\varphi(\varepsilon,\cdot)\cdot\gamma} - \mu_{\gamma'}\,,\ \mu_\gamma - \mu_{\gamma'}\big\rangle_{W^*} = 2\partial_\varepsilon\big\langle\mu_{\varphi(\varepsilon,\cdot)\cdot\gamma}\,,\ \mu_\gamma - \mu_{\gamma'}\big\rangle_{W^*},$$

the derivatives being computed at $\varepsilon = 0$. Introduce

$$\tilde E(\varphi) = \big\langle\mu_{\varphi\cdot\gamma}\,,\ \mu_\gamma - \mu_{\gamma'}\big\rangle_{W^*}$$

and let, for a given curve $\tilde\gamma$,

$$Z^{\tilde\gamma}(\cdot) = \int_{\tilde\gamma}\xi(\cdot,p)d\sigma_{\tilde\gamma}(p). \qquad (9.41)$$

Let also $\zeta = Z^\gamma - Z^{\gamma'}$ and, for further use, $\zeta^\varphi = Z^{\varphi\cdot\gamma} - Z^{\gamma'}$. With this notation, we have

$$\tilde E(\varphi) = \int_{\varphi(\gamma)}\zeta(p)d\sigma_{\varphi(\gamma)}(p)$$

and we can use Theorem 5.2 to derive, letting p_0 and p_1 be the extremities of γ,

$$\partial_\varepsilon \tilde{E}(\varphi(\varepsilon,\cdot))_{|\varepsilon=0} = \zeta(p_1)h(p_1)^T T^\gamma(p_1) - \zeta(p_0)h(p_0)^T T^\gamma(p_0) + \int_\gamma \big(\nabla\zeta^T N^\gamma - \zeta\kappa^\gamma\big)h^T N^\gamma dl.$$

Replacing γ by $\varphi\cdot\gamma$, this provides the expression of the Eulerian derivative of E at φ, namely

$$\frac{1}{2}\bar{\partial}E_{\gamma,\gamma'}(\varphi) = \zeta^\varphi T^\gamma(\delta_{p_1}-\delta_{p_0}) + \big((\nabla\zeta^\varphi)^T N^{\varphi\cdot\gamma} - \zeta^\varphi\kappa^{\varphi\cdot\gamma}\big)N^{\varphi\cdot\gamma}\mu_{\varphi\cdot\gamma}. \tag{9.42}$$

The Eulerian gradient on V therefore is

$$\frac{1}{2}\bar{\nabla}E_{\gamma,\gamma'}(\varphi) = K(\cdot,p_1)\zeta^\varphi(p_1)T^\gamma(p_1) - K(\cdot,p_0)\zeta^\varphi(p_0)T^\gamma(p_0) + \int_{\varphi\cdot\gamma}\Big(\nabla\zeta^\varphi(p)^T N^{\varphi\cdot\gamma}(p) - \zeta^\varphi(p)\kappa^{\varphi\cdot\gamma}(p)\Big)K(\cdot,p)N^{\varphi\cdot\gamma}(p)d\sigma_{\varphi\cdot\gamma}(p). \tag{9.43}$$

To write this expression, we have implicitly assumed that γ is C^2. In fact, we can give an alternative expression for the Eulerian gradient that does not require this assumption, by directly computing the variation of $\tilde{E}(\varphi(\varepsilon,\cdot))$ without applying Theorem 5.2. This yields, using the fact that, if z is a function of a parameter ε, then $\partial_\varepsilon|z| = (\dot{z}^T z)/|z|$,

$$\partial_\varepsilon|d\varphi(\varepsilon,\cdot)(\gamma)\dot{\gamma}|_{|\varepsilon=0} = (T^\gamma)^T dh(\gamma)\dot{\gamma} = (T^\gamma)^T dh(\gamma)T^\gamma|\dot{\gamma}|$$

$$\text{and } \partial_\varepsilon\tilde{E}(\varphi(\varepsilon,\cdot)) = \int_\gamma \big(\nabla\zeta^T h + \zeta(T^\gamma)^T dhT^\gamma\big)d\sigma_\gamma.$$

The term involving dh can be written in terms of V-dot products of h with derivatives of the kernel, K, since (we use the notation introduced in Sect. 8.1.3, Eq. (8.9))

$$a^T dh(x)b = \big\langle h\,,\,\partial_2 K(\cdot,x)(a,b)\big\rangle_V. \tag{9.44}$$

This gives

$$\partial_\varepsilon\tilde{E}(\varphi(\varepsilon,\cdot)) = \int_\gamma \Big(\big\langle K(\cdot,p)\nabla\zeta(p)\,,\,h\big\rangle_V + \big\langle\zeta(p)\partial_2 K(\cdot,p)(T^\gamma(p),T^\gamma(p))\,,\,h\big\rangle_V\Big)d\sigma_\gamma(p)$$

and a new expression of the Eulerian gradient

$$\frac{1}{2}\overline{\nabla}^V E_{\gamma,\gamma'}(\varphi) = \int_{\varphi\cdot\gamma} \Big(K(\cdot, p)\nabla\zeta^\varphi(p) + \zeta^\varphi(p)\partial_2 K(., p)(T^{\varphi\cdot\gamma}(p), T^{\varphi\cdot\gamma}(p)) \Big) d\sigma_\gamma(p). \quad (9.45)$$

To be complete, let us consider the variation of a discrete form of $E_{\gamma,\gamma'}(\varphi)$. If a curve γ is discretized with points $x_0, \ldots, x_N$ (with $x_N = x_0$ if the curve is closed), one can define the discrete measure, still denoted μ_γ

$$\big(\mu_\gamma \mid f\big) = \sum_{i=1}^{N} f(c_i)|\tau_i|$$

with $c_i = (x_i + x_{i-1})/2$ and $\tau_i = x_i - x_{i-1}$. Use a similar expression for the measure associated to a discretization of $\varphi\cdot\gamma$, with $c_i^\varphi = (\varphi(x_i) + \varphi(x_{i-1}))/2$ and $\tau_i^\varphi = \varphi(x_i) - \varphi(x_{i-1})$. Finally, let γ' be discretized in $x'_1, \ldots, x'_M$, and define

$$E_{\gamma,\gamma'}(\varphi) = \sum_{i,j=1}^{N} \xi(c_i^\varphi, c_j^\varphi)|\tau_i^\varphi|\,|\tau_j^\varphi| - 2\sum_{i=1}^{N}\sum_{j=1}^{N} \xi(c_i^\varphi, c'_j)|\tau_i^\varphi|\,|\tau'_j| + \sum_{i,j=1}^{M} \xi(c'_i, c'_j)|\tau'_i|\,|\tau'_j| \quad (9.46)$$

in which we identify indices 1 and $N+1$ or $M+1$ (assuming closed curves). Note that this functional depends on $\varphi\cdot x$ and x'. The computation of the differential proceeds as above. Define, for a point set $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_Q)$

$$Z^{\tilde{x}}(\cdot) = \sum_{j=1}^{Q} \xi(\cdot, \tilde{c}_j)|\tilde{\tau}_j|,$$

and $\zeta = Z^x - Z^{x'}$, $\zeta^\varphi = Z^{\varphi\cdot x} - Z^{x'}$. We then obtain

$$\frac{1}{2} dE_{\gamma,\gamma'}(\mathrm{id}) = \sum_{i=1}^{N} (\nabla\zeta(c_i)|\tau_i| + \nabla\zeta(c_{i+1})|\tau_{i+1}|)\delta_{x_i} - 2\sum_{i=1}^{N} \Big(\zeta(c_{i+1})\frac{\tau_{i+1}}{|\tau_{i+1}|} - \zeta(c_i)\frac{\tau_i}{|\tau_i|} \Big)\delta_{x_i}.$$

The Eulerian differential at $\varphi \neq \mathrm{id}$ is obtained by replacing ζ, c_i, τ_i by $\zeta^\varphi, c_i^\varphi, \tau_i^\varphi$ and the Eulerian gradient by applying the V-kernel to it.

9.7.2 Curve Matching with Vector Measures

Instead of describing a curve with a measure, which is a linear form on functions, it is possible to represent it by a vector measure, which is a linear form on vector fields. Given a parametrized curve $\gamma : [a, b] \to \mathbb{R}^d$, we define a vector measure ν_γ, which associates to each vector field f on $\mathbb{R}^d$ a number $\left(\nu_\gamma \mid f\right)$ given by

$$\left(\nu_\gamma \mid f\right) = \int_a^b \dot{\gamma}(u)^T f \circ \gamma(u) du,$$

i.e., $\nu_\gamma = T^\gamma \mu_\gamma$ where T^γ is the unit tangent to γ and μ_γ is the line measure along γ, as defined in the previous section. This definition is invariant under a change of parametrization, but depends on the orientation of γ. If φ is a diffeomorphism, we then have

$$\nu_{\varphi\cdot\gamma}(f) = \int_a^b (d\varphi(\gamma(u))\dot{\gamma}(u))^T f \circ \varphi(\gamma(u)) du.$$

As done with scalar measures, we can use a dual norm for the comparison of two vector measures. Such a norm is defined by

$$\|\nu\|_{W^*} = \sup\{(\nu \mid f) : \|f\|_W = 1\},$$

where W is now an RKHS of vector fields, and we still have $\|\nu\|^2_{W^*} = (\nu \mid \mathbb{K}_W \nu)$. Still letting ξ denote the kernel of W (which is now matrix-valued), we have

$$\|\nu_\gamma\|^2_{W^*} = \int_a^b \int_a^b \dot{\gamma}(u)^T \xi(\gamma(u), \gamma(v))\dot{\gamma}(v) du dv$$

and

$$\|\nu_{\varphi\cdot\gamma}\|^2_{W^*} = \int_a^b \int_a^b \dot{\gamma}(u)^T d\varphi(\gamma(u))^T \xi(\varphi(\gamma(u)), \varphi(\gamma(v))) d\varphi(\gamma(v))\dot{\gamma}(v) du dv.$$

Define $E_{\gamma,\gamma'}(\varphi) = \|\nu_{\varphi\cdot\gamma} - \nu_{\gamma'}\|^2_{W^*}$. We follow the same pattern as in the previous section and define

$$\tilde{E}(\varphi) = \left\langle \nu_{\varphi\cdot\gamma}\,,\ \nu_\gamma - \nu_{\gamma'} \right\rangle_{W^*},$$

which (introducing $\varphi(\varepsilon, \cdot)$ with $\varphi(0, \cdot) = \mathrm{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h$) is such that $\partial_\varepsilon E(\varphi(\varepsilon, \cdot)) = 2\partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot))$ at $\varepsilon = 0$. Define

$$Z^{\tilde{\gamma}}(\cdot) = \int_{\tilde{\gamma}} \xi(\cdot, p) N^{\tilde{\gamma}}(p) dp,$$

and $\zeta = Z^\gamma - Z^{\gamma'}$, $\zeta^\varphi = Z^{\varphi\cdot\gamma} - Z^{\gamma'}$, so that (using $(T^{\varphi\cdot\gamma})^T T^{\gamma'} = (N^{\varphi\cdot\gamma})^T N^{\gamma'}$)

$$\tilde{E}(\varphi) = \int_{\varphi\cdot\gamma} \zeta^T N^{\varphi\cdot\gamma} d\sigma_{\varphi\cdot\gamma}.$$

We can use Theorem 5.2, Eq. (5.4), to find

$$\partial_\varepsilon E(\varphi(\varepsilon,\cdot)) = -[\det(\zeta,h)]_0^A + \int_\gamma \mathrm{div}(\zeta)(N^\gamma)^T h dl.$$

This yields in turn (replacing γ by $\varphi\cdot\gamma$, and letting p_0 and p_1 be the extremities of γ)

$$\frac{1}{2}\bar{\partial} E_{\gamma,\gamma'}(\varphi) = -(R_{\pi/2}\zeta^\varphi)(\delta_{\varphi(p_1)} - \delta_{\varphi(p_0)}) + \mathrm{div}(\zeta^\varphi)\nu_{\varphi\cdot\gamma}, \tag{9.47}$$

where $R_{\pi/2}$ is a 90° rotation. This final expression is remarkably simple, especially for closed curves, for which the first term cancels. A discrete version of the matching functional can also be defined, namely, using the notation of the previous section:

$$\begin{aligned} E_{\gamma,\gamma'}(\varphi) = &\sum_{i,j=1}^{N} \xi(c_i^\varphi, c_j^\varphi)(\tau_i^\varphi)^T \tau_j^\varphi \\ &-2\sum_{i=1}^{N}\sum_{j=1}^{N} \xi(c_i^\varphi, c_j')(\tau_i^\varphi)^T \tau_j' + \sum_{i,j=1}^{M} \xi(c_i', c_j')(\tau_i')^T \tau_j'. \end{aligned}$$

We leave the computation of the associated Eulerian differential (which is a slight variation of the one we made with measures) to the reader.

9.7.3 *Surface Matching*

We now extend to surfaces the matching functionals that we just studied for curves. The construction is formally very similar. If S is a surface in $\mathbb{R}^3$, one can compute a measure μ_S and a vector measure ν_S defined by

$$(\mu_S \mid f) = \int_S f(x) d\sigma_S(x) \text{ for a scalar } f \tag{9.48}$$

and

$$(\nu_S \mid f) = \int_S f(x)^T N(x) d\sigma_S(x) \text{ for a vector field } f, \tag{9.49}$$

where $d\sigma_S$ is the volume measure on S and N is the unit normal (S being assumed to be oriented in the definition of ν_S).

We state without proof the following result:

Proposition 9.9 *If S is a surface and φ a diffeomorphism of $\mathbb{R}^3$ that preserves the orientation (i.e., with positive Jacobian), we have*

$$\left(\mu_{\varphi(S)} \mid f\right) = \int_S f \circ \varphi(x) |d\varphi(x)^{-T} N| \det(d\varphi(x)) d\sigma_S(x)$$

for a scalar f and for a vector-valued f,

$$\left(\nu_{\varphi(S)} \mid f\right) = \int_S f \circ \varphi(x)^T d\varphi(x)^{-T} N \det(d\varphi(x)) d\sigma_S(x).$$

If $e_1(x), e_2(x)$ is a basis of the tangent plane to S at x, we have

$$d\varphi(x)^{-T} N \det(d\varphi(x)) = (d\varphi(x)e_1 \times d\varphi(x)e_2)/|e_1 \times e_2|. \tag{9.50}$$

The last formula implies in particular that if S is parametrized by $(u, v) \mapsto m(u, v)$, then (since $N = (\partial_1 m \times \partial_2 m)/|\partial_1 m \times \partial_2 m|$ and $d\sigma_S = |\partial_1 m \times \partial_2 m| \, du dv$)

$$\begin{aligned}(\nu_S \mid f) &= \int f(x)^T (\partial_1 m \times \partial_2 m) du dv \\ &= \int \det(\partial_1 m, \partial_2 m, f) du dv\end{aligned}$$

and

$$\left(\nu_{\varphi(S)} \mid f\right) = \int \det(d\varphi \, \partial_1 m, d\varphi \, \partial_2 m, f \circ \varphi) du dv.$$

If W is an RKHS of scalar functions or vector fields, we can compare two surfaces by using the norm of the difference of their associated measures on W^*. So define (in the scalar measure case)

$$E_{S,S'}(\varphi) = \|\mu_{\varphi \cdot S} - \mu_{S'}\|^2_{W^*} \tag{9.51}$$

and the associated

$$\tilde{E}(\varphi) = \left\langle \mu_{\varphi \cdot S}, \, \mu_S - \mu_{S'} \right\rangle_{W^*}$$

so that, for $\varphi(\varepsilon, \cdot)$ such that $\varphi(0, \cdot) = \mathrm{id}$ and $\partial_\varepsilon \varphi(0, \cdot) = h$

$$\partial_\varepsilon E_{\gamma,\gamma'}(\varphi(\varepsilon, \cdot)) = 2 \partial_\varepsilon \tilde{E}(\varphi(\varepsilon, \cdot)).$$

To a given surface $\tilde{S}$, associate the function

$$Z^{\tilde{S}}(\cdot) = \int_{\tilde{S}} \xi(\cdot, p) d\sigma_{\tilde{S}}(p)$$

and $\zeta = Z^S - Z^{S'}$, $\zeta^\varphi = Z^{\varphi\cdot S} - Z^{S'}$. Since

$$\tilde{E}(\varphi) = \int_{\varphi\cdot S} \zeta(p) d\sigma_{\varphi\cdot S}(p),$$

Theorem 5.4 yields

$$\partial_\varepsilon \tilde{E}(\varphi(\varepsilon,\cdot)) = -\int_{\partial S} \zeta\,(n^S)^T h d\sigma_{\partial_S} + \int_S \big(-2\zeta H^S + \nabla\zeta^T N^S\big)\,(N^S)^T h d\sigma_S$$

where H^S is the mean curvature on S. This implies

$$\frac{1}{2}\bar{\partial} E_{S,S'}(\varphi) = -\zeta^\varphi n^{\varphi\cdot S}\mu_{\varphi\cdot\partial S} + \big(-2\zeta^\varphi H^{\varphi\cdot S} + (\nabla\zeta^\varphi)^T N^{\varphi\cdot S}\big)\nu_{\varphi\cdot S}. \tag{9.52}$$

If we now use vector measures, so that

$$E_{S,S'}(\varphi) = \|\nu_{\varphi\cdot S} - \nu_{S'}\|^2_{W^*} \tag{9.53}$$

and

$$\tilde{E}(\varphi) = \big\langle \nu_{\varphi\cdot S}\,,\ \nu_S - \nu_{S'}\big\rangle_{W^*},$$

we need to define

$$Z^{\tilde{S}}(\cdot) = \int_{\tilde{S}} \xi(\cdot, p) N^{\tilde{S}} d\sigma_{\tilde{S}}(p)$$

and $\zeta = Z^S - Z^{S'}$, $\zeta^\varphi = Z^{\varphi\cdot S} - Z^{S'}$, so that

$$\tilde{E}(\varphi) = \int_{\varphi\cdot S} \zeta^T N^{\varphi\cdot S}\, d\sigma_{\varphi\cdot S}.$$

Variations derive again from Theorem 5.4, yielding

$$\partial_\varepsilon \tilde{E} = -\int_{\partial S} ((\zeta^T N^S)(h^T n^S) - (\zeta^T n^S)(h^T N^S))\, d\sigma_{\partial S} + \int_S \operatorname{div}(\zeta)(N^S)^T h\, d\sigma_S.$$

We therefore have

$$\frac{1}{2}\bar{\partial}E_{S,S'}(\varphi) = -\big((\zeta^{\varphi})^T N^{\varphi\cdot S}\, n^{\varphi\cdot S} - (\zeta^{\varphi})^T n^{\varphi\cdot S}\, N^{\varphi\cdot S}\big)\mu_{\varphi\cdot\partial S} + \operatorname{div}(\zeta^{\varphi})\nu_{\varphi\cdot S}. \tag{9.54}$$

Again the expression is remarkably simple for surfaces without boundary.

Consider now the discrete case and let S be a triangulated surface [289]. Let $x_1, \ldots, x_N$ be the vertices of S and $f_1, \ldots, f_Q$ be the faces (triangles) which are ordered triples of vertices $f_i = (x_{i1}, x_{i2}, x_{i3})$. Let c_i be the center of f_i, N_i its oriented unit normal and a_i its area. Define the discrete versions of the previous measures by

$$(\mu_S \mid h) = \sum_{i=1}^{Q} h(c_i)a_i, \text{ for a scalar } h \tag{9.55}$$

and

$$(\nu_S \mid h) = \sum_{i=1}^{Q} (h(c_i)^T N_i)a_i, \text{ for a vector field } h. \tag{9.56}$$

The previous formulae can be written as

$$(\mu_S \mid h) = \sum_{i=1}^{K} h\Big(\frac{x_{i1} + x_{i2} + x_{i3}}{3}\Big)|(x_{i2} - x_{i1}) \times (x_{i3} - x_{i1})|$$

and

$$(\nu_S \mid h) = \sum_{i=1}^{K} h\Big(\frac{x_{i1} + x_{i2} + x_{i3}}{3}\Big)^T (x_{i2} - x_{i1}) \times (x_{i3} - x_{i1}),$$

where the last formula requires that the vertices or the triangles are ordered consistently with the orientation (see Sect. 4.2). The transformed surfaces are now represented by the same expressions with x_{ik} replaced by $\varphi(x_{ik})$. If, given two triangulated surfaces, one defines $E_{S,S'}(\varphi) = \|\mu_{\varphi\cdot S} - \mu_{S'}\|^2_{W^*}$, then (leaving the computation to the reader)

$$\frac{1}{2}\bar{\partial}E_{S,S'}(\mathrm{id}) = \sum_{k=1}^{N}\Big(\sum_{i:x_k\in f_i}(\nabla\zeta(c_i)\frac{a_i}{3} - \zeta(c_i)e_{ik}\times N_i)\Big)\delta_{x_k},$$

where e_{ik} is the edge opposite x_k in f_i (oriented so that (x_k, e_{ik}) is positively ordered), and $\zeta = Z^S - Z^{S'}$, with

$$Z^{\tilde{S}}(\cdot) = \sum_{i=1}^{\tilde{K}} \xi(\cdot, \tilde{c}_i)\tilde{a}_i$$

for a triangulated surface $\tilde{S}$. The Eulerian differential at φ is obtained by replacing all x_k's by $\varphi(x_k)$.

For the vector-measure form, $E_{S,S'}(\varphi) = \|\nu_{\varphi\cdot S} - \nu_{S'}\|^2_{W^*}$, we get

$$\frac{1}{2}\bar{\partial}E_{S,S'}(\mathrm{id}) = \sum_{k=1}^{N}\Big(\sum_{i:x_k\in f_i}(d\zeta(c_i)N_i)\frac{a_i}{3} - e_{ik}\times\zeta(c_i)\Big)\delta_{x_k}$$

still with $\zeta = Z^S - Z^{S'}$, but with

$$Z^{\tilde{S}}(\cdot) = \sum_{i=1}^{\tilde{K}}\xi(\cdot,\tilde{c}_i)N_i a_i.$$

9.7.4 Induced Actions and Currents

We have designed the action of diffeomorphisms on measures by $(\varphi\cdot\mu\,|\,h) = (\mu\,|\,h\circ\varphi)$. Recall that we have the usual action of diffeomorphisms on functions defined by $\varphi\cdot h = h\circ\varphi^{-1}$, so that we can write $(\varphi\cdot\mu\,|\,h) = \big(\mu\,\big|\,\varphi^{-1}\cdot h\big)$. In the case of curves, we have seen that this action on the induced measure did not correspond to the image of the curve by a diffeomorphism, in the sense that $\mu_{\varphi\cdot\gamma}\neq\varphi\cdot\mu_\gamma$. Here, we discuss whether the transformations $\mu_\gamma\to\mu_{\varphi\cdot\gamma}$ or $\nu_\gamma\to\nu_{\varphi\cdot\gamma}$ (and the equivalent transformations for surfaces) can be described by a similar operation, e.g., whether one can write $(\varphi\cdot\mu\,|\,h) = \big(\mu\,\big|\,\varphi^{-1}\star h\big)$ where $\star$ would represent another action of diffeomorphisms on functions (or on vector fields for vector measures).

For μ_γ, the answer is negative. We have, letting $T(\gamma(u))$ be the unit tangent to γ,

$$\begin{aligned}\big(\mu_{\varphi\cdot\gamma}\,\big|\,h\big) &= \int_a^b h(\varphi(\gamma(u)))|d\varphi(\gamma(u))\dot{\gamma}(u)|du\\ &= \int_a^b h(\varphi(\gamma(u)))|d\varphi(\gamma(u))T(u)||\dot{\gamma}(u)|du,\end{aligned}$$

so that $\big(\mu_{\varphi\cdot\gamma}\,\big|\,h\big) = \big(\mu_\gamma\,\big|\,h\circ\varphi|d\varphi\,T|\big)$, with some abuse of notation in the last formula, since T is only defined along γ. The important fact here is that the function h is transformed according to a rule which depends not only on the diffeomorphism φ, but also on the curve γ, and therefore the result cannot be put in the form $\varphi^{-1}\star h$.

The situation is different for vector measures. Indeed, we have

$$\begin{aligned}\nu_{\varphi\cdot\gamma}(h) &= \int_a^b (d\varphi(\gamma(u))\dot{\gamma}(u))^T h\circ\varphi(\gamma(u))du\\ &= \big(\nu_\gamma\,\big|\,d\varphi^T h\circ\varphi\big).\end{aligned}$$

So, if we define $\varphi \star h = d(\varphi^{-1})^T h \circ \varphi^{-1}$, we have $\big(\nu_{\varphi\cdot\gamma} \,\big|\, h\big) = \big(\nu_\gamma \,\big|\, \varphi^{-1} \star h\big)$. The transformation $(\varphi, h) \mapsto \varphi \star h$ is a valid action of diffeomorphisms on vector fields, since $\mathrm{id} \star h = h$ and $\varphi \star (\psi \star h) = (\varphi \circ \psi) \star h$, as can easily be checked.

The same analysis can be made for surfaces; scalar measures do not transform in accordance to an action, but vector measures do. Let us check this last point by considering the formula in a local chart, where

$$\begin{aligned}\big(\nu_{\varphi(S)} \,\big|\, h\big) &= \int \det(d\varphi\, \partial_1 m, d\varphi\, \partial_2 m, h \circ \varphi) du dv \\ &= \int \det(d\varphi) \det(\partial_1 m, \partial_2 m, (d\varphi)^{-1} h \circ \varphi) du dv \\ &= \big(\nu_S \,\big|\, \det(d\varphi)(d\varphi)^{-1} h \circ \varphi\big).\end{aligned}$$

So, we need here to define

$$\varphi \star h = \det(d(\varphi^{-1}))(d\varphi^{-1})^{-1} h \circ \varphi^{-1} = (d\varphi\, h / \det(d\varphi)) \circ \varphi^{-1}.$$

Here again, a direct computation shows that this is an action.

We have just proved that vector measures are transformed by a diffeomorphism φ according to a rule $(\varphi \cdot \mu \,|\, h) = \big(\mu \,\big|\, \varphi^{-1} \star h\big)$, the action $\star$ being apparently different for curves and surfaces. In fact, all these actions (including the scalar one) can be placed within a single framework if one replaces vector fields by differential forms and measures by currents [126, 127, 289].

The reader may refer to Sects. B.7.1 and B.7.2 for basic definitions of linear and differential forms, in which the space of differential k-forms on $\mathbb{R}^d$ is denoted Ω_k, or Ω_k^d. We can consider spaces of smooth differential k-forms, and in particular, reproducing kernel Hilbert spaces of such forms: a space $W \subset \Omega_k$ is an RKHS if, for every $x \in \mathbb{R}^d$ and $e_1, \ldots, e_k \in \mathbb{R}^d$, the evaluation function

$$(e_1, \ldots, e_k)\delta_x : q \mapsto (q(x) \,|\, e_1, \ldots, e_k)$$

belongs to W^*. Introduce the duality operator, so that $\mathbb{K}_W((e_1, \ldots, e_k)\delta_x) \in W$. Introduce, for $x, y \in \mathbb{R}^d$, the $2k$-linear form $\xi(x, y)$ defined by

$$(\xi(x, y) \,|\, e_1, \ldots, e_k; f_1, \ldots, f_k) = (\mathbb{K}_W((e_1, \ldots, e_k)\delta_x)(y) \,|\, f_1, \ldots, f_k).$$

Notice that this form is skew-symmetric with respect to its first k and its last k variables and that

$$\big\langle \xi_x(e_1, \ldots, e_k)\,,\, \xi_y(f_1, \ldots, f_k) \big\rangle_W = (\xi(x, y) \,|\, e_1, \ldots, e_k; f_1, \ldots, f_k),$$

so that ξ may be called the reproducing kernel of W. Similar to vector fields, kernels for differential k-forms can be derived from scalar kernels by letting

$$(\xi(x,y) \mid e_1,\dots,e_k; f_1,\dots,f_k) = \xi(x,y)\big\langle e_1 \times \cdots \times e_k\,,\; f_1 \times \cdots \times f_k\big\rangle_{\Lambda_{d-k}}, \tag{9.57}$$

where the dot product on the space of k-linear forms, Λ_k, is the product of coefficients of the forms over a basis formed by all cross products of subsets of k elements of an orthonormal basis of $\mathbb{R}^d$, as described in Sect. B.7.1.

Elements of the dual space, W^*, to W are therefore linear forms over differential k-forms, and are special instances of k-currents [107, 210] (k-currents are bounded differential forms over C^∞ differential k-forms with compact support, which is less restrictive than being bounded on W). Important examples of currents are those associated to submanifolds of $\mathbb{R}^d$, and are defined as follows. Let M be an oriented k-dimensional submanifold of $\mathbb{R}^d$. To a differential k-form q, associate the quantity

$$(\eta_M \mid q) = \int_M (q(x) \mid e_1(x),\dots,e_k(x)), d\sigma_M(x)$$

where $e_1,\dots,e_k$ is, for all x, a positively oriented orthonormal basis of the tangent space to M at x (by Eq. (B.16), the result does not depend on the chosen basis).

If W is an RKHS of differential k-forms, η_M belongs to W^* and we can compute the dual norm of η_M, which is

$$\|\eta_M\|^2_{W^*} = \int_M \int_M (\xi(x,y) \mid e_1(x),\dots,e_k(x); e_1(y),\dots,e_k(y)) d\sigma_M(x) d\sigma_M(y)$$

or, for a scalar kernel defined by (9.57),

$$\|\eta_M\|^2_{W^*} = \int_M \int_M \xi(x,y)\big\langle e_1(x) \times \cdots \times e_k(x)\,,\; e_1(y) \times \cdots \times e_k(y)\big\rangle_{\Lambda_{d-k}} d\sigma_M(x) d\sigma_M(y).$$

The expressions of η_M and its norm in a local chart of M are quite simple. Indeed, if $(u_1,\dots,u_k)$ is the parametrization in the chart and $(\partial_1 m,\dots,\partial_k m)$ the associated tangent vectors (assumed to be positively oriented), we have, for a k-form q (using (B.16))

$$(q \mid \partial_1 m,\dots,\partial_k m) = (q \mid e_1,\dots,e_k) \det(\partial_1 m,\dots,\partial_k m)$$

which immediately yields

$$(q \mid \partial_1 m,\dots,\partial_k m) du_1 \dots du_k = (q \mid e_1,\dots,e_k) d\sigma_M.$$

We therefore have, in the chart,

$$(\eta_M \mid q) = \int (q \mid \partial_1 m,\dots,\partial_k m) du_1 \dots du_k$$

and similar formulas for the norm.

Now consider the action of diffeomorphisms. If M becomes $\varphi(M)$, the formula in the chart yields

$$\left(\eta_{\varphi(M)} \,\middle|\, q\right) = \int (q \circ \varphi \mid d\varphi \partial_1 m, \ldots, d\varphi \partial_k m) du_1 \ldots du_k$$

so that $\left(\eta_{\varphi(M)} \,\middle|\, q\right) = (\eta_M \mid \tilde{q})$ with

$$(\tilde{q}(x) \mid f_1, \ldots, f_k) = (q(\varphi(x)) \mid d\varphi f_1, \ldots, d\varphi f_k).$$

As we did with vector measures, we can introduce the left action on k-forms (also called the push-forward of the k-form):

$$(\varphi \star q \mid f_1, \ldots, f_k) = \left(q \circ \varphi^{-1} \,\middle|\, d(\varphi^{-1}) f_1, \ldots, d(\varphi^{-1}) f_k\right)$$

and the resulting action on p-currents

$$(\varphi \cdot \eta \mid q) = \left(\eta \,\middle|\, \varphi^{-1} \star q\right), \tag{9.58}$$

so that we can write $\eta_{\varphi(M)} = \varphi \cdot \eta_M$.

This is reminiscent of what we have obtained for measures, and for vector measures with curves and surfaces. We now check that these examples are particular cases of the previous discussion.

Measures are linear forms on functions, which are also differential 0-forms. The definition $(\varphi \cdot \mu \mid f) = (\mu \mid f \circ \varphi)$ is exactly the same as in (9.58).

Consider now the case of curves, which are 1D submanifolds, so that $k = 1$. If γ is a curve, and T is its unit tangent, we have

$$\left(\eta_\gamma \,\middle|\, q\right) = \int_\gamma (q(\gamma) \mid T) d\sigma_\gamma = \int_a^b (q(\gamma(u)) \mid \dot{\gamma}(u)) du.$$

To a vector field h on $\mathbb{R}^d$, we can associate the differential 1-form q_h defined by $(q_h(x) \mid v) = h(x)^T v$. In fact all differential 1-forms can be expressed as q_h for some vector field h. Using this identification and noting that $\left(\nu_\gamma \,\middle|\, h\right) = \left(\eta_\gamma \,\middle|\, q_h\right)$, we can see that the vector measure for curve matching is a special case of the currents that we have considered here.

For surfaces in three dimensions, we need to take $k = 2$, and if S is a surface, we have

$$(\eta_S \mid q) = \int_S (q(x) \mid e_1(x), e_2(x)) d\sigma_S(x).$$

Again, a vector field f on $\mathbb{R}^3$ induces a 2-form q_f, defined by $\left(q_f \,\middle|\, v_1, v_2\right) = \det(f, v_1, v_2) = f^T(v_1 \times v_2)$, and every 2-form can be obtained this way. Using the

fact that, if (e_1, e_2) is a positively oriented basis of the tangent space to the surface, then $e_1 \times e_2 = N$, we retrieve $(\nu_S \mid f) = \left(\eta_S \,\middle|\, q_f\right)$.

9.7.5 Varifolds

A differential k-form ω on $\mathbb{R}^d$ uniquely defines a function on the product space $\mathbb{R}^d \times \widetilde{\mathrm{Gr}}(d,k)$, the product space of $\mathbb{R}^d$ with the set of all oriented k-dimensional subspaces of $\mathbb{R}^d$ (called the oriented Grassmannian, on which a manifold structure similar to the one discussed in Sect. B.6.7 for the Grassmann manifold can be defined). One can indeed assign to any pair (x, α) in that set the scalar $F_q(x, \alpha) = (q(x) \mid e_1, \ldots, e_k)$ where $e_1, \ldots, e_k$ is any positively oriented orthonormal basis of α, and the value does not depend on the chosen basis. Given an oriented k-dimensional submanifold of $\mathbb{R}^d$, one can define the linear form on continuous functions F defined on $\mathbb{R}^d \times \widetilde{\mathrm{Gr}}(d,k)$, given by

$$(\tilde{\rho}_M \mid F) = \int_M F(p, T_p M) d\sigma_M,$$

where $T_p M$ is considered with its orientation. The current η_M defined in the previous section is such that $(\eta_M \mid q) = \left(\tilde{\rho}_M \,\middle|\, F_q\right)$.

When one wants to disregard orientation, which may be convenient, and sometimes necessary in practice, it is natural to replace $\widetilde{\mathrm{Gr}}(d,k)$ by $\mathrm{Gr}(d,k)$ (the Grassmannian) and define the same linear form (that we now call ρ_M) on functions defined on $\mathbb{R}^d \times \mathrm{Gr}(d,k)$. The linear form ρ_M is a special case of a *varifold*, where varifolds are defined as (Radon) measures on $\mathbb{R}^d \times \mathrm{Gr}(d,k)$.

From this point, and following [61], one can make a construction analogous to the one just described for measures on $\mathbb{R}^d$. Given a reproducing kernel Hilbert space W of functions defined on $\mathbb{R}^d \times \mathrm{Gr}(d,k)$, define the square distance between two k-dimensional submanifolds of $\mathbb{R}^d$ by

$$D(M, M') = \|\rho_M - \rho_{M'}\|^2_{W^*}.$$

For the approach to be practical, one needs to have explicit kernels on $\mathbb{R}^d \times \mathrm{Gr}(d,k)$. Referring to [61] for a complete discussion, we note here that a class of such kernels can be designed based on the following observations.

(i) The function defined for $\alpha, \beta \in \widetilde{\mathrm{Gr}}(d,k)$ by

$$\tilde{\xi}(\alpha, \beta) = \left\langle e_1(x) \times \cdots \times e_k(x)\,,\ f_1(y) \times \cdots \times f_k(y) \right\rangle_{\Lambda_{d-k}},$$

where $(e_1, \ldots, e_k)$ and $(f_1, \ldots, f_k)$ are positively oriented orthonormal bases of $\tilde{\alpha}$ and $\tilde{\beta}$, is a positive definite kernel. Hence, the function defined for $\alpha, \beta \in \mathrm{Gr}(d,k)$

$$\xi(\alpha, \beta) = \left\langle e_1(x) \times \cdots \times e_k(x)\,,\ f_1(y) \times \cdots \times f_k(y) \right\rangle^2_{\Lambda_{d-k}},$$

where $(e_1, \ldots, e_k)$ and $(f_1, \ldots, f_k)$ are orthonormal bases of α and β, is also definite positive. More generally, if $\tilde{\xi}$ is positive definite on $\widetilde{\mathrm{Gr}}(d, k)$, then $\xi = f(\tilde{\xi})$ is definite positive on $\mathrm{Gr}(d, k)$ for any even analytic function f whose derivatives at 0 are all non-negative, and at least one of them positive. These statements simply use the fact that products of positive kernels remain positive.

(ii) If η is a positive kernel on differential p-forms, then $\tilde{\xi}$ defined by

$$\tilde{\xi}(x, \alpha; y, \beta) = (\eta(x, y) \mid e_1, \ldots, e_k; f_1, \ldots, f_k)$$

is a positive kernel on $\mathbb{R}^d \times \widetilde{\mathrm{Gr}}(d, k)$.

(iii) If $\xi^{(1)}$ is a reproducing kernel on $\mathbb{R}^d$ and $\xi^{(2)}$ a reproducing kernel on $\mathrm{Gr}(d, k)$, then ξ defined by

$$\xi(x, \alpha; y, \beta) = \xi^{(1)}(x, y)\xi^{(2)}(\alpha, \beta)$$

is a reproducing kernel on $\mathbb{R}^d \times \mathrm{Gr}(d, k)$.

Applying this to surfaces, for example, and using the discussion at the end of the previous section, we find that taking

$$\left\langle \rho_S\,,\ \rho_{\tilde{S}} \right\rangle_{W^*} = \int_S \int_{\tilde{S}} \xi(x, \tilde{x}) \left(1 + a\left(N(x)^T \tilde{N}(\tilde{x})\right)^2\right) d\sigma_{\tilde{S}}(\tilde{x}) d\sigma_S(x),$$

where ξ is a reproducing kernel on $\mathbb{R}^d$, provides an RKHS dual inner-product on varifolds. The discretization of such a norm is similar to those detailed for scalar and vector measures and is left to the reader.

9.8 Matching Vector Fields

We now study vector fields as deformable objects. They correspond, for example, to velocity fields (that can be observed for weather data), or to gradient fields that can be computed for images. Orientation fields (that can be represented by unit vector fields) are also interesting. They can correspond, for example, to fiber orientations in tissues observed in medical images.

We want to compare two vector fields f and f', i.e., two functions from $\mathbb{R}^d$ to $\mathbb{R}^d$. To simplify, we restrict ourselves to $E_{f,f'}(\varphi)$ being the L^2 norm between $\varphi \cdot f$ and f', and focus our discussion on the definition of the action of diffeomorphisms on vector fields.

The simplest choice is to use the same action as in image matching and take $\varphi \cdot f = f \circ \varphi^{-1}$, where f is a vector field on $\mathbb{R}^d$. It is, however, natural (and more consistent with applications) to combine the displacement of the points at which f

is evaluated with a reorientation of f, also induced by the transformation. Several choices can be made for such an action and all may be of interest depending on the context.

For example, we can interpret a vector field as a velocity field, assuming that each point in Ω moves on to a trajectory $x(t)$ and that $f(x) = \dot{x}(t)$, say at time $t = 0$. If we make the transformation $x \mapsto x' = \varphi(x)$, and let f' be the transformed vector field, such that $\dot{x}'(0) = f'(x')$, we get: $\dot{x}'(0) = d\varphi(x)\dot{x}(0) = f' \circ \varphi(x)$ so that $f' = (d\varphi\, f) \circ \varphi^{-1}$. The transformation $f \mapsto (d\varphi\, f) \circ \varphi^{-1}$ is an important Lie group operation, called the adjoint representation ($\mathrm{Ad}_\varphi f$). This is anecdotal here, but we will use it again later as a fundamental tool. So, our first action is

$$\varphi * f = (d\varphi\, f) \circ \varphi^{-1}.$$

To define a second action, we now consider vector fields that are obtained as gradients of a function I: $f = \nabla I$. If I becomes $\varphi \cdot I = I \circ \varphi^{-1}$, then f becomes $d(\varphi^{-1})^T \nabla I \circ \varphi^{-1}$. This defines a new action

$$\varphi \star f = d(\varphi^{-1})^T f \circ \varphi^{-1} = (d\varphi^{-T}\, f) \circ \varphi^{-1}.$$

This action can be applied to any vector field, not only gradients, but one can check that the set of vector fields f such that $\mathrm{curl}\, f = 0$ is left invariant by this action.

Sometimes, it is important that the norms of the vector fields at each point remain invariant under the transformation, when dealing, for example, with orientation fields. This can be achieved in both cases by normalizing the result, and we define the following normalized actions:

$$\varphi \bar{*} f = \left(|f| \frac{d\varphi f}{|d\varphi f|} \right) \circ \varphi^{-1}$$
$$\varphi \bar{\star} f = \left(|f| \frac{d\varphi^{-T} f}{|d\varphi^{-T} f|} \right) \circ \varphi^{-1}$$

(taking, in both cases, the right-hand side equal to 0 if $|f| = 0$).

We now evaluate the differential of $E_{f,f'}(\varphi) = \|\varphi \cdot f - f'\|_2^2$, where $\varphi \cdot f$ is one of the actions above. We will make the computation below under the assumption that f is C^1 and compactly supported. For the $*$ action, we can observe that, for $\varphi = \mathrm{id} + h$,

$$\varphi * f - f = dh\, f \circ (\mathrm{id} + h)^{-1} + f \circ (\mathrm{id} + h)^{-1} - f,$$

so that

$$\varphi * f - f - dh\, f + df\, h = dh(f \circ (\mathrm{id} + h)^{-1} - f) + f \circ (\mathrm{id} + h)^{-1} - f \circ (\mathrm{id} - h) + f \circ (\mathrm{id} - h) - f + df\, h.$$

Using the fact that $\|(\mathrm{id}+h)^{-1}-\mathrm{id}\|_\infty = \|h\|_\infty$,

$$\|(\mathrm{id}+h)^{-1}-(\mathrm{id}-h)\|_\infty = \|h\circ(\mathrm{id}+h)-h\|_\infty \le \|h\|^2_{1,\infty}$$

and letting

$$\omega_f^{(1)}(\varepsilon) = \sup_{x\in\mathbb{R}^d}\sup_{|\delta|<\varepsilon} |f(x+\delta)-f(x)-df(x)\delta|$$

we find that

$$\|\varphi * f - f - dh\, f + df\, h\|_\infty \le \|f\|_\infty \|h\|^2_{1,\infty} + \|f\|_{1,\infty}\|h\|^2_{1,\infty} + \omega_f^{(1)}(\|h\|_\infty).$$

Noting that $\omega_f^{(1)}(\varepsilon) = o(\varepsilon)$, we find that

$$\|\varphi * f - f - dh\, f + df\, h\|_\infty = o(\|h\|_{1,\infty}). \tag{9.59}$$

Using this estimate, it is now easy to show that $E_{f,f'} : \mathrm{Diff}_0^{1,\infty} \to \mathbb{R}$ is differentiable at $\varphi = \mathrm{id}$ with derivative

$$\begin{aligned}\big(dE_{f,f'}(\mathrm{id}) \,\big|\, h\big) &= 2\big\langle dh\, f - df\, h\,,\; f - f'\big\rangle_2 \\ &= 2\int_\Omega (dh\, f - df\, h)^T (f-f')dx.\end{aligned}$$

For $f \mapsto \varphi * f$ to map compactly supported C^1 vector fields into vector fields with the same property, we need to take twice-differentiable diffeomorphisms, i.e., $\varphi \in \mathrm{Diff}_0^{2,\infty}$. Over this group, we find that $E_{f,f'}$ is differentiable everywhere, with $\big(dE_{f,f'}(\psi) \,\big|\, h\big) = \big(dE_{\psi * f,f'}(\mathrm{id}) \,\big|\, h\circ\psi^{-1}\big)$.

The Eulerian derivative is then given by

$$\begin{aligned}\big(\bar\partial E_{f,f'}(\psi) \,\big|\, v\big) &= \big(dE_{\psi*f,f'}(\mathrm{id}) \,\big|\, v\big) \\ &= 2\big\langle dv\,(\psi*f) - d(\psi*f)\,v\,,\; \psi*f - f'\big\rangle_2.\end{aligned}$$

This expression can be combined with (9.44) to obtain the Eulerian gradient of U, namely

$$\overline{\nabla}^V E_{f,f'}(\psi) = 2\int_\Omega \big(\partial_2 K(.,x)(\psi*f-f', \psi*f) - K(.,x)d(\psi*f)^T(\psi*f-f')\big)\,dx.$$

The Eulerian differential can be rewritten in another form to avoid the intervention of the differential of h. The following lemma is a consequence of the divergence theorem.

Lemma 9.10 *If Ω is a bounded open domain of $\mathbb{R}^d$ and v, w, h are smooth vector fields on $\mathbb{R}^d$, then*

$$\int_\Omega v^T dhw\, dx = \int_{\partial\Omega} (v^T h)(w^T N) d\sigma_{\partial\Omega} - \int_\Omega \big(w^T dv^T h + (\operatorname{div} w)(v^T h)\big) dx. \quad (9.60)$$

Equation (9.60) can be rewritten as

$$\big\langle dh\, w\,,\ v\big\rangle_2 = \big((w^T N) v \sigma_{\partial\Omega}\ \big|\ h\big) - \big\langle dv\, w + (\operatorname{div} w)\, v\,,\ h\big\rangle_2. \quad (9.61)$$

Proof To prove this, introduce the coordinates $h^1, \ldots, h^d$ for h and $v^1, \ldots, v^d$ for v so that

$$v^T dhw = \sum_{i=1}^d v^i (\nabla h^i)^T w.$$

Now, use the fact that

$$\begin{aligned}\operatorname{div}(v^T h\, w) &= \operatorname{div}\Big(\sum_{i=1}^d v^i h^i w\Big)\\ &= \sum_{i=1}^d \big(v^i (\nabla h^i)^T w + h^i (\nabla v^i)^T w + h_i v^i \operatorname{div} w\big)\\ &= v^T dhw + h^T dvw + (h^T v)\operatorname{div} w\end{aligned}$$

and the divergence theorem to obtain the result. □

Using this lemma with Ω large enough so that f vanishes on $\partial\Omega$, we find

$$\big(dE_{f,f'}(\mathrm{id})\ \big|\ h\big) = -2\big\langle (df - df')\, f + \operatorname{div} f\, (f - f') + df^T (f - f')\,,\ h\big\rangle_2,$$

which directly provides a new version of the Eulerian derivative at an arbitrary φ, with the corresponding new expression of the Eulerian gradient:

$$\overline{\nabla}^V E_{f,f'}(\varphi) = -2\int_\Omega K(\cdot, x)\Big(d(\varphi * f - f')(\varphi * f) + \operatorname{div}(\varphi * f)(\varphi * f - f') + d(\varphi * f)^T (\varphi * f - f')\Big) dx.$$

Let us now consider the normalized version of this action. We will make the computation under a few additional assumptions on f, namely, that f is compactly supported and $f/|f|$ can be replaced by a smooth unit vector field that can be extended to an open set that contains the support of f. More precisely, we will assume that there exists a scalar function ρ, continuously differentiable and supported by a compact set Q, and a vector field u such that $|u(x)| = 1$ for all x in an open set Ω containing Q, u is continuously differentiable on Ω and $f(x) = \rho(x)u(x)$. With this notation,

we have

$$\varphi \bar{*} f = \rho \circ \varphi^{-1} \frac{\varphi * u}{|\varphi * u|}.$$

Note that, for $z \neq 0$, the derivative of $z/|z|$ is

$$h \mapsto h/|z| - zz^T h/|z|^3 = \frac{1}{|z|}\pi_{z^\perp}(h),$$

where $\pi_{z^\perp}$ is the orthogonal projection on the space of vectors perpendicular to z.

Consider $\varphi = \mathrm{id} + h$ for some $h \in \mathrm{Diff}_0^{1,\infty}$. Let $\delta = \mathrm{dist}(Q, \Omega^c)$ and assume that

$$\|h\|_\infty = \|\varphi^{-1} - \mathrm{id}\|_\infty < \delta/2.$$

From (9.59), we have, letting Ω' be the set of $x \in \mathbb{R}^d$ such that $\mathrm{dist}(x, Q) < \delta$,

$$\sup_{x\in\Omega'} |(\varphi * u)(x) - u(x) - dh(x)u(x) + du(x)h(x)| = o(\|h\|_{1,\infty}),$$

from which we deduce

$$\sup_{x\in\Omega'} \left| \frac{(\varphi * u)(x)}{|(\varphi * u)(x)|} - u(x) - \pi_{u^\perp}\left(dh(x)u(x) - du(x)h(x)\right)\right| = o(\|h\|_{1,\infty}).$$

Since $\|\rho \circ \varphi^{-1} - \rho - \nabla\rho^T h\|_\infty = o(\|h\|_\infty)$, we obtain the fact that $E_{f,f'}$ is differentiable at $\varphi = \mathrm{id}$ with

$$\begin{aligned}\big(dE_{f,f'}(\mathrm{id}) \,\big|\, h\big) &= 2\big(-\nabla\rho^T h\, u + \rho\,\pi_{u^\perp}(dh\, u - du\, h) \,\big|\, f - f'\big)\\ &= -2\big(\nabla\rho^T h\, u \,\big|\, f - f'\big) - 2\big(\rho\,(dh\, u - du\, h) \,\big|\, \pi_{u^\perp}(f')\big).\end{aligned}$$

Assuming now that $\psi \in \mathrm{Diff}_0^{2,\infty}$, we obtain the fact that $E_{f,f'}$ is differentiable at ψ, with

$$\big(dE_{f,f'}(\psi) \,\big|\, h\big) = \big(dE_{\psi\bar{*}f,f'}(\mathrm{id}) \,\big|\, h \circ \psi^{-1}\big).$$

The Eulerian derivative is

$$\big(\bar{\partial}E_{f,f'}(\psi) \,\big|\, v\big) = \big(dE_{\psi\bar{*}f,f'}(\mathrm{id}) \,\big|\, v\big).$$

Finally, we note that after integration by parts, we can write

$$\begin{aligned}dE_{f,f'}(\mathrm{id}) = 2\big(&- u^T(f - f')\nabla\rho + \rho\, du^T(\pi_{u^\perp}(f'))\\ &+ d(\pi_{u^\perp}(f'))f + \mathrm{div}(f)\pi_{u^\perp}(f')\big)dx.\end{aligned}$$

The computations for $\varphi \star f = (d\varphi^{-T} f) \circ \varphi^{-1}$ and its normalized version are very similar. One only needs to note that (9.59) is now replaced by

$$\|\varphi \star f - f + dh^T f + df\, h\|_\infty = o(\|h\|_{1,\infty}). \tag{9.62}$$

As a consequence, the formulas for the differentials of the $\star$ and $\bar{\star}$ can be deduced from the $*$ and $\bar{*}$ actions by replacing $(dh\, f - df\, f)$ by $(-dh^T f - df\, h)$.

For the unnormalized action, this yields

$$\begin{aligned}\big(dE_{f,f'}(\mathrm{id}) \,\big|\, h\big) &= -2\big\langle dh(f - f')\,,\, f\big\rangle_2 - 2\big\langle df^T(f - f')\,,\, h\big\rangle_2 \\ &= 2\big\langle (df - df^T)(f - f') + \mathrm{div}(f - f')f\,,\, h\big\rangle_2\end{aligned}$$

and $\bar{\partial} E_{f,f'}(\varphi)$ is obtained by replacing f by $\varphi \star f$. To obtain the differential of $E_{f,f'}$ for the normalized $\star$ action, we get

$$\big(dE_{f,f'}(\mathrm{id}) \,\big|\, h\big) = -2\big(\nabla\rho^T h\, u \,\big|\, f - f'\big) + 2\big(\rho\,(dh^T u + du\, h) \,\big|\, \pi_{u^\perp}(f')\big),$$

where $f = \rho\, u$ as above, which can also be written as

$$dE_{f,f'}(\mathrm{id}) = 2\big(-u^T(f - f')\nabla\rho + (\rho du^T - df)(\pi_{u^\perp}(f')) - \mathrm{div}(\pi_{u^\perp}(f'))f\big)dx.$$

As an example of application of vector field matching, let us consider contrast-invariant image registration [90]. If $I : \Omega \to \mathbb{R}$ is an image, a change of contrast is a transformation $I \mapsto q \circ I$, where q is a scalar diffeomorphism of the image intensity range. The level sets $I_\lambda = \{x, I(x) \le \lambda\}$ are simply relabeled by a change of contrast, and one obtains a contrast-invariant representation of the image by considering the normals to these level sets, i.e., the vector field

$$f = \nabla I/|\nabla I|$$

with the convention that $f = 0$ when $\nabla I = 0$. Two images represented in this way can now be compared using vector field matching. Since we are using normalized gradients, the natural action is $(\varphi, f) \mapsto \varphi \bar{\star} f$. For our results to hold, some regularization needs to be applied, replacing f by $\rho \tilde{f}$ where $\rho = 1$ and $\tilde{f} = f$ when $|f| = 1$, $\tilde{f}$ is a unit vector field that smoothly extends f over a neighborhood of the domain over which $|f| = 1$, ρ is smooth and vanishes outside this neighborhood.

9.9 Matching Fields of Frames

We now extend vector field deformation models to define an action of diffeomorphisms on fields of positively oriented orthogonal matrices, or frames. We will restrict ourselves to dimension 3, so that the deformable objects considered in this section are mappings $x \mapsto R(x)$, with, for all $x \in \Omega$, $R(x) \in \mathrm{SO}_3(\mathbb{R})$ (the group of rotation matrices).

The $*$ and $\star$ actions we have just defined on vector fields have the nice property of conserving the Euclidean dot product when combined, that is

$$(\varphi * f)^T (\varphi \star g) = (f^T g) \circ \varphi^{-1}.$$

Since $\bar{*}$ and $\bar{\star}$ also conserve the norm, we find that $(\varphi \bar{*} f, \varphi \bar{\star} g)$ is orthonormal as soon as (f, g) is.

We now define an action of diffeomorphisms on fields on frames. Writing $R(x) = (f_1(x), f_2(x), f_3(x))$, we let

$$\varphi \cdot R = (\varphi \bar{*} f_1, (\varphi \bar{\star} f_3) \times (\varphi \bar{*} f_1), \varphi \bar{\star} f_3). \tag{9.63}$$

That this defines an action is a straightforward consequence of $\bar{*}$ and $\bar{\star}$ being actions.

The action can be interpreted as follows. Given a local chart in $\mathbb{R}^3$, which is a diffeomorphic change of coordinates $x = m(s, t, u)$, one uniquely specifies a positively oriented frame $R_m = (f_1, f_2, f_3)$ by $f_1 = \partial_1 m/|\partial_1 m|$ and $f_3 = (\partial_1 m \times \partial_2 m)/|\partial_1 m \times \partial_2 m|$. Then, the action we have just defined is such that $\varphi \cdot R$ is the frame associated to the change of coordinates $\varphi \circ m$, i.e.,

$$R_{\varphi \circ m} \circ \varphi = \varphi \cdot R_m.$$

The transformation $m \to R_m$ has in turn the following interpretation, which is relevant for some medical imaging modalities. Let the change of coordinates be adapted to the following stratified description of a tissue. Curves $s \mapsto m(s, t, u)$ correspond to tissue fibers, and surfaces $(s, t) \mapsto m(s, t, u)$ describe a layered organization. The cardiac muscle, for example, exhibits this kind of structure. Then f_1 in R_m represents the fiber orientation, and f_3 the normal to the layers; $\varphi \cdot R_m$ then corresponds to the tissue to which the deformation φ has been applied.

Frame fields are typically observed over some object-dependent subregion of the observation domain. To account for this, we assume that we are dealing with weighted fields of frames, taking the form $A = \rho R$, where the weight ρ vanishes outside a compact set and R is smooth over a neighborhood of this compact set. We will then consider the action

$$\varphi \cdot A = (\rho \circ \varphi^{-1})\, \varphi \cdot R.$$

The computations of the previous sections can now be applied to each column of A. In particular, letting $\varphi = \mathrm{id} + h$, we have $\varphi \cdot A = A + (w_1, w_2, w_3) + o(\|h\|_{1,\infty})$ with, writing $R = (f_1, f_2, f_3)$,

$$\begin{cases} w_1 = -(\nabla \rho^T h) f_1 + \rho \pi_{f_1^\perp}(dh f_1 - df_1 h), \\ w_3 = -(\nabla \rho^T h) f_3 - \rho \pi_{f_3^\perp}(dh^T f_3 + df_3 h), \\ w_2 = -(\nabla \rho^T h) f_2 - \rho \left(\pi_{f_3^\perp}(dh^T f_3 + df_3 h) \right) \times f_1 \\ \qquad\quad + \rho f_3 \times \left(\pi_{f_1^\perp}(dh f_1 - df_1 h) \right). \end{cases}$$

Noticing that, for any vector $u \in \mathbb{R}^3$,

$$\left(\pi_{f_3^\perp} u\right) \times f_1 = \left((u^T f_1) f_1 + (u^T f_2) f_2\right) \times f_1 = -(u^T f_2) f_3$$

and similarly $f_3 \times \left(\pi_{f_1^\perp} u\right) = -(u^T f_2) f_1$, we can simplify the expression of w_2, yielding

$$\begin{cases} w_1 = -(\nabla \rho^T h) f_1 + \rho \pi_{f_1^\perp}(dh f_1 - df_1 h), \\ w_2 = -(\nabla \rho^T h) f_2 + \rho((dh^T f_3 + df_3 h)^T f_2) f_3 \\ \qquad\quad - \rho((dh f_1 - df_1 h)^T f_2) f_1, \\ w_3 = -(\nabla \rho^T h) f_3 - \rho \pi_{f_3^\perp}(dh^T f_3 + df_3 h). \end{cases} \tag{9.64}$$

Consider the matching functional

$$E_{A,A'}(\varphi) = \int_{\mathbb{R}^3} |\varphi \cdot A - A'|^2 \, dx$$

with $|A|^2 = \text{trace}(A^T A)$. If $A = \rho R$ and $A' = \rho' R'$, then

$$\begin{aligned} |A - A'|^2 &= 3\rho^2 - 2\rho\rho' \text{trace}(R^T R') + 3\rho'^2 \\ &= 3(\rho - \rho')^2 + 2\rho\rho' \text{trace}(\text{Id} - R^T R'). \end{aligned}$$

Introducing the rotation angle, θ, from R to R', defined by

$$\text{trace}(R^T R') = 1 + 2\cos\theta, \tag{9.65}$$

we get

$$|A - A'|^2 = 3(\rho - \rho')^2 + 4\rho\rho'(1 - \cos\theta).$$

Obviously, if $A = (u_1, u_2, u_3)$ and $A' = \rho(u_1', u_2', u_3')$, we also have

$$|A - A'|^2 = |u_1 - u_1'|^2 + |u_2 - u_2'|^2 + |u_3 - u_3'|^2.$$

Using this, one gets the expression of the differential of $E_{A,A'}$ at $\varphi = \text{id}$,

$$\left(dE_{A,A'}(\text{id}) \,\middle|\, h\right) = 2 \int_{\mathbb{R}^2} \text{trace}(W^T (A - A')) \, dx,$$

where $W = (w_1, w_2, w_3)$ is given by (9.64). In particular,

$$\begin{aligned} \text{trace}(W^T (A - A')) = &-\nabla \rho^T h \left(3\rho - \rho' \text{trace}(R^T R')\right) \\ &+ \rho\rho'(dh f_1 - df_1 h)^T \left(-\pi_{f_1^\perp}(f_1') + (f_1^T f_2') f_2\right) \\ &- \rho\rho'(dh^T f_3 + df_3 h)^T \left(-\pi_{f_3^\perp}(f_3') + (f_3^T f_2') f_2\right). \end{aligned}$$

Letting

$$u^1_{A,A'} = \rho\rho' \left(-\pi_{f_1^\perp}(f_1') + (f_1^T f_2') f_2\right)$$
$$u^3_{A,A'} = \rho\rho' \left(-\pi_{f_3^\perp}(f_3') + (f_3^T f_2') f_2\right)$$

and using Lemma 9.10 to eliminate dh, we find

$$dE_{A,A'}(\mathrm{id}) = 2\Big(- \left(3\rho - \rho' \mathrm{trace}(R^T R')\right) \nabla\rho - du^1_{A,A'} f_1 - \mathrm{div}(f_1) u^1_{A,A'} - df_1^T u^1_{A,A'} + \mathrm{div}(u^3_{A,A'}) f_3 \Big) dx. \tag{9.66}$$

9.10 Matching Tensors

The last class of deformable objects we will consider in this chapter are fields of matrices (or tensor fields). For general matrices, we can use the actions we have defined on vector fields, and apply them to each column of M, where M is a field of matrices. The differential of matching functionals is then computed as done in the previous two sections.

One sometimes needs to consider subclasses of tensors, and therefore define an action that leaves this subclass invariant. Here we consider symmetric matrices, which have especially been studied in diffusion tensor imaging (DTI) [6]. The previous actions applied to each column do not work, because they would break the symmetry. A simple choice to address this is to make the diffeomorphism also act on the right, in transpose form, defining, for a field $x \mapsto S(x)$ of symmetric matrices

$$\varphi * S = (d\varphi S d\varphi^T) \circ \varphi^{-1}$$
$$\varphi \star S = (d\varphi^{-T} S d\varphi^{-1}) \circ \varphi^{-1}.$$

We leave to the reader the computation of the differentials of objective functions derived from these actions.

These actions are not necessarily well adapted to DTI data, though, for which alternative options may be considered. DTI produces, at each point x in space, a symmetric positive definite matrix $S(x)$ that measures the diffusion of water molecules in the imaged tissue. Roughly speaking, the tensor $S(x)$ is such that if a water molecule is at s at time t, the probability of being at $x + dx$ at time $t + dt$ is centered Gaussian with variance $dt^2 dx^T S(x) dx$.

If we return to the structured tissue model discussed in the last section (represented by the parametrization $x = m(s, t, u)$), we can assume that molecules travel more easily along fibers, and with most difficulty across layers. So the direction of $\partial_1 m$ is the direction of largest variance, and $\partial_1 m \times \partial_2 m$ of smallest variance, so

that the frame $R_m = (f_1, f_2, f_3)$ associated to the parametrization is such that f_1 is an eigenvector of S for the largest eigenvalue, and f_3 for the smallest eigenvalue, which implies that f_2 is an eigenvector for the intermediate eigenvalue. According to our discussion in the last section, a diffeomorphism φ should transform S so that the frame R_S formed by the eigenbasis of S transforms according to the action of diffeomorphisms on frames, namely, $R_{\varphi\cdot S} = \varphi \cdot R_S$ defined in (9.63).

So, if we express the decomposition of S in the form

$$S = \lambda_1 f_1 f_1^T + \lambda_2 f_2 f_2^T + \lambda_3 f_3 f_3^T$$

with $\lambda_1 \geq \lambda_2 \geq \lambda_3$, we should take

$$\varphi \cdot S = \tilde{\lambda}_1 \tilde{f}_1 \tilde{f}_1^T + \tilde{\lambda}_2 \tilde{f}_2 \tilde{f}_2^T + \tilde{\lambda}_3 \tilde{f}_3 \tilde{f}_3^T \tag{9.67}$$

with $(\tilde{f}_1, \tilde{f}_2, \tilde{f}_3) = \varphi \cdot (f_1, f_2, f_3)$ and $\tilde{\lambda}_i = \lambda_i \circ \varphi^{-1}$, $i = 1, 2, 3$. The action on eigenvalues expresses that intrinsic tissue properties have not been affected by the deformation. If there are reasons to believe that variations in volume should affect the intensity of water diffusion, using the action of diffeomorphisms on densities may be a better option, namely $\tilde{\lambda}_i = \det d(\varphi^{-1})\lambda_i \circ \varphi^{-1}$.

The action with $\tilde{\lambda}_i = \lambda_i \circ \varphi^{-1}$ is identical to the eigenvector-based tensor reorientation discussed in [6]. One of the important (and required) features or the construction is that, although the eigen-decomposition of S is not unique (when two or three eigenvalues coincide) the transformation $S \mapsto \varphi \cdot S$ is defined without ambiguity. This will be justified below.

The following computations require that $\lambda_1, \lambda_2, \lambda_3$ are C^1 and vanish outside a compact set, and that R_S is smooth in a small neighborhood of this compact set. They are sketchily justified here, as they strongly resemble the computations that have been done before. It will be convenient to introduce the three-dimensional rotation $U_S(\varphi) = ((\varphi \cdot R_S) \circ \varphi)\, R_S^T$, so that

$$\varphi \cdot S = (U_S(\varphi) S U_S(\varphi)^T) \circ \varphi^{-1}.$$

Taking $\varphi = \mathrm{id} + h$, we have $U_S(\varphi) - S = \omega_S(h) + o(\|h\|_{1,\infty})$, where

$$\omega_S(h) = \pi_{f_1^\perp} dh f_1 f_1^T - ((\pi_{f_3^\perp} dh^T f_3) \times f_1) f_2^T + (f_3 \times (\pi_{f_1^\perp} dh f_1)) f_2^T - (\pi_{f_3^\perp} dh^T f_3) f_3^T$$

is a skew-symmetric matrix.With this notation, we can write

$$\varphi \cdot S - S = \omega_S(h) S - S \omega_S(h) - dS\, h + o(\|h\|_{1,\infty}).$$

(Here $dS\,h$ is the matrix with coefficients $(\nabla S^{ij})^T h$.) Letting

$$E_{S,S'}(\varphi) = \int_\Omega \mathrm{trace}((\varphi \cdot S - S')^2) dx,$$

we then get

$$\big(dE_{S,S'}(\mathrm{id}) \,\big|\, h\big) = 2\int_\Omega \mathrm{trace}\big((S - S')(\omega_S(h)S - S\omega_S(h) - dS\,h)\big)\, dx,$$

with, as usual, for $\psi \in \mathrm{Diff}_0^{2,\infty}$, $\big(dE_{S,S'}(\psi) \,\big|\, h\big) = \big(dE_{\psi\cdot S,S'}(\mathrm{id}) \,\big|\, h\circ\psi^{-1}\big)$ and $\big(\bar\partial E_{S,S'}(\psi) \,\big|\, v\big) = \big(dE_{\psi\cdot S,S'}(\mathrm{id}) \,\big|\, v\big)$.

Here again, the derivatives of h that are involved in $\omega_S(h)$ can be integrated by parts using the divergence theorem. Let us sketch this computation at $\psi = \mathrm{id}$, which leads to a vector measure form for the differential. We focus on the term

$$(\eta \mid h) := \int_\Omega \mathrm{trace}((S - S')(\omega_S(h)S - S\omega_S(h)))dx = \int_\Omega \mathrm{trace}(A\omega_S(h))dx,$$

where $A = S(S - S') - (S - S')S = SS' - S'S$, and want to express η as a vector measure. We have (using the fact that A is skew symmetric and that (f_1, f_2, f_3) is orthonormal)

$$\begin{aligned}
-\mathrm{trace}(A\omega_S(h)) &= (\omega_S(h)f_1)^T Af_1 + (\omega_S(h)f_2)^T Af_2 + (\omega_S(h)f_3)^T Af_3\\
&= (\pi_{f_1^\perp} dh f_1)^T Af_1 - ((\pi_{f_3^\perp} dh^T f_3)\times f_1)^T Af_2\\
&+ (f_3 \times (\pi_{f_1^\perp} dh f_1))^T Af_2 - (\pi_{f_3^\perp} dh^T f_3)^T Af_3\\
&= (dhf_1)^T u^1_{S,S'} - (dh^T f_3)^T u^3_{S,S'},
\end{aligned}$$

with

$$\begin{aligned}
u^1_{S,S'} &= \pi_{f_1^\perp}(Af_1 + (Af_2 \times f_3))\\
\text{and } u^3_{S,S'} &= \pi_{f_3^\perp}(Af_3 + (f_1 \times Af_2)).
\end{aligned}$$

It now remains to use Lemma 9.10 to identify η as

$$\eta = \big(du^1_{S,S'} f_1 + \mathrm{div}(f_1)u^1_{S,S'} - df_3 u^3_{S,S'} - \mathrm{div}(u^3_{S,S'})f_3\big)dx.$$

To write the final expression of $dE_{S,S'}(\mathrm{id})$, define $(S - S') \odot dS$ to be the vector

$$(S - S') \odot dS = \sum_{i,j=1}^{3} (S^{ij} - (S')^{ij})\nabla S^{ij},$$

so that we have

$$dE_{S,S'}(\mathrm{id}) = 2\big(du^1_{S,S'} f_1 + \mathrm{div}(f_1)u^1_{S,S'} - df_3 u^3_{S,S'} - \mathrm{div}(u^3_{S,S'})f_3 - (S - S') \odot dS\big)dx. \tag{9.68}$$

We now generalize this action to arbitrary dimensions, in a way that will provide a new interpretation of the three-dimensional case. Decompose a field of d by d symmetric matrices S in $\mathbb{R}^d$ in the form

$$S(x) = \sum_{k=1}^{d} \lambda_k(x) f_k(x) f_k(x)^T$$

with $\lambda_1 \geq \cdots \geq \lambda_d$ and $(f_1, \ldots, f_d)$ orthonormal. The matrices $f_k f_k^T$ represent the orthogonal projections on the one-dimensional space $\mathbb{R}\, f_k$ and, letting

$$W_k = \mathrm{span}(f_1, \ldots, f_k),$$

and noting that the projection on W_k, π_{W_k} is equal to $f_1 f_1^T + \cdots + f_k f_k^T$, we can obviously write

$$S(x) = \sum_{k=1}^{d} \lambda_k(x)(\pi_{W_k(x)} - \pi_{W_{k-1}(x)}),$$

where we have set $W_0 = \{0\}$.

Define the action $S \mapsto \varphi \cdot S$ by

$$\varphi \cdot S = \left(\sum_{k=1}^{d} \lambda_k \left(\pi_{d\varphi(W_k)} - \pi_{d\varphi(W_{k-1})} \right) \right) \circ \varphi^{-1}.$$

In three dimensions, because

$$(d\varphi f_2)^T \tilde{f}_3 \circ \varphi = (f_2^T f_3)/|d\varphi^{-T} f_3| = 0,$$

we see that $d\varphi f_2 \in \mathrm{span}(\tilde{f}_1 \circ \varphi, \tilde{f}_2 \circ \varphi)$. Since $\tilde{f}_1 \circ \varphi$ is proportional to $d\varphi\, f_1$, we can conclude that

$$d\varphi\, \mathrm{span}(f_1, f_2) = \mathrm{span}(\tilde{f}_1 \circ \varphi, \tilde{f}_2 \circ \varphi).$$

This proves that the action we have just defined coincides with the one we have considered for the case $d = 3$.

Returning to the general d-dimensional case, the definition we just gave does not depend on the choice made for the basis $f_1, \ldots, f_d$. Indeed, if we let $\mu_1 > \cdots > \mu_q$ denote the distinct eigenvalues of S, and $\Lambda_1, \ldots, \Lambda_q$ the corresponding eigenspaces, then, regrouping together the terms with identical eigenvalues in the decomposition

of S and $\varphi \cdot S$, and letting

$$\Gamma_k = \Lambda_1 + \cdots + \Lambda_k, \quad \Gamma_0 = \{0\},$$

we clearly have

$$S(x) = \sum_{k=1}^{q} \mu_k(x)(\pi_{\Gamma_k(x)} - \pi_{\Gamma_{k-1}(x)})$$

and

$$\varphi \cdot S = \left(\sum_{k=1}^{q} \mu_k \left(\pi_{d\varphi(\Gamma_k)} - \pi_{d\varphi(\Gamma_{k-1})} \right) \right) \circ \varphi^{-1}.$$

Since the decomposition of S in terms of its eigenspaces is uniquely defined, we obtain the fact that the definition of $\varphi \cdot S$ is non-ambiguous.

9.11 Pros and Cons of Greedy Algorithms

We have studied in this chapter a series of deformable objects, by defining the relevant action(s) that diffeomorphisms have on them and computing the variations of associated matching functionals.

This computation can be used, as we did with landmarks and images, to design "greedy" registration algorithms, which implement gradient descent to progressively minimize the functionals within the group of diffeomorphisms. These algorithms have the advantage of providing relatively simple implementations, and of requiring a relatively limited computation time.

Most of the time, however, this minimization is an ill-posed problem. Minimizers may fail to exist, for example. This has required, for image matching, the implementation of a suitable stopping rule that prevents the algorithm from running indefinitely. Even when a minimizer exists, it is generally not unique (see the example we gave with landmarks). Greedy algorithms provide the minimizer corresponding to the path of steepest descent from where they have been initialized (usually the identity). This solution does not have to be the "best one", and we will see that other methods can find much smoother solutions when large deformations are involved.

To design potentially well-posed problems, the matching functionals need to be combined with regularization terms that measure the smoothness of the registration. This will be discussed in detail in the next chapter.

Chapter 10
Diffeomorphic Matching

10.1 Linearized Deformations

A standard way to ensure the existence of a smooth solution of a matching problem is to add a penalty term in the matching functional. This term would complete (9.1) to form

$$E_{I,I'}(\varphi) = \rho(\varphi) + D(\varphi \cdot I, I'). \tag{10.1}$$

A large variety of such methods have been designed, in approximation theory, statistics and signal processing for solving ill-posed problems. The simplest (and typical) form of penalty function is

$$\rho(\varphi) = \|\varphi - \mathrm{id}\|_H^2$$

for some Hilbert (or Banach) space of functions. Some more complex functions of $\varphi - \mathrm{id}$ may also be designed, related to energies of non-linear elasticity (see, among others [13, 27, 28, 89, 123, 144, 237]). Such methods may be called "small deformation" methods because they work on the deviation of $u = \varphi - \mathrm{id}$, and controlling the size or smoothness of u alone is most of the time not enough to guarantee that φ is a diffeomorphism (unless u is small, as we have seen in Sect. 7.1). There is, in general, no way of proving the existence of a solution of the minimization problem within some group of diffeomorphisms G, unless some restrictive assumptions are made on the objects to be matched.

Our focus here is on diffeomorphic matching. Because of this, we shall not detail many of these methods. However, it is interesting to note that these functionals also have a Eulerian gradient within an RKHS of vector fields with a smooth enough kernel, and can therefore be minimized using (9.7). We illustrate this with the following example, in which we skip the proper justification of the existence of derivatives.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_10

Consider the function $\rho(\varphi) = \int_{\mathbb{R}^d} |d\varphi(x) - \mathrm{Id}|^2\, dx$, where the matrix norm is

$$|A|^2 = \mathrm{trace}(A^T A) = \sum_{i,j} a_{ij}^2$$

(Hilbert–Schmidt norm). Letting $u = \varphi - \mathrm{id}$, we have

$$(d\rho(\varphi) \mid h) = 2\int_{\mathbb{R}^d} \mathrm{trace}(du^T dh)dx = -2\int_{\mathbb{R}^d} \Delta u^T h dx,$$

where Δu is the vector formed by the Laplacian of the coordinates of u (recall that we assume that $u = 0$ at infinity). This implies that (given that $\Delta u = \Delta\varphi$)

$$\left(\bar{\partial}\rho(\varphi) \;\middle|\; h\right) = -2\int_{\Omega} \Delta\varphi^T h \circ \varphi dx$$

and

$$\overline{\nabla}^V_{\varphi}\rho(\cdot) = -2\int_{\Omega} K(\cdot, \varphi(x))\Delta\varphi(x)dx. \tag{10.2}$$

This provides a regularized greedy image-matching algorithm, which includes a regularization term (a similar algorithm may easily be written for point matching).

Algorithm 2 The following procedure is an Eulerian gradient descent, on V, for the energy

$$E_{I,I'}(\varphi) = \int_{\mathbb{R}^d} |d\varphi(x) - \mathrm{id}|^2\, dx + \frac{1}{\sigma^2}\int_{\mathbb{R}^d} \left|I \circ \varphi^{-1}(x) - I'(x)\right| dx.$$

Start with an initial $\varphi_0 = \mathrm{id}$ and solve the differential equation

$$\partial_t\varphi(t, y) = -2\int_{\Omega} K(\varphi(t, y), \varphi(t, x))\Delta\varphi(t, x)dx \tag{10.3}$$

$$+ \frac{2}{\sigma^2}\int_{\Omega} (J(t, x) - I'(x))K(\varphi(t, y), x)\nabla J(t, x)dx \tag{10.4}$$

with $J(t, \cdot) = I \circ \varphi(t)^{-1}(\cdot)$.

This algorithm, which, like the previous greedy procedures, has the fundamental feature of providing a smooth flow of diffeomorphisms to minimize the matching functional, suffers from the same limitations as its predecessors concerning its limit behavior, which are essentially due to the fact that the variational problem itself is not well-posed; minimizers may not exist, and when they exist they are not necessarily diffeomorphisms. In order to ensure the existence of, at least, homeomorphic solutions, the energy must include terms that must not only prevent $d\varphi$ from being too large, but also from being too small (or its inverse from being too large). In [90],

the following regularization is proved to ensure the existence of homeomorphic solutions:

$$\delta(\varphi) = \int_\Omega (a\|d\varphi\|^p + b\|\mathrm{Adj}(d\varphi)\|^q + c(\det d\varphi)^r + d(\det d\varphi)^{-s})dx \qquad (10.5)$$

under some assumptions on p, q, r and s, namely $p, q > 3$, $r > 1$ and $s > 2q/(q-3)$.

10.2 The Monge–Kantorovitch Problem

We briefly discuss in this section the mass transfer problem, which is, under some assumptions, a diffeomorphic method for matching probability densities, i.e., positive functions on $\mathbb{R}^d$ with integral equal to 1. Consider such a density, ζ, and a diffeomorphism φ on $\mathbb{R}^d$. If an object has density ζ, the mass included in an infinitesimal volume dx around x is $\zeta(x)dx$. Now, if each point x in the object is transported to the location $y = \varphi(x)$, the mass of a volume dy around y is the same as the mass of the volume $\varphi^{-1}(dy)$ around $x = \varphi^{-1}(y)$, which is $\zeta \circ \varphi^{-1}(y)|\det(d(\varphi^{-1}))(y)|dy$ (this provides a physical interpretation of Proposition 9.5).

Given two densities ζ and ζ', the optimal mass transfer problem consists in finding a diffeomorphism φ with minimal cost such that $\zeta' = \zeta \circ \varphi^{-1}|\det(d(\varphi^{-1}))|$. The cost associated to φ in this context is related to the distance along which the transfer is made, measured by a function $\rho(x, \varphi(x))$. The total cost comes after summing over the transferred mass, yielding

$$E(\varphi) = \int_\Omega \rho(x, \varphi(x))\zeta(x)dx.$$

The mass transfer problem now is to minimize E over all φ's such that $\zeta' = \zeta \circ \varphi^{-1}|\det(d(\varphi^{-1}))|$. The problem is slightly different from the matching formulations that we discuss in the other sections of this chapter, because the minimization is associated to exact matching.

It is very interesting that this apparently very complex and highly nonlinear problem can be reduced to linear programming, albeit infinite-dimensional. Let us first consider a more general formulation. Instead of looking for a one-to-one correspondence $x \mapsto \varphi(x)$, one can decide that the mass in a small neighborhood of x is dispatched over all Ω with weights $y \mapsto q(x, y)$, where $q(x, y) \geq 0$ and $\int_\Omega q(x, y)dy = 1$. We still have the constraint that the mass density arriving at y is $\tilde{\zeta}(y)$, which gives

$$\int_\Omega \zeta(x)q(x, y)dx = \tilde{\zeta}(y).$$

The cost now has the simple expression (linear in q)

$$E = \int_{\Omega^2} \rho(x, y)\zeta(x)q(x, y)dxdy.$$

The original formulation can be retrieved by letting $q(x, y)dy \to \delta_{\varphi(x)}(y)$ (i.e., pass to the limit $\sigma = 0$ with $q(x, y) = \exp(-|y - \varphi(x)|^2/2\sigma^2)/(2\pi\sigma^2)^{d/2}$).

If we write $g(x, y) = \zeta(x)q(x, y)$, this relaxed problem is clearly equivalent to minimizing

$$E(g) = \int_{\Omega^2} \rho(x, y)g(x, y)dxdy$$

subject to the constraints $g(x, y) \geq 0$, $\int g(x, y)dy = \zeta(x)$ and $\int g(x, y)dx = \tilde{\zeta}(y)$. In fact, the natural formulation of this problem uses measures instead of densities: given two probability measures μ and $\tilde{\mu}$ on Ω, minimize

$$E(\nu) = \int_{\Omega^2} \rho(x, y)\nu(dx, dy)$$

subject to the constraints that the marginals of ν are μ and $\tilde{\mu}$. This provides the Wasserstein distance between μ and $\tilde{\mu}$, associated to the transportation cost ρ. Note that this formulation generalizes the computation of the Wasserstein distance (9.24) between discrete measures.

This problem is much nicer than the original one, since it is a linear programming problem. The theory of convex optimization (that we only apply formally in this infinite-dimensional context; see [44] for rigorous proofs) implies that it has an equivalent dual formulation which is: maximize

$$F(h) = \int_{\Omega} hd\mu + \int_{\Omega} \tilde{h}\tilde{\mu}$$

subject to the constraint that, for all $x, y \in \Omega$, $h(x) + \tilde{h}(y) \leq \rho(x, y)$.

The duality equivalence means that the maximum of F coincides with the minimum of E. The solutions are, moreover, related by duality conditions (the KKT conditions) that imply that ν must be supported by the set

$$A = \left\{(x, y) : h(x) + \tilde{h}(y) = \rho(x, y)\right\}. \tag{10.6}$$

For the dual problem, one is obviously interested in making h and $\tilde{h}$ as large as possible. Given h, one should therefore choose $\tilde{h}$ as

$$\tilde{h}(y) = \sup_x(\rho(x, y) - h(x)),$$

so that the set in (10.6) is exactly the set of (y^*, y) where y^* is a point that achieves the maximum of $\rho(x, y) - h(x)$.

The situation is particularly interesting when $\rho(x, y) = |x - y|^2/2$. In this situation,

$$\tilde{h}(y) = \frac{y^2}{2} + \sup_x \left(x^T y + \frac{x^2}{2} - h(x) \right).$$

From this equation, it is natural to introduce the auxiliary functions $s(x) = h(x) - x^2/2$ and $\tilde{s}(y) = \tilde{h}(y) - y^2/2$. Using these functions, the set A in (10.6) becomes

$$A = \left\{ (x, y) : s(x) + \tilde{s}(y) = x^T y \right\},$$

with $\tilde{s}(y) = \sup_x (x^T y - s(x))$. Because the latter is a supremum of linear functions, we obtain the fact that $\tilde{s}$ is convex, and so is s by symmetry; $\tilde{s}$ is in fact what is called the convex conjugate of s, denoted $\tilde{s} = s^*$. Convex functions are almost everywhere differentiable, and, in order that $(x, y) \in A$, x must maximize $u \mapsto u^T y - s(u)$, which implies that $y = \nabla s(x)$. So, the conclusion is that, whenever s is the solution of the dual problem, the solution of the primal problem is provided by $y = \nabla s(x)$. This shows that the relaxed mass transport problem has the same solution as the initial one, with $\varphi = \nabla s$, s being a convex function. That φ is invertible is obvious by symmetry: $\varphi^{-1} = \nabla \tilde{s}$.

This result is fundamental, since it is the basis for the construction of a numerical procedure for the solution of the mass transport problem in this case. Introduce a time-dependent vector field $v(t, \cdot)$ and the corresponding flow of diffeomorphisms φ^v_{0t}. Let $h(t, \cdot) = \det(d\varphi^v_{t0})\, \zeta \circ \varphi^v_{t0}$. Then

$$\det(d\varphi^v_{0t})\, h(t) \circ \varphi^v_{0t} = \zeta.$$

The time derivative of this equation yields

$$\partial_t h + \operatorname{div}(hv) = 0. \tag{10.7}$$

We have the following theorem [34].

Theorem 10.1 *Consider the following energy:*

$$G(v) = \int_0^1 \int_\Omega h(t, x) |v(t, x)|^2 dx dt$$

and the variational problem: minimize G subject to the constraints $h(0) = \zeta$, $h(1) = \tilde{\zeta}$ and (10.7). If v is the solution of the above problem, then φ^v_{01} solves the optimal mass transport problem.

Proof Indeed, in G, we can make the change of variables $x = \varphi_{0t}(y)$, which yields

$$G(v) = \int_0^1 \int_\Omega \zeta(y) \left| v(t, \varphi^v_{0t}(y)) \right|^2 dy dt$$

$$= \int_\Omega \zeta(y) \int_0^1 \left|\partial_t \varphi_{0t}^v\right|^2 dt$$
$$\geq \int_\Omega \zeta(y) \left|\varphi_{01}^v(y) - y\right|^2 dy.$$

So the minimum of G is always larger than the minimum of E. If φ solves the mass transport problem, then one can take $v(t,x)$ such that $\varphi_{0t}^v(x) = (1-t)x + t\varphi(x)$, which is a diffeomorphism [190] and achieves the minimum of G. □

We refer to [34] for a numerical algorithm that computes the optimal φ. Note that $\rho(x,y) = |x-y|^2$ is not the only transportation cost that can be used in this context, but that others (like $|x-y|$, which is not strictly convex in the distance) may fail to provide diffeomorphic solutions. Important developments on this subject can be found in [49, 119, 296].

We now discuss methods that are both diffeomorphic and metric (i.e., they relate to a distance). They also rely on the representation of diffeomorphisms using flows of ordinary differential equations.

10.3 Optimizing Over Flows

We return in this section to the representation of diffeomorphisms with flows of ordinary differential equations (ODEs) and describe how this representation can be used for diffeomorphic registration. Instead of using a norm to evaluate the difference between φ and the identity mapping, we now consider, as a regularizing term, the distance d_V that was defined in Sect. 7.2.6. More precisely, we set

$$\rho(\varphi) = \frac{1}{2} d_V(\mathrm{id}, \varphi)^2$$

and henceforth restrict the matching to diffeomorphisms belonging to Diff_V.

In this context, we have the following important theorem:

Theorem 10.2 *Let V be a Hilbert space embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$ so that $\mathrm{Diff}_V \subset \mathrm{Diff}_0^{p,\infty}$. Assume that the functional $U : \mathrm{Diff}_0^{p,\infty} \mapsto \mathbb{R}$ is bounded from below and continuous for the (p,∞)-compact topology. Then, there exists a minimizer of*

$$E(\varphi) = \frac{1}{2} d_V(\mathrm{id}, \varphi)^2 + U(\varphi) \tag{10.8}$$

over Diff_V.

(The (p,∞)-compact topology is defined just after Theorem 7.13.)

Proof E has an infimum E_{min} over Diff_V, since it is bounded from below. We need to show that this infimum is also a minimum, i.e., that it is achieved at some $\varphi \in \mathrm{Diff}_V$.

We first use the following lemma (recall that we have denoted by $\mathcal{X}^1_V$ (resp. $\mathcal{X}^2_V$) the set of time-dependent vector fields on Ω with integrable (resp. square integrable) V-norm over $[0, 1]$):

Lemma 10.3 *Minimizing $E(\varphi) = d(\mathrm{id}, \varphi)^2/2 + U(\varphi)$ over Diff_V is equivalent to minimizing the function*

$$\tilde{E}(v) = \frac{1}{2}\int_0^1 \|v(t)\|_V^2 \, dt + U(\varphi_{01}^v) \tag{10.9}$$

over $\mathcal{X}^2_V$.

Let us prove this lemma. For $v \in \mathcal{X}^2_V$, we have, by definition of the distance

$$d_V(\mathrm{id}, \varphi_{01}^v)^2 \leq \int_0^1 \|v(t)\|_V^2 \, dt,$$

which implies $E(\varphi_{01}^v) \leq \tilde{E}(v)$. This obviously implies that $\inf_{\mathrm{Diff}_V} E(\varphi) \leq \tilde{E}(v)$, and since this is true for all $v \in \mathcal{X}^2_V$, we have $\inf E \leq \inf \tilde{E}$. Now, assume that φ is such that $E(\varphi) \leq \inf E + \varepsilon/2$. Then, by definition of the distance, there exists a v such that $\varphi = \varphi_{01}^v$ and

$$\int_0^1 \|v(t)\|_V^2 \, dt \leq d_V(\mathrm{id}, \varphi)^2 + \varepsilon,$$

which implies that

$$\tilde{E}(v) \leq E(\varphi) + \varepsilon/2 \leq \inf E + \varepsilon,$$

so that $\inf E \leq \inf \tilde{E}$.

We therefore have $\inf E = \inf \tilde{E}$. Moreover, if there exists a v such that $\tilde{E}(v) = \min \tilde{E} = \inf E$, then, since we know that $E(\varphi_{01}^v) \leq \tilde{E}$, we must have $E(\varphi_{01}^v) = \min E$. Conversely, if $E(\varphi) = \min E$, by Theorem 7.22, $E(\varphi) = E(\varphi_{01}^v)$ for some v and this v must achieve the infimum of $\tilde{E}$, which proves the lemma.

This lemma shows that it suffices to study the minimizers of $\tilde{E}$. Now, as done in the proof of Theorem 7.22, one can find, by taking a subsequence of a minimizing sequence, a sequence v^n in $\mathcal{X}^2_V$ which converges weakly to some $v \in \mathcal{X}^2_V$ and $\tilde{E}(v^n)$ tends to E_{min}. Because

$$\liminf \int_0^1 \left\|v^n(t)\right\|_V^2 \, dt \geq \int_0^1 \|v(t)\|_V^2 \, dt$$

and because weak convergence in $\mathcal{X}^1_2$ implies convergence of the flow in the (p, ∞)-compact topology (Theorem 7.13) we also have $U(\varphi_{01}^{v^n}) \to U(\varphi_{01}^v)$, so that $\tilde{E}(v) = E_{min}$ and v is a minimizer. □

The general problem of minimizing functionals such as (10.9) has been called "large deformation diffeomorphic metric mapping", or LDDMM. The first algorithms were introduced for this purpose in the case of landmark matching [159] and image matching [32] (these papers were preceded by theoretical developments in [93, 278, 283]). The following sections describe these algorithms, and other that were recently proposed.

10.4 Euler–Lagrange Equations and Gradient

10.4.1 Gradient: Direct Computation

We now detail the computation of the gradient for energies such as (10.8). As remarked in the proof of Theorem 10.2, the variational problem which has to be solved is conveniently expressed as a problem over $\mathcal{X}_V^2$. The function which is minimized over this space takes the form

$$E(v) = \frac{1}{2}\int_0^1 \|v(t)\|_V^2 \, dt + U(\varphi_{01}^v).$$

Assume that V is embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$ and that U is differentiable on $\mathrm{Diff}_0^{p,\infty}$. Then Theorem 7.12 and the chain rule implies that E is differentiable on $\mathcal{X}_V^2$ with

$$(dE(v) \mid h) = \langle v \, , \, h \rangle_{\mathcal{X}_V^2(\Omega)} + \big(dU(\varphi_{01}^v) \;\big|\; \partial_v \varphi_{01}^v \, h\big),$$

where $\partial_v \varphi_{01}^v \, h$ is given in Theorem 7.12.

We now identify the gradient of E for the Hilbert structure of $\mathcal{X}_V^2$. This gradient is a function, denoted $\nabla^V E : v \mapsto \nabla^V E(v) \in \mathcal{X}_V^2$, that satisfies

$$(dE(v) \mid h) = \langle \nabla^V E(v) \, , \, h \rangle_{\mathcal{X}_V^2} = \int_0^1 \langle \nabla^V E(v)(t) \, , \, h(t) \rangle_V dt$$

for all v, h in $\mathcal{X}_V^2$.

Since the set V is fixed in this section, we will drop the exponent from the notation, and simply refer to the gradient $\nabla E(v)$. Note that this is different from the Eulerian gradient we have dealt with before; ∇E now represents the usual gradient of a function defined over a Hilbert space. One important thing to keep in mind is that the gradient we define here is an element of $\mathcal{X}_V^2$, henceforth a time-dependent vector field, whereas the Eulerian gradient was an element of V (a vector field on Ω). Theorem 10.5 relates the two (and allows us to reuse the computations that were made in Chap. 9). For this, we need to introduce the following operation of diffeomorphisms acting on vector fields.

Definition 10.4 Let φ be a diffeomorphism of Ω and v a vector field on Ω. We denote by $\mathrm{Ad}_\varphi v$ the vector field on Ω defined by

$$\mathrm{Ad}_\varphi v(x) = (d\varphi\, v) \circ \varphi^{-1}(x). \tag{10.10}$$

Ad_φ is called the adjoint representation of φ.

If $\varphi \in \mathrm{Diff}^{p+1,\infty}(\Omega)$, then an application of Lemma 7.3 and the Leibnitz formula implies that $\mathrm{Ad}_\varphi v \in C_0^p(\Omega, \mathbb{R}^d)$ as soon as $v \in C_0^p(\Omega, \mathbb{R}^d)$ and more precisely that Ad_φ is a bounded linear operator from $C_0^p(\Omega, \mathbb{R}^d)$ to itself. We can therefore define its conjugate on $C_0^p(\Omega, \mathbb{R}^d)^*$, with $\mathrm{Ad}_\varphi^* \rho$ given by

$$\left(\mathrm{Ad}_\varphi^* \rho \;\middle|\; v\right) = \left(\rho \;\middle|\; \mathrm{Ad}_\varphi v\right) \tag{10.11}$$

for $\rho \in C_0^p(\Omega, \mathbb{R}^d)^*$, $v \in C_0^p(\Omega, \mathbb{R}^d)$. Note that $\mathrm{Ad}_\varphi^* \rho$ is, a fortiori, in V^*, because V is continuously embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$.

Let $\mathbb{L} : V \to V^*$ denote the duality operator on V and $V^{(r)}$ denote the set of vector fields $v \in V$ such that $\mathbb{L}v \in C_0^r(\Omega, \mathbb{R}^d)^*$ (for $r \le p+1$). Then, for $v \in V^{(p)}$, we can define, with $\mathbb{K} = \mathbb{L}^{-1}$,

$$\mathrm{Ad}_\varphi^T v = \mathbb{K}(\mathrm{Ad}_\varphi^* \mathbb{L} v). \tag{10.12}$$

This is well-defined, because, by construction, $\mathrm{Ad}_\varphi^* \mathbb{L} v \in C_0^p(\Omega, \mathbb{R}^d)^* \subset V^*$. We have in particular, for $v \in V^{(p)}$ and $w \in V$,

$$\left\langle \mathrm{Ad}_\varphi^T v\,,\; w \right\rangle_V = \left(\mathrm{Ad}_\varphi^* \mathbb{L} v \;\middle|\; w\right) = \left(\mathbb{L} v \;\middle|\; \mathrm{Ad}_\varphi w\right).$$

Recall that the Eulerian derivative of U is defined by

$$\left(\bar\partial U(\varphi) \;\middle|\; w\right) = (dU(\varphi) \mid w \circ \varphi).$$

Using Theorem 7.12, we have

$$\partial_v \varphi_{01}^v\, h = \int_0^1 (d\varphi_{u1}^v h(u)) \circ \varphi_{0u}^v\, du = \int_0^1 (\mathrm{Ad}_{\varphi_{u1}^v} h(u)) \circ \varphi_{01}^v\, du$$

so that

$$\begin{aligned}\left(dU(\varphi_{01}^v) \;\middle|\; \partial_v \varphi_{01}^v\, h\right) &= \left(\bar\partial U(\varphi_{01}^v) \;\middle|\; \int_0^1 (\mathrm{Ad}_{\varphi_{u1}^v} h(u))\, du\right) \\ &= \int_0^1 \left(\bar\partial U(\varphi_{01}^v) \;\middle|\; \mathrm{Ad}_{\varphi_{u1}^v} h(u)\right)\right) du.\end{aligned}$$

With this notation, we have the following theorem.

Theorem 10.5 *Assume that V is continuously embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$ and that U is continuously differentiable on* $\mathrm{Diff}_0^{p,\infty}$. *Then, the $\mathcal{X}_V^2$ gradient of $\tilde{U} : v \mapsto U(\varphi_{01}^v)$ is given by the formula*

$$\nabla \tilde{U}(v)(t) = \mathbb{K}\mathrm{Ad}^*_{\varphi_{t1}^v} \bar{\partial} U(\varphi_{01}^v) = \mathrm{Ad}^T_{\varphi_{t1}^v} \overline{\nabla} U(\varphi_{01}^v). \tag{10.13}$$

This important result has the following simple consequences.

Proposition 10.6 *Let U satisfy the assumptions of Theorem 10.5. If $v \in \mathcal{X}_V^2$ is a minimizer of*

$$\tilde{E}(v) = \frac{1}{2}\int_0^1 \|v(t)\|_V^2 dt + U(\varphi_{01}^v), \tag{10.14}$$

then, for all t

$$v(t) = \mathrm{Ad}^T_{\varphi_{t1}^v} v(1), \tag{10.15}$$

with $v(1) = -\overline{\nabla}^V U(\varphi_{01}^v)(x)$. In particular, v is a continuous function of t and $v(t) \in V^{(p)}$ for all t.

Corollary 10.7 *Under the same conditions on U, if $v \in \mathcal{X}_V^2$ is a minimizer of*

$$\tilde{E}(v) = \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + U(\varphi_{01}^v)$$

then, for all t,

$$v_t = \mathrm{Ad}^T_{\varphi_{t0}^v} v_0, \tag{10.16}$$

with $v_0 \in V^{(p)}$.

Proposition 10.6 is a direct consequence of Theorem 10.5. For the corollary, we need to use the fact that $\mathrm{Ad}_\varphi \mathrm{Ad}_\psi = \mathrm{Ad}_{\varphi\circ\psi}$, which can be checked by direct computation, and write

$$v_t = \mathrm{Ad}^T_{\varphi_{t1}^v} v_1 = \mathrm{Ad}^T_{\varphi_{t1}^v} \mathrm{Ad}^T_{\varphi_{10}^v} v_0 = (\mathrm{Ad}_{\varphi_{10}^v} \mathrm{Ad}_{\varphi_{t1}^v})^T v_0 = \mathrm{Ad}^T_{\varphi_{t0}^v} v_0.$$

Equations $v_t = \mathrm{Ad}^T_{\varphi_{t0}^v} v_0$ and $v_1 = -\overline{\nabla}^V U(\varphi_{01}^v)(x)$ together are equivalent to the Euler–Lagrange equations for $\tilde{E}$ and will lead to interesting numerical procedures. Equation (10.16) is a cornerstone of the theory. It describes a general mechanical property called the conservation of momentum, to which we will return later.

10.4.2 Derivative Using Optimal Control

We can also apply the Pontryagin maximum principle (see Appendix D) to obtain an alternative expression of the optimality conditions and gradient. Indeed, we can

repeat the construction made in Sect. 7.2.2 with a slightly different notation, letting $f(\omega, v) = v \circ (\mathrm{id} + \omega)$, defined over $C_0^p(\Omega, \mathbb{R}^d) \times V$. With $g(\omega, v) = \|v\|_V^2$, we are in the framework described in Sect. D.3.1, leading to Theorem D.7, where ω represents the state and v is the control. Introducing a co-state μ, define the Hamiltonian

$$H_v(\mu, \omega) = (\mu \mid v \circ (\mathrm{id} + \omega)) - \|v\|_V^2/2.$$

Letting $\xi_\varphi : v \mapsto v \circ \varphi$ from V to $C_0^p(\Omega, \mathbb{R}^d)$, we obtain the fact that an optimal solution must satisfy (with $\varphi = \mathrm{id} + \omega$), for some $\mu : [0, 1] \to C_0^p(\Omega, \mathbb{R}^d)^*$

$$\begin{cases} \partial_t \varphi(t) = v(t) \circ \varphi(t) \\ (\partial_t \mu(t) \mid h) = -(\mu(t) \mid dv(t) \circ \varphi(t)\, h), \quad \forall h \in C_0^p(\Omega, \mathbb{R}^d) \\ \mathbb{L}v(t) = \xi^*_{\varphi(t)} \mu(t) \end{cases} \tag{10.17}$$

with $\varphi(0) = \mathrm{id}$ and $\mu(1) = -dU(\varphi(1))$. One can check that the second equation is equivalent to

$$(\mu(t) \mid h) = \big(\mu(0) \;\big|\; (d\varphi(t))^{-1} h\big),$$

which is Corollary 10.7 expressed in terms of the co-state μ. Applying Eq. (D.12), we obtain

$$d\tilde{E}(v)(t) = -\xi^*_{\varphi(t)} \mu(t) + 2\mathbb{L}v(t), \tag{10.18}$$

where φ and μ satisfy the first two equations of (10.17).

10.4.3 An Alternative Form Using the RKHS Structure

The conjugate of the adjoint can be put into a form explicitly involving the reproducing kernel of V. Before detailing this, we introduce a notation that will be used throughout this chapter. If ρ is a linear form on function spaces, we have been denoting by $(\rho \mid v)$ the result of ρ applied to v. In the formulas that will come, we will need to emphasize the variable on which v depends, and we will use the alternative notation $(\rho \mid v(x))_x$ to denote the same quantity. Thus,

$$\rho(v) = (\rho \mid v) = (\rho \mid v(x))_x.$$

In particular, when v depends on two variables, the notation $(\rho \mid v(x, y))_x$ will represent ρ applied to the function $x \mapsto v(x, y)$ with y considered as constant.

We still assume that V is continuously embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$. Then, the following theorem holds.

Theorem 10.8 *Assume that $\varphi \in C_0^{q+1}(\Omega, \mathbb{R}^d)$ and $\rho \in C_0^r(\Omega, \mathbb{R}^d)^*$, with $r = \min(p + 1, q)$. Let $v = \mathbb{K}\rho$ and $(e_1, \ldots, e_d)$ be an orthonormal basis of $\mathbb{R}^d$. Then, for $y \in \Omega$, we have*

$$\mathrm{Ad}_\varphi^T v(y) = \sum_{i=1}^{d} \left(\rho \,\middle|\, \mathrm{Ad}_\varphi(K(x,y)e_i)\right)_x e_i, \tag{10.19}$$

where K is the reproducing kernel of V.

Proof For $b \in \mathbb{R}^d$, we have

$$\begin{aligned} b^T \mathrm{Ad}_\varphi^T v(y) &= \left\langle \mathrm{Ad}_\varphi^T v\,,\ K(\cdot,y)b \right\rangle_V \\ &= \left\langle v\,,\ \mathrm{Ad}_\varphi(K(\cdot,y)b) \right\rangle_V \\ &= \left(\rho \,\middle|\, \mathrm{Ad}_\varphi(K(x,y)b)\right)_x. \end{aligned}$$

Theorem 10.8 is now a consequence of the decomposition

$$\mathrm{Ad}_\varphi^T v(y) = \sum_{i=1}^{d} e_i^T \mathrm{Ad}_\varphi^T v(y) e_i.$$

□

Recall that $K(\cdot,\cdot)$ is a matrix, so that $K(\cdot,y)e_i$ is the ith column of $K(\cdot,y)$, which we can denote by K^i. Equation (10.19) states that the ith coordinate of $\mathrm{Ad}_\varphi^T v$ is $\left(\rho \,\middle|\, \mathrm{Ad}_\varphi K^i(x,y)\right)_x$.

Using Proposition 10.6 and Theorem 10.8, we obtain another expression of the V-gradient of E:

Corollary 10.9 *Under the hypotheses of Proposition 10.6, the V-gradient of*

$$\tilde{E}(v) = \frac{1}{2}\int_0^1 \|v(t)\|_V^2 dt + U(\varphi_{01}^v)$$

is equal to

$$\nabla^V \tilde{E}(v)(y) = v(t,y) + \sum_{i=1}^{d} \left(\rho(1) \,\middle|\, d\varphi_{t1}(\varphi_{1t}^v(x)) K^i(\varphi_{1t}^v(x), y)\right)_x e_i \tag{10.20}$$

with $\rho(1) = \bar{\partial} U(\varphi_{01}^v)(x)$.

10.5 Conservation of Momentum

10.5.1 Interpretation

Equation (10.16) can be interpreted as a momentum conservation equation. The justification of the term momentum comes from the analogy of $E_{\mathrm{kin}} := (1/2)\|v(t)\|_V^2$

with the total kinetic energy at time t of a dynamical system. In fluid mechanics, this energy is usually defined as (introducing a mass density, z)

$$E_{\text{kin}} = \frac{1}{2}\int z(t,y)|v(t,y)|^2 dy,$$

the momentum here being $\rho(t) = z(t,y)v(t,y)dy$ with $E_{\text{kin}} = (1/2)(\rho \mid v)$. In our case, taking $\rho(t) = \mathbb{L}v(t)$, we still have $E_{\text{kin}} = (1/2)(\rho \mid v)$, so that $\rho(t)$ is also interpreted as a momentum.

To interpret (10.16) as a conservation equation, we need to understand how a change of coordinate system affects the momentum. Indeed, interpret $v(t,y)$ as the velocity of a particle located at coordinates y, so $v = dy/dt$. Now assume that we want to use a new coordinate system, and replace y by $x = \varphi(y)$. In the new coordinates, the same particle moves with velocity

$$\partial_t x = d\varphi(y)\partial_t y = d\varphi(y)\, v(t,y) = (d\varphi\, v(t)) \circ \varphi^{-1}(x)$$

so that the translation from the old to the new expression of the velocity is precisely given by the adjoint operator: $v(y) \to \tilde{v}(x) = \text{Ad}_{\varphi} v(x)$ if $x = \varphi(y)$. To obtain the correct transformation of the momentum, it suffices to notice that the energy of the system must remain the same if we just change the coordinates, so that, if ρ and $\tilde{\rho}$ are the momenta before and after the change of coordinates, we must have

$$(\tilde{\rho} \mid \tilde{v}) = (\rho \mid v)$$

which yields $\text{Ad}^*_{\varphi}\tilde{\rho} = \rho$ or $\tilde{\rho} = \text{Ad}^*_{\varphi^{-1}}\rho$.

Now, we return to Eq. (10.16). Here, $v(t,y)$ is the velocity at time t of the particle that was at $x = \varphi^v_{t0}(y)$ at time 0. So it is the expression of the velocity in a coordinate system that evolves with the flow, and $\mathbb{L}v(t)$ is the momentum in the same system. By the previous argument, the expression of the momentum in the fixed coordinate system, taken at time $t = 0$, is $\text{Ad}^*_{\varphi^v_{0t}}\mathbb{L}v(t)$. Equation (10.16) simply states that this expression remains constant over time, i.e., the momentum is conserved when measured in a fixed coordinate system.

The conservation of momentum equation, described in Corollary 10.7, is a fundamental equation in *Geometric Mechanics* [149, 187], which appears in a wide variety of contexts. It has been described in abstract form by Arnold [18, 19] in his analysis of invariant Riemannian metrics on Lie groups. This equation also derives from an application of the Euler–Poincaré principle, as described in [149, 150, 188]. Combined with a volume-preservation constraint, this equation is equivalent to the Euler equation for incompressible fluids, in the case when $\|v(t)\|_V = \|v(t)\|_2$, the L^2 norm. Another type of norm on V (called the H^1_α norm) relates to models of waves in shallow water, and provides the Camassa–Holm equation [50, 116, 149]. A discussion of (10.16) in the particular case of template matching is provided in [205], and a parallel with the solitons emerging from the Camassa–Holm equation is discussed in [151].

10.5.2 Properties of the Momentum Conservation Equation

Combining Eq. (10.19) and the fact that $\partial_t \varphi^v_{0t} = v(t, \varphi^v_{0t})$, we get, for the optimal v (letting $v_0 = \mathbb{K}\rho_0$)

$$\partial_t \varphi(t, y) = \sum_{i=1}^{d} \left(\rho_0 \mid (d\varphi(t, x))^{-1} K^i(\varphi(t, x), \varphi(t, y))\right)_x e_i.$$

Letting $\varphi = \mathrm{id} + \omega$, we consider the equation

$$\partial_t \omega(t, y) = \sum_{i=1}^{d} \left(\rho_0 \mid (\mathrm{Id} + d\omega(t, x))^{-1} K^i(x + \omega(t, x), y + \omega(t, y))\right)_x e_i. \tag{10.21}$$

We now consider this equation as an ODE over $C^p_0(\Omega, \mathbb{R}^d)$ and discuss conditions on ρ_0 ensuring the existence and uniqueness of solutions. We will make the following assumptions.

(I) V is continuously embedded in $C^{p+1}_0(\Omega, \mathbb{R}^d)$ and its kernel, K, is such that all derivatives $\partial_1^k \partial_2^k K(y, y)$ are bounded over Ω for $k \leq p + 1$.
(II) $\rho_0 \in C^r(\Omega, \mathbb{R}^d)^*$ for some $r \leq p - 1$.
(III) ρ_0 is compactly supported: there exists a compact subset $Q' \subset \mathbb{R}^d$ such that $(\rho_0 \mid f) = 0$ for all $f \in C^r_0(\Omega, \mathbb{R}^d)$ such that $f(x) = 0$ for all $x \in Q'$.

Assumption (I) is true in particular when $\Omega = \mathbb{R}^d$ and K is translation-invariant.

Taking Q slightly larger than Q' in assumption (III), and choosing a C^∞ function ε such that $\varepsilon = 1$ on Q' and $\varepsilon = 0$ on Q^c, we have $(\rho_0 \mid f) = (\rho_0 \mid \varepsilon f)$ for all $f \in C^r_0(\Omega, \mathbb{R}^d)$, from which we can deduce that, for some constant C

$$(\rho_0 \mid f) \leq C \|f\|_{r,Q},$$

where

$$\|f\|_{r,Q} = \max_{x \in Q} \max_{|J| \leq r} |\partial_J f(x)|.$$

The following lemma provides the required properties for the well-posedness of (10.21).

Let $\mathcal{O} = \mathrm{Diff}^p_0 - \mathrm{id}$, an open subset of $C^p_0(\Omega, \mathbb{R}^d)$.

Lemma 10.10 *Let*

$$\mathscr{V}(\omega)(y) = \sum_{i=1}^{d} \left(\rho_0 \mid (\mathrm{Id} + d\omega(t, x))^{-1} K^i(x + \omega(t, x), y + \omega(t, y))\right)_x e_i. \tag{10.22}$$

Under assumptions (I), (II), (III) above, $\mathscr{V}$ is a differentiable mapping from $\mathcal{O}$ into $C^p_0(\Omega, \mathbb{R}^d)$ and, letting $\varphi = \mathrm{id} + \omega$,

$$\|d\mathscr{V}\|_{op(C_0^p(\Omega,\mathbb{R}^d))} \leq C(\|d\varphi(\cdot)^{-1}\|_\infty, \|\varphi\|_{p,\infty}) \tag{10.23}$$

for some continuous function C.

Proof Step 1. We first check that the right-hand side of (10.21) is well defined. Since we assume that V is embedded in $C_0^{p+1}(\Omega,\mathbb{R}^d)$, we know that, for all $0 \leq r, s, \leq p+1$, $\partial_1^r\partial_2^s K^i$ is in $C_0(\Omega,\mathbb{R}^d)$ with respect to each of its variables. In particular, $x \mapsto (\mathrm{Id} + d\omega(t,x))^{-1}K^i(x+\omega(t,x), y+\omega(t,y))$ is in $C_0^{p-1}(\Omega,\mathbb{R}^d)$ as soon as $\omega \in C_0^p(\Omega,\mathbb{R}^d)$, so that ρ_0 can be evaluated on it.
Step 2. We now prove that the right-hand side of (10.21) is in $C_0^p(\Omega,\mathbb{R}^d)$, which ensures that (10.21) forms an ODE in this space. Let

$$v^\varphi(y) = \sum_{i=1}^d \left(\rho_0 \,\middle|\, d\varphi(x)^{-1}K^i(\varphi(x), y\right)_x e_i$$

so that (10.21) can be written as $\partial_t\omega = v^{\mathrm{id}+\omega} \circ (\mathrm{id}+\omega)$. We want to show that $v^\varphi \in C_0^p(\Omega,\mathbb{R}^d)$ when $\varphi = \mathrm{id}+\omega$ and $\omega \in C_0^p(\Omega,\mathbb{R}^d)$. It is obviously sufficient to prove that each coordinate

$$v_i^\varphi(y) = \left(\rho_0 \,\middle|\, d\varphi(x)^{-1}K^i(\varphi(x), y)\right)_x$$

belongs to $C_0^p(\Omega,\mathbb{R})$. We first justify the fact that v_i^φ is p-times differentiable, with

$$d^r v_i^\varphi(y) = \left(\rho_0 \,\middle|\, d\varphi(x)^{-1}\partial_2^r K^i(\varphi(x), y)\right)_x$$

for $r \leq p$. Using a Taylor expansion, we can write (letting $h^{(k)}$ denote the k-tuple $(h,\dots,h)$)

$$\begin{aligned} K^i(\varphi(x), y+h) &= \sum_{k=0}^{p+1} \frac{1}{k!}\partial_2^k K^i(\varphi(x), y)h^{(k)} \\ &\quad + \frac{1}{p!}\int_0^1 (\partial_2^{p+1}K^i(\varphi(x), y+th) - \partial_2^{p+1}K^i(\varphi(x), y))h^{(p+1)}(1-t)^p\,dt \end{aligned}$$

so that

$$\begin{aligned} v_i^\varphi(y+h) &= \sum_{k=0}^{p+1}\frac{1}{k!}\left(\rho_0 \,\middle|\, d\varphi(x)^{-1}\partial_2^k K^i(\varphi(x), y)h^{(k)}\right)_x \\ &+ \frac{1}{p!}\left(\rho_0 \,\middle|\, d\varphi(x)^{-1}\int_0^1 (\partial_2^{p+1}K^i(\varphi(x), y+th) - \partial_2^{p+1}K^i(\varphi(x), y))h^{(p+1)}(1-t)^p\,dt\right)_x \end{aligned}$$

and it suffices to prove that the remainder is an $o(|h|^{p+1})$. This will be true provided

$$\lim_{y' \to y} \|\partial_2^{p+1} K^i(\cdot, y') - \partial_2^{p+1} K^i(\cdot, y)\|_{r,\infty} = 0.$$

For $k_1 \leq r$, we have, using Eq. (8.8),

$$\begin{aligned}
&\|\partial_1^{k_1} \partial_2^{p+1} K^i(\cdot, y') - \partial_1^{k_1} \partial_2^{p+1} K^i(\cdot, y)\|_{r,\infty} \\
&\quad \leq C \max_{h:|h|=1} \|\partial_2^{p+1} K^i(\cdot, y')(h^{(p+1)}) - \partial_2^{p+1} K^i(\cdot, y)(h^{(p+1)})\|_V \\
&\quad = C \left| \partial_1^{p+1} \partial_2^{p+1} K^{ii}(y', y') + \partial_1^{p+1} \partial_2^{p+1} K^{ii}(y, y) - 2\partial_1^{p+1} \partial_2^{p+1} K^{ii}(y, y') \right|^{1/2}
\end{aligned}$$

for some constant C, where K^{ij} denotes the i, j entry of K. This proves the desired result, since $\partial_1^{p+1} \partial_2^{p+1} K$ is continuous. A similar argument can be made to prove the continuity of $y \mapsto d^p v(y)$.

To prove that $v^\varphi \in C_0^p(\Omega, \mathbb{R}^d)$, it suffices to show that, for all $k \leq p+1$, $\|\partial_2^k K^i(\cdot, y)\|_{r,Q}$ goes to 0 when y goes to infinity. (This is where we use the fact that ρ_0 has compact support.)

To reach a contradiction, assume that there exists sequences (x_n), (y_n) with $x_n \in Q$ and y_n tending to infinity or $\partial\Omega$ such that $|\partial_1^{k_1} \partial_2^{k_2} K^i(x_n, y_n)| > \varepsilon$, for some fixed $\varepsilon > 0$ and $k_1 \leq r, k_2 \leq p+1$. Replacing (x_n) by a subsequence if needed, we can assume that x_n converges to some $x \in Q$. Note that $\partial_1^{k_1} \partial_2^{k_2} K^{ij}(x, y_n) = \partial_1^{k_2} \partial_2^{k_1} K^{ji}(y_n, x)$. Since $\partial_2^{k_1} K^j(\cdot, x) \in V$, we can conclude that $\partial_1^{k_2} \partial_2^{k_1} K^j(y_n, x) \to 0$ for all j, implying that $\partial_1^{k_1} \partial_2^{k_2} K^i(x, y_n) \to 0$ for all i, too.

Similarly, $\partial_1^{k_1} \partial_2^{k_2} K^{ij}(x_n, y_n) - \partial_1^{k_1} \partial_2^{k_2} K^{ij}(x, y_n)$ is the ith entry of $\partial_1^{k_2} \partial_2^{k_1} K^j(y_n, x_n) - \partial_2^{k_1} \partial_1^{k_2} K^j(y_n, x)$ and

$$\begin{aligned}
&\sup_y |\partial_1^{k_2} \partial_2^{k_1} K^j(y, x_n) - \partial_1^{k_2} \partial_2^{k_1} K^j(y, x)| \\
&\quad \leq C \max_{h:|h|=1} \|\partial_2^{k_1} K^j(\cdot, x_n)(h^{(k_1)}) - \partial_2^{k_1} K^j(\cdot, x)(h^{(k_1)})\|_V \\
&\quad \leq C \left| \partial_1^{k_1} \partial_2^{k_1} K^{jj}(x_n, x_n) - 2\partial_1^{k_1} \partial_2^{k_1} K^{jj}(x_n, x) + \partial_1^{k_1} \partial_2^{k_1} K^{jj}(x, x) \right|^{1/2},
\end{aligned}$$

which goes to 0. This is our contradiction.

Step 3: We now study the differentiability of the mapping $\mathscr{V} : \omega \mapsto v^{\mathrm{id}+\omega} \circ (\mathrm{id} + \omega)$ from $C_0^p(\Omega, \mathbb{R}^d)$ into itself. The candidate for $d\mathscr{V}(\omega)\eta$ is

$$
\begin{aligned}
(\mathscr{W}(\omega)\eta)(y) = &- \sum_{i=1}^{d} \big(\rho_0 \mid d\varphi(x)^{-1} d\eta(x)^{-1} d\varphi(x)^{-1} K^i(\varphi(x), \varphi(y))\big)_x e_i \\
&+ \sum_{i=1}^{d} \big(\rho_0 \mid d\varphi(x)^{-1} \partial_1 K^i(\varphi(x), \varphi(y))\eta(x)\big)_x e_i \\
&+ \sum_{i=1}^{d} \big(\rho_0 \mid d\varphi(x)^{-1} \partial_2 K^i(\varphi(x), \varphi(y))\eta(y)\big)_x e_i,
\end{aligned}
$$

still with $\varphi = \mathrm{id} + \omega$. We can decompose $\mathcal{V}(\omega + \eta)(y) - \mathcal{V}(\omega)(y) - (\mathscr{W}(\omega)\eta)(y)$ as the sum of five terms

$$
\sum_{k=1}^{5} \sum_{i=1}^{d} \big(\rho_0 \mid A^i(x, y)\big)_x
$$

(described below), which we will study separately. For each term, we need to prove that, for $k_1 \le r$, $k_2 \le p$, one has

$$
\sup_{x,y} |\partial_1^{k_1} \partial_2^{k_2} A_k^i(x, y)| = o(\|\eta\|_{p,\infty}).
$$

The important point in the following discussion is that none of the estimates will require more than p derivatives in φ and η, and no more than $p + 1$ in K.

(i) We let

$$
\begin{aligned}
A_1^i(x, y) = ((d\varphi(x) + d\eta(x))^{-1} - d\varphi(x)^{-1} + d\varphi(x)^{-1} d\eta(x) d\varphi(x)^{-1}) \\
K^i(\varphi(x) + \eta(x), \varphi(y) + \eta(y)).
\end{aligned}
$$

We first note that $Inv : M \mapsto M^{-1}$ is infinitely differentiable on $\mathrm{GL}_d(\mathbb{R})$ with

$$
d^q Inv(M)(H_1, \ldots, H_q) = (-1)^q \sum_{\sigma \in \mathfrak{S}_q} M^{-1} H_{\sigma(1)} M^{-1} \cdots M^{-1} H_{\sigma(q)} M^{-1},
$$

where $\mathfrak{S}_q$ is the set of permutations of $\{1, \ldots, q\}$. In particular, $\|d^q Inv(M)\| = O(\|M^{-1}\|^{q+1})$. Writing

$$
\begin{aligned}
(d\varphi(x) + d\eta(x))^{-1} - d\varphi(x)^{-1} + d\varphi(x)^{-1} d\eta(x) d\varphi(x)^{-1} = \\
\int_0^1 d^2 Inv(d\varphi(x) + t\, d\eta(x))(d\eta(x), d\eta(x))(1 - t)\, dt,
\end{aligned}
$$

we see that

$$
\big\| d^{k_1}((d\varphi(x) + d\eta(x))^{-1} - d\varphi(x)^{-1} + d\varphi(x)^{-1} d\eta(x) d\varphi(x)^{-1}) \big\|_\infty
$$

will be less than $C(\varphi)\|d\varphi^{-1}\|_\infty^{k_1+3}\|\eta\|_{k_1+1,\infty}^2$. Using the bound

$$\|\partial_2^{k_2}K^i(\cdot,y)\|_{p+1,\infty}^2 \le C\partial_1^{k_2}\partial_2^{k_2}K^{ii}(y,y),$$

applying Lemma 7.3 and the product formula, we see that the desired conclusion holds for A_1^i.
(ii) Let

$$A_2^i(x,y) = d\varphi(x)^{-1}(K^i(\varphi(x)+\eta(x),\varphi(y)+\eta(y)) - K^i(\varphi(x),\varphi(y)+\eta(y)) \\ - \partial_1 K^i(\varphi(x),\varphi(y)+\eta(y))\eta(x)).$$

Writing the right-hand side in the form

$$d\varphi(x)^{-1}\int_0^1 \partial_1^2 K^i(\varphi(x)+t\eta(x),\varphi(y)+\eta(y))(\eta(x),\eta(x))(1-t)\,dt,$$

the same estimate on the derivative of K can be used, based on the fact that $k_1+2 \le p+1$.
(iii) The third term is

$$A_3^i(x,y) = d\varphi(x)^{-1}(K^i(\varphi(x),\varphi(y)+\eta(y)) - K^i(\varphi(x),\varphi(y)) \\ - \partial_2 K^i(\varphi(x),\varphi(y))\eta(y)).$$

It can be handled similarly, requiring $k_2+1 \le p+1$ derivatives of K^i in the second variable.
(iv) These were the three main terms in the decomposition and the remaining two are just bridging gaps. The first one is

$$A_3^i(x,y) = d\varphi(x)^{-1}d\eta(x)d\varphi(x)^{-1} \\ (K^i(\varphi(x)+\eta(x),\varphi(y)+\eta(y)) - K^i(\varphi(x),\varphi(y))).$$

Here, we note that, for some constants C and $\tilde{C}$,

$$\begin{aligned}&\sup_x |\partial_1^{k_1}\partial_2^{k_2}K^i(x,y') - \partial_1^{k_1}\partial_2^{k_2}K^i(x,y)|^2 \\ &\le C\left|\partial_1^{k_2}\partial_2^{k_2}(K(y',y') - 2K(y',y)+K(y,y))\right| \\ &\le \tilde{C}(\partial_1^{k_2+1}\partial_2^{k_2+1}K(y,y) + \partial_1^{k_2+1}\partial_2^{k_2+1}K(y',y'))|y-y'|\end{aligned}$$

(with a similar inequality when the roles of x and y are reversed) and these estimates can be used to check that

$$\partial_1^{k_1}\partial_2^{k_2}(K^i(\varphi(x)+\eta(x),\varphi(y)+\eta(y)) - K^i(\varphi(x),\varphi(y)))$$

tends to 0 uniformly in x and y.
(v) The last term is

$$A_5^i(x, y) = d\varphi(x)^{-1}(\partial_1 K^i(\varphi(x), \varphi(y) + \eta(y)) - \partial_1 K^i(\varphi(x), \varphi(y)))\eta(x)$$

and can be handled similarly.

Step 4. It remains to check that $\mathscr{W}(\omega)$ maps $C_0^p(\Omega, \mathbb{R}^d)$ to itself. This can be done in the same way we proved that $\mathscr{V}(\omega) \in C_0^p(\Omega, \mathbb{R}^d)$, using Taylor expansions and the fact that $d^k(\mathscr{W}(\omega)\eta)(y)$ will involve no more than k derivatives of ω and η, and $k + 1$ of K. This shows that $\mathscr{W} = d\mathscr{V}$. The bound (10.23) can also be shown using the same techniques. We leave the final details to the reader. □

Lemma 10.10 implies that (10.21) has unique local solutions (unique solutions over small enough time intervals). If we can prove that $\|(d\varphi)^{-1}\|_\infty$ and $\|\varphi\|_{p,\infty}$ remains bounded over solutions of the equation, inequality (10.23) will be enough to ensure that solutions exist over arbitrary times intervals. This fact will be obtained at the end of the next section.

10.5.3 Time Variation of the Eulerian Momentum

Assume that φ satisfies $\partial_t\varphi(t) = v(t) \circ \varphi(t)$ with $v \in \mathcal{X}^{p+1,1}(\Omega)$. If $\rho_0 \in C^{p-1}(\Omega, \mathbb{R}^d)^*$, we can apply the chain rule to the equation

$$(\rho(t) \mid w) = \big(\rho_0 \,\big|\, \mathrm{Ad}_{\varphi(t)^{-1}} w\big) = \big(\rho_0 \,\big|\, d\varphi(t)^{-1} w \circ \varphi(t)\big),$$

in which we assume that $w \in C_0^p(\Omega, \mathbb{R}^d)$. We have (with $\partial_t d\varphi = dv \circ \varphi\, d\varphi$)

$$\begin{aligned}\partial_t \mathrm{Ad}_{\varphi(t)^{-1}} w &= -d\varphi(t)^{-1} dv(t) \circ \varphi(t)\, w \circ \varphi(t) + d\varphi(t)^{-1} dw \circ \varphi(t)\, v(t) \circ \varphi(t) \\ &= -\mathrm{Ad}_{\varphi(t)^{-1}}(dv(t)\, w - dw\, v(t)).\end{aligned}$$

The term in the right-hand side involves the adjoint representation of $v(t)$, as expressed in the following definition.

Definition 10.11 If v is a differentiable vector field on Ω, we denote by ad_v the mapping that transform a differentiable vector field w into

$$\mathrm{ad}_v w = dv\, w - dw\, v. \tag{10.24}$$

Observe that $dv\, w - dw\, v = -[v, w]$, where the latter is the Lie bracket between right-invariant vector fields over the group of diffeomorphisms. Note that ad_v continuously maps $C_0^p(\Omega, \mathbb{R}^d)$ to $C_0^{p-1}(\Omega, \mathbb{R}^d)$. With this notation, we therefore have, for $w \in C_0^p(\Omega, \mathbb{R}^d)$:

$$\partial_t \mathrm{Ad}_{\varphi(t)^{-1}} w = -\mathrm{Ad}_{\varphi(t)^{-1}} \mathrm{ad}_{v(t)} w$$

so that

$$\partial_t(\rho(t) \mid w) = -\big(\rho(t) \mid \mathrm{ad}_{v(t)} w\big).$$

This yields the equation, called EPDiff, in which we let $\tilde{\rho}(t)$ denote the restriction of $\rho(t)$ to $C_0^p(\Omega, \mathbb{R}^d)$,

$$\partial_t \tilde{\rho}(t) + \mathrm{ad}^*_{v(t)} \rho(t) = 0. \tag{10.25}$$

Equation (10.25) can be used to prove the following proposition.

Proposition 10.12 *Let $\varphi(t) = \mathrm{id} + \omega(t)$, where ω is a solution of (10.21). Let $v_0 = K\rho_0$ and $v(t) = \mathrm{Ad}^T_{\varphi(t)^{-1}} v_0$. Then $\|v(t)\|_V$ is independent of time.*

Proof Indeed, we have, for $\varepsilon > 0$,

$$\frac{1}{\varepsilon}(\|v(t+\varepsilon)\|^2 - \|v(t)\|_V^2) = \frac{2}{\varepsilon}(\rho(t+\varepsilon) - \rho(t) \mid v(t)) + \frac{1}{\varepsilon}\|v(t+\varepsilon). - v(t)\|_V^2$$

Since $v(t) \in V \subset C_0^p(\Omega, \mathbb{R}^d)$, (10.25) implies that the first term on the right-hand side converges to

$$-2\big(\rho(t) \mid \mathrm{ad}_{v(t)} v(t)\big) = 0.$$

For the second term, we have

$$\begin{aligned} \|v(t+\varepsilon) - v(t)\|_V &= \sup_{\|w\|_V \le 1} (\rho(t+\varepsilon) - \rho(t) \mid w) \\ &= \sup_{\|w\|_V \le 1} \int_0^\varepsilon \big(\rho(t+s) \mid \mathrm{ad}_{v(t+s)} w\big) ds, \end{aligned}$$

which tends to 0 with ε. □

We can now prove that (10.21) has a unique solution over arbitrary time intervals.

Theorem 10.13 *Under the hypotheses of Lemma 10.10, Eq. (10.21) has solutions over all times, uniquely specified by its initial conditions.*

Proof As already mentioned, Lemma 10.10 implies that solutions exist over small time intervals. Inequality (10.23) implies that these solutions can be extended as long as $\|d\varphi(t)^{-1}\|_\infty$ and $\|\varphi(t)\|_{p,\infty}$ remain finite. However, both these quantities are controlled by $\int_0^t \|v(t)\|_V \, dt$. For the latter, this is a consequence of (C.6). For $d\varphi(t)^{-1}$, we can note that

$$\partial_t(d\varphi(t)^{-1}) = -d\varphi(t)^{-1} dv(t) \circ \varphi(t)$$

and use Gronwall's lemma to ensure that

$$\|d\varphi(t)^{-1}\|_\infty \le \exp\left(C \int_0^1 \|v(s)\|_V \, ds\right)$$

for some constant C. □

10.5.4 Explicit Expression

The assumption that $\rho_0 \in C_0^{p-1}(\Omega, \mathbb{R}^d)^*$ "essentially" expresses the fact that the evaluation of $(\rho_0 \mid w)$ will involve no more than $p-1$ derivatives of w. This implies that the evaluation of the right-hand side of (10.21) will involve derivatives up to order p in $\varphi = \mathrm{id} + \omega$. In numerical implementations, it is often preferable to track the evolution of these derivatives over time, rather than approximate them using, e.g., finite differences. It often happens, for example, that the evaluation of ρ_0 only requires the evaluation of φ and its derivatives over a submanifold of lower dimension, and tracking their values on a dense grid becomes counter-productive.

The evolution of the derivatives of φ can easily be computed by differentiating (10.21) with respect to the y variable. This is summarized in the system

$$\begin{cases} \partial_t \varphi(t,y) = \displaystyle\sum_{i=1}^{d} \big(\rho_0 \;\big|\; (d\varphi(t,x))^{-1} K^i(\varphi(t,x),\varphi(t,y))\big)_x e_i \\ \partial_t d\varphi(t,y)a = \displaystyle\sum_{i=1}^{d} \big(\rho_0 \;\big|\; (d\varphi(t,x))^{-1} \partial_2 K^i(\varphi(t,x),\varphi(t,y))(d\varphi(t,y)a)\big)_x e_i \\ \quad \vdots \\ \partial_t d^p \varphi(t,y)(a_1,\dots,a_p) \\ \qquad = \displaystyle\sum_{i=1}^{d} \big(\rho_0 \;\big|\; (d\varphi(t,x))^{-1} \partial_2^p K^i(\varphi(t,x),\varphi(t,y))(a_1,\dots,a_p)\big)_x e_i . \end{cases} \tag{10.26}$$

It should be clear from this system that, if the computation of $(\rho_0 \mid w)$ only requires the evaluation of w and its derivatives on some subset of $\mathbb{R}^d$, then φ and its derivatives only need to be tracked for y belonging to the same subset.

10.5.5 The Hamiltonian Form of EPDiff

We now provide an alternative form of (10.26), using the optimal control formulation discussed in Sect. 10.4.2, in which we introduced the co-state

$$(\mu(t) \mid w) = \big(\rho_0 \;\big|\; d\varphi(t)^{-1} w\big) = \big(\rho(t) \;\big|\; w \circ \varphi(t)^{-1}\big). \tag{10.27}$$

Let $M(t) = (d\varphi(t))^{-1}$ so that $\partial_t M = -M\,(\partial_t d\varphi)\,M$. The second equation of (10.26) then becomes

$$\partial_t M(t,y)a = -\sum_{i=1}^{d} \big(\rho_0 \,\big|\, M(t,x)\partial_2 K^i(\varphi(t,x),\varphi(t,y))a\big)_x M(t,y)e_i.$$

This implies that, for any $w \in V$

$$\begin{aligned}\Big(\partial_t\mu(t)\,\Big|\,w\Big) &= -\Big(\rho_0\,\Big|\,\partial_t M\, w\Big)\\ &= \sum_{i=1}^{d}\big(\rho_0 \,\big|\, \big(\rho_0 \,\big|\, M(t,x)\partial_2 K^i(\varphi(t,x),\varphi(t,y))w(y)\big)_x M(t,y)e_i\big)_y\\ &= -\sum_{i=1}^{d}\big(\mu(t) \,\big|\, \big(\mu(t) \,\big|\, \partial_2 K^i(\varphi(t,x),\varphi(t,y))w(y)\big)_x e_i\big)_y.\end{aligned}$$

We therefore have the system

$$\begin{cases} \partial_t\varphi(t,y) = \displaystyle\sum_{i=1}^{d}\big(\mu(t) \,\big|\, K^i(\varphi(t,x),\varphi(t,y))\big)_x e_i\\ \Big(\partial_t\mu(t)\,\Big|\,w\Big) = -\displaystyle\sum_{i=1}^{d}\big(\mu(t) \,\big|\, \big(\mu(t) \,\big|\, \partial_2 K^i(\varphi(t,x),\varphi(t,y))w(y)\big)_x e_i\big)_y. \end{cases} \tag{10.28}$$

Note that this system is an alternative expression of the first two equations of (10.17). When $(\rho_0 \mid w)$ does not depend on the derivatives of w (more precisely, $\rho_0 \in C_0^0(\Omega, \mathbb{R}^d)^*$), this provides an ordinary differential equation in the variables (φ, μ) (of the form $(d/dt)(\varphi,\mu) = F(\varphi,\mu)$). The initial conditions are $\varphi_0 = \mathrm{id}$ and $\mu_0 = \rho_0$.

10.5.6 The Case of Measure Momenta

An interesting feature of (10.28) is that it can easily be reduced to a smaller number of dimensions when ρ_0 takes specific forms. As a typical example, we perform the computation in the case

$$\rho_0 = \sum_{k=1}^{N} z_k(0,\cdot)\gamma_k, \tag{10.29}$$

where γ_k is an arbitrary measure on Ω and $z_k(0)$ a vector field. (We recall the notation $(z\gamma \mid w) = \int z(x)^T w(x)\,\gamma(dx)$.) Most of the Eulerian differentials that we have computed in Chap. 9 have been reduced to this form. From the definition of $\mu(t)$, we have $\mu(t) = \sum_{k=1}^{N} z_k(t,.)\gamma_k$ (where $z_k(t,x) = d\varphi_{0t}(x)^{-T} z_k(0,x)$). The first equation in (10.28) is

$$\partial_t \varphi(t, y) = \sum_{i=1}^{d} \sum_{k=1}^{N} \int_{\Omega} z_k(t, x)^T K^i(\varphi(t, x), \varphi(t, y)) e_i d\gamma_k(x).$$

For a matrix A with ith column vector A^i, and a vector z, $z^T A^i$ is the ith coordinate of $A^T z$. Applying this to the previous equation yields

$$\partial_t \varphi(t, y) = \sum_{k=1}^{N} \int_{\Omega} K(\varphi(t, y), \varphi(t, x)) z_k(t, x) d\gamma_k(x), \tag{10.30}$$

where we have used the fact that $K(\varphi(t, x), \varphi(t, y))^T = K(\varphi(t, y), \varphi(t, x))$. The second equation in (10.28) becomes

$$\begin{aligned}
&\left(\partial_t \mu(t) \,\middle|\, w\right) \\
&= -\sum_{i=1}^{d} \left(\mu(t) \,\middle|\, \left(\mu(t) \,\middle|\, \partial_2 K^i(\varphi(t, x), \varphi(t, y)) w(y)\right)_x e_i\right)_y \\
&= -\sum_{k,l=1}^{N} \int_{\Omega} \int_{\Omega} \sum_{i=1}^{d} z_l^T(t, x) \partial_2 K^i(\varphi(t, x), \varphi(t, y)) w(y)\, z_k(t, y)^T e_i d\gamma_l(x) d\gamma_k(y) \\
&= -\sum_{k=1}^{N} \int_{\Omega} \left(\int_{\Omega} \sum_{l=1}^{N} \sum_{i=1}^{d} z_k^i(t, y) z_l(t, x)^T \partial_2 K^i(\varphi(t, x), \varphi(t, y)) d\gamma_l(x) \right) w(y) d\gamma_k(y),
\end{aligned}$$

where z_k^i is the ith coordinate of z_k. From the expression of $\mu(t)$, we also have

$$\partial_t \mu = \sum_{k=1}^{N} (\partial_t z_k) \gamma_k.$$

Letting K^{ij} denote the entries of K, we can identify $\partial_t z_k$ as

$$\begin{aligned}
&\partial_t z_k(t, y) = \\
&\quad - \int_{\Omega} \sum_{l=1}^{N} \sum_{i,j=1}^{d} z_k^i(t, y) z_l^j(t, x) \nabla_2 K^{ij}(\varphi(t, x), \varphi(t, y)) d\gamma_l(x) \\
&= - \int_{\Omega} \sum_{l=1}^{N} \sum_{i,j=1}^{d} z_l^i(t, y) z_k^j(t, x) \nabla_1 K^{ij}(\varphi(t, y), \varphi(t, x)) d\gamma_l(x).
\end{aligned} \tag{10.31}$$

This equation is somewhat simpler when K is a scalar kernel, in which case $K^{ij}(x, y) = \Gamma(x, y)$ if $i = j$ and 0 otherwise, where Γ takes real values. We get, in this case

$$\partial_t z_k(t,y) = -\sum_{l=1}^{N} \int_\Omega \nabla_2 \Gamma(\varphi(t,x),\varphi(t,y)) z_k(t,y)^T z_l(t,x) d\gamma_l(x)$$
$$= -\sum_{l=1}^{N} \int_\Omega \nabla_1 \Gamma(\varphi(t,y),\varphi(t,x)) z_k(t,y)^T z_l(t,x) d\gamma_l(x).$$

In all cases, we see that the evolution of μ can be completely described using the evolution of $z_1, \ldots, z_N$. In the particular case when the z_k's are constant vectors (which corresponds to most of the point-matching problems), this provides a finite-dimensional system on the μ part.

10.6 Optimization Strategies for Flow-Based Matching

We have formulated flow-based matching as an optimization problem over time-dependent vector fields. We discuss here other possible optimization strategies that take advantage of the different formulations that we obtained for the EPDiff equation. They will correspond to taking different control variables with respect to which the minimization is performed, and we will in each case provide the expression of the gradient of E with respect to a suitable metric. Optimization can then be performed by gradient descent, conjugate gradient or higher-order optimization algorithms when feasible (see Appendix D or [221]).

After discussing the general formulation of each of these strategies, we will provide the specific expression of the gradients for point-matching problems, in the following form: minimize

$$E(\varphi) = \frac{1}{2} d_V(\mathrm{id}, \varphi)^2 + F(\varphi(x_1), \ldots, \varphi(x_N)) \tag{10.32}$$

with respect to φ, where $x_1, \ldots, x_N$ are fixed points in Ω. These problems are important because, in addition to the labeled and unlabeled point matching problems we have discussed, other problems, such as curve and surface matching, end up being discretized in this form (we will discuss algorithms for image matching in the next section). The following discussion describes (and often extends) several algorithms that have been proposed in the literature, in [32, 159, 203, 204, 289, 309] among other references.

10.6.1 Gradient Descent in $\mathcal{X}_V^2$

The original problem having been expressed in this form, Corollary 10.9 directly provides the expression of the gradient of E considered as a function defined over

$\mathcal{X}_V^2$, with respect to the metric in this space. Using $t \mapsto v(t, \cdot)$ as an optimization variable has some disadvantages, however. The most obvious is that it results in solving a huge dimensional problem (over a $d + 1$-dimensional variable) even if the original objects are, say, collections of N landmarks in $\mathbb{R}^d$.

When the matching functional U is only a function of the deformation of a fixed object, i.e.,

$$U(\varphi) = F(\varphi \cdot I),$$

then some simplifications can be made. To go further, we will need to compute derivatives in the object space, and henceforth assume that $\mathcal{I}$ is an open subset of a Banach space $\boldsymbol{I}$. We assume that Diff_0^{p+1} acts on $\mathcal{I}$ and that the mapping $\Phi_I : \varphi \mapsto \varphi \cdot I$ is differentiable on Diff_0^{p+1} for all $I \in \mathcal{I}$, so that an infinitesimal action is defined by (see Sect. B.5.3)

$$h \cdot I = d\Phi_I(\mathrm{id})\, h \in \boldsymbol{I}$$

for $h \in C_0^{p+1}(\Omega, \mathbb{R}^d)$. We assume as usual that V is continuously embedded in $C_0^{p+1}(\Omega, \mathbb{R}^d)$ so that $v \cdot I$ is well defined for $v \in V$ and $d\Phi_I(\mathrm{id})$ restricted to V is also bounded with respect to $\| \cdot \|_V$.

Let $v \in \mathcal{X}_V^2$. If $\partial_t \varphi = v \circ \varphi$, let $J(t) = \varphi(t) \cdot I$ be the deforming object. Then $\partial_t J(t) = v(t) \cdot J(t)$. With this in mind, we can write, when $\tilde{E}$ is given by (10.14)

$$\min_{v(t,\cdot)} \tilde{E}(v) = \min_{J(t,\cdot), J(0)=I} \left(\min_{v:\, \partial_t J = v(t) \cdot J(t)} \tilde{E}(v) \right).$$

The iterated minimization first minimizes with respect to v for fixed object trajectories, then optimizes over the object trajectories.

When $J(t, \cdot)$ is given, the inner minimization is

$$\begin{aligned} \min_{v:\, \partial_t J = v(t) \cdot J(t)} \tilde{E}(v) &= \min_{v:\, \partial_t J = v(t) \cdot J(t)} \left(\frac{1}{2} \int_0^1 \|v(t)\|_V^2 \, dt + F(J(1)) \right) \\ &= \frac{1}{2} \int_0^1 \left(\inf_{w:\, \partial_t J = w \cdot J(t)} \|w\|_V^2 \right) dt + F(J(1)) \end{aligned} \tag{10.33}$$

since the constraints apply separately to each $v(t)$. This expression only depends on the trajectory $J(t)$. One can therefore try to compute its gradient with respect to this object trajectory and run a minimization algorithm accordingly. One difficulty with this approach is that, given an object trajectory $J(t)$, there may exist no $w \in V$ such that $\partial_t J = w \cdot J(t)$ (which results in the minimum in the integral being infinite), so that the possibility of expressing the trajectory as evolving according to a flow is a constraint of the problem. This may be intractable in the general case, but always satisfied for point-matching problems as long as the points remain distinct. We will discuss this in the next section.

However, what (10.33) tells us is that, if a time-dependent vector field $\tilde{v}(t, \cdot)$ is given, one always reduces the value of $\tilde{E}(\tilde{v})$ by replacing $\tilde{v}(t, \cdot)$ by

$$v(t, \cdot) = \underset{w:\, w \cdot J(t) = \tilde{v} \cdot J(t)}{\operatorname{argmin}} \|w\|_V^2 \tag{10.34}$$

with $J(t) = \varphi_{0t}^{\tilde{v}} \cdot I$. Introduce the space

$$N_J = Null(d\Phi_J(\mathrm{id})) = \{u \in V : u \cdot J = 0\}$$

and its perpendicular $V_J = N_J^\perp = \left\{u \in V : \langle u\,,\, \tilde{u}\rangle_V = 0 \text{ for all } \tilde{u} \in N_J\right\}$. Then we have the following lemma.

Lemma 10.14 *Let $I \in \mathcal{I}$ and $\tilde{v} \in V$. Then, the minimizer of $\|w\|_V^2$ over all $V \in V$ such that $w \cdot J = \tilde{v} \cdot J$ is given by $\pi_{V_J}(\tilde{v})$, the orthogonal projection of $\tilde{v}$ on V_J.*

Proof Let $v = \pi_{V_J}(\tilde{v})$. We want to prove that v is a minimizer of $\|\cdot\|_V^2$ over the set of all $w \in V$ such that $w = \tilde{v} + u$ with $u \in N_J$. For such a w, we have

$$\pi_{V_J}(w) = v + \pi_{V_J}(u) = v$$

and $\|w\|_V^2 \geq \|\pi_{V_J}(w)\|_V^2$. Moreover, from the characteristic properties of an orthogonal projection, we have $\tilde{v} - v \in V_J^\perp = N_J$, the inequality holding because N_J is closed (because it is the null set of a bounded linear map). □

The numerical computation of this orthogonal projection is not always easy, but when it is, it generally has a form which is more specific than a generic time-dependent vector field, and provides an improved gradient descent algorithm in $\mathcal{X}_V^2$ as follows. Assume that, at time τ in the algorithm, the current vector field v^τ in the minimization of E is such that $v^\tau(t) \in V_{J^\tau(t)}$ at all times t. Then define a vector field at the next step $\tau + \delta\tau$ by

$$\tilde{v}^{\tau+\delta\tau}(t, y) = v^\tau(t, y) - \delta\tau \left(v(t, y) + \sum_{i=1}^{d} \left(\rho(1) \,\middle|\, d\varphi_{t1}(\varphi_{1t}^v(x)) K^i(\varphi_{1t}^v(x), y) \right)_x e_i \right),$$

which corresponds to one step of gradient descent, as specified in (10.20), then compute $J(t) = \varphi_{0t}^{\tilde{v}^{\tau+\delta\tau}} \cdot I$ and define

$$v^{\tau+\delta\tau}(t) = \pi_{V_{J(t)}}(\tilde{v}^{\tau+\delta\tau})$$

at all times t.

Application to Point Matching

Consider the point-matching energy. In this case, letting

$$U(\varphi) = F(\varphi(x_1), \ldots, \varphi(x_N)),$$

we have

$$\rho(1) = \bar{\partial} U(\varphi_{01}^v) = \sum_{i=1}^N \partial_i F(\varphi_{01}^v(x)) \delta_{\varphi_{01}^v(x_k)}.$$

We therefore have, by Corollary 10.9, with $\tilde{U}(x) = U(\varphi_{01}^v)$,

$$\begin{aligned}
\nabla^V \tilde{U}(v)(t,y) &= \sum_{i=1}^d \left(\rho(1) \,\middle|\, d\varphi_{t1}^v(\varphi_{1t}^v(x)) K(\varphi_{1t}^v(x), y) e_i\right)_x e_i \\
&= \sum_{i=1}^d \sum_{q=1}^N \left(\partial_q F(\varphi_{01}^v(x))^T d\varphi_{t1}^v(\varphi_{0t}^v(x_q)) K(\varphi_{0t}^v(x_q), y) e_i\right) e_i \\
&= \sum_{q=1}^N K(y, \varphi_{0t}^v(x_q)) d\varphi_{t1}^v(\varphi_{0t}^v(x_q))^T \partial_q F(\varphi_{01}^v(x_q)),
\end{aligned}$$

so that

$$\nabla^V \tilde{E}(v)(t,y) = v(t,y) + \sum_{q=1}^N K(y, \varphi_{0t}^v(x_q)) d\varphi_{t1}^v(\varphi_{0t}^v(x_q))^T \partial_q F(\varphi_{01}^v(x_q)). \tag{10.35}$$

So, a basic gradient descent algorithm in $\mathcal{X}_V^2$ would implement the evolution (letting τ denote the algorithm time)

$$\partial_\tau v^\tau(t,y) = -\gamma \left(v^\tau(t,y) + \sum_{q=1}^N K(y, \varphi_{0t}^{v^\tau}(x_q)) d\varphi_{t1}^{v^\tau}(\varphi_{0t}^{v^\tau}(x_q))^T \partial_q F(\varphi_{01}^{v^\tau}(x_q)) \right). \tag{10.36}$$

The two-step algorithm defined in the previous section is especially efficient with point sets. When $x = (x_1, \ldots, x_N)$, $v \cdot x = (v(x_1), \ldots, v(x_N)$, the projection on

$$V_x = \{v : v \cdot x = 0\}^\perp = \{v : v(x_1) = \cdots = v(x_N) = 0\}^\perp$$

is given by spline interpolation with the kernel, as described in Theorem 8.8, i.e.,

$$V_x = \left\{ v = \sum_{k=1}^N K(., x_k) a_k, a_1, \ldots, a_N \in \mathbb{R}^d \right\}. \tag{10.37}$$

More precisely, define $x_i^v(t) = \varphi_{0t}^v(x_i)$. We assume that, at time τ, we have a time-dependent vector field v^τ which takes the form

$$v^\tau(t,y) = \sum_{i=1}^N K(y, x_i^{v^\tau}(t)) \alpha_i^\tau(t). \tag{10.38}$$

Using (10.36), we define

$$\tilde{v}(t,y) = v^\tau - \delta\tau\Big(v^\tau(t,y) + \sum_{q=1}^{N} K(y, \varphi_{0t}^{v^\tau}(x_q)) d\varphi_{t1}^{v^\tau}(\varphi_{0t}^{v^\tau}(x_q))^T \partial_q F(\varphi_{01}^{v^\tau}(x_q))\Big).$$

The values of $\tilde{v}(t,\cdot)$ are in fact only needed at the points $x_i^{\tilde{v}}(t) = \varphi_{0t}^{\tilde{v}}(x_i)$. These points are obtained by solving the differential equation

$$\partial_t x = v^\tau(t,x) - \delta\tau\left(v^\tau(t,x) + \sum_{i=1}^{N} K(x_i^{v^\tau}(t), x) d\varphi_{t1}^{v^\tau}(x_i^{v^\tau}(t))^T F_i(x^{v^\tau}(1))\right) \tag{10.39}$$

with $x(0) = x_i$. Solving this equation provides both $x_i^{\tilde{v}}(t)$ and $\tilde{v}(x_i^{\tilde{v}}(t))$ for $t \in [0,1]$.

Once this is done, define $v^{\tau+\delta\tau}(t,\cdot)$ to be the solution of the approximation problem $\inf_w \|w\|_V$ with $w(x_i^{\tilde{v}}(t)) = v(x_i^{\tilde{v}}(t))$, which will therefore take the form

$$v^{\tau+\delta\tau}(t,y) = \sum_{i=1}^{N} K(y, x_i^{v^{\tau+\delta\tau}}(t)) \alpha_i^{\tau+\delta\tau}(t).$$

Solving (10.39) requires evaluating the expression of v^τ, which can be done exactly using (10.38). It also requires computing the expression of $d\varphi_{t1}^{v^\tau}(x_i^{v^\tau}(t))$, which can be obtained from the expression

$$\partial_t d\varphi_{t1}^{v} \circ \varphi_{1t} = \partial_t (d\varphi_{1t})^{-1} = -(d\varphi_{1t})^{-1}(\partial_t(d\varphi_{1t}))(d\varphi_{1t})^{-1},$$

which yields:

$$\partial_t d\varphi_{t1}^{v}(x_i^v(t)) = -d\varphi_{t1}^{v}(x_i^v(t)) dv(t, x_i^v(t)).$$

Thus, $d\varphi_{t1}^{v}(x_i^v(t))$ is a solution of $\partial_t M = -M dv(x_i^v(t))$ with initial condition $M = \mathrm{Id}$. The matrix $dv(t, x_i^v(t))$ can be computed explicitly as a function of the point trajectories $x_j^v(t)$, $j = 1, \ldots, N$, using the explicit expression (10.38). This algorithm was introduced in [31].

10.6.2 Gradient in the Hamiltonian Form

As we have seen, one can use the optimal control formalism with the Pontryagin principle to compute the gradient of $\tilde{E}$ in v. Given $v \in \mathcal{X}_V^2$, this gradient can be computed by solving (10.28) with boundary conditions $\varphi(0) = \mathrm{id}$ and $\mu(1) = -dU(\varphi(1))$ (which can be achieved by solving the first equation in (10.28) from $t = 0$ to $t = 1$, then the second one backward in time, from $t = 1$ to $t = 0$) and, using (10.18), letting

$$\nabla\tilde{E}(v)(t) = \mathbb{K}(d\tilde{E}(v)(t)) = -\mathbb{K}\xi^*_{\varphi(t)}\mu(t) + v(t).$$

This equation (or the maximum principle) implies that the optimal v must be such that $v(t) = \mathbb{K}\xi^*_{\varphi(t)}\mu(t)$ for some $\mu \in C_0^{p-1}(\Omega, \mathbb{R}^d)^*$ and there is therefore no loss of generality in restricting the optimization problem to v's taking this form. With this constraint, we have

$$\|v(t)\|_V^2 = \left\langle \mathbb{K}\xi^*_{\varphi(t)}\mu(t)\,,\ \mathbb{K}\xi^*_{\varphi(t)}\mu(t)\right\rangle_V = \left(\mu(t)\ \middle|\ \xi_{\varphi(t)}\mathbb{K}\xi^*_{\varphi(t)}\mu(t)\right).$$

Let $\mathbb{K}_\varphi = \xi_\varphi \mathbb{K}\xi^*_\varphi$ so that $\|v(t)\|_V^2 = \left(\mu(t)\ \middle|\ \mathbb{K}_{\varphi(t)}\mu(t)\right)$. One has

$$\begin{aligned} a^T(\mathbb{K}_\varphi\mu)(y) &= a^T(\mathbb{K}\xi^*_\varphi\mu)(\varphi(y)) \\ &= \left\langle K(\cdot,\varphi(y))a\,,\ \mathbb{K}\xi^*_\varphi\mu\right\rangle_V \\ &= (\mu \mid K(\varphi(x),\varphi(y))a)_x \end{aligned}$$

so that

$$(\mathbb{K}_\varphi\mu)(y) = \sum_{i=1}^d (\mu \mid K(\varphi(x),\varphi(y))e_i)_x e_i.$$

With this notation, the state equation $\partial_t\varphi = v\circ\varphi$ becomes $\partial_t\varphi = \mathbb{K}_\varphi\mu$ and the original optimal control problem is reformulated as minimizing

$$E(\varphi,\mu) = \frac{1}{2}\int_0^1 \left(\mu(t)\ \middle|\ \mathbb{K}_{\varphi(t)}\mu(t)\right)dt + U(\varphi_{01})$$

subject to $\partial_t\varphi = \mathbb{K}_\varphi\mu$.

Expressing the problem in this form slightly changes the expression of the differential. The computation of the gradient (and its justification) based on a co-state α and the Hamiltonian

$$H_\mu(\alpha,\varphi) = \left(\alpha\ \middle|\ \mathbb{K}_\varphi\mu\right) - \frac{1}{2}\left(\mu\ \middle|\ \mathbb{K}_\varphi\mu\right)$$

are obtained using the same methods as in Sect. 10.4.2, so we skip the details. Let φ^μ be the solution of $\partial_t\varphi = \mathbb{K}_\varphi\mu$ with $\varphi(0) = \mathrm{id}$. Then, with $\tilde{E}(\mu) = E(\varphi^\mu,\mu)$, we have

$$d\tilde{E}(\mu) = \mathbb{K}_\varphi\alpha - \mathbb{K}_\varphi\mu,$$

where φ and α are given by the system

$$\left\{\begin{aligned}
\partial_t \varphi &= \mathbb{K}_\varphi \mu \\
\left(\partial_t \alpha(t) \,\middle|\, w\right) &= -\sum_{i=1}^{d} \left(\alpha(t) \,\middle|\, \left(\mu(t) \,\middle|\, \partial_2 K^i(\varphi(t,x),\varphi(t,y))w(y)\right)_x e_i\right)_y \\
&\quad - \sum_{i=1}^{d} \left(\mu(t) \,\middle|\, \left(\alpha(t) \,\middle|\, \partial_2 K^i(\varphi(t,x),\varphi(t,y))w(y)\right)_x e_i\right)_y \\
&\quad + 2\sum_{i=1}^{d} \left(\mu(t) \,\middle|\, \left(\mu(t) \,\middle|\, \partial_2 K^i(\varphi(t,x),\varphi(t,y))w(y)\right)_x e_i\right)_y
\end{aligned}\right. \tag{10.40}$$

with $\varphi(0) = \mathrm{id}$ and $\alpha(1) = -dU(\varphi(1))$. Unsurprisingly, this system boils down to (10.28) when $d\tilde{E}(\mu) = 0$, i.e., when $\alpha = \mu$.

The gradient of $\tilde{E}$ expressed with respect to the inner product

$$\left\langle \mu\,,\, \tilde{\mu} \right\rangle_\varphi = \left(\mu \,\middle|\, K_\varphi \mu\right).$$

(this choice will be justified in Sect. 11.4 as the dual Riemannian metric on the diffeomorphism group) is

$$\nabla \tilde{E}(\mu) = \alpha - \mu,$$

a remarkably simple expression.

Consider now the case in which $U(\varphi) = F(\varphi \cdot I)$ where I is a fixed object. With the notation and assumptions made in Sect. 10.6.1, we found in Lemma 10.14 that there was no loss of generality in restricting the minimization to $v(t) \in V_{J(t)}$ at all times. This often entails additional constraints on the momentum $\rho(t) = \mathbb{L}v(t)$, or on $\mu(t) = \xi^*_{\varphi(t)^{-1}}\rho(t)$, that can be leveraged to reduce the dimension of the control variable. For example, we have seen that for point sets (in which we let $J = x$) V_x was given by (10.37), so that $\rho(t)$ must take the form

$$\rho(t) = \sum_{k=1}^{N} z_k(t)\delta_{x_k(t)}$$

for some $z_1(t), \ldots, z_N(t) \in \mathbb{R}^d$, from which we can deduce (using $x_k(t) = \varphi(t, x_k(0))$) that $\mu(t)$ must take the form

$$\mu(t) = \sum_{k=1}^{N} z_k(t)\delta_{x_k(0)}.$$

One can then use $z_1, \ldots, z_N$ as a new control, as described below.

Another interesting special case is when $\mu(t)$ can be expressed as a vector measure, because, as discussed in Sect. 10.5.6, one can then assume that

$$\mu(t) = \sum_{k=1}^{N} z_k(t, \cdot)\gamma_k,$$

where $\gamma_1, \ldots, \gamma_N$ are fixed measures. One can then use the vector fields $z_1, \ldots, z_N$ to parametrize the problem. This leads to a simplification when the measures have a sparse support. They are, for example, Dirac measures for point matching. We now review this case in more detail.

Application to Point Matching

When $U(\varphi) = F(\varphi(x_1), \ldots, \varphi(x_N))$, we have

$$(dU(\varphi) \mid h) = \sum_{k=1}^{N} \partial_k F(\varphi(x_1), \ldots, \varphi(x_N))^T h(x_k)$$

so that

$$\mu(1) = -\sum_{k=1}^{N} \partial_k F(\varphi(x_1), \ldots, \varphi(x_N))\delta_{x_k}$$

is a vector measure. We can therefore look for a solution in the form

$$\mu(t) = \sum_{k=1}^{N} z_k(t)\delta_{x_k}$$

at all times, for some coefficients $z_1, \ldots, z_N$.

In order to obtain α in (10.40) given a current μ, it suffices to solve the first equation only for the values of $y_k(t) = \varphi(t, x_k)$, $k = 1, \ldots, N$, which requires us to solve the system

$$\partial_t y_k = \sum_{l=1}^{N} K(y_k, y_l)\xi_l.$$

One then sets

$$\alpha(1) = -\sum_{k=1}^{N} \partial_k F(y_1(1), \ldots, y_N(1))\delta_{x_k}$$

and solves the second equation backward in time, knowing that the solution will take the form

$$\alpha(t) = \sum_{k=1}^{N} \eta_k(t)\delta_{x_k}$$

with $\eta_k(1) = -\partial_k F(y_1(1), \ldots, y_n(1))$ and

$$\partial_t \eta_k = \sum_{l=1}^{N} \sum_{i,j=1}^{d} \eta_l^i z_k^j \nabla_1 K^{ij}(y_k, y_l) + \sum_{l=1}^{N} \sum_{i,j=1}^{d} z_l^i \eta_k^j \nabla_1 K^{ij}(y_k, y_l)$$
$$- 2 \sum_{l=1}^{N} \sum_{i,j=1}^{d} z_l^i z_k^j \nabla_1 K^{ij}(y_k, y_l).$$

Given this, we have

$$\nabla \tilde{E}(\mu) = \sum_{k=1}^{N} (\eta_k - z_k) \delta_{x_k}.$$

10.6.3 Gradient in the Initial Momentum

We now use the fact that Eq. (10.16) implies that the optimal $v(t)$ is uniquely constrained by its value at $t = 0$ for formulating the variations of the objective function in terms of these initial conditions. We therefore optimize with respect to v_0, or equivalently with respect to $\mu_0 = \rho_0$. This requires finding ρ_0 such that

$$\frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt + U(\varphi(1))$$

is minimal under the constraints $\partial_t \varphi(t) = v(t) \circ \varphi(t)$, with

$$v(t) = \sum_{i=1}^{d} \left(\rho_0 \,\middle|\, d\varphi(t) K^{(i)}(x, y) \right)_x e_i.$$

Proposition 10.12 helps us to simplify this expression, since it implies that $\int_0^1 \|v(t)\|^2 dt = (\rho_0 \mid \mathbb{K}\rho_0)$ and the minimization problem therefore is to find ρ_0 such that

$$E(\rho_0) = \frac{1}{2} (\rho_0 \mid \mathbb{K}\rho_0) + U(\varphi(1))$$

is minimal, where (φ, μ) is a solution of system (10.28) with initial conditions $\varphi(0) = \mathrm{id}$ and $\mu(0) = \rho_0$. Writing (10.28) as

$$\partial_t \begin{pmatrix} \varphi \\ \mu \end{pmatrix} = \begin{pmatrix} \mathscr{V}_1(\varphi, \mu) \\ \mathscr{V}_2(\varphi, \mu) \end{pmatrix} = \mathscr{V}(\varphi, \mu)$$

and applying Proposition D.12, we have

$$dE(\rho_0) = \mathbb{K}\rho_0 - p_\mu(0),$$

where the pair $\begin{pmatrix} p_\varphi(t) \\ p_\mu(t) \end{pmatrix}$ satisfies $p_\varphi(1) = -dU(\varphi(1))$, $p_\mu(0) = 0$ and

$$\begin{cases} \partial_t p_\varphi = -\partial_1 \mathscr{V}_1^* \, p_\varphi - \partial_1 \mathscr{V}_2^* \, p_\mu, \\ \partial_t p_\mu = -\partial_2 \mathscr{V}_1^* \, p_\varphi - \partial_2 \mathscr{V}_2^* \, p_\mu. \end{cases} \tag{10.41}$$

The gradient of E with respect to the metric on V^* is then given by $\nabla E(\rho_0) = \rho_0 - \mathbb{K}^{-1} p_\mu$.

The practical application of these formulas requires us to make explicit the expressions of $\partial_i \mathscr{V}_j^*$ for $i, j = 1, 2$. Returning to (10.28), we have

$$\begin{aligned} \big(\partial_1 \mathscr{V}_1^* \, p_\varphi \,\big|\, h\big) &= \sum_{i=1}^d \big(p_\varphi \,\big|\, \big(\mu \,\big|\, \partial_1 K^i(\varphi(x), \varphi(y)) h(x)\big)_x e_i\big)_y \\ &\quad + \sum_{i=1}^d \big(p_\varphi \,\big|\, \big(\mu \,\big|\, \partial_2 K^i(\varphi(x), \varphi(y)) h(y)\big)_x e_i\big)_y, \end{aligned}$$

$$\big(\partial_2 \mathscr{V}_1^* \, p_\varphi \,\big|\, \eta\big) = \sum_{i=1}^d \big(p_\varphi \,\big|\, \big(\eta \,\big|\, K^i(\varphi(x), \varphi(y))\big)_x e_i\big)_y,$$

$$\begin{aligned} \big(\partial_1 \mathscr{V}_2^* \, p_\mu \,\big|\, h\big) &= -\sum_{i=1}^d \big(\mu \,\big|\, \big(\mu \,\big|\, \partial_1 \partial_2 K^i(\varphi(x), \varphi(y))(h(x), p_\mu(y))\big)_x e_i\big)_y \\ &\quad - \sum_{i=1}^d \big(\mu \,\big|\, \big(\mu \,\big|\, \partial_2^2 K^i(\varphi(x), \varphi(y))(h(y), p_\mu(y))\big)_x e_i\big)_y, \text{ and} \end{aligned}$$

$$\begin{aligned} \big(\partial_2 \mathscr{V}_2^* \, p_\mu \,\big|\, \eta\big) &= -\sum_{i=1}^d \big(\eta \,\big|\, \big(\mu \,\big|\, \partial_2 K^i(\varphi(x), \varphi(y)) p_\mu(y)\big)_x e_i\big)_y \\ &\quad - \sum_{i=1}^d \big(\mu \,\big|\, \big(\eta \,\big|\, \partial_2 K^i(\varphi(x), \varphi(y)) p_\mu(y)\big)_x e_i\big)_y. \end{aligned}$$

Forming explicit expressions of $\partial_i \mathscr{V}_j^*$ requires isolating h or η from the right-hand sides. To do this, we will need to change the order in which linear forms are applied to the x and y coordinates. This issue is addressed in the following lemma.

Lemma 10.15 *Assume that $\mu \in C^r(\Omega, \mathbb{R}^d)^*$ and $\nu \in C^{r'}(\Omega, \mathbb{R}^d)^*$. Let $g : \Omega \times \Omega \to \mathbb{R}$ be a function such that $\partial_1^k \partial_2^{k'} g \in C_0(\Omega \times \Omega, \mathbb{R})$ for all $k \leq r$ and $k' \leq r$. Then, for all $a, b \in \mathbb{R}^d$, $(\mu \,|\, g(x, \cdot)a)_x \in C^{r'}(\Omega, \mathbb{R}^d)$ and $(\nu \,|\, g(\cdot, y)b)_y \in C^r(\Omega, \mathbb{R})$, with*

$$\big(\mu \,\big|\, (\nu \,|\, g(x, y)b)_y a\big)_x = \big(\nu \,\big|\, (\mu \,|\, g(x, y)a)_x b\big)_y. \tag{10.42}$$

Proof Let $f(y) = (\mu \mid g(x,y)a)_x$. Using Taylor's formula, we can write

$$g(x,y+h) = \sum_{k=0}^{r'} \frac{1}{k!}\partial_2^k g(x,y)h^{(k)} + \frac{1}{(r'-1)!}\int_0^1 (\partial_2^{r'} g(x,y+th) - \partial_2^{r'} g(x,y))h^{(r')}(1-t)^{r'-1}\,dt$$

so that

$$f(y+h) = \sum_{k=0}^{r'} \frac{1}{k!}\big(\mu \mid \partial_2^k g(x,y)h^{(k)}a\big) + \frac{1}{(r'-1)!}\Big(\mu \Big| \int_0^1 (\partial_2^{r'} g(x,y+th) - \partial_2^{r'} g(x,y))h^{(r')})(1-t)^{r'-1}a\,dt\Big).$$

The last term (call it R) is such that

$$|R| \le \frac{\|\mu\|_{r,\infty,*}|h|^{r'}|a|}{r'!} \max_{t\in[0,1]} \|\partial_2^{r'} g(\cdot, y+th) - \partial_2^{r'} g(\cdot,y)\|_{r,\infty}.$$

The uniform continuity of $\partial_1^k\partial_2^{r'} g$ for $k \le r$ implies that $R = o(|h|^{r'})$ so that $f \in C^{r'}(\Omega,\mathbb{R})$. Similarly, letting $f'(x) = (\mu \mid g(x,y)b)_y$, one has $f' \in C^r(\Omega,\mathbb{R})$.

The computation also shows that, for some constant C,

$$\max\Big(\big(\mu \mid (\nu \mid g(x,y)b)_y a\big)_x, \big(\nu \mid (\mu \mid g(x,y)a)_x b\big)_y\Big) \le C\|\mu\|_{r,\infty,*}\|\nu\|_{r',\infty,*}\|g\|_{r,r',\infty}$$

with

$$\|g\|_{r,r',\infty} = \max_{k\le r, k'\le r'} \|\partial_1^k\partial_2^{k'} g\|_\infty,$$

so that both sides of (10.42) are continuous in g with respect to this norm. To conclude, it suffices to notice that (10.42) is true when g takes the form

$$g(x,y) = \sum_{k=1}^n c_k f_k(x) f_k'(y)$$

and that these functions form a dense set for $\|g\|_{r,r',\infty}$, so that the identity extends by continuity. □

Let us use this lemma to identify the first term in $\big(\partial_1 \mathscr{V}_1^* p_\varphi \mid h\big)$ as a linear form acting on h. Write, letting $\partial_{i,k}$ denote the derivative with respect to the kth coordinate of the ith variable,

$$
\begin{aligned}
&\sum_{i=1}^{d} \big(p_\varphi \,\big|\, \big(\mu \,\big|\, \partial_1 K^i(\varphi(x),\varphi(y))h(x)\big)_x e_i\big)_y \\
&= \sum_{i,j,k=1}^{d} \big(p_\varphi \,\big|\, \big(\mu \,\big|\, \partial_{1,k} K^{ij}(\varphi(x),\varphi(y))h^k(x)e_j\big)_x e_i\big)_y \\
&= \sum_{i,j,k=1}^{d} \Big(\mu \,\Big|\, \big(p_\varphi \,\big|\, \partial_{1,k} K^{ij}(\varphi(x),\varphi(y))h^k(x)e_i\big)_y e_j\Big)_x \\
&= \sum_{i,j,k=1}^{d} \Big(\mu \,\Big|\, \big(p_\varphi \,\big|\, \partial_{2,k} K^{ji}(\varphi(y),\varphi(x))e_i\big)_y h^k(x)e_j\Big)_x \\
&= \sum_{j,k=1}^{d} \Big(\mu \,\Big|\, \big(p_\varphi \,\big|\, \partial_{2,k} K^{j}(\varphi(y),\varphi(x))\big)_y h^k(x)e_j\Big)_x \\
&= (\mu \mid \mathscr{U}_1 h),
\end{aligned}
$$

where $\mathscr{U}_1(x)$ is the matrix with coefficients

$$
\mathscr{U}_1^{j,q}(x) = \big(p_\varphi \,\big|\, \partial_{2,q} K^j(\varphi(y),\varphi(x))\big)_y.
$$

Write $\big(\mathscr{U}_1^T \mu \,\big|\, h\big) = (\mu \mid \mathscr{U}_1 h)$, a notation generalizing the one introduced for vector measures. After a similar computation for the second term of $\big(\partial_1 \mathscr{V}_1^* \, p_\varphi \,\big|\, h\big)$ (which does not require Lemma 10.15), we get

$$
\partial_1 \mathscr{V}_1^* \, p_\varphi = \mathscr{U}_1^T \mu + \mathscr{U}_2^T p_\varphi
$$

with

$$
\mathscr{U}_2^{j,q}(x) = \big(\mu \,\big|\, \partial_{2,q} K^j(\varphi(y),\varphi(x))\big)_y.
$$

Consider now $\partial_2 \mathscr{V}_1^* \, p_\varphi$, writing

$$
\begin{aligned}
\big(\partial_2 \mathscr{V}_1^* \, p_\varphi \,\big|\, \eta\big) &= \sum_{i,j=1}^{d} \Big(\eta \,\Big|\, \big(p_\varphi \,\big|\, K^{ij}(\varphi(x),\varphi(y))e_i\big)_y e_j\Big)_x \\
&= \sum_{j=1}^{d} \Big(\eta \,\Big|\, \big(p_\varphi \,\big|\, K^{j}(\varphi(y),\varphi(x))\big)_y e_j\Big)_x,
\end{aligned}
$$

so that

$$
\partial_2 \mathscr{V}_1^* \, p_\varphi(x) = \sum_{j=1}^{d} \big(p_\varphi \,\big|\, K^j(\varphi(y),\varphi(x))\big)_y e_j.
$$

With similar computations for $\mathscr{V}_2$, and skipping the details, we find

$$\partial_1 \mathscr{V}_2^* \, p_\mu = -\mathscr{U}_3^T \mu - \mathscr{U}_4^T \mu$$

where

$$\mathscr{U}_3^{jq}(x) = \left(\mu \;\middle|\; \partial_{2,q}(\partial_1 K^j(\varphi(y), \varphi(x)) p_\mu(y))\right)_y, \text{ and}$$
$$\mathscr{U}_4^{jq}(x) = \left(\mu \;\middle|\; \partial_{2,q}(\partial_2 K^j(\varphi(y), \varphi(x)) p_\mu(x))\right)_y.$$

Finally,

$$\partial_2 \mathscr{V}_2^* \, p_\mu(x) = -\sum_{i=1}^{d} \left(\mu \;\middle|\; \partial_2 K^i(\varphi(y), \varphi(x)) p_\mu(x)\right)_y e_i \\ - \sum_{i=1}^{d} \left(\mu \;\middle|\; \partial_1 K^i(\varphi(y), \varphi(x)) p_\mu(y)\right)_y e_i.$$

Let us take the special case of vector measures, assuming that $\mu(t) = \sum_{k=1}^{N} z_k(t, \cdot)\gamma_k$. We will look for p_φ in the form

$$p_\varphi(t) = \sum_{k=1}^{N} \alpha_k(t, \cdot)\gamma_k,$$

p_μ being a function defined over the support of μ.

With these assumptions, we have

- $\partial_1 \mathscr{V}_1^* \, p_\varphi = \sum_{k=1}^{N} \zeta_k^{1,1} \gamma_k$ with

$$\zeta_k^{1,1}(x) = \sum_{l=1}^{N} \sum_{i,j=1}^{d} \left(\gamma_l \;\middle|\; \alpha_l^i(y) z_k^j(x) \nabla_1 K^{ij}(\varphi(x), \varphi(y))\right)_y \\ + \sum_{l=1}^{N} \sum_{i,j=1}^{d} \left(\gamma_l \;\middle|\; z_l^i(y) \alpha_k^j(x) \nabla_1 K^{ij}(\varphi(x), \varphi(y))\right)_y.$$

- $\partial_2 \mathscr{V}_1^* \, p_\varphi(x) = \sum_{k=1}^{N} (\gamma_k \mid K(\varphi(x), \varphi(y))\alpha_k(y))_y.$
- $\partial_1 \mathscr{V}_2^* \, p_\mu = \sum_{k=1}^{N} \zeta_k^{2,1} \gamma_k$ with

$$\zeta_k^{2,1}(x) = -\sum_{i,j=1}^{d}\sum_{l=1}^{N}\left(\gamma_l \;\middle|\; z_l^i(y) z_k^j(x) \partial_1\partial_2 K^{ij}(\varphi(x),\varphi(y)) p_\mu(y)\right)_y - \sum_{i,j=1}^{d}\sum_{l=1}^{N}\left(\gamma_l \;\middle|\; z_l^i(y) z_k^j(x) \partial_1^2 K^{ij}(\varphi(x),\varphi(y)) p_\mu(x)\right)_y.$$

- $\partial_2 \mathscr{V}_2^*\, p_\mu(x) = -\sum_{i=1}^{d}\sum_{k=1}^{N}\left(\gamma_k \;|\; z_k^i(y)\partial_1 K^i(\varphi(x),\varphi(y)) p_\mu(x)\right)_y - \sum_{i=1}^{d}\sum_{k=1}^{N}\left(\gamma_k \;|\; z_k^i(y)\partial_2 K^i(\varphi(x),\varphi(y)) p_\mu(y)\right)_y.$

System (10.41) can now be simplified as

$$\begin{cases} \partial_t \alpha_k = \zeta_k^{1,1} + \zeta_k^{2,1} \\ \partial_t p_\mu = \displaystyle\sum_{k=1}^{N}(\gamma_k \mid K(\varphi(x),\varphi(y))\alpha_k(y))_y \\ \qquad - \displaystyle\sum_{i=1}^{d}\sum_{k=1}^{N}\left(\gamma_k \;|\; z_k^i(y)\partial_1 K^i(\varphi(x),\varphi(y)) p_\mu(x)\right)_y \\ \qquad - \displaystyle\sum_{i=1}^{d}\sum_{k=1}^{N}\left(\gamma_k \;|\; z_k^i(y)\partial_2 K^i(\varphi(x),\varphi(y)) p_\mu(y)\right)_y. \end{cases} \tag{10.43}$$

Application to Point Matching

We now apply this approach to point-matching problems. Since ρ_0 takes the form

$$\rho_0 = \sum_{k=1}^{N} a_{0,k}\delta_{x_{0,k}}$$

we are in the vector measure case with $\gamma_k = \delta_{x_{0,k}}$. The densities z_k and α_k for μ and p_φ can therefore be considered as vectors in $\mathbb{R}^d$, and p_μ being defined on the support of μ is also a collection of vectors $p_{\mu,k} = p_\mu(x_k)$. Given this, we can therefore immediately rewrite

- $\partial_1 \mathscr{V}_1^*\, p_\varphi = \sum_{k=1}^{N} \zeta_k^{1,1}\delta_{x_{0,k}}$ with

$$\zeta_k^{1,1} = \sum_{l=1}^{N}\sum_{i,j=1}^{d}\left(\alpha_l^i \nabla_1 K^{ij}(x_k,x_l) z_k^j + z_l^i \nabla_1 K^{ij}(x_k,x_l)\alpha_k^j\right).$$

- $\partial_2 \mathscr{V}_1^* \, p_\varphi(x_{0,k}) = \sum_{l=1}^N K(x_k, x_l)\alpha_l.$
- $\partial_1 \mathscr{V}_2^* \, p_\mu = \sum_{k=1}^N \zeta_k^{2,1} \delta_{x_{0,k}}$ with

$$\zeta_k^{2,1} = -\sum_{i,j=1}^d \sum_{l=1}^N z_l^i z_k^j (\partial_1 \partial_2 K^{ij}(x_k, x_l)\, p_{\mu,l} + \partial_1^2 K^{ij}(x_k, x_l)\, p_{\mu,k}).$$

- $\partial_2 \mathscr{V}_2^* \, p_\mu(x_k) = -\sum_{i=1}^d \sum_{l=1}^N z_l^i (\partial_1 K^i(x_k, x_l) p_{\mu,k} + \partial_2 K^i(x_k, x_l) p_{\mu,l}).$

This algorithm is illustrated in Fig. 10.1. In the same figure, we also provide (for comparison purposes) the results provided by spline interpolation, which computes $\varphi(x) = x + v(x)$, where v is computed (using Theorem 8.9) in order to minimize

$$\|v\|_V^2 + C \sum_{i=1}^N |v(x_i) - (y_i - x_i)|^2 .$$

Although this is a widely spread registration method [42], Fig. 10.1 shows that it is far from being diffeomorphic for large deformations.

10.6.4 Shooting

The optimality conditions for our problem are $\mu(1) = -dU(\varphi(1))$ with $\mu(t)$ given by (10.28). The shooting approach in optimal control consists in finding an initial momentum $\rho_0 = \mu(0)$ such that these conditions are satisfied. Root finding methods, such as Newton's algorithm, can be used for this purpose. At a given step of Newton's algorithm, one modifies the current value of ρ_0, by letting $\rho_0 \to \rho_0 + \eta$ such that, letting $F(\rho_0) := \mu(1) + dU(\varphi(1))$, one has

$$F(\rho_0) + dF(\rho_0)\eta = 0.$$

One therefore needs to solve this linear equation in order to update the current ρ_0. One has

$$dF(\rho_0) = W_{\mu\mu}(1) + W_{\varphi\mu}(1)^* d^2 U(\varphi(1)),$$

where

$$W = \begin{pmatrix} W_{\varphi\varphi} & W_{\varphi\mu} \\ W_{\mu\varphi} & W_{\mu\mu} \end{pmatrix}$$

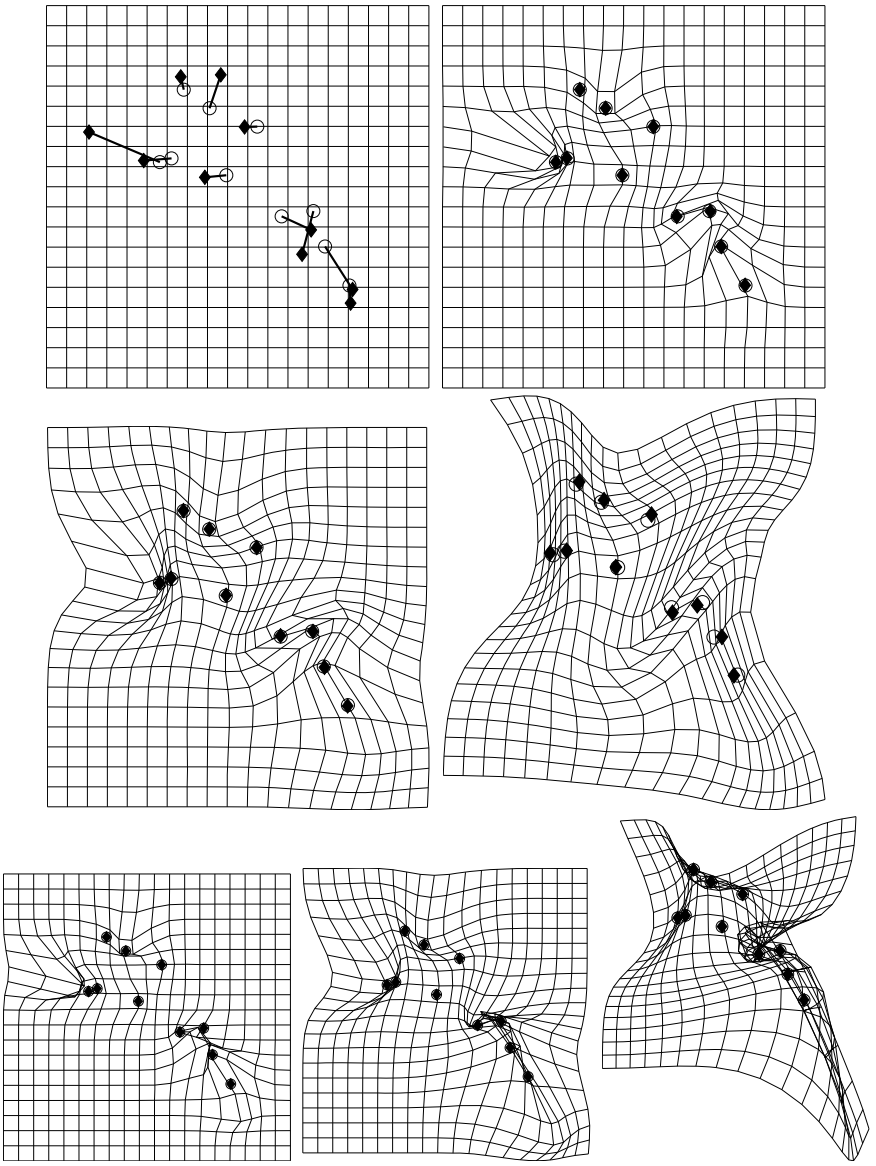

Fig. 10.1 Metric point matching. The first two rows provide results obtained with gradient descent in the initial momentum for point matching, with the same input as in Fig. 9.1, using Gaussian kernels $K(x, y) = \exp(-|x - y|^2/2\sigma^2)$ with $\sigma = 1, 2, 4$ in grid units. The impact of the diffeomorphic regularization on the quality of the result is particularly obvious in the last experiment. The last row provides the output of Gaussian spline registration with the same kernels, exhibiting singularities and ambiguities in the registration

is (using the notation of the previous section) the differential of the solution of the equation

$$\partial_t \begin{pmatrix} \varphi \\ \mu \end{pmatrix} = \mathscr{V}(\varphi, \mu),$$

with respect to its initial condition, i.e., the solution of

$$\partial_t W = d\mathscr{V}(\varphi, \mu) W$$

with initial condition $W(0) = \mathrm{Id}$.

Because one needs to compute the solution of this differential equation at every step of the algorithm, then solve for a linear system, the shooting method is feasible only for problems that can be discretized into a relatively small number of dimensions. One can use it, for example, in point matching problems with no more than a few hundred landmarks (see [290] for an application to labeled point matching), in which case the algorithm can be very efficient. Another issue is that root finding algorithms are not guaranteed to converge. Usually, a good initial solution must be found, using, for example, a few preliminary steps of gradient descent.

10.6.5 Gradient in the Deformable Object

Finally, we consider the option of using the time derivative of the deformable object as a control variable, using the fact that, by (10.33), the objective function can be reduced to

$$E(J) = \int_0^1 L(\partial_t J(t), J(t))dt + F(J(1))$$

with $L(\eta, J) = \min_{w:\, \eta = w \cdot J(t)} \|w\|_V^2$. This formulation is limited, in that $L(\eta, J)$ is not always defined for all (η, J), resulting in constraints in the minimization that are not always easy to handle. Even if well-defined, the computation of L may be numerically demanding. To illustrate this, consider the image-matching case, in which $v \cdot J = -\nabla J^T v$. An obvious constraint is that, in order for

$$\nabla J^T w = -\eta$$

to have at least one solution, the variation η must be supported by the set $\nabla J \neq 0$. To compute this solution when it exists, one can write, for $x \in \Omega$,

$$\nabla J(x)^T w(x) = \left\langle K(\cdot, x)\nabla J(x)\,,\ w \right\rangle_V,$$

and it is possible to look for a solution in the form

$$w(y) = \int_\Omega \lambda(x) K(y,x) \nabla J(x) dx,$$

where $\lambda(x)$ can be interpreted as a continuous family of Lagrange multipliers. This results in a linear equation in λ, namely

$$\int_\Omega \lambda(x) K(y,x) \nabla J(y)^T \nabla J(x)) dx = -\eta(y),$$

which is numerically challenging.

For point sets, however, the approach is feasible [159] because L can be made explicit. Given a point-set trajectory $x(t) = (x^{(1)}(t), \ldots, x^{(N)}(t))$, let $S(x(t))$ denote the block matrix with (i,j) block given by $K(x^{(i)}(t), x^{(j)}(t))$. The constraints are $x_t = S(x(t))\xi(t)$ so that $\xi(t) = S(x(t))^{-1}\dot{x}_t$ and the minimization reduces to

$$E(x) = \frac{1}{2}\int_0^1 \dot{x}(t)^T S(x(t))^{-1}\dot{x}(t)dt + U(x(1)).$$

Minimizing this function with respect to x by gradient descent is possible, and has been described in [158, 159] for labeled landmark matching. The basic computation is as follows: if $s_{pq,r} = \partial_r s_{pq}$, we can write (using the fact that $\partial_r(S^{-1}) = -S^{-1}(\partial_r S)S^{-1}$)

$$\begin{aligned}(dE(x) \mid h) = &\int_0^1 \dot{x}(t)^T S(x(t))^{-1}\dot{h}(t)dt \\ &- \int_0^1 \sum_{p,q,r} \xi^{(p)}(t)\xi^{q)}(t) s_{pq,r}(x(t))h^{(r)}(t)dt + \nabla U(x(1))^T h(1).\end{aligned}$$

After an integration by parts in the first integral, we obtain

$$dE(x) = -\partial_t\left(S(x(t))^{-1}\dot{x}\right) - z(t) + \left(S(x(1))^{-1}\dot{x}(1) + \nabla U(x(1))\right)\delta_1(t),$$

where $z_r(t) = \sum_{p,q} \xi_p(t)\xi_q(t)s_{pq,r}(x(t))$ and δ_1 is the Dirac measure at $t = 1$.

This singular part can be dealt with by computing the gradient in a Hilbert space in which the evaluation function $x(\cdot) \mapsto x(1)$ is continuous. This method has been suggested, in particular, in [129, 161]. Let H be the space of all trajectories $x : t \mapsto x(t) = (x^{(1)}(t), \ldots, x^{(N)}(t))$, with fixed starting point $x(0)$, free end-point $x(1)$ and square integrable time derivative. This is a space of the form $x(0) + H$ where H is the Hilbert space of time-dependent functions $t \mapsto h(t)$, considered as column vectors of size Nk, with $h(0) = 0$ and

$$\left\langle h\,,\,\tilde{h}\right\rangle_H = \int_0^1 \dot{h}^T\dot{\tilde{h}}dt + h(1)^T\tilde{h}(1).$$

To compute the gradient for this inner product, we need to write $(dE(x) \mid h)$ in the form $\left\langle \nabla^H E(x)\,,\, h\right\rangle_H$. We will make the assumption that

$$\int_0^1 \left|S(x(t))^{-1}\dot{x}(t)\right|^2 dt < \infty,$$

which implies that

$$\int_0^1 \dot{x}(t)^T S(x(t))^{-1}\dot{h}(t)dt \leq \sqrt{\int_0^1 \left|S(x(t))^{-1}\dot{x}(t)\right|^2 dt \int_0^1 \left|\dot{h}(t)\right|^2 dt}$$

is continuous in h. Similarly, the linear form $\xi \mapsto \nabla U(x(1))^T h(1)$ is continuous since

$$\nabla U(x(1))^T h(1) \leq |\nabla U(x(1))|\,|h(1)|\,.$$

Finally, $h \mapsto \int_0^1 z(t)^T \dot{h}dt$ is continuous provided that we assume that

$$\eta(t) = \int_0^t z(s)dt$$

is square integrable over [0, 1], since this yields

$$\int_0^1 z(t)^T h(t)dt = \eta(1)h(1) - \int_0^1 \eta(t)\dot{h}(t)dt,$$

which is continuous in h with respect to the H norm.

Thus, under these assumptions, $h \mapsto (dE(x) \mid h)$ is continuous over H, and the Riesz representation theorem implies that $\nabla^H E(x)$ exists as an element of H. We now proceed to its computation. Letting

$$\mu(t) = \int_0^t S(x(s))^{-1}\dot{h}(s)ds$$

and $a = \nabla U(x(1))$, the problem is to find $\zeta \in H$ such that, for all $h \in H$,

$$\left\langle \zeta\,,\, h\right\rangle_H = \int_0^1 \dot{\mu}^T \dot{h}dt + \int_0^1 z(t)^T h(t)dt + a^T \xi(1).$$

This expression can also be written

$$\int_0^1 \left(\dot{\zeta} + \zeta(1)\right)^T \dot{h}dt = \int_0^1 (\dot{\mu} + \eta(1) - \eta(t) + a)^T \dot{h}dt.$$

This suggests selecting ζ such that $\zeta(0) = 0$ and

$$\dot{\zeta} + \zeta(1) = \dot{\mu} + \eta(1) - \eta(t) + a,$$

which implies

$$\zeta(t) + t\zeta(1) = \mu(t) - \int_0^t \eta(s)ds + t(\eta(1) + a).$$

At $t = 1$, this yields

$$2\zeta(1) = \mu(1) - \int_0^1 \eta(s)ds + \eta(1) + a$$

and we finally obtain

$$\zeta(t) = \mu(t) - \int_0^t \eta(s)ds + \frac{t}{2}\left(\int_0^1 \eta(s)ds - \mu(1) + \eta(1) + a\right).$$

We summarize this in an algorithm, in which τ is again the computation time.

Algorithm 3 (*Gradient descent algorithm for landmark matching*) Start with initial landmark trajectories $x(t, \tau) = (x^{(1)}(t, \tau), \ldots, x^{(N)}(t, \tau))$.

Solve

$$\begin{aligned}\partial_\tau x(t, \tau) = -\gamma\Big(&\mu(t, \tau) - \int_0^t \eta(s, \tau)ds \\ &+ \frac{t}{2}\Big(\int_0^1 \eta(s, \tau)ds - \mu(1, \tau) + \eta(1, \tau) + a(\tau)\Big)\Big)\end{aligned}$$

with $a(\tau) = \nabla U(x(1, \tau))$, $\mu(t, \tau) = \int_0^t \xi(s, \tau)ds$, $\eta(t, \tau) = \int_0^t z(s, \tau)dt$ and

$$\begin{aligned}\xi(t, \tau) &= S(x(t, \tau))^{-1}\dot{x}(t, \tau) \\ z^{(q)}(t, \tau) &= \sum_{p,r} \xi^{(p)}(t, \tau)\xi^{(r)}(t, \tau)s_{pq,r}(x(t, \tau)).\end{aligned}$$

10.6.6 Image Matching

We now take an infinite-dimensional example to illustrate some of the previously discussed methods and focus on the image-matching problem. We therefore consider

$$U(\varphi) = \frac{\lambda}{2}\int_\Omega (I \circ \varphi^{-1} - \tilde{I})^2 dx,$$

where I, $\tilde{I}$ are functions $\Omega \to \mathbb{R}$, I being differentiable. The Eulerian differential of U is given by (9.21):

$$\bar{\partial} U(\varphi) = -\lambda (I \circ \varphi^{-1} - I') \nabla (I \circ \varphi^{-1}) dx.$$

So, according to (10.19), and letting $\tilde{U}(v) = U(\varphi_{01}^v)$,

$$\begin{aligned}
\nabla \tilde{U}(v)(t,y) &= \sum_{i=1}^{d} \left(\bar{\partial} U(\varphi_{01}^v) \,\middle|\, d\varphi_{t1}^v(\varphi_{1t}^v(.)) K(y, \varphi_{1t}^v(.)) e_i \right) e_i \\
&= -\lambda \sum_{i=1}^{d} e_i \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) \\
&\qquad \nabla (I \circ \varphi_{10}^v)(x)^T d\varphi_{t1}^v(\varphi_{1t}^v(x)) K(y, \varphi_{1t}^v(x)) e_i dx \\
&= -\lambda \sum_{i=1}^{d} e_i \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) \\
&\qquad \nabla I \circ \varphi_{10}^v(x)^T d\varphi_{10}^v(x) d\varphi_{t1}^v(\varphi_{1t}^v(x)) K(y, \varphi_{1t}^v(x)) e_i dx \\
&= -\lambda \sum_{i=1}^{d} e_i \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) \\
&\qquad \nabla I \circ \varphi_{10}^v(x)^T d\varphi_{t0}^v(\varphi_{1t}^v(x)) K(y, \varphi_{1t}^v(x)) e_i dx \\
&= -\lambda \sum_{i=1}^{d} e_i \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) \nabla (I \circ \varphi_{t0}^v)(\varphi_{1t}^v(x))^T K(y, \varphi_{1t}^v(x)) e_i dx \\
&= -\lambda \sum_{i=1}^{d} e_i \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) e_i^T K(\varphi_{1t}^v(x), y) \nabla (I \circ \varphi_{t0}^v)(\varphi_{1t}^v(x)) dx \\
&= -\lambda \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) K(\varphi_{1t}^v(x), y) \nabla (I \circ \varphi_{t0}^v)(\varphi_{1t}^v(x)) dx.
\end{aligned}$$

This provides the expression of the V-gradient of $\tilde{E}$ for image matching, namely

$$\begin{aligned}
(\nabla^V E(v))(t,y) &= v(t,y) \\
&\quad - \lambda \int_\Omega (I \circ \varphi_{10}^v(x) - \tilde{I}(x)) K(\varphi_{1t}^v(x), y) \nabla (I \circ \varphi_{t0}^v)(\varphi_{1t}^v(x)) dx.
\end{aligned} \tag{10.44}$$

Using a change of variable in the integral, the gradient may also be written as

$$\begin{aligned}
(\nabla^V E)(t,y) &= v(t,y) \\
- \lambda \int_\Omega (I \circ \varphi_{t0}^v(x) - \tilde{I} \circ \varphi_{t1}^v(x)) K(x,y) \nabla (I \circ \varphi_{t0}^v)(x) \det(d\varphi_{t1}^v(x)) dx,
\end{aligned} \tag{10.45}$$

the associated gradient descent algorithm having been proposed in [32].

Let us now consider an optimization with respect to the initial ρ_0. First notice that, by (9.20), $\mu(1) = \lambda \det(d\varphi(1))(I - I' \circ \varphi(1))d\varphi(1)^{-T}\nabla I\, dx$ is a vector measure. Also, we have

$$\begin{aligned}(\rho_0 \mid w) &= (\mu(1) \mid d\varphi(1)w) \\ &= \lambda\big(dx \;\big|\; \det(d\varphi(1))(I - I' \circ \varphi(1))\nabla I^T w\big),\end{aligned}$$

which shows that one can assume that $\rho_0 = z_0 dx$ for some vector-valued function z_0 (with $z_0 = \det(d\varphi(1))(I - I' \circ \varphi(1))\nabla I$ for an optimal control).

We now make explicit the computation of the differential of the energy with respect to ρ_0. We have $\mu(t) = z(t, \cdot)dx$, with $z(0) = z_0$ and

$$\begin{cases} \partial_t \varphi(t, y) = \displaystyle\int_{\mathbb{R}^d} K(\varphi(t, y), \varphi(t, x))z(t, x)dx \\ \partial_t z(t, y) = -\displaystyle\int_{\mathbb{R}^d} z^i(t, y)z^j(t, x)\nabla_1 K^{ij}(\varphi(t, y), \varphi(t, x))dx. \end{cases} \tag{10.46}$$

The differential $dE(\rho_0) = K\rho_0 - p_\mu(0)$ is computed by solving, using $\alpha(1) = \lambda \det(d\varphi(1))(I - I' \circ \varphi(1))d\varphi(1)^{-T}\nabla I$ and $p_\mu(1) = 0$,

$$\begin{cases} \partial_t \alpha = \zeta^{1,1} + \zeta^{2,1} \\ \partial_t p_\mu = \displaystyle\int_{\mathbb{R}^d} K(\varphi(x), \varphi(y))\alpha(y)dy \\ \qquad\qquad - \displaystyle\sum_{i=1}^d \int_{\mathbb{R}^d} z^i(y)\partial_1 K^i(\varphi(x), \varphi(y))p_\mu(x)dy \\ \qquad\qquad - \displaystyle\sum_{i=1}^d \int_{\mathbb{R}^d} z^i(y)\partial_2 K^i(\varphi(x), \varphi(y))p_\mu(y)dy, \end{cases} \tag{10.47}$$

in which

$$\zeta^{1,1}(x) = \sum_{i,j=1}^d \int_{\mathbb{R}^d} (\alpha^i(y)z^j(x) + z^i(y)\alpha^j(x))\nabla_1 K^{ij}(\varphi(x), \varphi(y))dy$$

and

$$\begin{aligned}\zeta^{2,1}(x) = &- \sum_{i,j=1}^d \int_{\mathbb{R}^d} z^i(y)z^j(x)\partial_1\partial_2 K^{ij}(\varphi(x), \varphi(y))p_\mu(y)dy \\ &- \sum_{i,j=1}^d \int_{\mathbb{R}^d} z_l^i(y)z_k^j(x)\partial_1^2 K^{ij}(\varphi(x), \varphi(y))p_\mu(x)dy.\end{aligned}$$

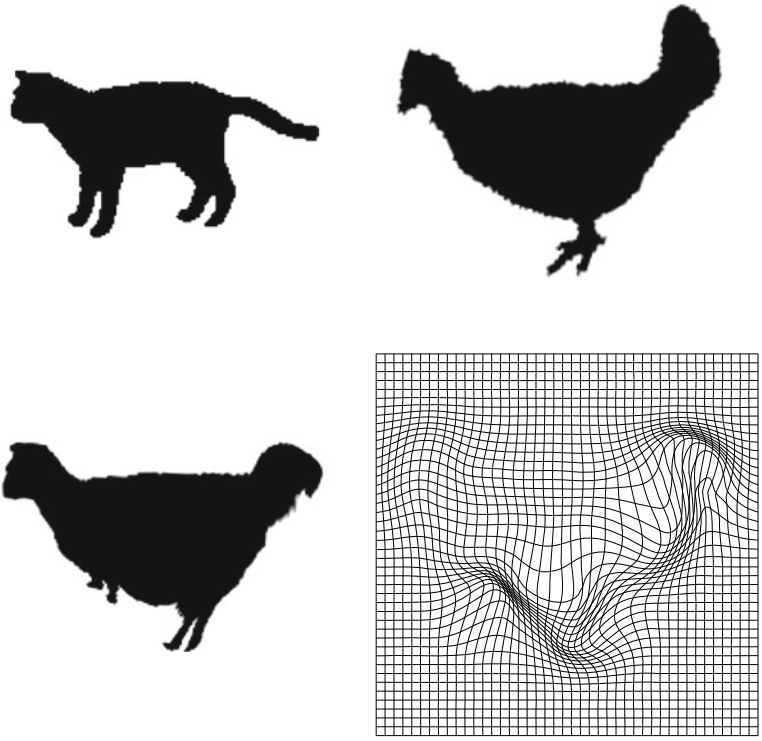

Fig. 10.2 Metric image matching. Output of Algorithm 4 when estimating a deformation of the first image to match the second one (compare to Fig. 9.2). The third image is the obtained deformation of the first one and the last provides the deformation applied to a grid

We summarize the computation of the gradient of the image-matching functional with respect to z_0 such that $\rho_0 = z_0 dx$:

Algorithm 4

1. Solve (10.46) with initial conditions $\varphi(0) = \mathrm{id}$ and $z(0) = z_0$ and compute $dU(\varphi(1)) = -\lambda(I - I' \circ \varphi(1)) \det(d\varphi(1)) d\varphi(1)^{-T} \nabla I$.
2. Solve, backwards in time, until time $t = 0$ the system (10.47) with boundary conditions $\alpha(1) = -dU(\varphi(1))$ and $p_\mu(1) = 0$.
3. Set $\nabla E(z_0) = 2z_0 - \mathbb{K}^{-1} p_\mu(0)$.

The gradient is computed with the metric $\langle z\,,\, z' \rangle = \int_{\mathbb{R}^d} z(y)^T \mathbb{K} z(y) dy$. Results obtained with this algorithm are presented in Fig. 10.2.

One can also use the fact that $z_0 = f_0 \nabla I$ for a scalar-valued f_0. Since we have

$$\Big(dE(z_0) \,\Big|\, h_0 \Big) = \int_\Omega (\mathbb{K} z_0 - p_\mu(0))^T h_0 dy,$$

we can write, with $\tilde{E}(f_0) = E(f_0 \nabla I)$:

$$\left(d\tilde{E}(f_0) \,\middle|\, u_0\right) = \int_\Omega (\mathbb{K}(f_0 \nabla I) - p_\mu(0))^T \nabla I u_0 dy,$$

which leads to replacing the last step in Algorithm 4 by

$$\nabla E(f_0) = -\nabla I^T (\mathbb{K}(f_0 \nabla I) - p_\mu(0)),$$

which corresponds to using the L^2 metric in f_0 for gradient descent. However, a more natural metric, in this case, is the one induced by the kernel, i.e.,

$$\langle f \,,\, f' \rangle_I = \int_\Omega \int_\Omega \mathbb{K}(f \nabla I)(y)(f'(y) \nabla I(y) dy = \int_\Omega K_I(x, y) f(x) f'(y) dx dy$$

with $K_I(x, y) = \nabla I(x)^T K(x, y) \nabla I(y)$. With this metric, z_0 is updated with

$$\nabla E(f_0) = z_0 - \mathbb{K}_I^{-1} \nabla I^T p_\mu(0).$$

Although this metric is more satisfactory from a theoretical viewpoint, the inversion of $\mathbb{K}_I$ might be difficult, numerically.

10.6.7 Pros and Cons of the Optimization Strategies

In the previous sections we have reviewed several possible choices of control variables with respect to which the optimization of the matching energy can be performed. For all but the shooting method, this results in specific expressions of the gradient that can then be used in optimization procedures such as those discussed in Appendix D.

All these procedures have been implemented in the literature to solve a diffeomorphic-matching problem in at least one specific context, but no extensive study has ever been made to compare them. Even if the outcome of such a study is likely to be that the best method depends on the specific application, one can still provide a few general facts that can help a user decide which one to use.

When feasible (that is, when the linear system it involves at each step can be efficiently computed and solved), the shooting method is probably the most efficient. If the initialization is not too far from the solution, convergence can be achieved in a very small number of iterations. One cannot guarantee, however, that the method will converge starting from any initial point, and shooting needs to be combined with some gradient-based procedure in order to find a good starting position.

Since they optimize with respect to the same variable, the most natural procedure to combine with shooting is optimization with respect to the initial momentum. Even when shooting is not feasible (e.g., for large-scale problems), this specific choice of control variable is important, because it makes sure that the final solution satisfies the

EPDiff equation, which guarantees the consistency of the momentum representation, which will be discussed in Sect. 11.5.2. The limitation is that, with large and complex deformations, the sensitivity of the solution to small changes in the control variable can be large, which may result in an unstable optimization procedure.

The other methods, which optimize with respect to time-dependent quantities, are generally more able to compute very large deformations. Beside the obvious additional burden in computer memory that they require, one must be aware that the discrete solution can sometimes be far from satisfying the EPDiff equation unless the time discretization is fine enough (which may be impossible to achieve within a feasible implementation for large-scale problems). Therefore, these methods do not constitute the best choice if obtaining a reliable momentum representation is important. Among the three time-dependent control variables that we have studied (velocity, momentum and deformable object), one may have a slight preference for the representation using the time-dependent momenta, even if the computation it involves is slightly more complex than the others. There are at least two reasons for this. First, the momenta are generally more parsimonious in the space variables, because they incorporate normality constraints to transformations that leave the deformable objects invariant. Second, because the forward and backward equations solved at each iteration immediately provide a gradient with respect to the correct metric, so that the implementation does not have to include the solution of a possibly large-dimensional linear system which is required by other representations.

10.7 Numerical Aspects

10.7.1 Discretization

The implementation of the diffeomorphic matching algorithms that were just discussed requires a proper discretization of the different variables that are involved. The discretization in time of optimal control problems is discussed in Sect. D.4. This discussion directly applies here and we refer the reader to the relevant pages in the chapter for more details. If the deformed objects are already discrete (e.g., points sets), this suffices in order to design a numerical implementation.

When the deformed objects are continuous, some discrete approximation must obviously be made. One interesting feature of the problems that we have discussed is that they all derive from the general formulation (10.8), but can be reduced, using Sect. 10.6.2, to a situation in which the state and controls are finite dimensional after discretization. Typically, starting from (10.8), the discretization implies that only the end-point cost function is modified, replacing $U(\varphi) = F(\varphi \cdot I_0)$ by an approximation taking the form $U^{(n)}(\varphi) = F^{(n)}(\varphi, I_0^{(n)})$. For example, when matching curves, one may replace the objective function $F(\varphi \cdot I_0) = \|\mu_{\varphi \cdot I_0} - \mu_{I'}\|^2_{W^*}$ in (9.40) by the discrete approximation in (9.46), in which the curves I_0 and I' are approximated by point sets. Similar approximations can be made for the other types of cost functions

discussed for curves and surfaces. In such cases, the following proposition can be applied to compare solutions of the original problem with their discrete approximations.

Proposition 10.16 *Assume that V is continuously embedded in $C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$. Consider a family of optimal control problems minimizing*

$$E^{(n)}(v) = \frac{1}{2}\int_0^1 \|v\|_V^2\, dt + U^{(n)}(\varphi_{01}^v), \tag{10.48}$$

with $U^{(n)}$ continuous for the (p,∞)-compact topology. Let U be continuous with respect to the same topology and assume that, for some $p > 0$, the following uniform convergence is true: for all $A > 0$ and $\varepsilon > 0$, there exists an n_0 such that, for all $n \geq n_0$, for all $\varphi \in \mathrm{Diff}_0^{p,\infty}$ such that $\max(\|\varphi\|_{p,\infty}, \|\varphi^{-1}\|_{p,\infty}) < A$, one has $|U^{(n)}(\varphi) - U(\varphi)| < \varepsilon$.

Then, given a sequence $v^{(n)}$ of minimizers of (10.48), one can extract a subsequence $v^{(n_k)}$ that weakly converges to v in $\mathcal{X}_V^2$, with v minimizing

$$E(w) = \frac{1}{2}\int_0^1 \|w\|_V^2\, dt + U(\varphi_{01}^w). \tag{10.49}$$

Proof Let w be a minimizer of (10.49). Our assumptions implying that $U^{(n)}(\varphi_{01}^w)$ converges to $U(\varphi_{01}^w)$ (so that their difference is bounded), we see that $E^{(n)}(w) \leq E(w) + C$ for some constant C, so that, letting $v^{(n)}$ be a minimizer of $E^{(n)}$, we have $\|v^{(n)}\|_{\mathcal{X}_V^2}^2 \leq 2E^{(n)}(v^{(n)}) \leq 2E(w) + 2C$. From this we find that $v^{(n)}$ is a bounded sequence in $\mathcal{X}_V^2$, so that, replacing it with a subsequence if needed, we can assume that it weakly converges to some $v \in \mathcal{X}_V^2$. Applying Theorem 7.13, we find that $\varphi_{01}^{v^{(n)}}$ converges to φ^v in the (p,∞)-compact topology. Moreover, Theorem 7.10 implies that the sequences $(\|\varphi_{01}^{v^{(n)}}\|_{p,\infty}, \|\varphi_{10}^{v^{(n)}}\|_{p,\infty})$ are bounded. Applying the uniform convergence of $U^{(n)}$ to U on bounded sets and the continuity of U, we see that $U^{(n)}(\varphi_{01}^{v^{(n)}})$ converges to $U(\varphi_{01}^v)$ as n tends to infinity. Since, in addition

$$\|v\|_{\mathcal{X}_V^2} \leq \liminf \|v^{(n)}\|_{\mathcal{X}_V^2}$$

we obtain the fact that $E(v) \leq \liminf E^{(n)}(v^{(n)})$. We also have

$$E^{(n)}(v^{(n)}) \leq E^{(n)}(w) = E(w) + U(\varphi_{01}^w) - U^{(n)}(\varphi_{01}^w) \to E(w),$$

so that $E(v) = E(w)$ and v is also a minimizer of (10.49). □

Curves and Surfaces. We can apply this theorem to curve and surface matching according to the following discussion, in which we focus on surface matching using currents, but which can, with very little modification, be applied to curves, and to measure or varifold matching terms. Let Σ and $\tilde{\Sigma}$ be regular surfaces and $S^{(n)}$, $\tilde{S}^{(n)}$ be

sequences of triangulated surfaces that converge to them as defined before Theorem 4.3. Let (fixing an RKHS W with kernel ξ)

$$U(\varphi) = \|\nu_{\varphi\cdot\Sigma} - \nu_{\tilde{\Sigma}}\|^2_{W^*},$$

using the vector measures defined in Eq. (9.49), and

$$U^{(n)}(\varphi) = \|\nu_{\varphi\cdot S^{(n)}} - \nu_{\tilde{S}^{(n)}}\|^2_{W^*},$$

using the discrete version as in (9.56). Then, Theorem 4.3, slightly modified to account for double integrals, can be used to check that the assumptions of Proposition 10.16 are satisfied.

Images. The image matching problem can be discretized using finite grids, assuming that the considered images are supported by the interval $[0, 1]^d$. Consider the cost function

$$U(\varphi) = \|I \circ \varphi^{-1} - \tilde{I}\|^2_2,$$

in which we assume, to simplify, that I and $\tilde{I}$ are compactly supported (say, on $\mathcal{K} = [-M, M]^d$) and bounded. We first start with a discretization that can be applied to general L^2 functions. Let $\mathcal{G}_n = \{-M + 2^{-n+1}kM, k = 0, \ldots, 2^n\}^d$ provide a discrete grid on $\mathcal{K}$ and associate to each point $z \in \mathcal{G}_n$ its Voronoï cell, $\Gamma_n(z)$, provided by the set of points in $\mathcal{K}$ that are closer to x than to any other point in the grid (i.e., $\Gamma_n(z)$ is the intersection of $\mathcal{K}$ and the cube of size 2^{-n} centered at x). Define

$$I^{(n)}(x) = \sum_{z\in\mathcal{G}_n} \bar{I}^{(n)}(z)\mathbf{1}_{\Gamma_n(z)}(x),$$

where

$$\bar{I}^{(n)}(z) = \frac{1}{|\Gamma_n(z)|}\int_{\Gamma_n(z)} I(x)\,dx$$

is the average value of I over $\Gamma_n(z)$.

Define $\tilde{I}^{(n)}$ similarly and consider the approximation of U given by $U^{(n)}(\varphi) = \|I^{(n)} \circ \varphi^{-1} - \tilde{I}^{(n)}\|^2_2$. Then $U^{(n)}$ and U satisfy the hypotheses of Proposition 10.16.

Indeed, assume that $\max(\|\varphi\|_{1,\infty}, \|\varphi^{-1}\|_{1,\infty}) < A$. We have

$$\begin{aligned}|U^{(n)}(\varphi) - U(\varphi)| &\le 2\|I \circ \varphi^{-1} - I^{(n)} \circ \varphi^{-1}\|^2_2 + 2\|\tilde{I} - \tilde{I}^{(n)}\|^2_2\\ &\le 2C(A)\|I - I^{(n)}\|^2_2 + 2\|\tilde{I} - \tilde{I}^{(n)}\|^2_2,\end{aligned}$$

where the second inequality is obtained after a change of variable in the first L^2 norm and $C(A)$ is an upper bound for the Jacobian determinant of φ depending only on A. As a consequence, Proposition 10.12 will be true as soon as one shows that $I^{(n)}$ and $\tilde{I}^{(n)}$ converge in L^2 to I and $\tilde{I}$ respectively (and will also be true for any sequence of approximations of I and $\tilde{I}$ that satisfies this property). The L^2 convergence is

true in our case because $I^{(n)}$ is the orthogonal projection of I on the space W_n of L^2 functions that are constant on each set $\Gamma_n(z)$, $z \in \mathcal{G}_n$. This implies that $I^{(n)}$ converges in L^2 to the projection of I on $W_\infty = \overline{\bigcup_{n\geq 1} W_n}$ (see Proposition A.11), but one has $W_\infty = L^2$, because any function J orthogonal to this space would have its integral vanish on any dyadic cube, which is only possible for $J = 0$.

Note that, with this approximation, one can write

$$\|I^{(n)} \circ \varphi^{-1} - \tilde{I}\|_2^2 = \sum_{z\in\mathcal{G}_n} I(z)^2 |\varphi(\Gamma_n(z))| + \sum_{z\in\mathcal{G}_n} \tilde{I}(z)^2 |\Gamma_n(z)| - 2 \sum_{z,z'\in\mathcal{G}_n} I(z)\tilde{I}(z')|\varphi(\Gamma_n(z)) \cap \Gamma_n(z')|,$$

where $|A|$ denotes the volume of $A \subset \mathbb{R}^d$. To make this expression computable, one needs to approximate the sets $\varphi(\Gamma_n(z))$, where the simplest approximation is to take the polyhedron formed by the image of the vertices of $\Gamma_n(z)$ by φ (which will retain the same topology as the original cube is n is large enough). The verification that this approximation is valid (in the sense of Proposition 10.16) is left to the reader.

However, even with this approximation, the numerical problem is still highly computational, since it becomes a point set problem over $\mathcal{G}_n$, which is typically a huge set. Most current implementations use a simpler scheme, in which $I^{(n)}$ is interpolated between the values $(I(z), z \in \mathcal{G}_n)$, who are therefore assumed to be well defined, and the cost function is simply approximated by

$$U^{(n)}(\varphi) = \sum_{z\in\mathcal{G}_n} (I(\varphi^{-1}(z)) - \tilde{I}(z))^2 |\Gamma_n(z)|.$$

Here again, we leave to the reader to check that this provides a valid approximation in the sense of Proposition 10.16 as soon as, say, I and $\tilde{I}$ are continuous and one uses a linear interpolation scheme, as described below.

Using this approximation (for a fixed n that we will remove from the notation), we now work the implementation in more detail, starting with the computation of the gradient in (10.45). Assume that time is discretized at $t_k = kh$ for $h = 1/Q$ and that $v_k(\cdot) = v(t_k, \cdot)$ is discretized over a regular grid $\mathcal{G}$.

It will be convenient to introduce the momentum and express v_k in the form

$$v_k(y) = \sum_{z\in\mathcal{G}} K(y, z)\rho_k(z). \tag{10.50}$$

We can consider $(\rho_k(z), z \in \mathcal{G})$ as new control variables, noting that (10.45) directly provides the gradient of the energy in V^*, namely

$$(\nabla^{V^*} E)(t, y) = 2\rho(t) - 2\det(d\varphi_{t1}^v)(I \circ \varphi_{t0}^v - \tilde{I} \circ \varphi_{t1}^v)\nabla(I \circ \varphi_{t0}^v)dx.$$

From this expression, we see that we can interpret the family $(\rho_k(z), z \in \mathcal{G})$ as discretizing a measure, namely

$$\rho_k = \sum_{z \in \mathcal{G}} \rho_k(z) \delta_z.$$

Given this, the gradient in V^* can be discretized as

$$\xi_k(z) = 2\rho_k(z) - 2\det(d\varphi^v_{t_k 1}(z))(I \circ \varphi^v_{t_k 0}(z) - \tilde{I} \circ \varphi^v_{t_k 1}(z))\nabla(I \circ \varphi^v_{t_k 0}(z))\delta_z,$$

which can be used to update $\rho_k(z)$.

The last requirement in order to obtain a fully discrete procedure is to select interpolation schemes for the computation of the diffeomorphisms φ^v and for the compositions of I and I' with them. Interpolation algorithms (linear, or cubic, for example) are standard procedures that are included in many software packages [234]. In mathematical representation, they are linear operators that take a discrete signal f on a grid $\mathcal{G}$ (i.e., $f \in \mathbb{R}^{\mathcal{G}}$) and return a function, that we will denote by $\mathcal{R}f$, defined everywhere. By linearity, we must have

$$(\mathcal{R}f)(z) = \sum_{z \in \mathcal{G}} r_z(x) f(z)$$

for some "interpolants" $r_z(\cdot), z \in \mathcal{G}$. In the approximation of the data attachment term, one can then replace I by $\mathcal{R}(I_{|_{\mathcal{G}}})$, the interpolation of the restriction of I to $\mathcal{G}$.

Linear interpolation, for example, corresponds, in one dimension, to $r_z(x) = 1 - 2^n|z - x|$ if $|z - x| < 2^{-n}$ and 0 otherwise. In dimension d, one takes

$$r_z(x) = \prod_{i=1}^{d} (1 - 2^n|z_i - x_i|)$$

if $\max_i(|z_i - x_i|) < 2^{-n}$ and 0 otherwise (where $z = (z_1, \ldots, z_d)$ and $x = (x_1, \ldots, x_d)$).

Given an interpolation operator $\mathcal{R}$, one can replace, say, $I \circ \varphi_{t_k 0}(z)$ in the expression of the gradient by

$$(\mathcal{R}I)(\varphi_{t_k 0}(z)) = \sum_{z' \in \mathcal{G}} r_{z'}(\varphi_{t_k 0}(z)) I(z').$$

For computational purposes, it is also convenient to replace the definition of v_k in (10.50) by an interpolated form

$$v_k(x) = \sum_{z \in \mathcal{G}} r_z(x) \sum_{i \in \mathcal{G}} K(z, z') \rho_k(z') \tag{10.51}$$

because the inner sum can be computed very efficiently using Fourier transforms (see the next section).

To complete the discretization, introduce

$$\psi_{lk} = (\mathrm{id} - hv_l) \circ \cdots \circ (\mathrm{id} - hv_{k-1}),$$

where an empty product of compositions is equal to the identity, so that ψ_{lk} is an approximation of $\varphi_{t_k t_l}$. Define the cost function, which is explicitly computable as a function of $\rho_0, \ldots, \rho_{Q-1}$:

$$E(\rho) = \sum_{k=0}^{Q-1} \sum_{z,z' \in \mathcal{G}} K(z,z') \rho_k(z)^T \rho_k(z') + \sum_{z \in \mathcal{G}} ((\mathcal{R}I)(\psi_{0Q}(z)) - \tilde{I}(z))^2.$$

If we make a variation $\rho \mapsto \rho + \varepsilon\delta\rho$, then $v \mapsto v + \varepsilon\delta v$ with (using the interpolated expression of v)

$$\delta v_k(y) = \sum_{z \in \mathcal{G}} r_z(y) \sum_{z' \in \mathcal{G}} K(z,z') \delta\rho_k(z')$$

and letting $\delta\psi_{lk} = \partial_\varepsilon \psi_{lk}$, we have, by direct computation

$$\delta\psi_{lk} = -h \sum_{q=l}^{k-1} d\psi_{lq} \circ \psi_{qk}\, \delta v_q \circ \psi_{q+1k}.$$

Using this, we can compute the variation of the E, yielding

$$\begin{aligned}(\partial_\varepsilon E \mid \delta\rho) = 2 &\sum_{k=0}^{Q-1} \sum_{z,z',\in \mathcal{G}} K(z,z')\, \rho_k(z')^T \delta\rho_k(z) \\ - 2h &\sum_{k=0}^{Q-1} \sum_{z,z',y \in \mathcal{G}} K(z,z')\, r_z(\psi_{k+1\,Q}(y))\, ((\mathcal{R}I)(\psi_{0Q}(y)) - \tilde{I}(y)) \\ &\nabla(\mathcal{R}I)(\psi_{0Q}(y))^T (d\psi_{0k} \circ \psi_{k\,Q}(y)\, \delta\rho_k(z'))\end{aligned}$$

This provides the expression of the gradient of the discretized E in V^*, namely

$$\begin{aligned}&(\nabla^{V^*} E(\rho))_k(z) = 2\rho_k(z) \\ &- 2h \sum_{z' \in \mathcal{G}} r_z(\psi_{k+1\,Q}(z'))((\mathcal{R}I)(\psi_{0Q}(z')) - \tilde{I}(z'))\nabla(\mathcal{R}I \circ \psi_{0k})(\psi_{kQ}(z')).\end{aligned}$$

10.7.2 Kernel-Related Numerics

Most of the previously discussed methods included repeated computations of linear combination of the kernel. A basic such step is to compute, given points $y_1, \ldots, y_M$, $x_1, \ldots, x_N$ and vectors (or scalars) $\alpha_1, \ldots, \alpha_N$, the sums

$$\sum_{k=1}^{N} K(y_j, x_k)\alpha_k, \quad j = 1, \ldots, M.$$

Such sums are involved when deriving velocities from momenta, for example, or when evaluating dual RKHS norms in curve or surface matching.

Computing these sums explicitly requires NM evaluations of the kernel (and this probably several times per iteration of an optimization algorithm). When N or M are reasonably small (say, less than 1,000), such a direct evaluation is not a problem. But for large-scale methods, such as triangulated surface matching, where the surface may have tens of thousands of nodes, or image matching, where a three-dimensional grid typically has millions of nodes, this becomes unfeasible (the feasibility limit has however been pushed further by recent efficient implementations on GPUs [59, 157, 247]).

If $x = y$ is supported by a regular grid $\mathcal{G}$, and K is translation invariant, i.e., $K(x, y) = \Gamma(x - y)$, then, letting $x_k = hk$ where k is a multi-index ($k = (k_1, \ldots, k_d)$) and h the discretization step, we see that

$$\sum_{k \in \mathcal{G}} \Gamma(h(k - l))\alpha_l$$

is a convolution that can be implemented with $O(N \log N)$ operations, using fast Fourier transforms (with $N = |\mathcal{G}|$). The same conclusion holds if K takes the form $K(x, y) = A(x)^T \Gamma(x - y)A(y)$ for some matrix A (which can be used to censor the kernel at the boundary of a domain), since the resulting operation is

$$A(x_k)^T \left(\sum_{k \in \mathcal{G}} \Gamma(h(k - l))(A(x_l)\alpha_l) \right),$$

which can still be implemented in $O(N \log N)$ operations.

The situation is less favorable when x and y are not regularly spaced. In such cases, feasibility must come with some approximation.

Still assuming a translation-invariant kernel $K(x, y) = \Gamma(x - y)$, we can associate to a grid $\mathcal{G}$ in $\mathbb{R}^d$ the interpolated kernel

$$K_{\mathcal{G}}(x, y) = \sum_{j, j' \in \mathcal{G}} r_z(x)\Gamma(h(z - z'))r_{z'}(y),$$

where the r_z's are interpolants adapted to the grid. This approximation provides a non-negative kernel, with null space equal to the space of functions with vanishing interpolation on $\mathcal{G}$. With such a kernel, we have

$$\sum_{k=1}^{N} K(y_j, x_k)\alpha_k = \sum_{z\in\mathcal{G}} r_z(y_j) \sum_{z'\in\mathcal{G}} \Gamma(h(z-z')) \sum_{k=1}^{N} r_{z'}(x_k)\alpha_k.$$

The computation of this expression therefore requires using the following sequence of operations:

1. Compute, for all $z' \in \mathcal{G}$, the quantity

$$a_{z'} = \sum_{k=1}^{N} r_{z'}(x_k)\alpha_k.$$

   Because, for each x_k, only a fixed number of $r_{z'}(x_k)$ are non-vanishing, this requires an $O(N)$ number of operations.
2. Compute, for all $z \in \mathcal{G}$,

$$b_z = \sum_{z'\in\mathcal{G}} \Gamma(h(z-z'))a_{z'},$$

   which is a convolution requiring $O(|\mathcal{G}| \log |\mathcal{G}|)$ operations.
3. Compute, for all $j = 1, \ldots, M$, the interpolation

$$\sum_{z\in\mathcal{G}} r_z(y_j)b_z,$$

   which requires $O(M)$ operations.

So the resulting cost is $O(M + N + |\mathcal{G}| \log |\mathcal{G}|)$, which must be compared to the original $O(MN)$, the comparison being favorable essentially when MN is larger than the number of nodes in the grid, $|\mathcal{G}|$. This formulation (which has been proposed in [156]) has the advantage that the resulting algorithm is quite simple, and that the resulting $K_{\mathcal{G}}$ remains a non-negative kernel, which is important.

Another class of methods, called "fast multipole", computes sums such as

$$\sum_{k=1}^{N} K(y, x_k)\alpha_k$$

by taking advantage of the fact that $K(y, x)$ varies slowly as x varies in a region which is far away from y. By grouping the x_k's in clusters, assigning centers to these clusters and approximating the kernel using asymptotic expansions valid at a large enough distance from the clusters, fast multipole methods can organize the computation of the sums with a resulting cost of order $M + N$ when M sums over

N terms are computed. Even if it is smaller than a constant times $(M + N)$, the total number of operations increases (via the size of the constant) with the required accuracy. The interested reader may refer to [30, 140] for more details.

Another important operation involving the kernel is the inversion of the system of equations (say, with a scalar kernel)

$$\sum_{k=1}^{N} K(x_k, x_l)\alpha_l = u_k, \quad k = 1, \ldots, N. \tag{10.52}$$

This is the spline interpolation problem, but it is also part of several of the algorithms that we have discussed, including for example the projection steps that have been introduced to obtain a gradient in the correct metric.

Such a problem is governed by an uncertainty principle [258] between accuracy of the approximation, which is given by the distance between a smooth function $x \mapsto u(x)$ and its interpolation

$$x \mapsto \sum_{k=1}^{N} K(x, x_l)\alpha_l,$$

where $\alpha_1, \ldots, \alpha_N$ are given by (10.52) with $u_k = u(x_k)$, and the stability of the system (10.52) measured by the condition number (the ratio of the largest to the smallest eigenvalue) of the matrix $S(x) = (K(x_i, x_j), i, j = 1, \ldots, N)$, evaluated as a function of the smallest distance between two distinct x_k's ($S(x)$ is singular if two x_k's coincide).

When $K(x, y) = \Gamma(x - y)$, the trade-off is measured by how fast $\xi \mapsto \hat{\Gamma}(\xi)$ (the Fourier transform of Γ) decreases at infinity. One extreme is given by the Gaussian kernel, for which $\hat{\Gamma}$ decreases like $e^{-c|\xi|^2}$, which is highly accurate and highly unstable. On the other side of the range are Laplacian kernels, which decrease polynomially in the Fourier domain. In this dilemma, one possible rule is to prefer accuracy for small values of N, therefore using a kernel like the Gaussian, and go for stability for large-scale problems (using a Laplacian kernel with high enough degree).

For the numerical inversion of system (10.52), iterative methods, such as conjugate gradient, should be used (especially for large N). Methods using preconditioned conjugate gradient have been introduced, for example, in [105, 141] and the interested reader may refer to these references for more details.

Chapter 11
Distances and Group Actions

11.1 General Principles

In this chapter we discuss metric comparisons between deformable objects and their relation to the registration methods that we have studied in the previous chapters. We start with a general discussion on the interplay between distances on a set and transformation groups acting on it.

11.1.1 Distance Induced by a Group Action

Transformation groups acting on sets can help in defining or altering distances on these sets. We will first give a generic construction, based on a least action principle. We will then develop the related differential point of view, when a Lie group acts on a manifold.

A distance on a set M is a mapping $d : M^2 \mapsto [0, +\infty)$ such that: for all $m, m', m'' \in M$,

$$\begin{aligned} &\text{D1.} \quad d(m, m') = 0 \Leftrightarrow m = m', \\ &\text{D2.} \quad d(m.m') = d(m', m), \\ &\text{D3.} \quad d(m, m'') \leq d(m, m') + d(m', m''). \end{aligned}$$

If D1 is not satisfied, but only the fact that $d(m, m) = 0$ for all m, one says (still assuming D2 and D3) that d is a *pseudo-distance*.

If G is a group acting on M, we will say that a distance d on M is G-equivariant if and only if for all $g \in G$, for all $m, m' \in M$, $d(g \cdot m, g \cdot m') = d(m, m')$. A mapping $d : M^2 \mapsto \mathbb{R}_+$ is a G-invariant distance if and only if it is a pseudo-distance such that $d(m, m') = 0 \Leftrightarrow \exists g \in G,\ g \cdot m = m'$. This is equivalent to stating that d is a distance on the coset space M/G, composed of cosets, or orbits,

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_11

$$[m] = \{g \cdot m, g \in G\},$$

with the identification $d([m], [m']) = d(m, m')$. The next proposition shows how a G-equivariant distance can induce a G-invariant pseudo-distance.

Proposition 11.1 *Let d be equivariant under the left action of G on M. The function $\tilde{d}$, defined by*

$$\tilde{d}([m], [m']) = \inf\{d(g \cdot m, g' \cdot m') : g, g' \in G\}$$

is a pseudo-distance on M/G.

If, in addition, the orbits $[m]$ are closed subsets of M (in the topology associated to d), then $\tilde{d}$ is a distance.

Note that, because d is G-equivariant, $\tilde{d}$ in the previous proposition is also given by

$$\tilde{d}([m], [m']) = \inf\{d(g \cdot m, m') : g \in G\}.$$

Proof The symmetry of $\tilde{d}$ is obvious, as is the fact that $\tilde{d}((m], [m]) = 0$ for all m. For the triangle inequality, D3, it suffices to show that, for all $g_1, g_1', g_2', g_1'' \in G$, there exists $g_2, g_2'' \in G$ such that

$$d(g_2 \cdot m, g_2'' \cdot m'') \leq d(g_1 \cdot m, g_1' \cdot m') + d(g_2' \cdot m', g_1'' \cdot m''). \tag{11.1}$$

Indeed, if this is true, the minimum of the right-hand term in g_1, g_1', g_2', g_1'', which is $\tilde{d}([m], [m']) + \tilde{d}([m'], [m''])$, is larger than the minimum of the left-hand term in g_2, g_2'', which is $\tilde{d}([m], [m''])$.

To prove (11.1), write $d(g_2' \cdot m', g_1'' \cdot m'') = d(g_1' \cdot m', g_1'(g_2')^{-1} g_1'' \cdot m'')$, take $g_2 = g_1$ and $g_2'' = g_1'(g_2')^{-1} g_1''$; (11.1) is then a consequence of the triangle inequality for d.

We now make the additional assumption that the orbits are closed and prove that D1 is true. Take $m, m' \in M$ such that $\tilde{d}([m], [m']) = 0$. This implies that there exists a sequence $(g_n, n \geq 0)$ in G such that $d(g_n \cdot m, m') \to 0$ when $n \to \infty$, so that m' belongs to the closure of the orbit of m. Since the latter is assumed to be closed, this yields $m' \in [m]$, which is equivalent to $[m] = [m']$. □

The same statement can clearly be made with G acting on the right on m, writing $m \mapsto m \cdot g$. We state it without proof.

Proposition 11.2 *Let d be equivariant under the right action of G on M. The function $\tilde{d}$, defined by*

$$\tilde{d}([m], [m']) = \inf\{d(m \cdot g, m' \cdot g') : g, g' \in G\}$$

is a pseudo-distance on $G\backslash M$.

If, in addition, the orbits $[m]$ are closed subsets of M (in the topology associated to d), then $\tilde{d}$ is a distance.

Here $G\backslash M$ denotes the coset space for the right action of G.

11.1.2 Distance Altered by a Group Action

In this section, G is still a group acting on the left on M, but we consider the product space $\mathcal{M} = G \times M$ and project on M a distance defined on $\mathcal{M}$. The result of this analysis will be to allow a distance on $\mathcal{M}$ to incorporate a component that accounts for possible group transformations partially accounting for the difference between the compared objects.

The left action of G on M induces a right action of G on $\mathcal{M}$, defined, for $k \in G$, $z = (h, m) \in \mathcal{M}$, by

$$z \cdot k = (hk, k^{-1} \cdot m).$$

For $z = (h, m) \in \mathcal{M}$, we define the projection $\pi(z) = h \cdot m$, taking values in M. This projection is constant on the orbits $z \cdot G$ for a given z, i.e., for all $k \in G$, $\pi(z \cdot k) = \pi(z)$.

Let $d_{\mathcal{M}}$ be a distance on $\mathcal{M}$. We let, for $m, m' \in M$

$$d(m, m') = \inf\{d_{\mathcal{M}}(z, z') : z, z' \in \mathcal{M}, \pi(z) = m, \pi(z') = m'\}. \tag{11.2}$$

We have the following proposition:

Proposition 11.3 *If $d_{\mathcal{M}}$ is equivariant by the right action of G, then, the function d defined by (11.2) is a pseudo-distance on M.*

If, in addition, the orbits $[z] = \{z \cdot k, k \in G\}$ are closed in $\mathcal{M}$ in the topology associated to $d_{\mathcal{M}}$, then d is a distance.

This is in fact a corollary of Proposition 11.2. One only has to observe that the quotient space $G\backslash\mathcal{M}$ can be identified with M via the projection π, and that the distance in (11.2) then becomes the projection distance introduced in Proposition 11.2.

11.1.3 Transitive Action

Induced Distance

In this section, we assume that $\mathcal{G}$ is a group that acts transitively on M. The action being transitive means that for any m, m' in M, there exists an element $z \in \mathcal{G}$ such that $m' = z \cdot m$.

We fix a reference element m_0 in M, and define the group G by

$$G = \mathrm{Iso}_{m_0}(\mathcal{G}) = \{z \in \mathcal{G}, z \cdot m_0 = m_0\}.$$

This group is the isotropy group, or stabilizer, of m_0 in $\mathcal{O}$. We show that $\mathcal{G}$ can be identified with $\mathcal{M} := G \times M$, which will allow us to define a distance in M by projecting a distance on $\mathcal{G}$ as in Sect. 11.1.2.

Assume that a function $\rho : M \to \mathcal{G}$ has been defined, such that for all $m \in M$, $m = \rho(m) \cdot m_0$. This is possible, because the action is transitive (using the axiom of choice). Define

$$\begin{aligned} \Psi : \quad & G \times M \to \mathcal{G} \\ & (h, m) \mapsto \rho(h \cdot m)h. \end{aligned}$$

Ψ is a bijection: if $z \in \mathcal{G}$, we can compute a unique (h, m) such that $z = \Psi(h, m)$; this (h, m) must satisfy

$$z \cdot m_0 = \rho(h \cdot m)h \cdot m_0 = \rho(h \cdot m) \cdot m_0 = h \cdot m,$$

which implies that $\rho(h \cdot m) = \rho(z \cdot m_0)$ and therefore $h = \rho(z \cdot m_0)^{-1}z$, which is uniquely specified; but this also specifies $m = h^{-1}z \cdot m_0$. This proves that Ψ is one-to-one and onto and provides the identification we were looking for.

The right action of G on $\mathcal{M}$, which is $(h, m) \cdot k = (hk, k^{-1} \cdot m)$, translates to $\mathcal{G}$ via Ψ with

$$\Psi((h, m) \cdot k) = \rho(hkk^{-1} \cdot m)hk = \Psi(h, m) \cdot k$$

so that the right actions (of G on $\mathcal{M}$ and of G on $\mathcal{G}$) "commute" with Ψ. Finally, the constraint $\pi(h, m_1) = m$ in Proposition 11.3 becomes $z \cdot m_0 = m$ via the identification. All this provides a new version of Proposition 11.3 for transitive actions, given by:

Corollary 11.4 *Let $d_{\mathcal{G}}$ be a distance on $\mathcal{G}$ which is equivariant under the right action of the isotropy group of $m_0 \in M$. Define, for all $m, m' \in M$,*

$$d(m, m') = \inf\{d_{\mathcal{G}}(z, z') : z \cdot m_0 = m, z' \cdot m_0 = m'\}. \tag{11.3}$$

Then d is a pseudo-distance on M.

Note that, if $d_{\mathcal{G}}$ is right equivariant under the action of $\mathrm{Iso}_{m_0}(\mathcal{G})$, the distance

$$\tilde{d}_{\mathcal{G}}(z, z') = d_{\mathcal{G}}(z^{-1}, (z')^{-1})$$

is left equivariant, which yields the symmetric version of the previous corollary.

Corollary 11.5 *Let $d_{\mathcal{G}}$ be a distance on $\mathcal{G}$ which is equivariant under the left action of the isotropy group of $m_0 \in M$. Define, for all $m, m' \in M$,*

$$d(m, m') = \inf\{d_{\mathcal{G}}(z, z') : z \cdot m = m_0, z' \cdot m' = m_0\}. \tag{11.4}$$

Then d is a pseudo-distance on M.

From Propositions 11.1 and 11.2, d in Corollaries 11.4 and 11.5 is a distance as soon as the orbits $g \cdot \mathrm{Iso}_{m_0}(\mathcal{G})$ (assuming, for example, a left action) are closed for $d_{\mathcal{G}}$. If the left translations $h \mapsto g \cdot h$ are continuous, this is true as soon as $\mathrm{Iso}_{m_0}(\mathcal{G})$

is closed. This last property is itself true as soon as the action $g \mapsto g \cdot m_0$ is continuous, from G to M, given some topology on M.

Finally, if $d_{\mathcal{G}}$ is left- or right-invariant under the action of the whole group, $\mathcal{G}$, on itself, then the distances in (11.3) and (11.4) both reduce to

$$d(m, m') = \inf\{d_{\mathcal{G}}(\mathrm{id}, z) : z \cdot m = m'\}.$$

Indeed, assume right invariance (the left-invariant case is similar): then, if $z \cdot m_0 = m$ and $z' \cdot m_0 = m'$, then $z'z^{-1} \cdot m = m'$ and $d_{\mathcal{G}}(\mathrm{id}, z'z^{-1}) = d_{\mathcal{G}}(z, z')$. Conversely, assume that $\zeta \cdot m = m'$. Since the action is transitive, we know that there exists a z such that $z \cdot m_0 = m$, in which case $\zeta z \cdot m_0 = m'$ and $d_{\mathcal{G}}(\mathrm{id}, \zeta) = d_{\mathcal{G}}(z, \zeta z)$. We summarize this in the following, in which we take $\mathcal{G} = G$:

Corollary 11.6 *Assume that G acts transitively on M. Let d_G be a distance on G that is left or right equivariant. Define, for all $m, m' \in M$,*

$$d(m, m') = \inf\{d_G(\mathrm{id}, g) : g \cdot m = m'\}. \tag{11.5}$$

Then d is a pseudo-distance on M.

Effort Functionals

As formalized in [135], one can build a distance on M on which a group acts transitively using the notion of effort functionals. The definition we give here is slightly more general than in [135], to take into account a possible influence of the deformed object on the effort. We also make a connection with the previous, distance based, formulations.

We let $\mathcal{G}$ be a group acting transitively on M. Assume that a cost $\Gamma(z, m)$ is assigned to a transformation $m \to z \cdot m$. If m and m' are two objects, we define $d(m, m')$ as the minimal cost (effort) required to transform m to m', i.e.,

$$d(m, m') = \inf\{\Gamma(z, m) : z \in \mathcal{G}, z \cdot m = m'\}. \tag{11.6}$$

The proof of the following proposition is almost obvious.

Proposition 11.7 *If Γ satisfies:*

C1. $\Gamma(z, m) = 0 \Leftrightarrow z = \mathrm{id}_{\mathcal{G}}$,
C2. $\Gamma(z, m) = \Gamma(z^{-1}, z \cdot m)$,
C3. $\Gamma(zz', m) \leq \Gamma(z, m) + \Gamma(z', m)$,

then d defined by (11.6) is a pseudo-distance on M.

In fact, this is equivalent to the construction of Corollary 11.5. To see this, let G be the isotropy group of m_0 for the action of $\mathcal{G}$ on M. We have the following proposition.

Proposition 11.8 *If Γ satisfies C1, C2 and C3, then, for all $m_0 \in M$, the function $d_{\mathcal{G}}$ defined by*

$$d_{\mathcal{G}}(z, z') = \Gamma(z'z^{-1}, z \cdot m_0) \tag{11.7}$$

is a distance on $\mathcal{G}$ which is equivariant under the right action of G. Conversely, given such a distance $d_{\mathcal{G}}$, one builds an effort functional Γ satisfying C1, C2, C3 letting

$$\Gamma(h, m) = d_{\mathcal{G}}(z, h \cdot z)$$

where z is any element of $\mathcal{G}$ with the property $z \cdot m = m_0$.

The proof of this proposition is straightforward and left to the reader.

11.1.4 The Riemannian Viewpoint

The previous sections have demonstrated the usefulness of building distances on a space $\mathcal{M}$ that are equivariant to the actions of a group G. Probably the easiest way to construct such a distance (at least when $\mathcal{M}$ is a differential manifold and G is a Lie group) is to design a right-invariant Riemannian metric on $\mathcal{M}$ and use the associated geodesic distance. (See Appendix B.)

Recall that a Riemannian metric on $\mathcal{M}$ requires, for all $z \in \mathcal{M}$, an inner product $\langle \cdot\,, \cdot \rangle_z$ on the tangent space, $T_z\mathcal{M}$, to $\mathcal{M}$ at z, which depends smoothly on z. With such a metric, one defines the energy of a differentiable path $z(\cdot)$ in $\mathcal{M}$ by

$$E(z(\cdot)) = \int_0^1 \|\partial_t z\|^2_{z(t)}\, dt. \tag{11.8}$$

The associated Riemannian distance on $\mathcal{M}$ is

$$d_{\mathcal{M}}(z_0, z_1) = \inf\{\sqrt{E(z(\cdot))} : z(0) = z_0, z(1) = z_1\}. \tag{11.9}$$

To obtain a right-invariant distance, it suffices to ensure that the metric has this property. For $h \in G$, let R_h denote the right action of h on $\mathcal{M}$: $R_h : z \mapsto z \cdot h$. Let $dR_h(z) : T_z\mathcal{M} \to T_{z\cdot h}\mathcal{M}$ be its differential at $z \in \mathcal{M}$. The right invariance of the metric is expressed by the identity, true for all $z \in \mathcal{M}$, $A \in T_z\mathcal{M}$ and $h \in G$,

$$\|A\|_z = \|dR_h(z) \cdot A\|_{z\cdot h}\,. \tag{11.10}$$

When $\mathcal{M} = G \times M$, condition (11.10) implies that it suffices to define $\langle \cdot\,, \cdot \rangle_z$ at elements $z \in \mathcal{M}$ of the form $z = (\mathrm{id}, m)$ with $m \in M$. The metric at a generic point (h, m) can then be computed, by right invariance, from the metric at $(h, m) \cdot h^{-1} = (\mathrm{id}, h^{-1} \cdot m)$. Because the metric at (id, m) can be interpreted as a way to attribute a cost to a deformation $(\mathrm{id}, h(t) \cdot m)$ with $h(0) = \mathrm{id}$ and small t, defining it corresponds to an analysis of the cost of an infinitesimal perturbation of m by elements of G.

Of course, an identical construction could be made with left actions and left-invariant distances.

11.2 Invariant Distances Between Point Sets

11.2.1 Introduction

The purpose of this section is to present the construction provided by Kendall [166] on distances between landmarks, taking the infinitesimal point of view that we have just outlined. Here, configurations of landmarks are considered up to similitude (translation, rotation, scaling). Since its introduction, this space has led to a rich literature that specially focuses on statistical data analysis on landmark data. The reader interested in further developments can refer to [91, 167, 266] and to the references therein.

We only consider the two-dimensional case, which is also the simplest. For a fixed integer $N > 0$ let $\mathcal{P}_N$ denote the set of configurations of N points $(z^{(1)}, \dots, z^{(N)}) \in (\mathbb{R}^2)^N$ such that $z^{(i)} \neq z^{(j)}$ for $i \neq j$. We assume that the order in which the points are listed matters, which means that we consider labeled landmarks. The set $\mathcal{P}_N$ can therefore be identified with an open subset of $\mathbb{R}^{2N}$.

Two configurations $(z^{(1)}, \dots, z^{(N)})$ and $(\tilde{z}^{(1)}, \dots, \tilde{z}^{(N)})$ will be identified if one can be deduced from the other by the composition, say g, of a translation and a plane similitude, i.e., $\tilde{z}^{(k)} = g \cdot z^{(k)}$ for $k = 1, \dots, N$. The objects of interest are therefore equivalence classes of landmark configurations, which will be referred to as N-shapes.

It will be convenient to identify the plane $\mathbb{R}^2$ with the set of complex numbers $\mathbb{C}$, a point $z = (x, y)$ being represented as $x + iy$. A plane similitude composed with a translation can then be written in the form $z \mapsto az + b$ with $a, b \in \mathbb{C}$, $a \neq 0$.

For $Z = (z^{(1)}, \dots, z^{(N)}) \in \mathcal{P}_N$, we let $c(Z)$ be the center of inertia

$$c(Z) = (z^{(1)} + \cdots + z^{(N)})/N.$$

We also let $\|Z\|^2 = \sum_{k=1}^{N} |z^{(k)} - c(Z)|^2$.

11.2.2 The Space of Planar N-Shapes

Construction of a Distance

Let Σ_N be the quotient space of $\mathcal{P}_N$ by the equivalence relation: $Z \sim Z'$ if there exist $a, b \in \mathbb{C}$ such that $Z' = aZ + b$. We denote by $[Z]$ the equivalence class of Z for this relation. We want to define a distance between two equivalence classes $[Z]$ and $[Z']$.

Following Sect. 11.1.4, we define a Riemannian metric on $\mathcal{P}_N$ which is invariant under the action. We therefore must define, for all $Z \in \mathcal{P}_N$, a norm $\|A\|_Z$ over all $A = (a_1, \ldots, a_N) \in \mathbb{C}^N$ such that for all $a, b \in \mathbb{C}$:

$$\|A\|_Z = \|a \cdot A\|_{aZ+b},$$

and it suffices to define such a norm for Z such that $\|Z\| = 1$ and $c(Z) = 0$, since we have, for all Z,

$$\|A\|_Z = \left\| \frac{A}{\|Z\|} \right\|_{\frac{Z - c(Z)}{\|Z\|}}. \tag{11.11}$$

Once the metric has been chosen, the distance $D(W, Y)$ is defined by

$$D(W, Y)^2 = \inf \int_0^1 \|\partial_t Z\|^2_{Z(t)}\, dt, \tag{11.12}$$

the infimum being taken over all paths $Z(\cdot)$ such that $Z(0) = W$ and $Z(1) = Y$.

When $c(Z) = 0$ and $\|Z\| = 1$, we take

$$\|A\|_Z^2 = \sum_{k=1}^N |a^{(k)}|^2.$$

From (11.11) and (11.12), computing $D(W, Y)$ requires us to minimize, among all paths between W and Y,

$$\int_0^1 \frac{\sum_{k=1}^N |\partial_t Z^{(k)}|^2}{\sum_{k=1}^N |Z^{(k)}(t) - c(Z(t))|^2} dt.$$

Let $\bar{z}(t) = c(Z(t))$, $v^{(k)}(t) = (Z^{(k)}(t) - C(Z(t)))/\|Z(t)\|$ and $\rho(t) = \|Z(t)\|$. The path $Z(\cdot)$ is uniquely characterized by $(v(\cdot), \rho(\cdot), \bar{z}(\cdot))$. Moreover, we have

$$\partial_t Z^{(k)} = \partial_t \bar{z} + \rho \partial_t v + v \partial_t \rho$$

so that we need to minimize

$$\int_0^1 \sum_{k=1}^N \left| \frac{\partial_t \bar{z}}{\rho} + \frac{\partial_t \rho}{\rho} \cdot v^{(k)} + \partial_t v^{(k)} \right|^2 dt.$$

This is equal (using $\sum_k v^{(k)} = 0$ and $\sum_k |v^{(k)}|^2 = 1$, together with the differentials of these expressions) to

$$N\int_0^1\left(\frac{\partial_t\bar{z}}{\rho}\right)^2dt+\int_0^1\left(\frac{\partial_t\rho}{\rho}\right)^2dt+\int_0^1\sum_{k=1}^N\left|\partial_t v^{(k)}\right|^2dt\,. \tag{11.13}$$

The end-point conditions are directly deduced from those initially given in terms of W and Y.

The last term in (11.13), which only depends on v, can be minimized explicitly, under the constraints $\sum_k v^{(k)}=0$ and $\sum_k|v^{(k)}|^2=1$, which imply that v varies on a $(2N-3)$-dimensional real sphere. The geodesic distance is therefore given by the length of great circles, which yields the expression of the minimum: $\arccos(v(0)^*v(1))^2$, where the "$*$" exponent refers to the conjugate transpose.

Using this, we have

$$\begin{aligned}D(W,Y)^2=\inf\left(N\int_0^1\left(\frac{\partial_t\bar{z}}{\rho}\right)^2+\int_0^1\left(\frac{\partial_t\rho}{\rho}\right)^2dt\right)\\ +\arccos\left(\left(\frac{W-c(W)}{\|W\|}\right)^*\left(\frac{Y-c(Y)}{\|Y\|}\right)\right)^2,\end{aligned} \tag{11.14}$$

where the first infimum is over functions $t\mapsto(\bar{z}(t),\rho(t))\in\mathbb{C}\times[0,+\infty[$, such that $\bar{z}(0)=c(W)$, $\bar{z}(1)=c(Y)$, $\rho(0)=\|W\|$, $\rho(1)=\|Y\|$.

The induced distance on Σ_N is then given by

$$d([Y],[W])=\inf\{D(Y,aW+b),a,b\in\mathbb{C}\}.$$

Writing $a=\lambda h$ with $\lambda=|a|>0$ and $|h|=1$. Then

$$\frac{aW+b-c(aW+b)}{\|aW+b\|}=h\frac{W-c(W)}{\|W\|}.$$

Moreover, given h, we can take $\lambda=\|Y\|/\|W\|$ and $b=c(Y)-\lambda hc(W)$ so that $\|Y\|=\|aW+b\|$ and $c(Y)=c(aW+b)$, for which the infimum in the right-hand side of (11.14) is zero. We therefore have

$$d([Y],[W])=\inf_{h:|h|=1}\left(\arccos\left(\left(\frac{W-c(W)}{\|W\|}\right)^*\left(\frac{Y-c(Y)}{\|Y\|}\right)\right)\right).$$

Finally, optimizing this over the unit vector h, we get

$$d([Y],[W])=\arccos\left|\left(\frac{W-c(W)}{\|W\|}\right)^*\left(\frac{Y-c(Y)}{\|Y\|}\right)\right|. \tag{11.15}$$

Denote by S^{2N-3} the set of $v=(v_1,\ldots,v_{N-1})\in\mathbb{C}^{N-1}$ such that $\sum_i|v_i|^2=1$ (this can be identified with a real sphere of dimension $2N-3$). The complex projective space, denoted $\mathbb{C}P^{N-2}$, is defined as the space S^{2N-3} quotiented by the

equivalence relation: $v \mathcal{R} v'$ if and only if $\exists \nu \in \mathbb{C}$ such that $v' = \nu v$; in other words, $\mathbb{C}P^{N-2}$ contains all sets

$$S^1 \cdot v = \{\nu v, \nu \in \mathbb{C}, |\nu| = 1\}$$

when v varies in S^{2N-3}. This set has the structure of an $(N-2)$-dimensional complex manifold, which means that it can be covered with an atlas of open sets that are in bijection with open subsets of $\mathbb{C}^{N-2}$ (with analytic changes of coordinates). Such an atlas is provided, for example, by the family $(\mathcal{O}_k, \Psi_k)$, $k = 1, \ldots, N$, where $\mathcal{O}_k$ is the set of all $S^1 \cdot v \in \mathbb{C}P^{N-2}$ with $v^{(k)} \neq 0$, and

$$\Psi_k(S^1 \cdot v) = (v^{(1)}/v^{(k)}, \ldots, v^{(k-1)}/v^{(k)}, v^{(k+1)}/v^{(k)}, \ldots, v^{(N-1)}/v^{(k)}) \in \mathbb{C}^{N-2}$$

for all $S^1 \cdot v \in \mathcal{O}_i$. In fact, Σ_N is also an analytic complex manifold that can be identified with $\mathbb{C}P^{N-2}$.

Let us be more explicit with this identification [166]. Associate to $Z = (z^{(1)}, \ldots, z^{(N)})$ the family $(\zeta^{(1)}, \ldots, \zeta^{(N-1)})$ defined by

$$\zeta^{(k)} = (kz^{(k+1)} - (z^{(1)} + \cdots + z^{(k)}))/\sqrt{k^2 + k}.$$

One can verify that $\sum_{k=1}^{N-1} |\zeta^{(k)}|^2 = \|Z\|^2$ (similar decompositions are used, for example, for the analysis of large-dimensional systems of particles [305]). Denote by $F(Z)$ the element $S^1 \cdot (\zeta/\|Z\|)$ in $\mathbb{C}P^{N-2}$. One can check that $F(Z)$ only depends on $[Z]$ and that $[Z] \mapsto F(Z)$ is an isometry between Σ_N and $\mathbb{C}P^{N-2}$.

The Space of Triangles

This construction, applied to the case $N = 3$ (which corresponds to triangles with labeled vertices), yields a quite interesting result. For a triangle $Z = (z^{(1)}, z^{(2)}, z^{(3)})$, the previous function $F(Z)$ can be written

$$\begin{aligned} F(Z) &= S^1 \cdot \left(\frac{\left[\frac{z^{(2)} - z^{(1)}}{\sqrt{2}}, \frac{2z^{(3)} - z^{(1)} - z^{(2)}}{\sqrt{6}} \right]}{\sqrt{|z^{(2)} - z^{(1)}|^2/2 + |2z^{(3)} - z^{(1)} - z^{(2)}|^2/6}} \right) \\ &= S^1 \cdot [v^{(1)}, v^{(2)}]. \end{aligned}$$

On the set $v^{(1)} \neq 0$ (i.e., the set $z^{(1)} \neq z^{(2)}$) we have the local chart

$$Z \mapsto v^{(2)}/v^{(1)} = \frac{1}{\sqrt{3}} \left(\frac{2z^{(3)} - z^{(2)} - z^{(1)}}{z^{(2)} - z^{(1)}} \right) \in \mathbb{C}.$$

If we let $v^{(2)}/v^{(1)} = \tan \frac{\theta}{2} e^{i\varphi}$, and $M(Z) = (\sin\theta \cos\varphi, \sin\theta \sin\psi, \cos\theta) \in \mathbb{R}^3$, we obtain a correspondence between the triangles and the unit sphere S^2.

This correspondence is isometric: the distance between two triangles $[Z]$ and $[\tilde{Z}]$, which has been defined above by

$$d([Z],[\tilde{Z}]) = \arccos \left| \frac{\sum_{k=1}^{3} z^{(k)} \left(\tilde{z}^{(k)}\right)^*}{\|Z\| \, \|\tilde{Z}\|} \right|,$$

gives, after passing to coordinates θ and φ, exactly the length of the great circle between the images $M(Z)$ and $M(\tilde{Z})$. We therefore obtain a representation of labeled triangular shapes as points on the sphere S^2, with the possibility of comparing them using the standard metric on S^2.

11.3 Parametrization-Invariant Distances Between Plane Curves

We now describe distances between two-dimensional shapes when they are defined as plane curves modulo changes of parameter. Such distances have been the subject of extensive and detailed mathematical studies, [193, 201], but here we only give an overview of the main ideas and results.

Simple parametrization-free distances can be defined directly from the ranges of the curves. For example, it is possible to use standard norms applied to arc-length parametrizations of the curves, like L^p or Sobolev norms of the difference. With simple closed curves, one can measure the area of the symmetric difference between the interiors of the curves. A more advanced notion, the Hausdorff distance, is defined by

$$d(m, \tilde{m}) = \inf \{\varepsilon > 0, m \subset \tilde{m}^\varepsilon \text{ and } \tilde{m} \subset m^\varepsilon\},$$

where m^ε is the set of points at a distance less than ε from m (and similarly for $\tilde{m}^\varepsilon$). The same distance can be used with the interiors for simple closed curves. In fact, the Hausdorff distance is a distance between closed sets as stated in the following proposition.

Proposition 11.9 *For $\varepsilon > 0$ and a subset A of $\mathbb{R}^d$, let A^ε be the set of points $x \in \mathbb{R}^d$ such that there exists an $a \in A$ with $|a - x| < \varepsilon$. Let*

$$d_H(A, B) = \inf \{\varepsilon > 0 : A \subset B^\varepsilon \text{ and } B \subset A^\varepsilon\}.$$

Then d_H is a distance on the set of closed subsets of $\mathbb{R}^d$.

Proof Symmetry is obvious, and we leave to the reader the proof of the triangular inequality, which is a direct consequence of the fact that $(A^\varepsilon)^{\varepsilon'} \subset A^{\varepsilon+\varepsilon'}$.

Assume that $d_H(A, B) = 0$. Then $A \subset B^\varepsilon$ for all $\varepsilon > 0$. But $\bigcap_\varepsilon B^\varepsilon = \bar{B}$, the closure of B. We therefore have

$$d_H(A, B) = 0 \Rightarrow A \subset \bar{B} \text{ and } B \subset \bar{A},$$

which implies that $A = B$ if both sets are closed.

One can also proceed similarly to Sect. 11.1.1. First define equivariant distances over parametrized curves, then optimize them with respect to changes of parameters. Let $\mathcal{C}$ be a set of parametrized curves, defined as functions $m : [0, 1] \mapsto \mathbb{R}^2$, subject to additional properties (such as smoothness, closedness, etc.), and G a group of changes of parameter over $[0, 1]$ (including changes of offset for closed curves). Consider the quotient space $\mathcal{S} = \mathcal{C}/G$ for the action $\varphi \cdot m = m \circ \varphi^{-1}$, which are curves modulo change of parameter (one may also want to quotient out rotation, scaling). Based on our discussion in Sect. 11.1.1, a pseudo-distance on $\mathcal{S}$ can be defined from a distance on $\mathcal{C}$ that is equivariant by changes of parameter.

L^p norms between parametrized curves are not equivariant, unless $p = \infty$, with

$$d_\infty(m, \tilde{m}) = \sup_u |m(u) - \tilde{m}(u)|.$$

The distance obtained after reduction by diffeomorphism is called the Fréchet distance, defined by

$$d_F(m, \tilde{m}) = \inf_\varphi d_\infty(m \circ \varphi, \tilde{m}).$$

We note that if, for some diffeomorphism φ, $d_\infty(m \circ \varphi, \tilde{m}) \leq \varepsilon$, then $m \subset \tilde{m}^\varepsilon$ and $\tilde{m} \subset m^\varepsilon$. So we get the relation

$$\varepsilon > d_F(m, \tilde{m}) \Rightarrow \varepsilon > d_H(m, \tilde{m}),$$

which implies $d_H \leq d_F$. This and Proposition 11.9 prove that d_F is a distance between curves.

We now consider equivariant distances on $\mathcal{C}$ based on Riemannian metrics derived from invariant norms on the tangent space. We only give an informal discussion, ignoring the complications that arise from the infinite dimension of the space of curves (see [199, 200] for a rigorous presentation). Tangent vectors to $\mathcal{C}$ are derivatives of paths in $\mathcal{C}$, which are time-dependent parametrized curves $t \mapsto m(t, \cdot)$. Tangent vectors therefore take the form $v = \partial_t m(t, \cdot)$, which are functions $v : [0, 1] \to \mathbb{R}^2$. Since a change of parameter in a time-dependent curve induces the same change of parameter on the time derivative, a norm on the tangent space to $\mathcal{C}$ is equivariant under the action of changes of parameter, if, for any m, v, φ,

$$\|v \circ \varphi^{-1}\|_{m \circ \varphi^{-1}} = \|v\|_m. \tag{11.16}$$

It is therefore sufficient to define $\|v\|_m$ for curves parametrized by arc length, since (11.16) then defines the metric for any parametrized curve.

We now want to define tangent vectors to "plane curves modulo change of parameters." We know that we can modify the tangential component of the time derivative of a time-dependent parametrized curve $t \mapsto m(t, \cdot)$ without changing the geometry of the evolving curve. It follows from this that tangent vectors to $\mathcal{S}$ at a curve m are equivalent classes of vector fields along m that share the same normal component

to m, and can therefore be identified with this normal component itself, i.e., a scalar function along m. The induced metric on $\mathcal{S}$ is

$$\|a\|_m := \inf \left\{ \|v\|_m : v^T N = a \right\}.$$

The associated pseudo-distance on $\mathcal{S}$ is

$$d(m, \tilde{m})^2 = \inf \left\{ \int_0^1 \left\| \partial_t \mu^T N \right\|^2_{\mu(t)} dt, \mu(0, \cdot) = m, \mu(1, \cdot) = \tilde{m} \right\}. \tag{11.17}$$

The fact that we only get a pseudo-distance in general is interestingly illustrated by the following simple example. Define

$$\|v\|_m^2 = \int_0^{L_m} |v(s)|^2 ds. \tag{11.18}$$

This is the L^2 norm in the arc-length parametrization. Then, as stated in the following theorem, we have $d(m, \tilde{m}) = 0$.

Theorem 11.10 (Mumford–Michor) *The distance defined in* (11.17) *with the norm given by* (11.18) *vanishes between any smooth curves m and* $\tilde{m}$.

A proof of this result can be found in [199, 200]. It relies on the observation that one can grow thin protrusions ("teeth") on the curve at a cost which is negligible compared to the size of the tooth. It is an easy exercise to compute the geodesic length of a path that starts with a horizontal segment and progressively grows an isosceles triangle of width ε and height t (at time t) on the segment until $t = 1$. This length is $o(\varepsilon)$ (in fact, $O(\varepsilon^2 \ln \varepsilon)$). This implies that one can transform a curve into $O(1/\varepsilon)$ thin non-overlapping teeth at almost no cost. A repeated application of this concept is the basic idea in the construction made in [200] to create almost-zero-length paths between two arbitrary curves.

To prevent the distance from vanishing, one needs to penalize the curve length more than (11.18) does. For example, the distance associated with the metric

$$\|v\|_m^2 = L_m \int_0^{L_m} |v(s)|^2 ds, \tag{11.19}$$

introduced in [193, 259], does not give a degenerate distance on $\mathcal{S}$. The resulting distance is the area swept by the path relating the compared curves [259].

Another way to control degeneracy is to penalize high-curvature points, using for example

$$\|v\|_m^2 = \int_0^{L_m} (1 + a\kappa_m(s)^2) |v(s)|^2 ds. \tag{11.20}$$

This metric has been studied in [200], where it is shown (among other results) that the distance between distinct curves is positive. Finally, one can add derivatives of v (with respect to arc length) in the definition of the metric; this provides *Sobolev metrics* [193, 201] that we have already described for curve evolution.

11.4 Invariant Metrics on Diffeomorphisms

We discuss here the construction of a right-invariant distance between diffeomorphisms. We will see, in particular, that it coincides with the direct construction made in Chap. 7.

Here also, we only make an informal (non-rigorous) discussion. We consider a group G of diffeomorphisms of Ω, and define (to fix our ideas) the tangent space to G at $\varphi \in G$ by the set of $u : \Omega \to \mathbb{R}^d$ such that $\mathrm{id} + t\, u \circ \varphi^{-1} \in G$ for small enough t. Since the group product on G is the composition, $\varphi\psi = \varphi \circ \psi$, the right translation $R_\varphi : \psi \mapsto \psi \circ \varphi$ is linear, and therefore "equal" to its differential: for $u \in T_\psi G$,

$$dR_\varphi(\psi)u = u \circ \varphi.$$

A metric on G is right-invariant if, for all $\varphi, \psi \in G$ and for all $u \in T_\psi G$,

$$\|dR_\varphi(\psi)u\|_{\psi\circ\varphi} = \|u\|_\psi,$$

which yields, taking $\varphi = \psi^{-1}$:

$$\|u\|_\psi = \|u \circ \psi^{-1}\|_{\mathrm{id}}.$$

This implies that the energy of a path $(t \mapsto \varphi(t, \cdot))$ in G must be defined by

$$E(\varphi(\cdot)) = \int_0^1 \left\| (\partial_t \varphi)(t, \varphi^{-1}(t, \cdot)) \right\|_{\mathrm{id}}^2 dt.$$

If we let

$$v(t, x) = (\partial_t \varphi)(t, \varphi^{-1}(t, x)),$$

the energy can be written

$$E(\varphi) = \int_0^1 \|v(t, \cdot)\|_{\mathrm{id}}^2 \, dt$$

with the identity

$$\partial_t \varphi(t, x) = v(t, \varphi(t, x)).$$

This implies that φ is the flow associated to the velocity field $v(t, \cdot)$. We therefore retrieve the construction given in Chap. 7, with $\|\cdot\|_{\mathrm{id}} = \|\cdot\|_V$. Thus, V in Chap. 7 has a role similar to that of the tangent space to G at id here. Because of this, we let $V = T_{\mathrm{id}}G$ and $\|\cdot\|_{\mathrm{id}} = \|\cdot\|_V$ in the remaining discussion to homogenize the notation.

Assume that V is admissible, according to Definition 7.14. The right-invariant distance on G is

$$d(\varphi_0, \varphi_1) = \inf \sqrt{\int_0^1 \|v(t, \cdot)\|_V^2 \, dt}, \tag{11.21}$$

where the minimum is taken over all v such that, for all $x \in \Omega$, the solution of the ordinary differential equation

$$\partial_t y = v(t, y)$$

with initial conditions $y(0) = \varphi_0(x)$ is such that $y(1) = \varphi_1(x)$, consistently with Sect. 7.2.6.

We point out, however, that if G is, say, $\mathrm{Diff}_0^{p,\infty}$, which has its own structure of infinite-dimensional differential manifold, then V is a proper subspace of $T_{\mathrm{id}}G$, resulting in a "sub-Riemannian" metric.

11.4.1 The Geodesic Equation

The geodesic equation on G is equivalent to the Euler–Lagrange equation associated to the variational problem (11.21). This is similar to what we have computed in Sect. 10.4, except that here we have a fixed end-point condition. One may address this with a method called the Euler–Poincaré reduction [150, 188], and the presentation we make here is related to it. The energy

$$E(v) = \frac{1}{2} \int_0^1 \|v(t)\|_V^2 dt$$

is minimized over all v such that $\varphi_{01}^v = \varphi_1$ (without loss of generality, because the distance is right invariant, we can assume that $\varphi_0 = \mathrm{id}$).

Applying Theorem D.8 with

$$H_v(\varphi, p) = (p \mid v \circ \varphi) - \frac{\rho}{2} \|v\|_V^2$$

we obtain the fact that, if a trajectory is not "elusive," there exists $\rho \in \{0, 1\}$ and a co-state $p(\cdot)$ taking values in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)^*$ such that

$$\begin{cases} \partial_t \varphi = v \circ \varphi \\ (\partial_t p \mid h) + (p \mid dv \circ \varphi \, h) = 0 \\ \rho v = \xi_\varphi^* p, \end{cases}$$

where $\xi_\varphi : v \mapsto v \circ \varphi$. The first two equations do not depend on whether $\rho = 0$ or 1 (i.e., whether the geodesic is normal or abnormal), and are identical to the momentum conservation equation (10.17) and EPDiff studied in the previous chapter. When $\rho = 1$, the third equation describes v, and provides the same necessary conditions for optimality as those found for the diffeomorphic problem in that chapter. Abnormal solutions are such that $\xi_\varphi^* p = 0$ along the trajectory.

Note that the solutions of "soft registration" problems, minimizing

$$\int_0^1 \|v(t)\|_V^2 \, dt + U(\varphi_{01}^v)$$

for a differentiable function U, always provides normal geodesics.

11.4.2 A Simple Example

An explicit computation of the geodesic distance is generally impossible, but here is an exception, in one dimension. Take $\Omega = [0, 1]$ and

$$\|u\|_{\mathrm{id}}^2 = \int_0^1 |\partial_x u|^2 \, dx.$$

Note that this norm is not admissible, because it cannot be used to control the supremum norm of $\partial_x u$. The associated energy of a path of diffeomorphisms $\varphi(t, \cdot)$ is

$$U(\varphi(\cdot)) = \int_0^1 \int_0^1 \left|\partial_x \left(\varphi_t \circ \varphi^{-1}(t, \cdot)\right)\right|^2 dx dt.$$

This gives, after expanding the derivative and making the change of variables $x = \varphi(t, y)$:

$$U(\varphi(\cdot)) = \int_0^1 \int_0^1 |\partial_t \partial_x \varphi|^2 \, |\partial_x \varphi|^{-1} \, dy dt.$$

Define $q(t, y) = \sqrt{\partial_x \varphi(t, y)}$. We have

$$U(\varphi(\cdot)) = 4 \int_0^1 \int_0^1 |\partial_t q|^2 \, dy dt,$$

which yields

$$U(\varphi(\cdot)) = 4 \int_0^1 \|\partial_t q(t, \cdot)\|_2^2 \, dt.$$

If the problem were to minimize this energy under the constraints $q(0, \cdot) = \sqrt{\partial_x \varphi(0, \cdot)}$ and $q(1, \cdot) = \sqrt{\partial_x \varphi(1, \cdot)}$, the solution q would be given by the line segment

$$q(t, x) = t q(1, x) + (1 - t) q(0, x).$$

There is, however, an additional constraint that comes from the fact that $q(t, \cdot)$ must provide a homeomorphism of $[0, 1]$ for all t, which implies $\varphi(t, 1) = 1$, or, in terms of q

$$\|q(t, \cdot)\|_2^2 = \int_0^1 q(t, x)^2 dx = 1.$$

We therefore need to minimize the length of the path q under the constraint that it remains on a Hilbert (L^2) sphere. Similar to the finite-dimensional case, geodesics on Hilbert spheres are great circles. This implies that the optimal q is given by

$$q(t, \cdot) = \frac{1}{\sin \alpha}(\sin(\alpha(1-t))q_0 + \sin(\alpha t)q_1)$$

with $\alpha = \arccos \langle q_0 , q_1 \rangle_2$. The length of the geodesic is precisely given by α, which provides a closed-form expression of the distance on G [231]

$$d(\varphi, \tilde{\varphi}) = 2 \arccos \int_0^1 \sqrt{\partial_x \varphi \, \partial_x \tilde{\varphi}} dx.$$

11.4.3 Gradient Descent

Assume that a function $\varphi \mapsto U(\varphi)$ is defined over diffeomorphisms. Take C^1 h and ε_0 small enough so that $\varphi + \varepsilon h$ is a diffeomorphism if $|\varepsilon| \leq \varepsilon_0$, and assume that the Gâteaux derivative $\partial_\varepsilon U(\varphi + \varepsilon h)$ exists at $\varepsilon = 0$, denoting it, as in Sect. 9.2, by

$$\partial_\varepsilon U(\varphi + \varepsilon h) = \Big(dU(\varphi) \,\Big|\, h \Big).$$

If a right-invariant metric is given, in the form

$$\langle h , h' \rangle_\varphi = \langle h \circ \varphi^{-1} , h' \circ \varphi^{-1} \rangle_V$$

as above, the gradient of U at φ is computed by identifying

$$\begin{aligned} \Big(dU(\varphi) \,\Big|\, h \Big) &= \langle \nabla U(\varphi) , h \rangle_\varphi \\ &= \langle \nabla U(\varphi) \circ \varphi^{-1} , h \circ \varphi^{-1} \rangle_V \\ &= \big(\mathbb{L}(\nabla U(\varphi) \circ \varphi^{-1}) \,\big|\, h \circ \varphi^{-1} \big), \end{aligned}$$

where $\mathbb{L} = \mathbb{K}^{-1}$ is the duality operator on V. Since (with the notation of Sect. 9.2)

$$\Big(dU(\varphi) \,\Big|\, h \Big) = \big(\bar{\partial} U(\varphi) \,\big|\, h \circ \varphi^{-1} \big),$$

we see that, using $\overline{\nabla}^V U = \mathbb{K}\bar{\partial} U$,

$$\nabla U(\varphi) = \overline{\nabla}^V U(\varphi) \circ \varphi$$

and the evolution equation introduced in (9.7) is nothing but a Riemannian gradient descent for U for the considered metric.

11.4.4 Diffeomorphic Active Contours

As a new example of an application of this formalism, we provide a Riemannian version of the active contours algorithm discussed in Sect. 5.4. Let, for a curve m,

$$E(m) = \int_m F(p) d\sigma_m + \int_{\Omega_m} \tilde{F}(x) dx. \tag{11.22}$$

We can, fixing a template curve m_0, define the functional

$$U(\varphi) = E(\varphi(m_0)).$$

Letting $m = \varphi(m_0)$, a straightforward computation gives

$$\left(\bar{\partial} U(\varphi) \,\middle|\, v\right) = -\int_m (\kappa F - F^T N + \tilde{F}) v^T N d\sigma_m,$$

from which we deduce

$$\nabla U(\varphi)(x) = -\int_m (\kappa F - F^T N + \tilde{F}) K(\varphi(x), \cdot) N d\sigma_m.$$

This defines the continuous time gradient descent algorithm,

$$\partial_t \varphi(t, x) = \int_{m(t)} (\kappa F - F^T N + \tilde{F}) K(\varphi(t, x), \cdot) N d\sigma_{m(t)}$$

with $m(t) = \varphi(t, \cdot) \circ m_0$.

This algorithm also be expressed as an evolution equation in terms of $m(t)$ only, yielding the *diffeomorphic active contours evolution equation* [21, 310]

$$\partial_t m(t, u) = \int_{m(t)} (\kappa F - F^T N + \tilde{F}) K(m(t, u), \cdot) N d\sigma_{m(t)}. \tag{11.23}$$

A similar discussion can be made for surfaces instead of curves.

Examples of segmentations using this equation are provided in Fig. 11.1.

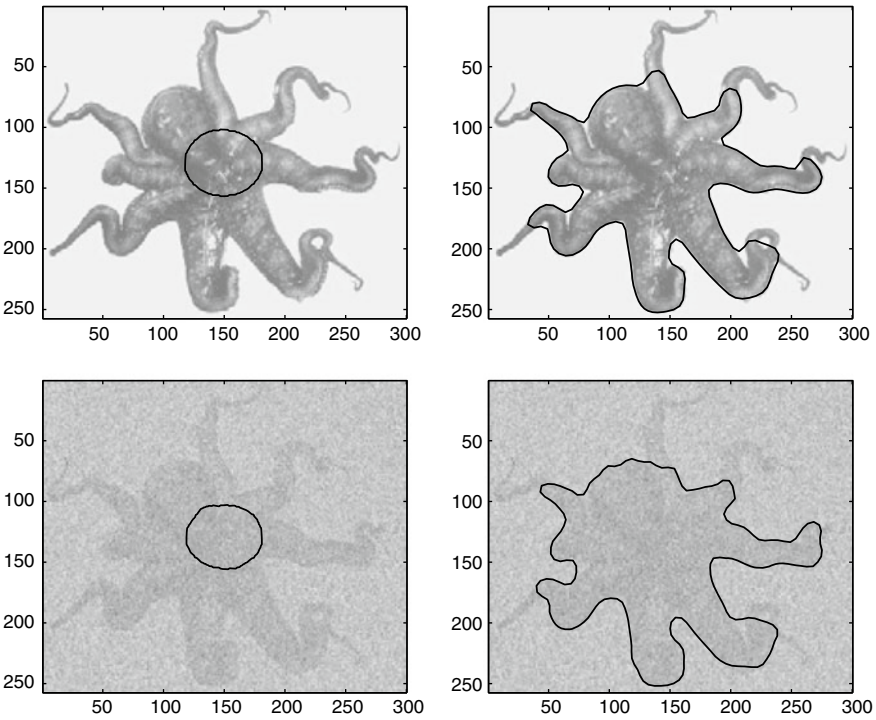

Fig. 11.1 Diffeomorphic active contours (compare with Fig. 5.4). On each row, the left image is the initial contour, and the right one is the solution obtained with diffeomorphic active contours. The first row presents a clean image and the second a noisy one

11.5 Group Actions and Riemannian Submersion

11.5.1 Riemannian Submersion

We temporarily switch to general (but finite-dimensional) Lie groups before returning to diffeomorphisms. Let $\mathcal{G}$ be a Lie group acting transitively (and smoothly) on a manifold, M. Fixing a reference element $m_0 \in M$, Corollary 11.5 and Eq. (11.4) show how a distance that is left-equivariant under the action of $G = \mathrm{Iso}_{m_0}(\mathcal{G})$ can be projected to a pseudo-distance on M. We now provide the infinitesimal version of this result, which involves the notion of Riemannian submersion discussed in Sect. B.6.7.

Define, as done in Sect. 11.1.3,

$$\begin{aligned} \pi : \quad & \mathcal{G} \to M \\ & g \mapsto g \cdot m_0 \,. \end{aligned}$$

This map is onto because the action is transitive and one can show [142] that it is a submersion, i.e., that $d\pi(g)$ has full rank (the dimension of M) for $g \in \mathcal{G}$.

For this submersion, the fiber over $m \in M$ is the set $\pi^{-1}(m) = \{g \in \mathcal{G} : g \cdot m_0 = m\}$. Fix $g \in \pi^{-1}(m)$. Another group element $\tilde{g}$ belongs to $\pi^{-1}(m)$ if and only if $g^{-1}\tilde{g} \in G$, so that $\pi^{-1}(m) = gG$ is a coset. Furthermore [142], the mapping

$$\begin{aligned} [\pi]: \quad \mathcal{G}/G &\to M \\ [g] &\mapsto \pi(g) \end{aligned}$$

is an isomorphism.

Assume that $\mathcal{G}$ is equipped with a Riemannian metric that is right-invariant for the action of G, so that, for any $g \in \mathcal{G}$, $h \in G$ and $w \in T_g\mathcal{G}$,

$$\|w\|_g = \|dR_h(g)w\|_{gh}.$$

In other terms $dR_h(g)$ is an isometry between $T_g\mathcal{G}$ and $T_{gh}\mathcal{G}$ (or R_h is a *Riemannian isometry*). Then one can build a Riemannian metric on M such that π is a Riemannian submersion. Indeed, if $g, \tilde{g} \in \pi^{-1}(m)$, then there exists an $h \in G$ such that $\tilde{g} = gh$ so that $T_{\tilde{g}}\mathcal{G} = dR_h T_g\mathcal{G}$. Moreover, $\pi(g'h) = \pi(g')$ for all $g' \in \mathcal{G}$ implies that $d\pi(gh)dR_h(g) = d\pi(g)$. This shows that $\mathcal{V}_{gh} = dR_h(g)\mathcal{V}_g$, where

$$\mathcal{V}_g = \left\{w \in T_g\mathcal{G} : d\pi(g)w = 0\right\}$$

is the vertical space at g. Letting $\mathcal{H}_g = \mathcal{V}_g^{\perp}$ be the horizontal space at g, this and the fact that $dR_h(g)$ is an isometry implies that $dR_h(g)\mathcal{H}_g = \mathcal{H}_{gh}$, so that the restriction of $dR_h(g)$ to the horizontal space provides an isometry between these spaces. This allows us to define, for any $m \in M$ and tangent vector $\xi \in T_mM$:

$$\|\xi\|_m = \|w\|_g$$

for any $g \in \pi^{-1}(m)$, where w is uniquely defined by $d\pi(g)w = \xi$ and $w \in \mathcal{H}_g$. Using the minimizing property of the orthogonal projection, an equivalent definition is that

$$\|\xi\|_m = \min\left\{\|w\|_g : w \in T_g\mathcal{G}, d\pi(g)w = \xi\right\}. \tag{11.24}$$

This is the infinitesimal counterpart of Eq. 11.3.

Using the Lie group structure, this construction can also be analyzed solely on the group's Lie algebra, $\mathfrak{g} = T_{\mathrm{id}}\mathcal{G}$. Notice that, if $\pi(g) = m$ and $\tilde{g}(t)$ is a curve on $\mathcal{G}$ such that $\tilde{g}(0) = g$ and $\partial_t\tilde{g}(0) = w$, then, taking derivatives at $t = 0$,

$$d\pi(g)w = \partial_t(\tilde{g}(t) \cdot m_0) = \partial_t(\tilde{g}(t)g^{-1}) \cdot m = v \cdot m,$$

where $v = dR_{g^{-1}}(g)w$ and $v \cdot m$ refers to the infinitesimal action (cf. Sect. B.5.3). From this, we deduce that, letting

$$V_m = \{v \in \mathfrak{g} : v \cdot m = 0\},$$

one has $\mathcal{V}_g = dR_g(\mathrm{id})V_m$. Moreover, we can rewrite (11.24) as

$$\|\xi\|_m = \min\left\{\|dR_g(\mathrm{id})v\|_g : v \in \mathfrak{g},\, v \cdot m = \xi\right\}, \tag{11.25}$$

for any $g \in \pi^{-1}(m)$, with $\mathfrak{g} = T_{\mathrm{id}}\mathcal{G}$. The mapping $v \mapsto \|dR_g(\mathrm{id})v\|_g$ provides a Euclidean norm on $\mathfrak{g}$ that does not depend on which g is chosen in $\pi^{-1}(m)$, therefore only depending on m. Denoting this by $\|\cdot\|_m$, $dR_g(\mathrm{id})$ is, by construction, an isometry from $\mathfrak{g}$ to $T_g\mathcal{G}$ that maps V_m onto $\mathcal{V}_g$, and therefore maps H_m onto $\mathcal{H}_g$, where H_m is the space perpendicular to V_m with respect to the dot product associated with $\|\cdot\|_m$ (which we shall denote by $V_m^{\perp_m}$). For $\xi \in T_mM$, there is therefore a unique vector, $v^\xi \in H_m$, such that $v^\xi \cdot m = \xi$ and (11.24) and (11.25) simply become

$$\|\xi\|_m = \|v^\xi\|_{\mathrm{id}} = \min\{\|v\|_m : v \in \mathfrak{g},\, v \cdot m = \xi\}. \tag{11.26}$$

Under the stronger assumption that the metric on $\mathcal{G}$ is right-invariant, so that R_h is a Riemannian isometry for all $h \in \mathcal{G}$, we have $\|dR_g(\mathrm{id})v\| = \|v\|_{\mathrm{id}}$ for all $g \in \mathcal{G}$ and $v \in \mathfrak{g}$, so that $\|v\|_m = \|v\|_{\mathrm{id}}$ for all $m \in M$ and one has:

$$\|\xi\|_m = \min\{\|v\|_{\mathrm{id}} : v \cdot m = \xi\}. \tag{11.27}$$

One also defines *horizontal linear forms*, or *horizontal covectors*, which are linear forms $z \in T_{\mathrm{id}}G^*$ such that $(z \mid v) = 0$ for all $v \in \mathcal{V}_m$.

If $\xi \in T_mM$, we have defined v^ξ as the vector in $T_{\mathrm{id}}G$ that minimizes $\|v\|_{\mathrm{id}}$ among all v such that $v \cdot m = \xi$, i.e., the orthogonal projection on H_m of any v_0^ξ such that $v_0^\xi \cdot m = \xi$. This leads to the following definitions, in which we let $v^\xi = h_m(\xi)$.

Definition 11.11 Let $\mathcal{G}$ be a Lie group acting transitively on a manifold M. If $m \in M$ and $\xi \in T_mM$, the horizontal lift of ξ is the vector $h_m(\xi) \in H_m = V_m^{\perp_m}$ such that $h_m(\xi) \cdot m = \xi$.

If $v \in T_{\mathrm{id}}G$, we call $\pi_{\mathcal{H}_m}(v)$ the horizontal part of v at m and $v - \pi_{\mathcal{H}_m}(v)$ its vertical part at m, where π_{H_m} is the orthogonal projection for $\|\cdot\|_m$, so that

$$\pi_{H_m}(v) = h_m(v \cdot m). \tag{11.28}$$

The projection on M of the Riemannian metric on G is defined by

$$\langle \xi\,,\, \eta\rangle_m = \langle h_m(\xi)\,,\, h_m(\eta)\rangle_{\mathrm{id}}. \tag{11.29}$$

In the full right-invariant case, geodesics for the projected metric are immediately deduced from those on $\mathcal{G}$, as stated in the following proposition.

Proposition 11.12 *Assume that the metric on $\mathcal{G}$ is right-invariant. Then the geodesic on M starting at m in the direction ξ is deduced from horizontal geodesics on $\mathcal{G}$ by*

$$\mathrm{Exp}_m(t\xi) = \mathrm{Exp}_{\mathrm{id}}(th_m(\xi)) \cdot m. \tag{11.30}$$

Proof This is a direct consequence of Proposition B.27. Let $\mu^*(t) = \mathrm{Exp}_m(t\xi)$ and $\hat{\mu}(t)$ be its horizontal lift starting at some $g \in \pi^{-1}(m)$. Then $\hat{\mu}g^{-1}$ is also a geodesic on $\mathcal{G}$ and necessarily takes the form given in (11.30). □

11.5.2 The Momentum Representation

We now apply these general results to groups of diffeomorphisms. While our previous discussion was limited to finite dimensions, some of the concepts introduced there can be generalized to the infinite-dimensional case. As before, we let V be a Hilbert space of vector fields continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$. To simplify the definition of derivatives, we assume that the shape space M is an open subset of a Banach space Q, and that the mapping $\mathcal{A} : \varphi \mapsto \varphi \cdot m$ is differentiable from $\mathrm{Diff}_0^p \times M$ to Q. Note that this imposes some restrictions on M and Q. For example, if M is the space of C^q embeddings from the unit circle to $\mathbb{R}^2$, then one needs $q \leq p$ for the action $\varphi \cdot q = \varphi \circ q$ to take values in M, and $q \leq p - 1$ to ensure its differentiability.

We will consider the action of Diff_V, the group of attainable diffeomorphisms (Definition 7.15), on M. One of our basic assumptions in finite dimensions was the transitivity of the action, which will not hold in general. We will however fix a reference shape $\bar{m}$, and define the space of attainable shapes as the orbit $M_V = \mathrm{Diff}_V \cdot \bar{m}$ of $\bar{m}$ through the action of Diff_V. For $m \in M_V$, we define

$$Q_m = \{v \cdot m : v \in V\} \subset Q$$

and the norm

$$\|\xi\|_m = \min\{\|v\|_V : \xi = v \cdot m\}$$

for $\xi \in Q_m$.

Notice that the infinitesimal action $v \cdot m = \partial_1 \mathcal{A}(\mathrm{id}, m)v$ is a bounded linear map from $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ to Q, and so is its restriction to V for $(V, \|\cdot\|_m)$, the Hilbert space topology. This implies that the space

$$V_m = \{v \in V : v \cdot m = 0\}$$

is closed in V, and we still denote $V_m^\perp$ by H_m. In particular, we have

$$\|\xi\|_m = \|v^\xi\|_V,$$

where $v^\xi = \pi_{H_m}(v)$, for any vector field v satisfying $v \cdot m = \xi$. This implies that the mapping $v \mapsto v \cdot m$ is an isometry between H_m and Q_m, which incidentally proves that the latter is a Hilbert space. Given this, we can define the variational problem of minimizing

$$\int_0^1 \|\partial_t m(t)\|^2_{m(t)}\, dt$$

subject to $m(0) = m_0$ and $m(1) = m_1$, $m_0, m_1 \in M_V$ and see that this problem is equivalent to minimizing

$$\int_0^1 \|v(t)\|^2_V\, dt$$

subject to $m(0) = 0, m(1) = m_1$ and $\partial_t m = v \cdot m$. It is also equivalent to minimizing the same energy subject to the constraint $\varphi_{01}^v \cdot m_0 = m_1$.

In finite dimensions, we showed that, if the acting group $\mathcal{G}$ is equipped with a right-invariant Riemannian metric and M with the associated projected metric, then solutions of this problem (which are geodesics in M) can be identified with geodesics in $\mathcal{G}$ that start horizontally. So, fixing $m_0 \in M$, the representation $w \mapsto \mathrm{Exp}_{\mathrm{id}}(w) \cdot m_0$ provided a local chart of M around m_0 when defined over a neighborhood of 0 in H_{m_0}.

With diffeomorphisms, we know that we can find elusive, abnormal and normal geodesics, with only the latter associated with an equation that can be solved given initial conditions (m_0, w). We therefore restrict the "local chart" representation to only those solutions, and express it in terms of co-tangent vectors instead of tangent vectors, which is equivalent in theory, but, as we will see, is much more parsimonious in practice.

As we have seen, this geodesic equation is characterized by momentum conservation, namely

$$\mathbb{L}v(t) = Ad_{\varphi(t)^{-1}}(\mathbb{L}v(0))$$

with $\partial_t \varphi = v \circ \varphi$ and $\mathbb{L}$ is the duality operator of V. Given $m \in M_V$, we define the space of horizontal momenta at m simply by $\mathbb{L}H_m \subset V^*$.

Definition 11.13 Let D_m be the subset of $\mathbb{L}H_m$ consisting of initial momenta for which the conditions of Theorem 10.13 on the existence of solutions of the geodesic equation hold. The momentum representation of a deformable template $\bar{m}$ is the map

$$\begin{aligned} \mathrm{Exp}^{\flat}_{\bar{m}} : \quad & D_m \to \mathrm{Diff}_V \cdot \bar{m} \\ & \rho \mapsto \mathrm{Exp}_{\bar{m}}(\mathbb{K}\rho) \cdot \bar{m} \end{aligned} \tag{11.31}$$

which associates to a horizontal momentum ρ the position at time 1 of the geodesic initialized at $(\bar{m}, (\mathbb{K}\rho) \cdot \bar{m})$ in M.

In finite dimensions, we have proved that horizontality is preserved along geodesics. We retrieve this fact directly in this infinite-dimensional case, as a consequence of the conservation of momentum.

Proposition 11.14 *Let $\bar{m}$ be a deformable object and $\rho_0 \in D_{\bar{m}}$. Let $(\rho(t), \varphi(t))$ be the evolving momentum and diffeomorphism provided by EPDiff initialized with*

$\rho(0) = \rho_0$. *Let* $m(t) = \varphi(t) \cdot \bar{m}$ *be the evolution of the deformable object. Then, at all times* t, $\rho(t) \in \mathbb{L}H_{m(t)}$.

Proof We will prove that if $w \in V_m$ and φ a diffeomorphism, then $\mathrm{Ad}_\varphi w \in V_{\varphi \cdot m}$.

Before proving this fact, we verify that it implies the proposition, which requires us to check that $(w \in V_{m(t)}) \Rightarrow ((\rho(t) \mid w) = 0)$. From the conservation of momentum we have

$$(\rho(t) \mid w) = \left(\rho_0 \mid \mathrm{Ad}_{\varphi(t)^{-1}} w\right)$$

and $\mathrm{Ad}_{\varphi(t)^{-1}} w \in V_{\bar{m}}$ if $w \in \mathcal{V}_{m(t)}$, which implies that $(\rho(t) \mid w) = 0$.

We now prove our claim. Let $\psi(\varepsilon)$ be such that $\partial_\varepsilon \psi(0) = w$, and $\partial_\varepsilon(\psi \cdot m)(0) = 0$ at $\varepsilon = 0$. By definition, $\mathrm{Ad}_\varphi w = \partial_\varepsilon(\varphi \circ \psi \circ \varphi^{-1})(0)$. But we have

$$\partial_\varepsilon(\varphi \circ \psi \circ \varphi^{-1}) \cdot (\varphi \cdot m) = dA_\varphi(\psi \cdot m)\, \partial_\epsilon(\psi \cdot m) = 0,$$

which implies that $\mathrm{Ad}_\varphi w \in V_{\varphi \cdot m}$. (Here, A_φ denotes the action $A_\varphi : m \mapsto \varphi \cdot m$.)□

We now describe horizontal momenta in a few special cases. First assume that deformable objects are point sets, so that

$$M = \left\{(x_1, \ldots, x_N) \in (\mathbb{R}^d)^N, x_i \neq x_j \text{ for } i \neq j\right\}$$

and $Q = (\mathbb{R}^d)^N$.

If $m = (x_1, \ldots, x_N)$, we have

$$V_m = \{v \in V : v(x_1) = \cdots = v(x_N) = 0\}.$$

Letting $e_1, \ldots, e_d$ be the canonical basis of $\mathbb{R}^d$, V_m is therefore defined as the set of v's such that $\left(e_j \delta_{x_k} \mid v\right) = 0$ for all $j = 1, \ldots, d$ and $k = 1, \ldots N$. So $V_m = W^\perp$, where W is the vector space generated by the $d \times N$ vector fields $K(\cdot, x_k)e_j$. Because W is finite-dimensional, it is closed and $H_m = V_m^\perp = (W^\perp)^\perp = W$. Switching to momenta, we obtain the fact that, for point sets $m = (x_1, \ldots, x_N)$

$$\mathbb{L}H_m = \left\{\sum_{k=1}^N z_k \delta_{x_k}, z_1, \ldots, z_N \in \mathbb{R}^d\right\}.$$

In particular, we see that the momentum representation is parametrized by the Nd-dimensional set $(z_1, \ldots, z_N)$ and therefore has the same dimension as the considered objects. Finally, we note that, in this finite-dimensional shape space, one has $M_V = M$ and $\|\cdot\|_m$ provides a Riemannian metric on this space.

The description of V_m is still valid when m is a general parametrized subset of $\mathbb{R}^d$: $m : u \mapsto m(u) = x_u \in \mathbb{R}^d$, defined for u in a, so far, arbitrary set U. Then

$$V_m = \{v \in V : v(x_u) = 0, u \in U\} \tag{11.32}$$

and we still have $V_m = W^\perp$, where W is the vector space generated by the vector fields $K(\cdot, x_u)e_j$, $j = 1, \ldots, d$, $u \in U$. The difference, now, is that W is not finite-dimensional if m is infinite, and not necessarily a closed subspace of V, so that

$$H_m = (W^\perp)^\perp = \bar{W},$$

the closure of W in V. Turning to the momenta, this says that

$$\mathbb{L}H_m = \overline{\left\{ \sum_{k=1}^{n} z_k \delta_{x_{u_k}}, n \geq 0, z_1, \ldots, z_n \in \mathbb{R}^d, u_1, \ldots, u_n \in U \right\}},$$

where the closure is now in V^*.

This argument applies to parametrized curves and surfaces, but must be adapted for geometric objects, that is, curves and surfaces seen modulo a change of parametrization. In this case, deformable objects are equivalent classes of parametrized manifolds. One way to address this is to use infinite-dimensional local charts that describe the equivalence classes in a neighborhood of a given object m. We will not detail this rigorously here, but the interested reader can refer to [201] for such a construction with plane curves.

Intuitively, however, the resulting description of V_m is clear. In contrast to the parametrized case, for which vector fields in V_m were not allowed to move any point in m, it is now possible to do so, provided the motion happens within m, i.e., the vector fields are tangent to m. This leads to the following set:

$$V_m = \{v \in V : v(x) \text{ is tangent to } m \text{ for all } x \in m\}.$$

Since $v(x)$ being tangent to m is equivalent to $N^T v(x) = 0$ for all N normal to m at x, we see that $V_m = W^\perp$, where W is the vector space generated by vector fields $K(\cdot, x)N$, with $x \in m$ and N normal to m at x. Again, this implies that $H_m = \bar{W}$ and that

$$\mathbb{L}H_m = \overline{\left\{ \sum_{k=1}^{n} z_k \delta_{x_k}, n \geq 0, x_1, \ldots, x_n \in m, z_1, \ldots, z_n \in N_{x_k} m \right\}},$$

where $N_x m$ is the set of vectors that are normal to m at x.

Now, consider the example of smooth scalar functions (or images): $m : \mathbb{R}^d \to \mathbb{R}$. In this case, the action being $\varphi \cdot m = m \circ \varphi^{-1}$, the set V_m is

$$V_m = \left\{ v \in V : \nabla m^T v = 0 \right\},$$

which directly implies that $V_m = W^\perp$, where W is the vector space generated by $K(\cdot, x)\nabla m(x)$ for $x \in \mathbb{R}^d$. Horizontal momenta therefore span the set

$$\mathbb{L}H_m = \overline{\left\{ \sum_{k=1}^{n} \nabla m(x_k)\delta_{x_k}, n \geq 0, x_1, \ldots, x_n \in \mathbb{R}^d \right\}}.$$

We conclude with the action of diffeomorphisms on measures, for which:

$$(\varphi \cdot m \mid f) = (m \mid f \circ \varphi),$$

so that $v \in V_m$ if and only if $\left(m \;\middle|\; \nabla f^T v\right) = 0$ for all smooth f. So $V_m = W^\perp$, where

$$W = \left\{ \mathbb{K}(\nabla f\, m),\ f \in C^1(\Omega, \mathbb{R}) \right\}$$

so that

$$\mathbb{L}H_m = \overline{\left\{ \nabla f\, m,\ f \in C^1(\Omega, \mathbb{R}) \right\}}.$$

We point out that the horizontal spaces may be much larger than one would expect by formally extending the case of point sets. For example, if V is a Gaussian RKHS and m is a set with non-empty interior, then $V_m = \{0\}$ in (11.32) and $H_m = V$! This is because Gaussian RKHSs only contain analytic functions [207, 268]. For the same reason, if m is a curve that contain a line segment, all vector fields in V_m must vanish on the whole line containing the segment.

This behavior cannot happen when V is a space containing all compactly-supported smooth functions, such as Sobolev spaces. In this case, if m is a closed subset of $\mathbb{R}^d$, then any smooth vector field with support in $\mathbb{R}^d \setminus m$ must belong to V_m and any $\rho \in \mathbb{L}H_m$ must therefore vanish on such functions, which shows that ρ (as a generalized function) is supported by m.

However, even in such contexts, an explicit description of horizontal momenta is generally beyond reach, and one generally restrict the momentum representation to more "manageable" subsets of H_m, using, for example, measure momenta supported by m, as considered in Sect. 10.5.6. As we have seen, such measure momenta cover most of the cases of interest for diffeomorphic matching with a differentiable end-point cost, even when using Gaussian kernels.

The momentum representation provides a diffeomorphic version of the deformable template approach described for polygons in Sect. 6.3. As we have seen, it can be applied to a wide class of deformable objects. Applications to datasets of three-dimensional medical images can be found in [143, 236, 290, 300].

Chapter 12
Metamorphosis

12.1 Definitions

The Riemannian version of the construction of Sect. 11.1.2 provides a metric based on transformations in which objects can change under the action of diffeomorphisms but also under independent variations. We shall refer to such metrics as *metamorphoses* [152, 206, 281, 282]. They will result in formulations that enable both object registration and metric comparison.

We start with an abstract description of the construction. We consider the setting in which deformations belong to a Lie group G with Lie algebra denoted V, acting on a Riemannian manifold M. We assume that V is a Hilbert space with norm $\|\cdot\|_V$; the metric on M at a given point $a \in M$ is denoted $\langle \cdot, \cdot \rangle_a$ and the corresponding norm $|\cdot|_a$.

For $\varphi \in G$ and $a \in M$, define

$$\begin{array}{lll} A_\varphi : & M \to M & R_a : \quad G \to M \qquad R_\varphi : \quad G \to G \\ & b \mapsto \varphi \cdot b, & \varphi \mapsto \varphi \cdot a, \qquad \psi \mapsto \psi\varphi. \end{array} \tag{12.1}$$

The first two maps are the components of the action, and the third is the right translation on G. It will also be convenient, in the following, to have special notation for derivatives of these maps evaluated at the identity, so we will write $\xi_\varphi = dR_\varphi(\mathrm{id})$, $\xi_a = dR_a(\mathrm{id})$. These maps coincide with the infinitesimal actions, i.e.,

$$v \cdot \varphi = \xi_\varphi v \text{ and } v \cdot a = \xi_a v$$

and we will also use this notation.

If $(\varphi(t), t \in [0, 1])$ is a differentiable curve on G, we define its Eulerian velocity $v(t)$ (which is a curve in V) by the relation:

$$\partial_t \varphi = \xi_{\varphi(t)} v(t) = v(t) \cdot \varphi(t)\,. \tag{12.2}$$

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_12

Definition 12.1 A *metamorphosis* is a pair of curves $(\varphi(t), \alpha(t))$ respectively on G and M, with $\varphi(0) = \mathrm{id}$. Its image is the curve $a(t)$ on M defined by $a(t) = \varphi(t) \cdot \alpha(t)$. We will say that $\varphi(t)$ is the deformation part of the metamorphosis, and that $\alpha(t)$ is the residual part. When $\alpha(t)$ is constant, the metamorphosis is a pure deformation.

12.2 A New Metric on M

Metamorphoses, by the evolution of their images, provide a convenient representation of combinations of group actions and of variations on M. Let a metamorphosis $((\varphi(t), \alpha(t)), t \in [0, 1])$ be given and $a(t) = \varphi(t) \cdot \alpha(t)$ be its image. Then, we can write

$$\begin{aligned}\partial_t a(t) &= dA_{\varphi(t)}(\alpha(t))\partial_t\alpha(t) + dR_{\alpha(t)}(\varphi(t))\partial_t\varphi \\ &= dA_{\varphi(t)}(\alpha(t))\partial_t\alpha(t) + dR_{\alpha(t)}(\varphi(t))dR_{\varphi(t)}(\mathrm{id})v(t).\end{aligned}$$

Since $R_\alpha \circ R_\varphi = R_{\varphi\alpha}$, we get

$$\partial_t a(t) = dA_{\varphi(t)}(\alpha(t))\partial_t\alpha(t) + dR_{a(t)}(\mathrm{id})v(t). \tag{12.3}$$

In particular, when $t = 0$:

$$\partial_t a(0) = \partial_t\alpha(0) + v(0) \cdot a(0). \tag{12.4}$$

This expression provides a decomposition of a generic element $\eta \in T_aM$ in terms of an *infinitesimal metamorphosis*, represented by an element of $V \times T_aM$. Indeed, for $a \in M$, introduce the map

$$\begin{aligned}\Phi^{(a)} : \quad V \times T_aM &\to T_aM \\ (v, \rho) &\mapsto \rho + v \cdot a\,.\end{aligned}$$

Then (12.4) can be written as

$$\partial_t a(0) = \Phi^{(a_0)}\left(\partial_t\alpha(0), v(0)\right).$$

We now introduce the Riemannian metric associated to metamorphoses.

Proposition 12.2 *Assume that $v \mapsto v \cdot a$ is continuous on V. With $\sigma^2 > 0$, the norm*

$$\|\eta\|_a^2 = \inf\left\{\|v\|_V^2 + \frac{1}{\sigma^2}|\rho|_a^2 : \eta = \Phi^{(a)}(v, \rho)\right\} \tag{12.5}$$

defines a new Riemannian metric on M. (Note that we are using double bars instead of single ones to distinguish between the new metric on M and the initial one.)

Proof Because $\Phi^{(a)}$ is onto (e.g., $\Phi^{(a)}(0,\rho)=\rho$), $\|\eta\|_a$ is finite (and bounded by $|\eta|_a/\sigma^2$). Define $V_a=(\Phi^{(a)})^{-1}(0)$. It is a linear subspace of $V\times T_aM$ and $\|\eta\|_a^2$ is the norm of the linear projection of $(0,\eta)$ on V_a, for the Hilbert structure on $V\times T_aM$ defined by

$$\|(v,\rho)\|_{\mathrm{id},a}^2=\|v\|_V^2+\frac{1}{\sigma^2}\,|\rho|_a^2\,.$$

Thus $\|\cdot\|_a=\left\|\pi_{V_a}(0,\eta)\right\|_{\mathrm{id},a}$ is associated with an inner product. Since it is a projection on a closed subspace ($\Phi^{(a)}$ is continuous), the infimum is attained and by definition, cannot vanish unless $\eta=0$. This therefore provides a new Riemannian metric on M. □

With this metric, the energy of a curve is

$$\begin{aligned}E(a(t))&=\int_0^1\|\partial_t a(t)\|_{a(t)}^2\,dt\\&=\inf_v\left(\int_0^1\|v(t)\|_V^2\,dt+\frac{1}{\sigma^2}\int_0^1|\partial_t a(t)-v(t)\cdot a(t)|_{a(t)}^2\,dt\right),\end{aligned}\tag{12.6}$$

the infimum being over all curves $t\mapsto v(t)$ on V. It can also be written

$$E(a(t))=\inf_v\left(\int_0^1\|v(t)\|_V^2\,dt+\frac{1}{\sigma^2}\int_0^1\left|dA_{\varphi(t)}(\alpha(t))\,\partial_t\alpha(t)\right|_{a(t)}^2dt\right)\tag{12.7}$$

with $\partial_t\varphi=v\circ\varphi$.

The distance between two elements a_0 and a_1 in M can therefore be computed by minimizing

$$U(v,a)=\int_0^1\|v(t)\|_V^2\,dt+\frac{1}{\sigma^2}\int_0^1|\partial_t a(t)-v(t)\cdot a(t)|_{a(t)}^2\,dt\tag{12.8}$$

over all curves $((v(t),a(t)),t\in[0,1])$ on $V\times M$, with boundary conditions $a(0)=a_0$ and $a(1)=a_1$ (no condition on v). From (12.7), this may also be seen as finding an optimal metamorphosis, by minimizing

$$\tilde U(\varphi,\alpha)=\int_0^1\left\|(dR_{\varphi(t)}(\mathrm{id}))^{-1}\partial_t\varphi(t)\right\|_V^2dt+\frac{1}{\sigma^2}\int_0^1\left|dA_{\varphi(t)}(\alpha(t))\partial_t\alpha(t)\right|_{\alpha(t)}^2dt$$

with boundary conditions $\varphi(0)=\mathrm{id}_G$, $\alpha(0)=a_0$, $\varphi(1)\cdot\alpha(1)=a_1$.

This construction can also be interpreted using a Riemannian submersion (Sect. B.6.7). Indeed, the mapping

$$\begin{aligned} \pi : \quad & G \times M \to M \\ & (\varphi, \alpha) \mapsto \varphi \cdot \alpha \end{aligned}$$

is a submersion from $G \times M$ to M with $d\pi(\varphi, \alpha)(h, \rho) = dA_\varphi(\alpha)\rho + dR_\alpha(\varphi)h$. ($\pi$ is a submersion because $dA_\varphi(\alpha)$ is invertible, with inverse $dA_{\varphi^{-1}}(\varphi \cdot \alpha)$.) The vertical space at (φ, α) is the set of all pairs (h, ρ) satisfying

$$\rho = -dA_{\varphi^{-1}}(\varphi \cdot \alpha) dR_\alpha(\varphi)h. \tag{12.9}$$

Two pairs (φ, α) and $(\tilde{\varphi}, \tilde{\alpha})$ belong to the same fiber if and only if $\varphi \cdot \alpha = \tilde{\varphi} \cdot \tilde{\alpha}$, or, letting $\psi = \varphi^{-1}\tilde{\varphi}$, $(\tilde{\varphi}, \tilde{\alpha}) = (\varphi\psi, \psi^{-1} \cdot \alpha)$. This leads us to introduce the right action of G on $G \times M$ defined by

$$(\varphi, \alpha) \cdot \psi = (\varphi\psi, \psi^{-1} \cdot \alpha) =: \tilde{R}_\psi(\varphi, \alpha),$$

which provides a transitive action on the fiber over $a = \varphi \cdot \alpha$. Assume that a Riemannian metric $\|(h, \rho)\|_{(\varphi,\alpha)}$ is given on $G \times M$, and that this metric is invariant through this action, so that

$$d\tilde{R}_\psi(\varphi, \alpha) : (h, \rho) \mapsto (dR_\psi(\varphi)h, dA_{\psi^{-1}}(\alpha)\rho)$$

is an isometry between $T_{(\varphi,\alpha)}(G \times M)$ and $T_{(\varphi,\alpha)\cdot\psi}(G \times M)$. Notice that this isometry maps vertical spaces onto vertical spaces. Indeed, take $(h, \rho) \in T_{(\varphi,\alpha)}G \times M$, so that h and ρ satisfy (12.9). Then

$$\begin{aligned} dA_{\psi^{-1}}(\alpha)\rho &= -dA_{\psi^{-1}}(\alpha)\, dA_{\varphi^{-1}}(\varphi \cdot \alpha)\, dR_\alpha(\varphi)h \\ &= -dA_{\psi^{-1}\varphi^{-1}}(\varphi \cdot \alpha)\, dR_\alpha(\varphi)h \\ &= -dA_{\psi^{-1}\varphi^{-1}}(\varphi \cdot \alpha)\, dR_\alpha(\varphi)\, dR_{\psi^{-1}}(\varphi\psi)\, dR_\psi(\varphi)h \\ &= -dA_{(\varphi\psi)^{-1}}(\varphi \cdot \alpha)\, dR_{\psi^{-1}\alpha}(\varphi\psi)\, dR_\psi(\varphi)h, \end{aligned}$$

in which we have applied the chain rule to the identities $A_{\psi^{-1}}A_{\varphi^{-1}} = A_{\psi^{-1}\varphi^{-1}}$, $R_{\psi^{-1}}R_\psi = \mathrm{id}$ and $R_\alpha R_{\psi^{-1}} = R_{\psi^{-1}\alpha}$. This shows that $d\tilde{R}_\psi(\varphi, \alpha)(h, \rho)$ belongs to the vertical space at $(\varphi\psi, \psi^{-1}\alpha)$.

Because $d\tilde{R}_\psi$ is an isometry that maps vertical spaces to vertical spaces, it also maps horizontal spaces to horizontal spaces. These spaces being isometric shows that π is a Riemannian submersion, provided that the norm on M is defined by

$$\|\eta\|_a = \min\left\{\|(h, \rho)\|_{(\varphi,\alpha)} : (h, \rho) \in T_\varphi G \times T_\alpha M, dA_\varphi(\alpha)\rho + dR_\alpha(\varphi)h = \eta\right\}$$

and this definition does not depend on (φ, α) such that $\varphi \cdot \alpha = a$. One can, in particular, take $(\varphi, \alpha) = (\mathrm{id}, a)$, yielding

$$\|\eta\|_a = \min\left\{\|(v, \rho)\|_{(\mathrm{id},a)} : (v, \rho) \in V \times T_a M, \rho + v \cdot a = \eta\right\}.$$

The construction leading to (12.5) simply corresponds to the special case

$$\|(v,\rho)\|^2_{(\mathrm{id},a)} = \|v\|^2_v + \frac{1}{\sigma^2}|\rho|^2_a$$

(which specifies $\|(h,\rho)\|_{(\varphi,\alpha)}$ everywhere because right translations are isometries). There is no particular reason for restricting to this special case beside it leading to simpler formulas, and the discussion that follows can easily be extended to the general case of a right-invariant metric on $G \times M$.

12.3 Euler–Lagrange Equations

We provide below optimality conditions in a Lagrangian setting that include metamorphosis in the case when M is a vector space. When M is a general Riemannian manifold, the optimality conditions were worked out in [282] and are expressed as follows (we refer to this reference for a proof of the statement)

$$\begin{cases} \partial_t a - v \cdot a = z \\ \nabla_{\partial_t a} z + \nabla^\dagger_z X^v = 0 \\ v = \dfrac{1}{\sigma^2}\mathbb{K}\xi^*_a v. \end{cases} \tag{12.10}$$

Here, ∇ is the covariant derivative on M and $\nabla_{\partial_t a} = D/Dt$ is the evaluation of this derivative along the curve $t \mapsto a(t) \in M$ (recall that $z(t) \in T_{a(t)}M$); X^v is the vector field $a \mapsto v \cdot a$ on M, and $\nabla^\dagger_z$ is defined by

$$\big\langle Z\,,\ \nabla_{Z'} Z'' \big\rangle = -\big\langle \nabla^\dagger_Z Z''\,,\ Z' \big\rangle,$$

where $\big\langle X\,,\ Y \big\rangle$ is the function $a \mapsto \big\langle X(a)\,,\ Y(a) \big\rangle_a$ for vector fields X, Y on M. Finally, $\mathbb{K}$ is, as usual, the inverse duality operator on V.

We now assume that M is a vector space and consider a generalized version of (12.8) minimizing

$$U(v,a) = \int_0^1 F\big(v(t), a(t), \partial_t a - v(t)\cdot a(t)\big)dt \tag{12.11}$$

for some function F. Let

$$z(t) = \partial_t a - v(t)\cdot a(t). \tag{12.12}$$

We will denote by $\partial_v F$, $\partial_a F$ and $\partial_z F$ the partial differentials of F with respect to each of its variables. Computing the variation with respect to v, we get, for all $t \mapsto h(t) \in V$,

$$\int_0^1 \Big(\partial_v F \,\Big|\, h(t)\Big)dt - \int_0^1 \Big(\partial_z F \,\Big|\, h(t)\cdot a(t)\Big)dt = 0.$$

Denoting by $Q_a\partial_z F$ the linear form

$$\Big(Q_a\partial_z F \,\Big|\, \tilde{h}\Big) := \Big(\partial_z F \,\Big|\, \tilde{h}\cdot a\Big),$$

we obtain the optimality equation for v

$$\partial_v F - Q_{a(t)}\partial_z F = 0. \tag{12.13}$$

If we now make the variation with respect to a, the result is, for every $t \mapsto \alpha(t) \in M$,

$$\int_0^1 \Big(\partial_a F \,\Big|\, \alpha(t)\Big)dt + \int_0^1 \Big(\partial_z F \,\Big|\, \partial_t\alpha - v(t)\cdot\alpha(t)\Big)dt = 0.$$

Integrating by parts and using the notation

$$\Big(\tilde{Q}_v\partial_z F \,\Big|\, \tilde{\alpha}\Big) := \Big(\partial_z F \,\Big|\, v\cdot\tilde{\alpha}\Big),$$

we obtain the optimality equation

$$-\partial_t\partial_z F - \tilde{Q}_v\partial_z F + \partial_a F = 0. \tag{12.14}$$

Equations (12.12)–(12.14) provide the Euler–Lagrange equations for metamorphosis. They also provide the differentials of the energy with respect to v and a and can be used to design minimization algorithms.

12.4 Application to Labeled Point Sets

We consider here diffeomorphisms acting on collections of points

$$a = (y^{(1)}, \dots, y^{(N)}),$$

with $y^{(k)} \in \mathbb{R}^d$. We therefore have $M = (\mathbb{R}^d)^N$ (note that we are not assuming here that points are distinct). We consider the function [51, 189]

$$F(v, a, z) = \|v\|_V^2 + \frac{1}{\sigma^2}\sum_{k=1}^N |z^{(k)}|^2,$$

where V is an admissible RKHS of vector fields. (Note that F does not depend on a.) Here, because $\varphi \cdot a = (\varphi(y^{(1)}), \ldots, \varphi(y^{(N)}))$, we have

$$v \cdot a = (v(y^{(1)}), \ldots, v(y^{(N)})),$$

so that $z^{(k)} = \partial_t y^{(k)} - v(y^{(k)})$. We can therefore write

$$U(v, a) = \int_0^1 \|v(t)\|_V^2 \, dt + \frac{1}{\sigma^2} \sum_{k=1}^N \int_0^1 \left|\partial_t y^{(k)} - v(t, y^{(k)}(t))\right|^2 dt.$$

We have $\partial_v F = 2\mathbb{L}v$ (where $\mathbb{L}$ is the duality operator of V) and $\partial_z F = (2/\sigma^2)(z^{(1)}, \ldots, z^{(N)})$. Moreover,

$$\left(Q_a \partial_z F \,\middle|\, h\right) = \frac{2}{\sigma^2} \sum_{k=1}^N (z^{(k)})^T h(y^{(k)}),$$

so that the first Euler–Lagrange equation is

$$\mathbb{L}v - \frac{1}{\sigma^2} \sum_{k=1}^N z^{(k)} \delta_{y^{(k)}} = 0.$$

sFor the second equation, we write

$$\left(\tilde{Q}_v \partial_z F \,\middle|\, \alpha\right) = \frac{2}{\sigma^2} \sum_{k=1}^N (z^{(k)})^T dv(y^{(k)}) \alpha_k,$$

yielding

$$-\partial_t z^{(k)} - dv(y^{(k)})^T z^{(k)} = 0.$$

This provides the system of Euler–Lagrange equations for labeled point-set metamorphosis:

$$\begin{cases} \mathbb{L}v = \dfrac{1}{\sigma^2} \displaystyle\sum_{k=1}^N z^{(k)} \delta_{y^{(k)}}, \\ \partial_t z^{(k)} + dv(y^{(k)})^T z^{(k)} = 0, \\ \partial_t y^{(k)} - v(y^{(k)}) = z^{(k)}. \end{cases} \tag{12.15}$$

Note that, introducing the reproducing kernel of V, the first equation is equivalent to

$$v(t, x) = \sum_{k=1}^N K(x, y^{(k)}(t)) z^{(k)}(t).$$

This implies that minimizing E is equivalent to minimizing

$$\tilde{E}(a,z) = \sum_{k,l=1}^{N} \int_0^1 z^{(k)}(t)^T K(y^{(k)}(t), y^{(l)}(t)) z^{(l)}(t) dt + \frac{1}{\sigma^2} \sum_{k=1}^{N} \int_0^1 \left| \partial_t y^{(k)}(t) - \sum_{l=1}^{n} K(y^{(k)}(t), y^{(l)}(t)) z^{(l)}(t) \right|^2 dt.$$

The vectors $z^{(1)}, \ldots, z^{(N)}$ can be computed explicitly given the trajectories $y^{(1)}, \ldots, y^{(N)}$, namely

$$z = (S(y) + \lambda I)^{-1} \partial_t y.$$

This provides an expression of the energy in terms of $y^{(1)}, \ldots, y^{(N)}$ that can be minimized directly, as proposed in [51]. Most of the methods developed for point sets in Chap. 10 can in fact be adapted to this new framework. Figure 12.1 provides examples of deformations computed with this method.

12.5 Application to Images

12.5.1 Formal Analysis

Let M be a set of square integrable and differentiable functions $a : \mathbb{R}^d \to \mathbb{R}$. We let $G = \mathrm{Diff}_V$ act on M by $ga = a \circ g^{-1}$. We use the L^2 norm as the initial metric on M, and we start with a formal discussion. Since we assume that a is differentiable, we can write $v \cdot a = -\nabla a^T v$. We then define

$$F(v,a,z) = \|v\|_V^2 + \frac{1}{\sigma^2} \|z\|_2^2,$$

which, here again, does not depend on a.

Equation (12.12) is $z(t) = \partial_t a(t) + \nabla a(t)^T v(t)$. We have

$$\left(Q_a \partial_z F \,\middle|\, h \right) = -\frac{2}{\sigma^2} \int_{\mathbb{R}^d} z \nabla a^T h dx,$$

so that (12.13) is $\mathbb{L} v = -z \nabla a dx$. Also,

$$\left(\tilde{Q}_v \partial_z F \,\middle|\, \alpha \right) = -\frac{2}{\sigma^2} \int_{\mathbb{R}^d} z v^T \nabla a dx,$$

so that, using the divergence theorem, $\tilde{Q}_v \partial_z F = \mathrm{div}(zv)$ and (12.14) is

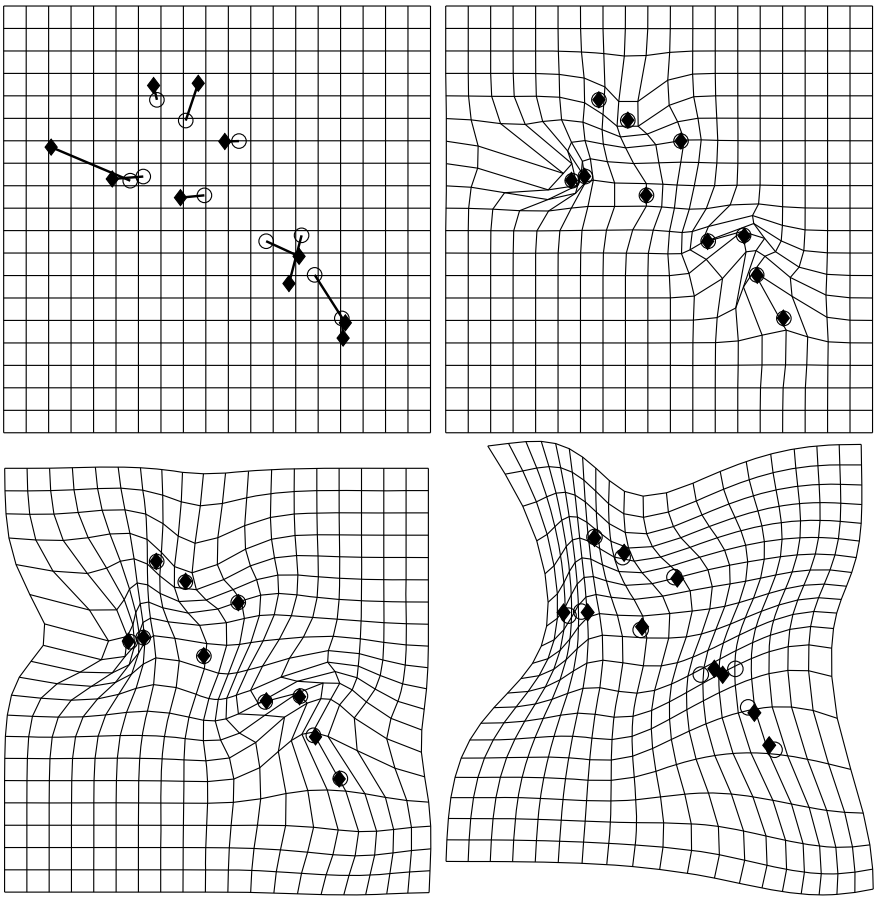

Fig. 12.1 Point-set matching using metamorphosis, with the same input as in Figs. 9.1 and 10.1, using Gaussian kernels $K(x, y)$=exp$(-|x-y|^2/2\sigma^2)$ with $\sigma = 1, 2, 4$ in grid units

$$-\partial_t z - \mathrm{div}(zv) = 0.$$

So the optimality equations for image metamorphosis are

$$\begin{cases} \partial_t a + \nabla a^T v = z, \\ \partial_t z + \mathrm{div}(zv) = 0, \\ \mathbb{L}v = -\dfrac{1}{\sigma^2}(z\nabla a)dx. \end{cases} \tag{12.16}$$

Figures 12.2, 12.3 and 12.4 provide examples of images matched using the associated energy. In these examples, the first and last images are given as input and two interpolated images are provided. The numerical scheme is described in Sect. 12.5.4.

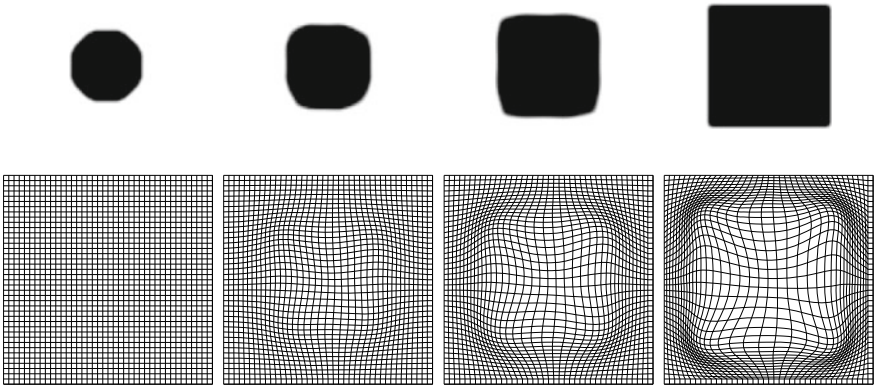

Fig. 12.2 Image metamorphosis. Estimation of a geodesic between a disc and a square. First row: image evolution in time. Second row: evolution of the diffeomorphism

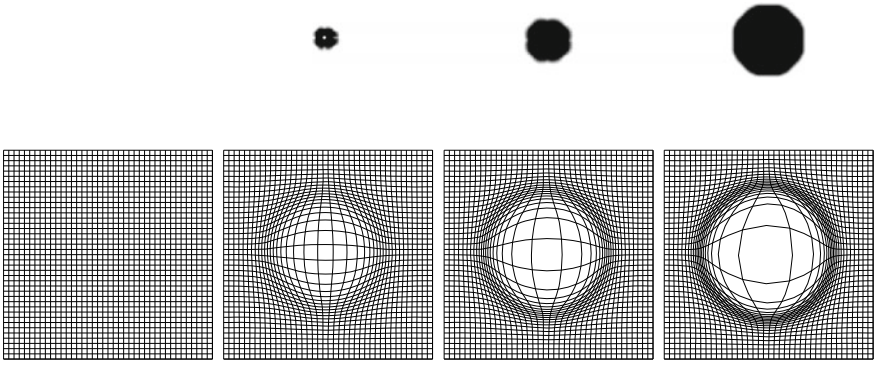

Fig. 12.3 Creation of new patterns with image metamorphosis. Geodesic between an empty image and a disc. First row: image evolution in time. Second row: evolution of the diffeomorphism

12.5.2 Some Rigorous Results

In this section, we provide a more rigorous treatment of image metamorphosis. We will extend the L^2 case considered above to Sobolev spaces, assuming that images (a or α) belong to a Hilbert space H, with norm equivalent to the $H^r(\mathbb{R}^d)$ norm for some integer $r \geq 0$, with notation for the H^r norm

$$\|u\|_{r,2}^2 = \sum_{|k|\leq r} \|\partial_k u\|_2^2,$$

where k denotes a d-dimensional multi-index $(k_1, \ldots, k_d)$, $|k| = k_1 + \cdots + k_d$,

$$\partial_k u = \partial_1^{k_1} \ldots \partial_d^{k_d} u$$

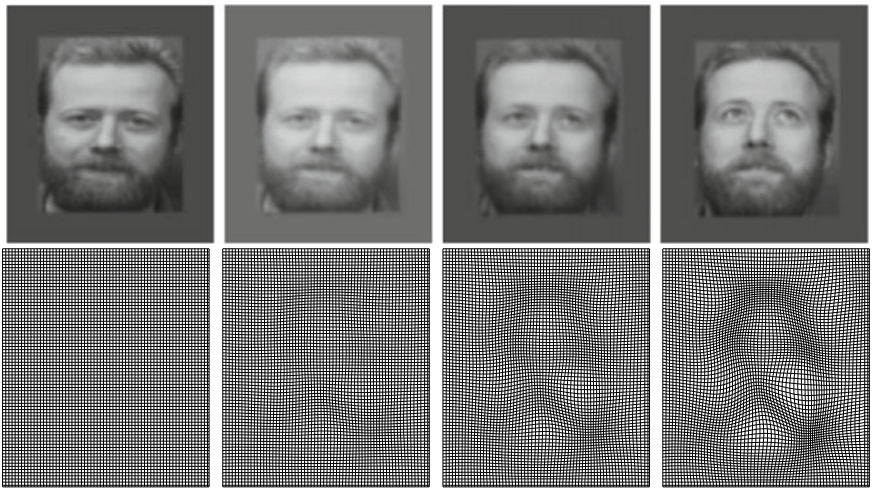

Fig. 12.4 Gray-level image metamorphosis. Geodesic between two gray-level face images. First row: image evolution in time. Second row: evolution of the diffeomorphism

and $\|\cdot\|_2$ is the L^2 norm. We will consider metamorphoses with cost

$$F(v, a, z) = \|v\|_V^2 + \frac{1}{\sigma^2}\|z\|_H^2$$

and provide a few results describing their properties. The proofs of these results being somewhat technical, we will skip them, referring the reader to [245] for more details.

We first reformulate the problem to make sure that it is also well defined for functions a that are not differentiable. We can write the advection equation $\partial_t a + \nabla a^T v = z$ as $\partial_t \alpha = z \circ \varphi$ with $\alpha = a \circ \varphi$ and $\partial_t \varphi = v \circ \varphi$. We can therefore consider the minimization of

$$\int_0^1 \|v\|_V^2 \, dt + \frac{1}{\sigma^2} \int_0^1 \|z(t)\|_H^2 \, dt \tag{12.17}$$

subject to the constraints $\partial_t \varphi = v \circ \varphi$, $\partial_t \alpha = z \circ \varphi$, $\varphi(0) = \mathrm{id}$, $\alpha(0) = a_0$ and $\alpha(1) = a_1 \circ \varphi(1)$.

We then have the following theorem.

Theorem 12.3 *Assume $r \geq 0$ and $p \geq \max(1, r)$. Then the image metamorphosis problem has at least one solution.*

The validity of the optimality conditions requires slightly more restrictive assumptions on the boundary conditions a_0 and a_1. We let $\mathbb{K}_V : V^* \to V$ and $\mathbb{K}_H : H^* \to H$ be the duality operators of the Hilbert spaces V and H and define $\xi_\varphi v = v \circ \varphi$ and $\tilde{\xi}_\varphi \alpha = \alpha \circ \varphi$. Then, the following result holds.

Theorem 12.4 *Assume that both a_0 and a_1 belong to $H^{r+1}(\mathbb{R}^d)$. Then, if (v, z, φ, α) is an optimal solution of the metamorphosis problem, there exist $\rho_\varphi(\cdot) \in C_0^p(\mathbb{R}^d, \mathbb{R}^d)^*$ and $\rho_\alpha \in H^*$ such that the system*

$$\begin{cases} \partial_t \varphi = v \circ \varphi \\ \partial_t \alpha = z \circ \varphi \\ \partial_t \rho_\varphi = -\partial_\varphi\big(\rho_\varphi \,\big|\, v \circ \varphi\big) - \partial_\varphi(\rho_\alpha \,|\, z \circ \varphi) \\ \partial_t \rho_\alpha = 0 \\ v = \mathbb{K}_V \xi_\varphi^* \rho_\varphi \\ z = \sigma^2 \mathbb{K}_H \tilde{\xi}_\varphi^* \rho_\alpha \end{cases} \tag{12.18}$$

is satisfied, with boundary conditions $\alpha(0) = a_0$, $\alpha(1) = a_1 \circ \varphi(1)$ and

$$\big(\rho_\varphi(1) \,\big|\, \delta\varphi\big) + \big(\rho_\alpha \,\big|\, \nabla a_1 \circ \varphi(1)^T \delta\varphi\big) = 0$$

for all $\delta\varphi \in C_0^p(\mathbb{R}^d, \mathbb{R}^d)$.

Moreover, the boundary condition propagates, so that

$$\big(\rho_\varphi(t) \,\big|\, \delta\varphi\big) + \big(\rho_\alpha \,\big|\, \nabla a(t) \circ \varphi(t)^T \delta\varphi\big) = 0 \tag{12.19}$$

for all $t \in [0, 1]$.

(Equation (12.19) is in fact the horizontality condition associated with geodesics obtained through the Riemannian submersion.)

Finally, the following theorem provides sufficient conditions for the existence of solutions of (12.18) with given initial conditions.

Theorem 12.5 *Assume that $p \geq 1 + d/2$ and $p \geq r + 1$. Then system (12.18) has a unique solution over any bounded interval as soon as $\rho_{\varphi,0} \in C_0^{p-2}(\mathbb{R}^d, \mathbb{R}^d)^*$ and $\rho_\alpha \in H^{r-1}(\mathbb{R}^d)^*$.*

Note that, with metamorphosis, the boundary condition requires that $\big(\rho_{\varphi,0} \,\big|\, w\big) = \big(\rho_\alpha \,\big|\, \nabla a_0^T w\big)$. Assuming that $a_0 \in H^1(\mathbb{R}^d)$ (which is restrictive only for $r = 0$), we see that $\rho_\alpha \in H^{r-1}(\mathbb{R}^d)^*$ implies that $\rho_{\varphi,0} \in C_0^{r-1}(\mathbb{R}^d, \mathbb{R}^d)^* \subset C_0^{p-2}(\mathbb{R}^d, \mathbb{R}^d)^*$, since $p \geq r + 1$, so that the regularity condition for $\rho_{\varphi,0}$ is automatically satisfied.

12.5.3 Remarks on the Optimality Conditions

System (12.18) corresponds to Pontryagin's maximum principle for (12.17) considered as an optimal control problem with state (φ, α), control (v, z) and co-state $(\rho_\varphi, \rho_\alpha)$. We first check that, under additional differentiability assumptions on the images, they are equivalent to those found in (12.16) in the L^2 case.

When $H = L^2(\mathbb{R}^d)$, system (12.18) gives $z = \sigma^2 \tilde{\xi}^*_\varphi \rho_\alpha$ with $\rho_\alpha \in (L^2(\mathbb{R}^d))^* = L^2(\mathbb{R}^d)$ (meaning that we identify $\rho_\alpha dx$ with ρ_α). We therefore have $\partial_t(z \mid \tilde{a} \circ \varphi) = 0$ for any $\tilde{a} \in \mathbb{R}^2$. When z is differentiable, this leads, after differentiating and applying the divergence theorem, to

$$\partial_t z + \operatorname{div}(zv) = 0.$$

We have already seen that, when a is differentiable, the equation $\partial_t \alpha = z \circ \varphi$ implies that $\partial_t a + \nabla a^T v = z$.

Equation (12.19) applied to $\delta\varphi = w \circ \varphi(t)$ gives $\big(\rho_\varphi(t) \,\big|\, w \circ \varphi(t)\big) = -\big(\rho_\alpha \,\big|\, (\nabla a(t)^T w) \circ \varphi(t)\big)$, or

$$\xi^*_{\varphi(t)} \rho_\varphi(t) = -(\tilde{\xi}^*_{\varphi(t)} \rho_\alpha) \nabla a(t) = -\frac{1}{\sigma^2} z(t) \nabla a(t).$$

This yields $\mathbb{L}_V v(t) = -z(t)\nabla a(t)/\sigma^2$ at all times (or $-(z(t)\nabla a(t))dx/\sigma^2$ if we relax the identification between L^2 and its dual), and we retrieve the last equation in (12.16).

We now return to the general case. The second and last equations in (12.18) imply that $\partial_t \alpha = \sigma^2 \tilde{\xi}_\varphi \mathbb{K}_H \tilde{\xi}^*_\varphi \rho_\alpha$, yielding

$$\alpha(t) = \alpha(0) + \sigma^2 \left(\int_0^t \tilde{\xi}_{\varphi(s)} \mathbb{K}_H \tilde{\xi}^*_{\varphi(s)} \, ds \right) \rho_\alpha.$$

From the boundary conditions, we therefore get

$$\rho_\alpha = \left(\int_0^1 \tilde{\xi}_{\varphi(s)} \mathbb{K}_H \tilde{\xi}^*_{\varphi(s)} \, ds \right)^{-1} (a_1 \circ \varphi(1) - a_0).$$

We therefore have

$$\alpha(t) = a_0 + \left(\int_0^t \tilde{\xi}_{\varphi(s)} \mathbb{K}_H \tilde{\xi}^*_{\varphi(s)} \, ds \right) \left(\int_0^1 \tilde{\xi}_{\varphi(s)} \mathbb{K}_H \tilde{\xi}^*_{\varphi(s)} \, ds \right)^{-1} (a_1 \circ \varphi(1) - a_0),$$

which provides a "closed-form" expression of the template part of the metamorphosis given the diffeomorphism part. When $H = L^2(\mathbb{R})$, for which $\mathbb{K}_H = \mathrm{Id}$, we have $\xi^*_\varphi \rho = \rho \circ \varphi^{-1} \,|\det(d(\varphi^{-1}))|$, so that

$$\xi_\varphi \xi^*_\varphi \rho = \rho \,|\det(d(\varphi^{-1}))| \circ \varphi = \rho \,|\det(d\varphi)|^{-1}.$$

It follows that

$$\alpha(t) = a_0 + \left(\frac{\int_0^t |\det(d\varphi(s))|^{-1} \, ds}{\int_0^1 |\det(d\varphi(s))|^{-1} \, ds} \right) (a_1 \circ \varphi(1) - a_0),$$

or

$$a(t) \circ \varphi(t) = \left(\frac{\int_t^1 |\det(d\varphi(s))|^{-1}\, ds}{\int_0^1 |\det(d\varphi(s))|^{-1}\, ds} \right) a_0 + \left(\frac{\int_0^t |\det(d\varphi(s))|^{-1}\, ds}{\int_0^1 |\det(d\varphi(s))|^{-1}\, ds} \right) a_1 \circ \varphi(1),$$

which describes how metamorphosis interpolates between the two images it compares.

When $r > d/2 + 1$, (12.18) for H^r metamorphoses has some interesting singular solutions [152, 245]. In this case, H is a reproducing kernel Hilbert space, and we will use K_H to denote its kernel. We look for solutions of (12.18) in which ρ_φ and ρ_α take the form

$$\rho_\varphi(t) = \sum_{k=1}^N \beta_k(t) \delta_{x_k^{(0)}}, \tag{12.20}$$

$$\rho_\alpha = \sum_{k=1}^N \gamma_k \delta_{x_k^{(0)}}, \tag{12.21}$$

where $x^{(0)} = \{x_k^{(0)}\}_{k=1}^N$ is a collection of points in $\mathbb{R}^d$, $\beta(t) = \{\beta_k(t)\}_{k=1}^N$ is a collection of time-dependent vectors in $\mathbb{R}^d$, and $\gamma = \{\gamma_k\}_{k=1}^N$ is a time-independent collection of scalars.

Introduce the trajectories $x_k(t) := \varphi(t, x_k^{(0)})$. Using this notation, we have

$$\left(\xi^*_{\varphi(t)} \rho_\varphi(t) \;\middle|\; w\right) = \left(\rho_\varphi(t) \;\middle|\; w \circ \varphi(t)\right) = \sum_{k=1}^N \beta_k(t)^T w(x_k(t)),$$

so that

$$\xi^*_{\varphi(t)} \rho_\varphi(t) = \sum_{k=1}^N \beta_k(t) \delta_{x_k(t)}$$

and (12.18) implies that

$$v(t, \cdot) = \sum_{k=1}^N K_V(\cdot, x_k(t)) \beta_k(t).$$

Similarly, one gets

$$z(t, \cdot) = \sigma^2 \sum_{k=1}^N K_H(\cdot, x_k(t)) \gamma_k.$$

The third equation in (12.18) gives, for $w \in C_0^p(\mathbb{R}^d, \mathbb{R}^d)$,

$$\sum_{k=1}^{N} \partial_t \beta_k(t)^T w(x_k^{(0)}) = \\ -\sum_{k=1}^{N} \beta_k(t)^T (dv(x_k(t))w(x_k^{(0)})) - \sum_{k=1}^{N} \gamma_k \nabla z(x_k(t))^T w(x_k^{(0)})$$

from which we get

$$\partial_t \beta_k(t) = -dv(x_k(t))^T \beta_k(t) - \gamma_k \nabla z(x_k(t)).$$

Using the expansions of v and z and the fact that $\partial_t x_k = v(t, x_k)$, we obtain the fact that (12.20) and (12.21) provide solutions of (12.18) as soon as x, α and z satisfy the coupled dynamical system

$$\begin{cases} \partial_t x_k(t) = \sum_{l=1}^{N} K_V(x_k(t), x_l(t))\beta_l(t) \\ \partial_t \alpha_k(t) = \sum_{l=1}^{N} K_H(x_k(t), x_l(t))\gamma_l \\ \partial_t \beta_k(t) = -\sum_{l=1}^{N} \nabla_1 K_V(x_k(t), x_l(t))\beta_l(t)^T \beta_k(t) \\ \qquad\qquad - \frac{1}{\sigma^2} \sum_{l=1}^{N} \nabla_1 K_H(x_k(t), x_l(t))\gamma_k \gamma_\ell. \end{cases} \tag{12.22}$$

Finally, we note that Eq. (12.19) applied to ρ_φ and ρ_α is

$$\sum_{k=1}^{N} \beta_k(t)^T \delta\varphi(x_k) = -\sum_{k=1}^{N} \gamma_k \nabla\alpha(t, x_k^{(0)})^T d\varphi(t, x_k^{(0)})^{-1} \delta\varphi(x_k)$$

for all $\delta\varphi$, yielding

$$\beta_k(t) = -\gamma_k d\varphi(t, x_k^{(0)})^{-T} \nabla\alpha(x_k^{(0)}) = -\gamma_k \nabla a(t, x_k(t)).$$

These special solutions have been used in [245] to provide approximations of solutions for metamorphoses between smooth images.

12.5.4 Discretization of Image Metamorphosis in the L^2 Case

The images in Figs. 12.2 and 12.3 have been obtained after minimization of a discretized version of the image metamorphosis energy,

$$E(a,v) = \int_0^1 \|v(t)\|_V^2 dt + \frac{1}{\sigma^2}\int_0^1 \|\partial_t a(t) + \nabla a(t)^T v(t)\|_2^2 dt.$$

The second integrand must be discretized with care. It represents the total derivative of a along the flow. Stable results can be obtained using the following scheme.

Introduce the auxiliary variable $w = \mathbb{L}^{1/2}v$, so that $v = \mathbb{K}^{1/2}w$ where $\mathbb{K} = \mathbb{L}^{-1}$ (the numerical implementation can in fact explicitly specify $\tilde{\mathbb{K}} = \mathbb{K}^{1/2}$ and use $\mathbb{K} = \tilde{\mathbb{K}}^2$). The following discretized energy has been used in the experiments of Figs. 12.2 and 12.3:

$$E = \sum_{t=1}^{T}\sum_{x\in\tilde{\Omega}} |w(x)|^2 + \lambda\delta t^{-2}\sum_{t=1}^{T}\sum_{x\in\tilde{\Omega}} \left|a(t+1, x+\delta t(\mathbb{K}^{1/2}w)(t,x)) - a(t,x)\right|^2,$$

where x and t now are discrete variables, $\tilde{\Omega}$ a discrete grid on Ω and δt the time discretization step (the space discretization is 1). The optimization algorithm alternates a few steps of nonlinear conjugate gradient in v, and a few steps of linear conjugate gradient in a [120, 121].

12.6 Applications to Densities

If one denotes $H^r(\mathbb{R}^d)^*$ by $H^{-r}(\mathbb{R}^d)$ for $r \geq 0$, one can consider the action of Diff_0^p on these spaces defined by

$$(\varphi \cdot a \mid f) = (a \mid f \circ \varphi),$$

which is well defined on $H^{-r}(\mathbb{R}^d)$ for $r \leq p$. For $r = 0$, in particular, this action boils down to the usual action on densities $\varphi \cdot a = a \circ \varphi^{-1} \,|\det(d(\varphi^{-1}))|$. Starting with this special case, for which the infinitesimal action is $v \cdot a = -\mathrm{div}(av)$, we can consider the metamorphosis problem associated here again with

$$F(v,a,z) = \|v\|_V^2 + \frac{1}{\sigma^2}\|z\|_2^2,$$

with $z(t) = \partial_t a(t) + \mathrm{div}(a(t)v(t))$. The counterpart of system (12.16) is (details being left to the reader)

$$\begin{cases} \partial_t a + \operatorname{div}(av) = z, \\ \partial_t z + \nabla z^T v = 0, \\ \mathbb{L}v = \dfrac{1}{\sigma^2}(a\nabla z)dx. \end{cases} \tag{12.23}$$

More generally, letting $H = H^r(\mathbb{R}^d)$, we can define metamorphoses in H^* using

$$F(v, a, z) = \|v\|_V^2 + \frac{1}{\sigma^2}\|z\|_{H^*}^2.$$

We can also consider the metamorphosis problem in optimal control form as in (12.17), minimizing

$$\int_0^1 \|v\|_V^2 \, dt + \frac{1}{\sigma^2}\int_0^1 \|z(t)\|_{H^*}^2 \, dt \tag{12.24}$$

subject to the constraints $\partial_t \varphi = v \circ \varphi$, $\partial_t \alpha = z \circ \varphi \,|\det d\varphi|$, $\varphi(0) = \mathrm{id}$, $\alpha(0) = a_0$ and $\alpha(1) = a_1 \circ \varphi(1)|\det d\varphi(1)|$. Defining, in this case $\tilde{\xi}_\varphi z = z \circ \varphi|\det d\varphi|$, the optimality conditions (12.18) become

$$\begin{cases} \partial_t \varphi = v \circ \varphi \\ \partial_t \alpha = z \circ \varphi \,|\det d\varphi| \\ \partial_t \rho_\varphi = -\partial_\varphi\big(\rho_\varphi \,\big|\, v \circ \varphi\big) - \partial_\varphi(\rho_\alpha \,|\, z \circ \varphi \,|\det d\varphi|\,) \\ \partial_t \rho_\alpha = 0 \\ v = \mathbb{K}_V \xi_\varphi^* \rho_\varphi \\ z = \sigma^2 \mathbb{K}_H \tilde{\xi}_\varphi^* \rho_\alpha \end{cases} \tag{12.25}$$

with boundary conditions $\alpha(0) = a_0, \alpha(1) = a_1 \circ \varphi(1)|\det d\varphi(1)|$ and horizontality condition

$$\big(\xi_\varphi^* \rho_\varphi \,\big|\, w\big) + \Big(\tilde{\xi}_\varphi^* \rho_\alpha \,\Big|\, \operatorname{div}(aw)\Big) = 0$$

for all $w \in C_0^p(\mathbb{R}^d, \mathbb{R}^d)$.

12.7 Application to Curves

12.7.1 Metamorphosis on Unit Tangents

We now consider the issue of comparing plane curves based on the orientation of their tangents [171, 307, 313]. If m is a plane curve parametrized by arc length, and L is it length, we define the normalized tangent T^m by

$$T^m : \quad [0,1] \to \mathbb{R}^d$$
$$s \mapsto \partial_s m(Ls).$$

The function T^m characterizes m up to translation and scaling, and the pair (L, T^m) characterizes m up to translation. In the following discussion, functions will depend on time $t \in [0,1]$ and normalized arc length s, also in $[0,1]$. For clarity, we will let $\Omega = [0,1]$ for the arc length, i.e., write $t \in [0,1]$, $s \in \Omega$. We will consider derivatives of functions h of $(t,s) \in [0,1] \times \Omega$, and we will use the notation $\dot{h}$ or $\partial_t h$ for derivatives with respect to the time variable, t, and dh for derivatives with respect to the space variable, s.

We consider the group of diffeomorphisms Diff of Ω, which acts on the set M of measurable functions $a : \Omega \to S^{d-1}$ (the unit sphere in $\mathbb{R}^d$) by $\varphi \cdot a = a \circ \varphi^{-1}$. Choosing an RKHS V of vector fields on Ω satisfying $v(0) = v(1) = 0$, we can consider the metamorphosis problem associated with

$$F(v,a,z) = \|v\|_V^2 + \frac{1}{\sigma^2}\|z\|_2^2. \tag{12.26}$$

If V is embedded in $C_0^1(\Omega, \mathbb{R})$, the discussion made in the previous sections applies, with the slight difference that one must take "$z \in T_a M$" (we will not try here to rigorously construct the shape space as a manifold), which means that $z(s) \perp a(s)$ at all times. If, in particular, the boundary conditions a_0 and a_1 are differentiable, then Eq. (12.10) applies, and leads to the system, for a metamorphosis $a(t,s)$

$$\begin{cases} \partial_t a + v da = z, \\ \partial_t \zeta + d(v\zeta) = 0, \\ v = -\dfrac{1}{\sigma^2}\mathbb{K}_V(da^T z) \end{cases}$$

in which $\zeta(,s) : \Omega \to \mathbb{R}$ is the normal coordinate of $z(t,s)$, defined by $z(t,s) = \zeta(t,s)a^\perp(t,s)$, where $a^\perp$ is a rotated by $\pi/2$.

However, the resulting metric takes an interesting form if one considers the Hilbert space V of functions $v : [0,1] \to \mathbb{R}$ such that $v(0) = v(1) = 0$ and

$$\|v\|_V^2 = \int_\Omega (dv)^2 \, ds. \tag{12.27}$$

Notice that this Hilbert space is not embedded in $C_0^1(\Omega, \mathbb{R})$ (one needs Sobolev spaces of order larger $d/2 + 1 = 3/2$ for this, i.e., one would need a second derivative in the norm). Functions $v \in V$ are continuous and satisfy a Hölder condition of order q for any $q < 1/2$, but are not necessarily Lipschitz continuous. While the general framework we have considered so far does not apply to this situation, one can directly formulate the metamorphosis problem in terms of time-dependent diffeomorphisms, $(t,s) \mapsto \varphi(t,s)$, of Ω, letting $v = \dot{\varphi} \circ \varphi^{-1}$ so that

$$dv = \frac{d\dot{\varphi} \circ \varphi^{-1}}{d\varphi \circ \varphi^{-1}},$$

which yields, after a change of variables

$$\int_\Omega (dv)^2 \, ds = \int_\Omega \frac{d\dot{\varphi}^2}{d\varphi} \, ds.$$

Writing the second term in template form, i.e., letting $z = \dot{\alpha} \circ \varphi^{-1}$ and making another change of variable, the metamorphosis objective function in this case becomes

$$U_\sigma(\varphi, \alpha) = \int_0^1 \int_\Omega \frac{(d\dot{\varphi})^2}{d\varphi} \, ds \, dt + \frac{1}{\sigma^2} \int_0^1 \int_\Omega |\dot{\alpha}|^2 d\varphi \, ds \, dt, \tag{12.28}$$

which needs to be minimized over all trajectories $t \mapsto \varphi(t)$ and $t \mapsto \alpha(t)$, such that φ is at all times an increasing diffeomorphism of Ω and α a function from $\Omega \to S^{d-1}$, with boundary conditions $\alpha(0) = a_0$ and $\alpha(1) = a_1 \circ \varphi(1)$. We point out (this will be useful later) that this energy can be minimized explicitly with respect to α when the φ trajectory is fixed. Indeed, first notice that in order to minimize the second term in U, it suffices to minimize separately the integrals

$$\int_0^1 |\dot{\alpha}(t, s)|^2 d\varphi(t, s) dt \tag{12.29}$$

for fixed s. Considering such an integral, we write $\alpha(t, s) = \tilde{\alpha}(\lambda(t, s), s)$, where

$$\lambda(t, s) := \frac{\int_0^t d\varphi(t, s)^{-1} \, dt}{\int_0^1 d\varphi(t, s)^{-1} \, dt}$$

is an increasing function satisfying $\lambda(0, s) = 0$ and $\lambda(1, s) = 1$. We have $\dot{\alpha} = \dot{\lambda}\dot{\tilde{\alpha}}(\lambda, s)$ and

$$\int_0^1 |\dot{\alpha}(t, s)|^2 d\varphi(t, s) dt = \frac{1}{c(s)} \int_0^1 |\dot{\tilde{\alpha}}(\lambda(t, s), s)|^2 \dot{\lambda}(t, s) \, dt = \frac{1}{c(s)} \int_0^1 |\dot{\tilde{\alpha}}(t, s)|^2 \, dt$$

with

$$c(s) = \int_0^1 d\varphi(t, s)^{-1} \, dt.$$

This integral must be minimized subject to $\tilde{\alpha}(0, s) = a_0(s)$, $\tilde{\alpha}(1, s) = a_1 \circ \varphi_1(s)$ and $|\tilde{\alpha}(t, s)| = 1$ for all t, and the solution is given by the circular arc between $\tilde{\alpha}(0, s)$ and $\tilde{\alpha}(1, s)$, which can be expressed as

$$\tilde{\alpha}(t,s) = \frac{\sin((1-t)\omega(s))}{\sin\omega(s)} a_0(s) + \frac{\sin(t\omega(s))}{\sin\omega(s)} a_1 \circ \varphi_1(s),$$

where

$$\omega(s) = \arccos(a_0(s)^T a_1 \circ \varphi_1(s)).$$

The optimal α is therefore given by

$$\alpha(t,s) = \frac{\sin((1-\lambda^{-1}(t,s))\omega(s))}{\sin\omega(s)} a_0(s) + \frac{\sin(\lambda^{-1}(t,s)\omega(s))}{\sin\omega(s)} a_1 \circ \varphi_1(s), \tag{12.30}$$

where $\lambda^{-1}(t,s)$ is defined by $\lambda(\lambda^{-1}(t,s),s) = t$. The optimal cost in (12.29) is then $\omega(s)^2/c(s)$. We note that, because $\omega(s) \in [0,\pi]$, the coefficients in (12.30) are non-negative. Moreover, $\alpha(t,s)$ is at all times in the plane generated by $a_0(s)$ and $\alpha_1 = a_1 \circ \varphi_1(s)$.

We now fix φ_1 and study optimal metamorphoses with $\varphi(1,\cdot) = \varphi_1$. Introduce the vector $a_0^\perp$ perpendicular to a_0 in the plane generated by a_0 and α_1, defined by

$$\alpha_1 = \cos\omega\, a_0 + \sin\omega a_0^\perp.$$

(This is well defined if $\omega \in (0,\pi)$ and we choose $a_0^\perp$ arbitrarily otherwise.) Without loss of generality, we can search for optimal metamorphoses taking the form

$$\alpha(t,s) = \cos\tau(t,s) a_0(s) + \sin\tau(t,s) a_0^\perp(s)\,.$$

Letting $\xi(t,s) = (\cos\tau(t,s), \sin\tau(t,s)) \in S^1$, we can write $U_\sigma(\varphi,\alpha) = \tilde{U}_\sigma(\varphi,\xi)$, where

$$\tilde{U}_\sigma(\varphi,\xi) = \int_0^1 \int_\Omega \frac{(d\dot{\varphi})^2}{d\varphi}\, ds\, dt + \frac{1}{\sigma^2} \int_0^1 \int_\Omega |\dot{\xi}|^2 d\varphi\, ds\, dt. \tag{12.31}$$

This function now has to be minimized subject to $\varphi(0,\cdot) = \mathrm{id}, \varphi(1,\cdot) = \varphi_1, \xi(0,\cdot) = (1,0)$ and $\xi(1,\cdot) = (\cos\omega, \sin\omega)$, with $\omega = \arccos(a_0^T a_1 \circ \varphi(1,\cdot))$. In other terms, we have reduced the S^{d-1}-valued metamorphosis problem to an S^1-valued problem, or, equivalently, our metric on d-dimensional curves to a two-dimensional case.

We now make a second reduction that will simplify the problem. Because $\xi(t,s)$ is differentiable in time, one can define uniquely a differentiable function $\tau(t,s)$ such that $\xi(0,s) = (1,0)$ and $\xi(t,s) = (\cos\tau(t,s), \sin\tau(t,s))$ at all times. Define $q(t,s)$ by

$$q(t,s) = \sqrt{d\varphi(t,s)}\,(\cos\eta(t,s), \sin\eta(t,s))\,,$$

with $2\sigma\eta(t,s) = \tau(t,s)$. Then, a straightforward computation yields

$$4|\dot{q}(t,s)|^2 = \frac{d\dot{\varphi}^2}{d\varphi} + \frac{1}{\sigma^2}|\dot{\xi}(t,s)|^2,$$

so that

$$\tilde{U}_\sigma(\varphi,\xi) = 4\int_0^1 \|\dot{q}(t,\cdot)\|_2^2\, dt. \tag{12.32}$$

We also note that

$$\|q(t,\cdot)\|_2^2 = \int_\Omega d\varphi(t,s)\, ds = 1,$$

so that $q(t,\cdot)$ is a curve on the unit sphere of $L^2(\Omega,\mathbb{R}^2)$. This implies that its energy, $\int_0^1 \|\dot{q}(t,\cdot)\|_2^2\, dt$, cannot be larger than that of the minimizing geodesic on this unit sphere, which is the shortest great circle connecting the functions $q_0 = (1,0)$ (which is constant) and $q_1 = \sqrt{d\varphi_1}\,(\cos(\eta(1,\cdot)), \sin(\eta(1,\cdot))$. Letting

$$\rho = \arccos\langle q_0\,,\, q_1\rangle_2 = \arccos\int_\Omega \sqrt{d\varphi_1}\cos\left(\frac{\omega(s)}{2\sigma}\right) ds,$$

this geodesic is given by

$$\gamma(t,s) = \frac{\sin((1-t)\rho)}{\sin\rho} q_0(s) + \frac{\sin(t\rho)}{\sin\rho} q_1(s) \tag{12.33}$$

with energy equal to ρ^2. We therefore find that

$$\tilde{U}_\sigma(\varphi,\xi) \geq 4\arccos^2\int_\Omega \sqrt{d\varphi_1(s)}\cos\left(\frac{\omega(s)}{2\sigma}\right) ds\,. \tag{12.34}$$

This provides a lower-bound for the metamorphosis energy. To prove that this lower-bound is achieved, one needs to show that the trajectory γ in (12.33) can be derived from a valid trajectory $(\psi(\cdot,\cdot), \mu(\cdot,\cdot))$ that connects $(\mathrm{id}, 0)$ to (φ_1, ξ_1).

We are therefore looking for representations of γ in the form

$$\gamma(t,s) = \sqrt{d\psi(t,s)}\,(\cos\tilde{\eta}(t,s), \sin\tilde{\eta}(t,s))$$

with $\tilde{\eta}(0,s) = 0$, which uniquely defines $\tilde{\eta}(t,s)$ by continuity in t. Notice that we automatically have $d\psi(0,\cdot) = 1$ and $d\psi(1,\cdot) = d\varphi_1$ by definition of q_0 and q_1. For $t \in (0,1)$, we have

$$d\psi(t,s) = |\gamma(t,s)|^2,$$

so that $\psi(t,\cdot)$ is non-decreasing and satisfies $\psi(t,0) = 0$, $\psi(t,1) = 1$. The function $s \mapsto d\psi(t,s)$ is positive if and only if $q(t,s)$ does not vanish, which requires

$$\sin^2((1-t)\rho) + \sin^2(t\rho)d\varphi(1) + 2\sin((1-t)\rho)\sin(t\rho)\sqrt{d\varphi_1}\cos\left(\frac{\omega}{2\sigma}\right) > 0. \tag{12.35}$$

A sufficient condition for this to hold for all t is that the cosine term is strictly larger than -1, which is equivalent to

$$\frac{\omega}{2\sigma} \not\equiv \pi \pmod{2\pi}. \tag{12.36}$$

Since $\omega \in [0, \pi]$, this condition will be automatically satisfied if $2\sigma > 1$.

Because $\gamma(1) = q_1$, the other end-point must satisfy

$$\tilde{\eta}(1, s) = \omega(s)/2\sigma + 2k(s)\pi, \tag{12.37}$$

where k is an integer-valued function. The right-hand side of (12.33) is a linear combination of q_0 and q_1 with positive coefficients, which implies that the time-continuous angular representation of $q(t)$ starting at 0 cannot deviate by more than π from its initial value, i.e.,

$$-\pi \leq \tilde{\eta}(1, s) \leq \pi, \tag{12.38}$$

and we also have

$$0 \leq \omega(s) \leq \pi. \tag{12.39}$$

We now (and, unless otherwise specified, for the rest of the discussion) make the assumption that $2\sigma \geq 1$. Under this assumption, $\tilde{\eta}(1, \cdot) = \omega/2\sigma$ satisfies both (12.37) and (12.38). Since it is clear that only one value of $\tilde{\eta}(1, s)$ can satisfy the two equations together, we find that, for all $s \in \Omega$, one has $\tilde{\eta}(1, s) = \omega(s)/\sigma$, and the curve γ is associated with a trajectory (ψ, β) between $(\mathrm{id}, 0)$ and (φ_1, ξ_1).

We have therefore proved that γ in (12.33) provides a valid improved solution to the original (φ, α) as soon as (12.36) is satisfied, which is true as soon as $2\sigma > 1$.

If $\sigma = 1/2$, then (12.36) may not hold for the curve in (12.33). However, the minimum of $U_\sigma(\varphi, \alpha)$ with given $\varphi(1) = \varphi_1$ is still given by the geodesic energy of this curve. To see this, it suffices to consider a small variation $\tilde{a}_0$ of a_0 such that $\tilde{a}_0^T a_1 \circ \varphi_1 > -1$, so that (12.36) is satisfied with $\tilde{a}_0$ instead of a_0, and the minimum energy when starting from $\tilde{a}_0$ is the geodesic energy of the associated great circle. One can then use the fact that U is a geodesic energy for a Riemannian metric on $\mathrm{Diff} \times M$, and combine this with the triangular inequality for the sequence of geodesics going from (id, a_0) to $(\mathrm{id}, \tilde{a}_0)$ then to $(\varphi(1), a_1)$. Indeed, the energy of the sequence is larger than the minimal energy between (id, a_0) and $(\varphi(1), a_1)$, but arbitrarily close to the energy of the minimal geodesic between $(\mathrm{id}, \tilde{a}_0)$ then to $(\varphi(1), a_1)$, itself arbitrarily close to the lower bound in (12.34).

We summarize this discussion in the following theorem.

Theorem 12.6 *Assume that* $2\sigma \geq 1$, *and let* $\varphi_1 : \Omega \to \Omega$ *satisfy* $\varphi_1(0) = 0$, $\varphi_1(1) = 1$ *and* $\partial_s \varphi_1 > 0$. *Then*

$$\inf \{U_\sigma(\varphi, \alpha) : \varphi(1) = \varphi_1, \alpha(0) = a_0, \alpha(1) = a_1 \circ \varphi_1\}$$
$$= 4 \arccos^2 \int_\Omega \sqrt{d\varphi_1} \cos\left(\frac{\arccos(a_0(s)^T a_1 \circ \varphi_1(s))}{2\sigma}\right) ds. \quad (12.40)$$

Moreover, if $2\sigma > 1$, the minimum is achieved and can be deduced from a geodesic curve γ on the unit sphere of $L^2(\mathbb{R})$.

This induces a distance on M, given by

$$d_\sigma(a_0, a_1) = 2 \inf_{\varphi_1} \left(\arccos \int_\Omega \sqrt{d\varphi_1} \cos\left(\arccos(a_0(s)^T a_1 \circ \varphi_1(s))/2\sigma\right) ds\right), \quad (12.41)$$

minimized over all strictly increasing diffeomorphisms of Ω.

This theorem is essentially proved in the discussion that precedes it, in which we have left a few loose ends, mostly regarding measurability and dealing with sets of measure 0, that can be tied up without too much effort by an interested reader.

In the two-dimensional case, one can represent functions $\alpha : \Omega \to S^1$ in the form $(\cos\theta, \sin\theta)$ for some angle function θ, which is only defined up to the addition of a multiple of 2π. Given such a representation, one can then consider the transform

$$\mathcal{G} : (\varphi, \theta) \mapsto q = \sqrt{d\varphi}(\cos\theta/2\sigma, \sin\theta/2\sigma)$$

that defines a mapping from Diff $\times L^2(\Omega, \mathbb{R})$ to the unit sphere of $L^2(\Omega, \mathbb{R}^2)$. Adding a time dependency, we find, using this transform, that

$$\int_0^1 \int_\Omega \frac{(d\dot{\varphi})^2}{d\varphi} ds\, dt + \frac{1}{\sigma^2} \int_0^1 \int_\Omega |\dot{\xi}|^2 d\varphi\, ds\, dt = 4 \int_0^1 \|\dot{q}\|_2\, dt,$$

so that minimizers on the left can be associated with geodesics on the unit sphere. One can then compute the metamorphosis distance by minimizing the lengths of great circles between, say, $\mathcal{G}(\mathrm{id}, \theta_0)$ and $\mathcal{G}(\varphi_1, \theta_1 \circ \varphi_1)$, for a given φ_1, and optimizing over all angle representations θ_0, θ_1 of a_0, a_1 that satisfy the constraint

$$-2\sigma\pi \le \theta_1 \circ \varphi_1 - \theta_0 \le 2\sigma\pi$$

because (12.38) still needs to hold for any time-continuous angle representation of q. This provides the same distance d_σ as the one obtained in Theorem 12.6. Notice, however, that this construction is special to the two-dimensional case. In dimension $d > 2$, our reduction to a unit sphere geodesic depended on the endpoints a_0 and a_1, and could not be deduced from a direct transformation applied to the curves themselves, such as $\mathcal{G}$. The only exception is the case $\sigma = 1/2$, for which $\mathcal{G}$ is equivalent to $(\varphi, \alpha) \mapsto \sqrt{d\varphi}\alpha$. This transform, called the "square root velocity transform", is clearly applicable to arbitrary dimensions. It has been extensively

studied in the literature, and we refer to [267] and references within for additional details and applications.

Returning to the two-dimensional case, alternative expressions of the distance can be derived for simple values of σ given angle representations θ_0 and θ_1 for a_0 and a_1. Indeed, the cosine in (12.41) is given by $\cos(\theta_1 \circ \varphi_1 - \theta_0)$ if $\sigma = 1/2$, by $|\cos((\theta_1 \circ \varphi_1 - \theta_0)/2)|$ if $\sigma = 1$, and by

$$\max(|\cos((\theta_1 \circ \varphi_1 - \theta_0)/4)|, |\sin((\theta_1 \circ \varphi_1 - \theta_0)/4)|)$$

if $\sigma = 2$.

The case $\sigma = 1$ for plane curves was first investigated in [307, 313]. It has interesting additional features, and we will see that it coincides with a geodesic distance that we have discussed previously in this chapter [311]. The optimization of φ_1 (which is still needed to compute the distance) can be done efficiently by dynamic programming, and the reader is referred to [279, 308] for more details.

To conclude this section, we notice that the metamorphosis metric is invariant under the action of rotations, so that one can optimize for a rotation parameter in all cases considered above. The rotation-invariant version of the distance between plane curves for $\sigma = 1$, for example, is

$$d_{1,\mathrm{rot}}(a_0, a_1) = 2 \inf_{\varphi_1, c} \left(\arccos \int_\Omega \sqrt{d\varphi_1} \left| \cos\left(\frac{\theta_1 \circ \varphi_1(s) - \theta_0(s) - c}{2} \right) \right| ds \right), \tag{12.42}$$

where c is a scalar. In higher dimensions, one needs to optimize (12.41) with a_0 replaced by Ra_0 when R varies over all rotations of the d-dimensional space.

12.7.2 One-Dimensional Case

In one dimension, the previous representation reduces to functions $a : \Omega \to \{-1, 1\}$, which does not leave much room to define time-continuous metamorphoses, and the previous approach cannot be extended to this case. One can however bypass this limitation by considering such functions as flat 2D functions, simply replacing the scalar-valued function a by $(a, 0)$. This defines a very special family of piecewise linear curves, to which the previous distance can nonetheless be applied. In this setting, given the functions a_0 and $a_1 \circ \varphi_1$, one has $\omega(s) = \pi$ if $a_1 \circ \varphi_1(s) \neq a_0(s)$ and 0 otherwise. One therefore gets

$$d_\sigma(a_0, a_1) = 2 \inf_{\varphi_1} \left(\arccos \int_\Omega \sqrt{d\varphi_1} \Big(\mathbf{1}_{a_0(s) = a_1 \circ \varphi_1(s)} + \cos\left(\frac{\pi}{2\sigma}\right) \mathbf{1}_{a_0(s) \neq a_1 \circ \varphi_1(s)} \Big) ds \right). \tag{12.43}$$

Notice that this compares functions modulo reparametrization, i.e., a_0 and $a_0 \circ \varphi$ are considered as identical for any increasing diffeomorphism of Ω. If this creates too much invariance, one can use the representation $\mathcal{G}$ for the curves $(a, 0)$ without optimizing for reparametrization, which provides the distance

$$\delta_\sigma(a_0, a_1) = 2\arccos \int_\Omega \sqrt{da_0(s)da_1(s)} \Big(\mathbf{1}_{a_0(s)=a_1(s)} + \cos\left(\frac{\pi}{2\sigma}\right) \mathbf{1}_{a_0(s)\neq a_1(s)} \Big) ds \,. \tag{12.44}$$

12.7.3 The Smooth Case

In this discussion, we placed few regularity conditions on the functions $a \in M$ beyond their measurability. The resulting class of curves includes, in particular, polygonal curves, for which a is piecewise constant. If one wants to restrict to spaces of smooth curves, then some modifications must be made. In particular, the integrals in (12.29) cannot be minimized independently for each s, because we need to ensure that the solution that one obtains is a continuous value of s. The optimal solution is however still given by (12.30) with, this time, ω being a continuous lift of $s \mapsto \arccos(a_0(s)^T a_1 \circ \varphi_1(s))$. One can therefore still reduce the d-dimensional setting to a two-dimensional one.

Taking the same definition for q, we find that the metamorphosis energy is no larger than four times the geodesic energy of q in the unit sphere of $L^2(\Omega, \mathbb{R}^2)$. When proving that the lower-bound is achieved, one finds that the function $k(s)$ in (12.37) must be continuous, hence constant. Here, we can use the fact that one can take $\omega(0) \in [0, \pi]$ and use the same argument as in the non-smooth case for $s = 0$, yielding $k(0) = 0$ and therefore $k(s) = 0$ for all s since k is constant. However, and regardless of the value of σ, one cannot ensure that (12.35) is satisfied unless the compared curves are close enough (so that their angles are at distance less than $2\sigma\pi$ after registration). When computed between curves that are too far apart, curves in M deduced from geodesics on the sphere will typically develop singularities (and therefore step out of M if this space is restricted to smooth curves).

The smooth case has also been studied in [29], in which a different transform is proposed, leading to a representation of plane curves in a three-dimensional space. More recently, [174] made a study of the smooth case for planar curves with an approach similar to the one we develop here. As pointed out in this reference, when correcting the distance for rotation invariance (similarly to (12.42)), the constant indetermination $2\pi k$ in (12.37) has no impact and one does not need the assumption that $2\sigma \geq 1$ in that case.

12.7.4 Existence of Optimal Metamorphoses

To complete the computation of the optimal metamorphosis, one must still optimize (12.40) with respect to the final diffeomorphism φ_1. The resulting variational problem is a special case of those studied in [280], which considered the maximization of functionals taking the form

$$F(\varphi) = \int_\Omega \sqrt{d\varphi} f(s, \varphi(s))\, ds$$

over the set Hom^+ of all strictly increasing functions $\varphi : \Omega \to \Omega$ satisfying $\varphi(0) = 0$ and $\varphi(1) = 1$, where f is a function defined on Ω^2. Following the notation there, we let

$$\Delta_f = \sup_{\psi \in \mathrm{Hom}^+} \int_\Omega \sqrt{d\psi} f(\psi(s), s)\, ds$$

and define the diagonal band

$$\Omega_c = \left\{(s, s') \in [0, 1]^2 \mid |s - s'| \le c\right\} .$$

We give without proof the following result, which is a consequence of Theorem 3.1 in [280].

Theorem 12.7 *Assume that $f \ge 0$ is continuous on $\bar{\Omega}^2$ except on a set G that can be decomposed as a union of a finite number of horizontal or vertical segments. Assume also that, for some c, with*

$$c > \sqrt{1 - \left(\frac{\Delta_f}{\|f\|_\infty}\right)^2},$$

there does not exist any non-empty open vertical or horizontal segment (a, b) such that $(a, b) \subset \Omega_c$ and f_l vanishes on (a, b), where

$$f_l(x) = \lim_{\delta \to 0} \inf_{|y-x|<\delta, y \notin G} f(y)$$

is the lower semi-continuous relaxation of f.

Then there exists a $\varphi^* \in Hom^+$ such that $F(\varphi^*) = \max\{F(\varphi), \varphi \in Hom^+\}$. Moreover, if φ is a maximizer of F, one has, for all $s \in \Omega$, $(\varphi(s), s) \in \Omega_c$.

Intuitively, f vanishing over vertical or horizontal segments allows for either very small or very large values of $d\varphi$ at very little cost, resulting in optimal solutions that may have stationary regions or jumps. In (12.41) (with $\sigma = 1/2$), this happens when the tangents of the compared curves are perpendicular. When $\sigma = 1$, this happens when θ_1 and θ_0 are oriented in opposite directions, i.e., their difference is equal to an odd multiple of π. For $\sigma > 1$, however, the cosine in (12.41) never vanishes.

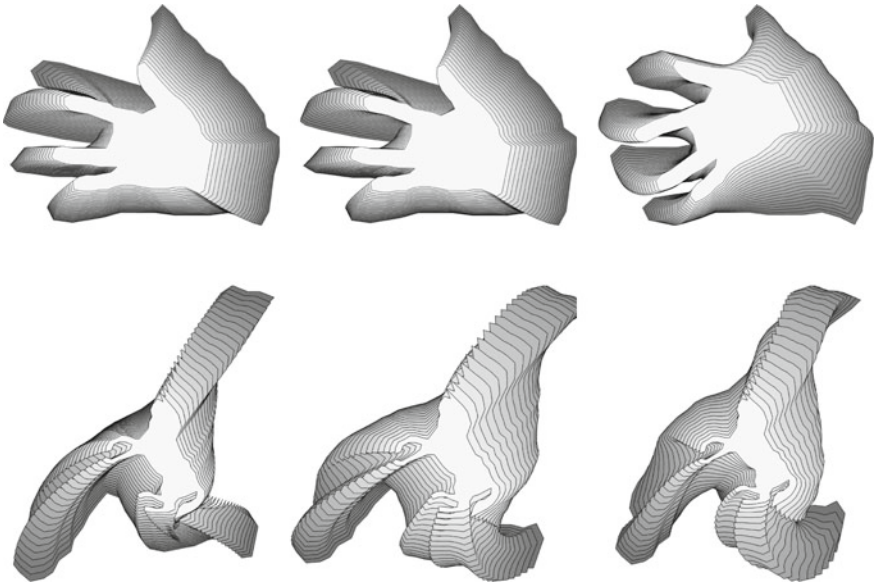

Fig. 12.5 Two sets of examples of optimal metamorphoses between curves with $\sigma = 0.5$ (left), 1 (center) and 2 (right). Note that, in the first row, the geodesics with $\sigma = 0.5$ and $\sigma = 1$ grow the missing finger out of the second finger, while that with $\sigma = 2$ grows a new thumb. In each image, the geodesic is computed between the outer curve and the inner curve. Scaling is for visualization only (the computation is scale-invariant)

There is no loss of generality in assuming that $f \geq 0$ in this discussion because if $f < 0$ on some rectangle, it is easy to check that any trajectory $(s, \varphi(s))$ that enters this rectangle can be improved if it is replaced by a trajectory that moves almost horizontally and/or almost vertically within the rectangle, with the new cost converging to 0 over this region. This can be used to show that there is no change in the minimizer if one replaces f by 0 within the rectangle.

One can efficiently maximize F by approximating f by a piecewise constant function taking the form

$$f(s, \tilde{s}) = \sum_{k=1}^{n} f_k \mathbf{1}_{R_k},$$

where $R_1, \ldots, R_n$ is a family of rectangles that partition the unit square and with $f_k \geq 0$, $k = 1, \ldots, n$. One can then show that the minimization can be performed over piecewise linear functions φ, which are furthermore linear whenever they cross the interior of a rectangle. The search for the optimal φ can then be organized as a dynamic program, and run very efficiently (see [279, 280] for details). This method is used in the experiments presented in Fig. 12.6 in which the optimal correspondence is drawn over an image representing the function $\max(f_\sigma, 0)$, where

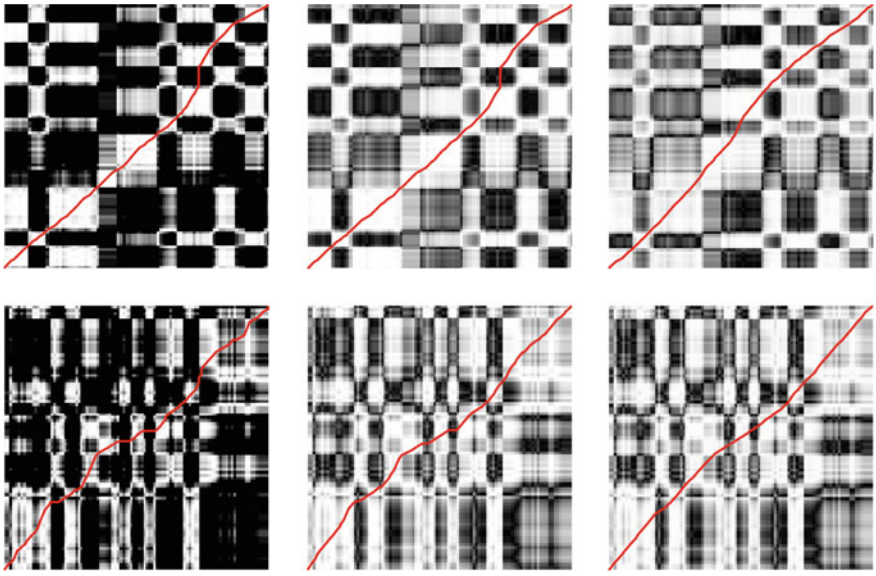

Fig. 12.6 Optimal correspondences estimated with the three metamorphoses in Fig. 12.5. Background images represent the function $\max(f_\sigma, 0)$ in (12.45), the zeros being indicated in black. The solution attempts to move across higher values of the function, cutting through zeros along vertical or horizontal lines if needed. Each panel corresponds to the same panel in Fig. 12.5

$$f_\sigma(s, \tilde{s}) = \cos\left(\frac{\arccos\cos(\theta_0(s) - \theta_1(\tilde{s}))}{2\sigma}\right). \tag{12.45}$$

Of course, if the compared curves are polygonal, f_σ is already piecewise constant.

12.7.5 The Case of Closed Curves

The previous developments assumed that curves were defined over open intervals, and therefore apply mostly to open curves. Closed curves are defined over $T^1 = [0, 1]_*$ (using the notation of Sect. 1.2.2), which is the open unit interval where the extremities are identified. The boundary condition on V, which was $v(0) = v(1) = 0$ for functions defined over Ω, now only requires $v(0) = v(1)$, offering a new degree of freedom, associated with a change of offset, or initial point of the parametrization, represented by the operation $s \mapsto s +_* \delta$ from T^1 to itself (where $+_*$ represents the translation along T^1, still using the notation introduced in Sect. 1.2.2). We restrict our discussion to the 2D case, in which we assume that the compared curves a_0 and a_1 have angle representations θ_0 and θ_1.

One can check easily that the distance in Theorem 12.6 is equivariant through this transformation, so that one can define a distance among closed curves that is invariant under rotations and changes of offset by (taking, for example, $\sigma = 1$)

$$\bar{d}_{1,\mathrm{rot}}(a_0, a_1)$$
$$= 2 \inf_{\varphi, c, \delta} \arccos \int_{T^1} \sqrt{\partial_s \varphi(s)}\, |\cos((\theta_0(x+\delta) - \theta_1 \circ \varphi(x) - c)/2)|\, dx. \quad (12.46)$$

Notice that, even when the resulting distance is still attained at a geodesic (or optimal metamorphosis) on the space of functions $a : T^1 \to S^1$, the corresponding curves at intermediate times, however, are not necessarily closed, because the associated closedness condition requires

$$\int_{T^1} a(s)\, ds = 0,$$

which is not enforced in this approach. Optimal trajectories are therefore not constrained to consist only of closed curves, and would typically become open for $t \in (0, 1)$, even though they start and end with closed curves. This method has been applied to obtain the geodesics shown in Fig. 12.5, to which an extra step has been added in order to close the intermediate curves for nicer visualization. This "closing" operation simply consisted in replacing a by $\tilde{a} = (a - \lambda)/|a - \lambda|$, where $\lambda \in \mathbb{R}^2$ was adjusted so that $\int_\Omega \tilde{a}\, ds = 0$.

To correctly define a geodesic distance on spaces of closed curves, one needs to consider the metric induced on the space M_c of functions $a : T^1 \to S^1$ such that $\int_{T^1} a\, ds = 0$. This space, however, is not invariant under a change of parameters, so that this induced metric is not associated with a metamorphosis. The expression of this constraint in terms of the q function is simple in the case $\sigma = 1/2$, for which $q = \sqrt{d\varphi}\, \alpha = \sqrt{d\varphi}\, a \circ \varphi$ so that, after a change of variable,

$$\int_{T^1} a\, ds = \int_{T^1} |q(u)|\, q(u)\, du\,.$$

Even in this case, there exists no closed form for the geodesic energy with fixed final reparametrization, but efficient algorithms have been designed to minimize

$$4 \int_0^1 \|\dot{q}(t, \cdot)\|_2^2\, dt$$

subject to the constraints that $\|q\|_2^2 = 1$, $\int_{T^1} |q|\, q\, du = 0$, $q(0) = q_0$ and $q(1) = q_1$ (see, for example, [161]).

The metric on closed curves also has a nice interpretation in the case $\sigma = 1$. In this case, let f and g denote the two coordinates of the representation $q = \mathcal{G}(\varphi, \theta)$ multiplied by $\sqrt{2}$, i.e., $f = \sqrt{2d\varphi} \cos\frac{\theta}{2}$ and $g = \sqrt{2d\varphi} \sin\frac{\theta}{2}$, where $\alpha = a \circ \varphi^{-1} = (\cos\theta, \sin\theta)$. The closedness constraint, which is

$$\int_\Omega \cos\theta \circ \varphi^{-1}\, ds = \int_\Omega \sin\theta \circ \varphi^{-1}\, ds = 0,$$

becomes

$$\int_\Omega d\varphi \cos\theta \, ds = \int_\Omega d\varphi \sin\theta \, ds = 0$$

after a change of variables. Writing $\cos\theta = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}$ and $\sin\theta = 2\cos\frac{\theta}{2}\sin\frac{\theta}{2}$, this is equivalent to

$$\|f\|_2^2 - \|g\|_2^2 = \langle f \, , \, g \rangle_2 = 0.$$

Because $\|f\|_2^2 + \|g\|_2^2 = 2\int_\Omega d\varphi = 2$, we find that the constraint is equivalent to $\|f\|_2^2 = \|g\|_2^2 = 1$ and $\langle f \, , \, g \rangle_2 = 0$, i.e., to (f, g) forming an orthonormal 2-frame in $L^2(\Omega)$, and we have

$$U_1(\varphi, \alpha) = 2\int_0^1 \left(\|\partial_t f\|^2 + \|\partial_t g\|^2 \right) ds,$$

where the left-hand side is two times the geodesic energy of the path (f, g) in the Stiefel manifold $St(\infty, 2)$ (see Sect. B.6.7). Repeating the arguments made in Sects. 12.7.1 or 12.7.3 in this setting shows that the optimal metamorphosis with fixed φ_1 is obtained from the shortest length geodesic in $St(\infty, 2)$ connecting the frames $(\sqrt{2}\cos\frac{\theta_0^\delta}{2}, \sqrt{2}\sin\frac{\theta_0^\delta}{2})$ and

$$\left(\varepsilon\sqrt{2\partial_s\varphi_1}\cos\frac{\theta_1 \circ \varphi_1}{2}, \varepsilon\sqrt{2\partial_s\varphi_1}\sin\frac{\theta_1 \circ \varphi_1}{2} \right),$$

where θ_0 and θ_1 are angle representations of a_0 and a_1 and the optimization is made over all possible measurable functions $\varepsilon : \Omega \to \{-1, 1\}$, and all possible offsets δ, with the notation $\theta_0^\delta(s) = \theta_0(s +_* \delta)$. (The optimization over ε results from optimizing over all possible angle representations of the two curves.) There is, however, no closed form expression for the geodesic distance on the Stiefel manifold (although equations for geodesics have been described in [99]), and no simple algorithm to solve this optimization problem. Notice that, if one restricts to smooth curves, the search for an optimal ε is only over constant functions $\varepsilon = \pm 1$ and optimal geodesics can be obtained using a root-finding algorithm over initial conditions of geodesics in $St(\infty, 2)$.

The rotation-invariant version of the distance also provides an interesting representation, because a rotation acting on curves simply induces a rotation of the frame (f, g), and the space of such frames modulo rotation is now the Grassmann manifold $Gr(\infty, 2)$ of two-dimensional subspaces of $L^2(\mathbb{R})$. The same analysis carries on, the only difference being that one uses now the geodesic distance on the Grassmannian. This geodesic distance can be computed in quasi closed form [216], and is given by $\sqrt{\arccos^2\lambda + \arccos^2\mu}$, where λ and μ are the singular values of the matrix

$$\begin{pmatrix} \langle f_0 \, , \, f_1 \rangle_2 & \langle f_0 \, , \, g_1 \rangle_2 \\ \langle g_0 \, , \, f_1 \rangle_2 & \langle g_0 \, , \, g_1 \rangle_2 \end{pmatrix},$$

(f_0, g_0), (f_1, g_1) being orthogonal bases of the two spaces that are compared. This closed form, however, does not lead to a simple version of the distance when one optimizes over changes of sign in (f_1, g_1). More analysis of this framework (in the smooth case), including explicit computations of the geodesic equation and of the scalar curvature, can be found in [311].

12.7.6 Alternative Interpretation ($\sigma = 1$)

Returning to the case of open curves, we consider the space of curves $m : \Omega \to \mathbb{R}^2$, not necessarily parametrized by arc-length. Recall the notation $\partial_s f = df/|dm|$ for the derivative with respect to the arc length. Consider the parametrization-invariant Sobolev metric on curves given by

$$\|\xi\|_m^2 = \frac{1}{L(m)} \int_\Omega |\partial_s \xi|^2 \, ds = \frac{1}{L(m)} \int_\Omega \frac{|d\xi|^2}{|dm|} \, du,$$

where $L(m) = \int_\Omega |dm(u')| \, du'$ is the length of the curve m.

Consider a trajectory $(t, u) \mapsto m(t, u)$. Let $\ell(t) = L(m(t))$ and let $\varphi(t, u)$ denote the normalized arc-length parametrization along this curve so that

$$\varphi(t, u) = \frac{1}{\ell(t)} \int_0^u |dm(t, u')| \, du'.$$

Let $\alpha(t, u) = \partial_s m(t, u)$ be the unit tangent. We want to obtain an alternative expression for

$$\int_0^1 \|\dot{m}(t)\|_{m(t)}^2 \, dt = \int_0^1 \int_\Omega \frac{|d\dot{m}|^2}{|dm|} \, du \, dt$$

in terms of φ and α. We first notice that

$$\dot{\alpha} = \partial_t \left(\frac{dm}{|dm|} \right) = \left(\left(\frac{\dot{m}}{|dm|} \right)^T \alpha^\perp \right) \alpha^\perp,$$

where $\alpha^\perp$ is the unit normal, obtained by rotating α by $\pi/2$. Moreover,

$$d\dot{\varphi} = \partial_t \left(\frac{|dm|}{\ell} \right) = \frac{1}{\ell} (d\dot{m})^T \alpha - \frac{|dm|}{\ell^2} \dot{\ell}$$

with

$$\dot{\ell} = \int_\Omega (d\dot{m})^T \alpha \, du.$$

A little computation yields

$$\int_\Omega \frac{(d\dot\varphi)^2}{d\varphi}\,du = \frac{1}{\ell}\int_\Omega \frac{((d\dot m)^T\alpha)^2}{|dm|}\,du - \frac{\dot\ell^2}{\ell^2}.$$

Given this, we can write

$$d\dot m = \left((d\dot m)^T\alpha\right)\alpha + \left((d\dot m)^T\alpha^\perp\right)\alpha^\perp$$

and

$$\|\dot m(t)\|^2_{m(t)} = \frac{\dot\ell^2}{\ell^2} + \int_\Omega \frac{(d\dot\varphi)^2}{d\varphi}\,du + \int_\Omega |\dot\alpha|^2 d\varphi\,du.$$

The last two terms integrated over time provide the function U_1 obtained in (12.28) with $\sigma = 1$. This expression of the metric treats the pair (φ, α) and the curve length independently, and the geodesic length can be optimized separately for ℓ and (φ, α), yielding a geodesic distance given by

$$d(m_0, m_1)^2 = \left(\log\frac{L(m_1)}{L(m_0)}\right)^2 + 4\arccos^2\int_\Omega \sqrt{d\varphi_0\, d\varphi_1}\left|\cos\left(\frac{\theta_1\circ\varphi_1 - \theta_0\circ\varphi_0}{2}\right)\right| du,$$

where φ_0, φ_1 are the arc-length reparametrizations of m_0, m_1 and θ_0, θ_1 are their angle representations expressed as functions of the arc lengths. Furthermore, the transformation $m \mapsto m/L(m)$ provides a Riemannian submersion onto the space of curves with length 1 (or the space of curves modulo scaling), and the resulting projected distance is exactly the one obtained via the metamorphosis approach.

Chapter 13
Analyzing Shape Datasets

We present in this chapter some "shape analysis" methods, among those that are mainly used in practice, where the goal is to provide a low-dimensional description and to perform statistical validations of hypotheses for datasets in which each object is a shape. Most recent applications of this framework have taken place in medical imaging, in which the shapes are provided by anatomical regions segmented by MRI or computer tomography scans. The analysis of the anatomy derived from such images is called *computational anatomy*, with a framework introduced in [137, 138], and has since generated a huge literature. Beside this important range of applications, shape analysis can also be used in computed vision, or in biology, which was, for example, the main focus of D'Arcy-Thompson's seminal treatise [276] on Growth and Form. We here focus on methods that derive from the analysis of diffeomorphisms developed in the previous chapters, leading to "morphometric" [23, 68], or "diffeomorphometric" [197] analyses.

13.1 Reference-Based Representation

The diffeomorphic matching methods that were described in this book can be seen to have a dual purpose. As a first goal, they provide a comparison tool between shapes, generally based on a formal or rigorous Riemannian paradigm. They also, by nature, provide an algorithm that aligns a target shape along a reference shape, i.e., that estimates a diffeomorphism φ, such that $\varphi \cdot (\text{reference}) \simeq (\text{target})$. This correspondence, φ, can be seen as a representation of the relationship between the reference and the target in the diffeomorphism group, i.e., a parametrization of the target relative to the reference.

In more formal terms, registration algorithms provide, given a reference $\bar{m}$, a mapping $m \mapsto \Phi(m)$ from a shape space to the diffeomorphism group such that $\Phi(m) \cdot \bar{m} \simeq m$. From a dataset $(m_1, \ldots, m_N)$ of shapes, one can then obtain a dataset $(\varphi_1, \ldots, \varphi_N)$ of diffeomorphisms, with $\varphi_k = \Phi(m_k)$. Even though diffeomorphisms may appear as more complex objects than many shapes, this representation actually

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5_13

simplifies the analysis of the dataset. It is certainly natural to restrict this analysis to the restriction of the diffeomorphism (or its derivative) to the reference, which (e.g., when dealing with landmarks) can represent a huge reduction of dimension. Image or shape morphometry, as described, for example, in [23], determines features, or descriptors, of shapes in the dataset based on these diffeomorphisms, using point displacements or Jacobian matrices. These features can then be used in a statistical learning framework to draw conclusions on properties of interest about the dataset.

When dealing with shape spaces, M, that do not involve diffeomorphisms, such as Kendall's space (see Sect. 11.2), other reference-based representations can be used, the most natural, in the Riemannian case, being to use exponential charts for the metric, such that $\Phi(m) = v \in T_{\bar{m}}M$ with $m = \mathrm{Exp}_{\bar{m}}(v)$. The statistical analysis can then be based on the vectors $v_1, \ldots, v_N$ that all belong to the same vector space. This point of view can actually be applied to the diffeomorphic representation, where one can use exponential charts, or, equivalently, the momentum representation, associated with right-invariant metrics on diffeomorphisms. Notice that the LDDMM algorithms described in Sect. 10.3 and later directly return such a representation while estimating an optimal correspondence.

Building such representations requires selecting a proper reference shape. While one can use any fixed shape for this purpose, it is, for many reasons, preferable to choose $\bar{m}$ close to the studied dataset. It is understandably easier to analyze diffeomorphisms when the deformation they define are not too severe. Also, tangent space representations linearize the shape space, and one wants to reduce as much as possible the metric distortions they induce. This is why one typically computes $\bar{m}$ as some kind of average of the dataset under study.

When using morphometric methods, one often estimates $\bar{m}$ and computes its optimal correspondences with the dataset in a single algorithm, which is often called *groupwise registration* [26, 35, 36, 160, 180, 181, 288]. In its simplest form, when the registration between $\bar{m}$ and m minimizes a cost function $U_{\bar{m},m}(\varphi)$, the associated groupwise registration minimizes $\sum_{k=1}^{N} U_{\bar{m},m_k}(\varphi_k)$ with respect to $\bar{m}, \varphi_1, \ldots, \varphi_N$. Some additional regularization constraints may also be used for the reference. For example, if one uses the LDDMM algorithm, one can define a groupwise registration method for image matching via the minimization of (using the notation Sect. 10.3, V being an admissible space)

$$\frac{1}{2}\sum_{k=1}^{N}\int_0^1 \|v_k(t)\|_V^2\, dt + \frac{1}{\sigma^2}\sum_{k=1}^{N}\|\bar{m}\circ\varphi_{10}^{v_k} - m_k\|_2^2 \tag{13.1}$$

with respect to $v_1, \ldots, v_N$ and $\bar{m}$. When $\bar{m}$ is fixed, this provides N independent image registration problems, and when $v_1, \ldots, v_N$ are fixed, the optimal $\bar{m}$ is given by

$$\bar{m} = \frac{1}{N}\sum_{k=1}^{N} m_k \circ \varphi_{01}^{v_k} \det(d\varphi_{01}^{v_k}).$$

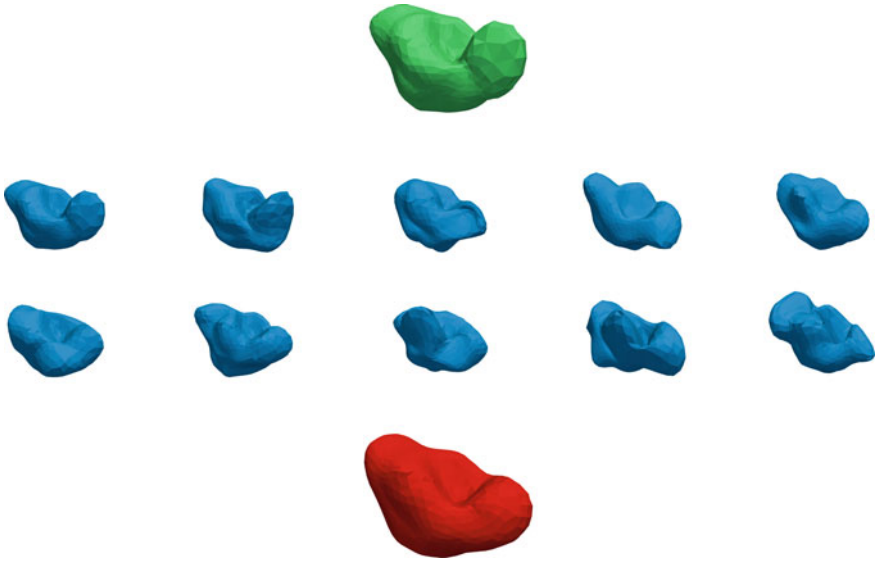

Fig. 13.1 Example of template estimation. Rows 2 and 3: ten randomly generated shapes. Row 1: hypertemplate, chosen as the first shape in the sequence. Row 4: estimated template

A modification of this method has been introduced in [180] where the reference is in the form $\bar{m} = m_0 \circ \psi_{10}^{v_0}$ where m_0 is fixed and v_0 is estimated jointly with the rest of the variables by minimizing

$$\frac{\lambda}{2}\int_0^1 \|v_0(t)\|_V^2\,dt + \frac{1}{2}\sum_{k=1}^N \int_0^1 \|v_k(t)\|_V^2\,dt + \frac{1}{\sigma^2}\sum_{k=1}^N \|m_0 \circ \varphi_{10}^{v_0} \circ \varphi_{10}^{v_k} - m_k\|_2^2$$

for some $\lambda > 0$. This constrains the topology of the estimated reference image to conform to that of the image m_0. A similar approach has been introduced in [181] for surface matching, in which m_0 is referred to as a *hypertemplate*. One interesting feature of this approach is that it can be represented as a family of branching optimal control problems, each with its own maximum principle that can also be branched backwards in time to compute the gradient of the objective function: one first uses v_0 as a control leading from the hypertemplate to the template, then $v_1, \ldots, v_n$ as controls driving the template to the targets. An example of template estimation with this method is provided in Fig. 13.1.

When one uses a tangent representation on a shape manifold, $\bar{m}$ is often estimated as a Fréchet mean, or Riemannian center of mass, of the collection $m_1, \ldots, m_N$. Such a mean is defined as a minimizer of

$$F : \bar{m} \mapsto \sum_{k=1}^N d_M(\bar{m}, m_k)^2,$$

where d_M is the geodesic distance on M. It is important to point out that this minimization problem does not always have a unique solution, i.e., that some datasets may have more than one Fréchet mean, even though this fact is often ignored in practice (or, one considers that any Fréchet mean is a good candidate for the reference shape). In finite dimensions, sufficient conditions for the uniqueness of such means involve the curvature of M and Rauch's comparison theorems. They are beyond the scope of this book, but we refer to [4, 62, 164, 168, 229] for additional details.

The computation of the gradient of F can be based on the following lemma, which we state without proof.

Lemma 13.1 *Let M be a Riemannian manifold. For $m_0 \in M$ define $f(m) = d_M(m_0, m)^2$. Let Exp_{m_0} be defined in some neighborhood of 0, $\Omega \subset T_{m_0}M$, and be a diffeomorphism onto its image. Then, for all $v_0 \in \Omega$, for all $m \in \mathrm{Exp}_{m_0}(\Omega)$ and $h \in T_m M$,*

$$df(m)h = 2\langle \dot\gamma(1)\,,\, h\rangle_m = -2\langle \dot{\tilde\gamma}(0)\,,\, h\rangle_m,$$

where $\gamma(t)$, $t \in [0, 1]$, is the geodesic joining m_0 and m and $\tilde\gamma(t) = \gamma(1-t)$ is the geodesic between m and m_0.

From this, we can deduce that if the dataset is fully included in a domain Ω that contains minimizing geodesics between any of its points (i.e., it is *geodesically convex*), is such that each of these geodesics is uniquely defined, and such that, for all $m \in \Omega$, Exp_m is a diffeomorphism from an open neighborhood of 0 in $T_m M$ onto Ω, then

$$dF(\bar m)\, h = -2 \sum_{k=1}^{N} \dot\gamma_k(0),$$

where γ_k is the geodesic between $\bar m$ and m_k, and the Fréchet mean must satisfy

$$\sum_{k=1}^{N} \dot\gamma_k(0) = 0.$$

This computation leads, in particular, to gradient descent algorithms designed to estimate the mean (see [177]).

It is also possible to define a reference shape through a stochastic shape model, in which $\bar m$ is deformed via random diffeomorphisms, possibly with additional noise, to generate $m_1, \ldots, m_N$. The estimation of $\bar m$ can then be performed using maximum likelihood. While describing in detail the associated statistical model and the estimation algorithm would take us too far from this discussion (and we refer to [7, 8, 173] for such details), it is important to note that minimizing (13.1) in this case may lead (when the noise level is high enough) to biased estimates of $\bar m$, in the sense that, even if N tends to infinity, minimizers of (13.1) will differ from $\bar m$ when the model is valid (which does not mean, however, that they cannot be used as reference shapes for subsequent morphometric analyses). See in particular [81] for a theoretical analysis of the issue.

13.2 Principal Component Analysis

Principal component analysis (PCA) is the simplest and most widely used method of performing dimension reduction for data analysis [92]. It is especially useful in shape analysis, which deals with virtually infinite-dimensional objects. The reader may refer, if needed, to the basic description of the method that is provided in Appendix E in the appendix for Hilbert spaces. In this section we focus on the specific adaptation of the approach to nonlinear shape spaces.

PCA is indeed a linear method, designed to be applied to vector spaces equipped with an inner product. On shape spaces, and more generally on Riemannian manifolds, a standard approach relies on a tangent-space linearization of the manifold using exponential charts (this is often referred to as tangent PCA). More precisely, given a dataset $(m_1, \dots, m_N)$ and a reference element $\bar{m}$ one computes normal coordinates $h_1, \dots, h_N \in T_{\bar{m}}M$ such that $m_k = \mathrm{Exp}_{\bar{m}}(h_k)$ for $k = 1, \dots, N$, and performs PCA on the collection $(h_1, \dots, h_N)$ using the Riemannian inner product $\langle \cdot, \cdot \rangle_{\bar{m}}$. The p first principal components then provide an orthonormal family $(e_1, \dots, e_p)$ spanning a subspace of $T_{\bar{m}}M$, and the PCA representation is given by

$$\Phi : (\lambda_1, \dots, \lambda_p) \mapsto \mathrm{Exp}_{\bar{m}}\left(\sum_{j=1}^{p} \lambda_j e_j \right) \in M.$$

When working with shape spaces with a metric induced by a right-invariant Riemannian metric on diffeomorphisms through a Riemannian submersion, it is easier, and formally equivalent, to reformulate the problem in terms of Diff rather than M. One can see, in particular, that the LDDMM registration algorithm minimizes

$$\frac{1}{2} \int_0^1 \|v(t)\|_V^2 \, dt + U(\varphi_{01}^v \cdot \bar{m}, m_k)$$

for some data attachment term U, which is equivalent to minimizing

$$\frac{1}{2} \|h\|_V^2 + U(\mathrm{Exp}_{\mathrm{id}}(h) \cdot \bar{m}, m_k),$$

where $\mathrm{Exp}_{\mathrm{id}}$ is the Riemannian exponential on the diffeomorphism group starting at $\varphi = \mathrm{id}$. One can use the optimal h, say $h_k \in V$, as a representation of m_k on which PCA can be applied, using the V inner product. Using the notation introduced in Sect. 11.5.2, Definition 11.13, one can replace h by $\rho = \mathbb{L}h$ and solve the equivalent problem of minimizing

$$\frac{1}{2} \|\rho\|_{V^*}^2 + U(\mathrm{Exp}_{\mathrm{id}}^{\flat}(\rho) \cdot \bar{m}, m_k)$$

with optimal solution given by $\rho_k = \mathbb{L}h_k$. The advantage of doing so is the parsimony of the momentum representation, as discussed in Sect. 11.5.2. PCA is then performed on the dataset $(\rho_1, \ldots, \rho_N)$ using the V^* inner product. Once a PCA basis, say $\xi_1, \ldots, \xi_p$, is computed, the representation is then

$$\Phi : (\lambda_1, \ldots, \lambda_p) \mapsto \mathrm{Exp}^{\flat}_{\mathrm{id}}\left(\sum_{j=1}^{p} \lambda_j \rho_j \right) \cdot \bar{m} \in M.$$

Notice that this is a representation of the "deformed templates" $(\varphi_{01}^{v_k} \cdot \bar{m}, k = 1, \ldots, N)$ rather than of the original data $(m_1, \ldots, m_N)$. This momentum PCA approach has been used multiple times in applications, starting with [290], in which it was introduced for landmark spaces. This formulation allows one to revisit the active shape model described in Sect. 6.2, in which shapes were represented by a decomposition in a linear basis, and develop a model based on the nonlinear representation associated with the function Φ above, which constrains the shape topology. Such approaches have been proposed in order to operate shape segmentation, or to regularize registrations in [285, 287].

Because tangent PCA is based on a linear representation of the manifold M, it necessarily suffers from the metric distortions that any linear representation must induce. The sum of residuals in the tangent space that is minimized by PCA may be quite different from the sum of squared distances of the actual shapes to their PCA representation provided by the mapping Φ. More precisely, one can formulate the search for p principal directions in tangent PCA as looking for a p-dimensional subspace $W \subset T_{\bar{m}}M$ such that

$$F(W) = \sum_{k=1}^{p} \min_{w \in W} \|h_k - w\|^2_{\bar{m}} \tag{13.2}$$

is minimized, with $\mathrm{Exp}_{\bar{m}}(h_k) = m_k$. However, in terms of approximating the dataset, one would probably be more interested in minimizing

$$F(W) = \sum_{k=1}^{p} \min_{w \in W} d_M(\mathrm{Exp}_{\bar{m}}(w), m_k)^2, \tag{13.3}$$

which measures how far each shape is from its representation in the manifold. The two criteria may be quite different when the dataset is spread out away from $\bar{m}$ and their solutions (the optimal W) may be quite different. Obviously, the first criterion is much easier to minimize than the second one, which represents a complex nonlinear optimization problem (with d_M usually non-explicit). One can make it slightly easier by building W one dimension at a time, starting with $p = 1$, in which one looks for the best geodesic approximating the data, progressively adding new directions without changing those that were found earlier. This procedure was introduced in [113] and called geodesic principal component analysis (GPCA). The non-incremental problem

requires a search within the space of all p-dimensional subspaces of $T_m M$, i.e., its Grassmann manifold of order p (cf. Sect. B.6.7).

Notice that one may opt for a simplified version of GPCA by replacing the Riemannian distance in M by some "extrinsic" discrepancy measure. For example, in the diffeomorphic framework, one can formulate the problem of finding a p-dimensional subspace W of V^* that minimizes

$$F(W) = \sum_{k=1}^{p} \min_{\rho \in W} U(\mathrm{Exp}_{\mathrm{id}}^{\flat}(\rho) \cdot \bar{m}, m_k), \tag{13.4}$$

in which U replaces the distance d_M and would be computationally more tractable. This problem can be rewritten in the form of finding $\rho_1, \ldots, \rho_N$ minimizing

$$G(\rho_1, \ldots, \rho_N) = \sum_{k=1}^{N} U(\mathrm{Exp}_{\mathrm{id}}^{\flat}(\rho_k) \cdot \bar{m}, m_k)$$

subject to $\mathrm{rank}(\rho_1, \ldots, \rho_N) = p$. The problem in this form is tackled in [60], in which a gradient descent algorithm over p-dimensional subspaces of V^* is proposed.

13.3 Time Series

13.3.1 Single Trajectory

We now assume that the dataset $(m_1, \ldots, m_N)$ is a time series, so that it describes the evolution of a given shape captured at times, say, $\tau_1 < \cdots < \tau_N$. We here study the regression problem of determining a function $\tau \mapsto m(\tau) \in M$ such that $m(\tau_k) \simeq m_k$.

Since geodesics are the Riemannian generalizations of straight lines in Euclidean spaces, one generalizes the standard linear regression model $m(\tau) = \bar{m} + \tau h$ to such spaces by looking for curves defined by $m(\tau) = \mathrm{Exp}_{\bar{m}}(\tau h)$ for fixed $\bar{m} \in M$ and $v \in T_{\bar{m}} M$, which both need to be estimated from data. Notice that, in this case, $\bar{m}$ is not an average of the considered dataset, but an "intercept", representing the estimated position at $\tau = 0$. The resulting "geodesic regression" model [114] can then be associated with the generalization of least-square estimation, minimizing

$$\sum_{k=1}^{N} d_M(\mathrm{Exp}_{\bar{m}}(\tau_k h), m_k)^2$$

with respect to $\bar{m}$ and v. Notice that this problem is similar, but distinct from the search for a geodesic principal direction, which would first choose $\bar{m}$ as a Fréchet mean, and then estimate h, with $\|h\|_{\bar{m}} = 1$, minimizing

$$\sum_{k=1}^{N} \min_{\tau} d_M(\mathrm{Exp}_{\bar{m}}(\tau h), m_k)^2 .$$

As discussed with PCA, the intrinsic error criterion using the Riemannian distance seldom leads to tractable optimization algorithms, and it is often replaced with a discrepancy measure that is more amenable to computation, minimizing

$$\sum_{k=1}^{N} U(\mathrm{Exp}_{\bar{m}}(\tau_k h), m_k) .$$

The exponential function being associated with a second-order differential equation, the derivative with respect to $\bar{m}$ and h of each term in the sum above can be computed using the formulas derived in Sect. C.4 for the variation of solutions of ODEs with respect to their initial conditions. Some regularization may be added to the objective function, to control, for example, the topology of the intercept, $\bar{m}$, which may be chosen in the form $\bar{m} = \mathrm{Exp}_{\tilde{m}}(h_0) \cdot m_0$ for some fixed shape m_0. Adding some penalty on the norms of h_0 and h, one can minimize

$$(h_0, h) \mapsto \lambda_0 \|h_0\|^2_{m_0} + \lambda \|h\|^2_{\tilde{m}} + \sum_{k=1}^{N} U(\mathrm{Exp}_{\tilde{m}}(\tau_k h), m_k) ,$$

with $\bar{m} = \mathrm{Exp}_{\mathrm{id}}(h_0) \cdot m_0$. This model was implemented in [155] on spaces of surfaces with a metric induced by diffeomorphisms, with a similar approach developed in [110]. Still in the diffeomorphic framework, a geodesic regression algorithm for images has been proposed in [219], and an approach using image metamorphosis has been proposed in [153].

Notice also that one can spare the estimation of $\bar{m}$ by assuming that $\tau_0 = 0$, considering the first observation as a baseline. This creates, however, an asymmetry in the data, in which the noise or variation from the geodesic is neglected for the baseline, which may sometimes be artificial [24, 242, 243].

It is not difficult to modify the previous framework to correct for the property that geodesics evolve at constant speed by making a time reparametrization of the trajectory. This corresponds to the model $m(\tau) = \mathrm{Exp}_{\tilde{m}}(f(\tau)h)$, where f is an increasing function from $[0, 1]$ to $[0, 1]$ that also needs to be estimated. This time reparametrization can be estimated using a method akin to LDDMM, modeling f as the result of a diffeomorphic flow [94], but simpler methods can be used, too, such as, for example, optimizing

$$(\bar{m}, h, \tilde{\tau}_1, \ldots, \tilde{\tau}_N) \mapsto \sum_{k=1}^{N} d_M(\mathrm{Exp}_{\bar{m}}(\tilde{\tau}_k h), m_k)^2$$

subject to $0 = \tilde{\tau}_1 < \tilde{\tau}_2 < \cdots < \tilde{\tau}_{N-1} < \tilde{\tau}_N = 1$, which corresponds to monotonic regression with respect to time [155]. The derivative of $\mathrm{Exp}_{\tilde{m}}(\tilde{\tau} h)$ with respect to

$\tilde{\tau}$ is straightforward to compute, since it is given by the speed of the geodesic, and is readily obtained from the differential equation that is integrated to compute this geodesic.

Aligning all shapes along a single geodesic may sometimes be too restrictive when the time series exhibits several modes of variation, and more flexible methods can be derived. At the extreme end of this range of methods, one can use a piecewise geodesic approach, which consists in estimating $h_1, \ldots, h_{N-1}$ such that $\mathrm{Exp}_{m_k}(h_k) = m_{k+1}$ with $h_k \in T_{m_k}M$. One can then define

$$m(\tau) = \mathrm{Exp}_{m_k}\left(\frac{\tau - \tau_k}{\tau_{k+1} - \tau_k} h_k\right)$$

for $\tau \in [\tau_k, \tau_{k+1}]$. This is just the Riemannian generalization of a piecewise linear curve interpolating the observed trajectory. One can modify this formulation by allowing for some error in the interpolation, thus taking into account possible measurement noise in the observed m_k's, minimizing, for example,

$$\frac{\lambda_0}{2}\|h_0\|^2_{\bar{m}_0} + \frac{\lambda}{2}\sum_{k=0}^{N-1}\|h_k\|^2_{\bar{m}_k} + \sum_{k=1}^{N} U(\bar{m}_k, m_k),$$

where $\bar{m}_k$ is defined recursively by $\bar{m}_{k+1} = \mathrm{Exp}_{\bar{m}_k}(h_k)$, $h_k \in T_{\bar{m}_k}M$ and $\bar{m}_0$ is a fixed shape. One obtains an equivalent formulation with the following time-continuous problem of minimizing

$$\frac{\lambda_0}{2}\|h_0\|^2_{\bar{m}_0} + \frac{\lambda}{2}\int_0^1 \|\dot{\gamma}(t)\|^2_{\gamma(t)} + \sum_{k=1}^{N} U(\gamma(\tau_k), m_k) \tag{13.5}$$

subject to $\gamma(0) = \mathrm{Exp}_{\bar{m}_0}(h_0)$. Indeed, the solution γ must be a minimizing geodesic between $\bar{m}_k := \gamma(\tau_k)$ and $\bar{m}_{k+1}$, with constant speed, which directly leads to a piecewise geodesic solution. In the LDDMM framework, this equivalent formulation reduces to minimizing

$$\frac{\lambda_0}{2}\|v_0\|^2_V + \frac{\lambda}{2}\int_0^1 \|v(t)\|^2_V\, dt + \sum_{k=1}^{N} U(\varphi^v_{0\tau_k} \circ \mathrm{Exp}_{\mathrm{id}}(v_0)\cdot \bar{m}_0, m_k)$$

with respect to $v_0 \in V$ and $t \mapsto v(t) \in L^2([0,1], V)$ (see [196, 197, 204]).

Piecewise geodesic interpolation is continuous in time, but not differentiable, and will certainly be too sensitive to noise, even when using inexact interpolation. Time differentiability of the solution can be obtained by controlling the second derivative of γ instead of the first derivative in (13.5), leading to a Riemannian generalization of interpolating splines. As discussed in Sect. B.6.4, curve acceleration in Riemannian manifolds involves the covariant derivative, and a formulation of the Riemannian spline problem can be obtained by replacing $\dot{\gamma}(t)$ by $\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t)$ in (13.5) [220].

The analysis of the new variational problem becomes more involved when studied intrinsically on the manifold, and here we restrict to situations in which one can work in a local chart, and take advantage of the Hamiltonian formulation of geodesics described in Sect. B.6.6. Using the notation of this section we let $S(m)$ be the representation of the metric in the chart, so that $\|h\|_m^2 = h^T S(m)h$, with the notation abuse of using h in order to represent a vector in $T_m M$ and its expression in a chart. Letting $H(m,a) = a^T S(m)^{-1}a/2$, with $a \in T_m M^*$, the geodesic equations in Hamiltonian form are

$$\begin{cases} \partial_t m = S(m)^{-1}a \\ \partial_t a + \frac{1}{2}\partial_m (a^T S(m)^{-1}a) = 0. \end{cases}$$

Moreover, given a curve γ on M (not necessarily a geodesic), one has, letting $a(t) = S(\gamma(t))\dot\gamma(t)$,

$$\nabla_{\dot\gamma}\dot\gamma = S(\gamma)^{-1}\left(\partial_t a + \frac{1}{2}\partial_\gamma (a^T S(\gamma)^{-1}a)\right).$$

One can then reformulate the Riemannian spline problem as an optimal control problem, with state (γ, a) and control u, minimizing

$$\frac{\lambda_0}{2}\|h_0\|_{\bar m_0}^2 + \frac{\lambda}{2}\int_0^1 u(t)^T S(\gamma(t))^{-1}u(t)\,dt + \sum_{k=1}^N U(\gamma(\tau_k), m_k)$$

subject to the state equation

$$\begin{cases} \partial_t \gamma = S(\gamma)^{-1}a \\ \partial_t a + \frac{1}{2}\partial_\gamma (a^T S(\gamma)^{-1}a) = u \end{cases}$$

with initial condition $\gamma(0) = \mathrm{Exp}_{\bar m_0}(h_0)$ and free condition for $a(0)$. One can consider higher-order Riemannian splines by iterating covariant derivatives (see, e.g., [122, 182]). This approach to the spline problem was introduced for shape spaces with a Riemannian metric induced by diffeomorphisms in [284], with further developments in [265]. In this case, one can take advantage of the right-invariance of the metric in the group to reformulate the problem as minimizing

$$\frac{\lambda_0}{2}\|v_0\|_V^2 + \frac{\lambda}{2}\int_0^1 \|u(t)\|^2\,dt + \sum_{k=1}^N U(\varphi^v_{0\tau_k}\circ \mathrm{Exp}_{\mathrm{id}}(v_0)\cdot \bar m_0, m_k)$$

subject to

$$\begin{cases} \partial_t \gamma = (\mathbb{K}\rho)\cdot\gamma \\ \partial_t \rho + \mathrm{ad}^*_{\mathbb{K}\rho}\rho = u \end{cases}$$

Fig. 13.2 Time series with four surfaces: left to right: $t = 1, 2, 3, 4$

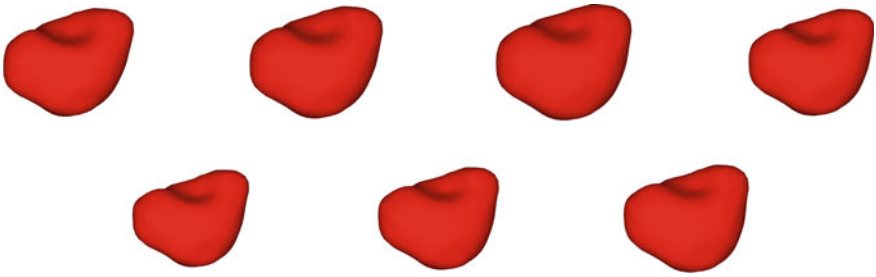

Fig. 13.3 Piecewise geodesic interpolation of sequence in Fig. 13.2 with seven time points (left to right and top to bottom): $t = 1, 1.5, 2, 2.5, 3, 3.5, 4$

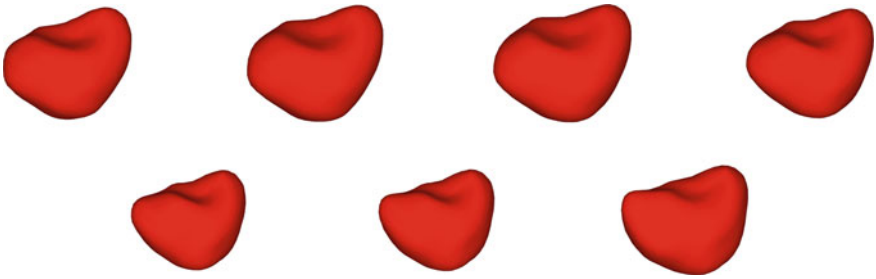

Fig. 13.4 Spline interpolation of sequence in Fig. 13.2 with seven time points (left to right and top to bottom): $t = 1, 1.5, 2, 2.5, 3, 3.5, 4$

with $\rho, u \in V^*$, $\mathbb{K}$ the inverse duality operator of V, and where the second equation in the system is the EPDiff equation. Notice that the norm on u in the integral is left unspecified, and there is much flexibility in choosing it, because u now belongs to a fixed space, V^*. One can take, in particular, any metric on a space W^* that is continuously embedded in V^* (so that V is embedded in W), bringing more regularity constraints to the control u. This includes, in particular, the L^2 norm, which significantly simplifies the implementation of the problem. Figures 13.2, 13.3, 13.4, 13.5 and 13.6 compares the interpolation schemes on a sequence of four target surfaces. The piecewise geodesic and spline methods interpolate the target almost exactly, with some small differences at intermediate points. The geodesic interpolation is more regular, but makes large errors interpolating the sequence. The difference between the methods is especially apparent when plotting the volumes of the interpolated surfaces over time (Fig. 13.6).

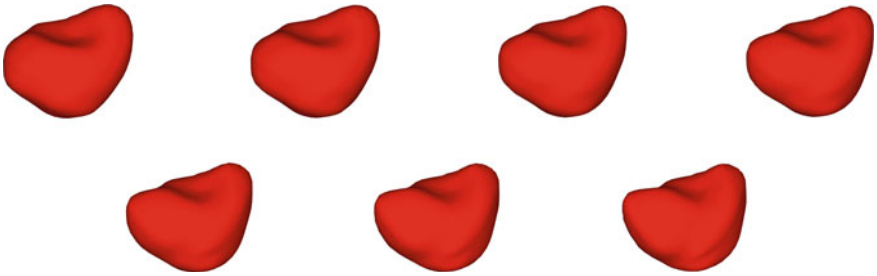

Fig. 13.5 Geodesic interpolation of sequence in Fig. 13.2 with seven time points (left to right and top to bottom): $t = 1, 1.5, 2, 2.5, 3, 3.5, 4$

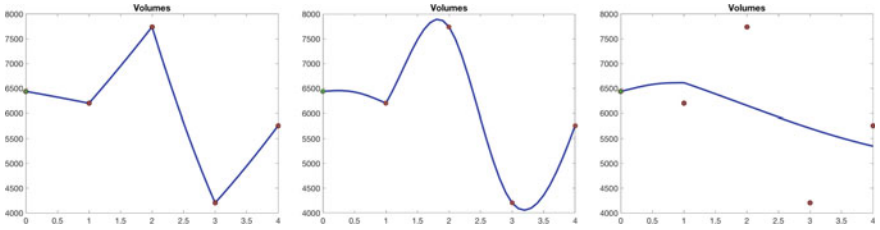

Fig. 13.6 Evolution of the volumes of the interpolated surfaces in Fig. 13.2. The first leg of the trajectory (from time 0 to 1) is the adjustment of the baseline starting with the volume of the "hypertemplate". The dots that follow are the volumes of the target surfaces. The time interpolation step is $\delta t = 0.1$. From left to right: piecewise geodesic, spline and geodesic

13.3.2 Multiple Trajectories

We now consider the situation in which several time series are observed, and start with the problem of computing an average trajectory from them. Assume, to begin with, that one observes full trajectories in the form of functions $m_k : [0, 1] \to M$ (assuming that the time interval has been rescaled to $[0, 1]$), for $k = 1, \ldots, n$. The goal is to compute an average trajectory $\bar{m}$.

The simplest and most direct approach is to apply one of the averaging methods that were discussed in Sect. 13.1 to each time coordinate separately. For example, one can define $\bar{m}(\tau)$ as a Fréchet mean, minimizing

$$\sum_{k=1}^{n} d_M(\bar{m}(\tau), m_k(\tau))^2$$

for each τ. This requires however that the observed trajectories are correctly aligned with each other, which may be valid in some contexts (e.g., for cardiac motion, which can be parametrized using well defined epochs in the cardiac cycle) but not always. In the general case, averaging has to be combined with some time realignment.

As a possible approach, let us consider this problem within the metamorphosis framework that was discussed in the previous chapter, which can be used to place a Riemannian metric on the space of trajectories, with respect to which Fréchet means can be computed while allowing for changes of parametrization. So, consider a metamorphosis metric in which the acting group is the space of diffeomorphisms of $\Omega = [0, 1]$ acting on curves via $g \cdot m = m \circ g^{-1}$ (reparametrization). Consider an RKHS H on functions defined over [0, 1] that vanish at 0 and at 1. We can associate a metamorphosis to the function

$$F(\xi, m, z) = \|\xi\|_H^2 + \frac{1}{\sigma^2}\|z\|_2^2$$

defined for $\xi \in H, m \in C^k(\Omega, M)$ and z a vector field along m (so that $z(\tau) \in T_{m(\tau)}$ for all $\tau \in \Omega$), with

$$\|z\|_2^2 = \int_0^1 \|z(\tau)\|_{m(\tau)}^2 \, d\tau.$$

The squared distance between $\bar{m}$ and m_k itself can be computed by minimizing

$$\int_0^1 \|\xi(t, \cdot)\|_H^2 \, dt + \frac{1}{\sigma^2} \int_0^1 \|z(t, \cdot)\|_2^2 \, dt$$

subject to $m(0, \cdot) = \bar{m}$, $m(1, \cdot) = m_k$ and $\partial_t m + \xi \partial_\tau m = z$. We use here the convention of denoting by $t \in [0, 1]$ the (numerical) metamorphosis time and $\tau \in \Omega$ $(= [0, 1])$ the "real" time associated with observed trajectories. Defining g as the flow of the equation $\partial_t g(t, \tau) = \xi(t, g(t, \tau))$, and letting $\alpha(t, \tau) = m(t, g(t, \tau))$, this objective function can be rewritten as

$$\int_0^1 \|\xi(t, \cdot)\|_H^2 \, dt + \frac{1}{\sigma^2} \int_0^1 \int_\Omega \|\partial_t \alpha(t, \tau)\|_{\alpha(t,\tau)}^2 \partial_\tau g \, d\tau \, dt,$$

which, after a change of variable in time, takes the form

$$\int_0^1 \|\xi(t, \cdot)\|_H^2 \, dt + \frac{1}{\sigma^2} \int_\Omega \left(\frac{1}{c_g(\tau)} \int_0^1 \|\partial_t \tilde{\alpha}(t, \tau)\|_{\tilde{\alpha}(t,\tau)}^2 \, dt \right) d\tau,$$

where

$$c_g(\tau) = \int_0^1 \partial_\tau g^{-1} \, dt$$

and $\tilde{\alpha}(t, \tau) = \alpha(\lambda(t, \tau), \tau)$ for some invertible time change $\lambda(\cdot, \tau)$ from [0,1] onto itself. (See the computation following Eq. (12.28).) This has to be minimized in ξ and α (or $\tilde{\alpha}$) with the constraints $\alpha(0) = \bar{m}$ and $\alpha(1) = m_k \circ g(1)$, $\partial_t g = \xi \circ g$. Using the fact that $\tilde{\alpha}(\cdot, \tau)$ minimizes the geodesic energy on M between $\bar{m}$ and $m_k \circ g(1)$, we finally find that computing the distance can be done by minimizing

$$\int_0^1 \|\xi(t,\cdot)\|_H^2\, dt + \frac{1}{\sigma^2}\int_\Omega \frac{d_M(\bar{m}(\tau), m_k(g(1,\tau)))^2}{c_g(\tau)}\, d\tau$$

with respect to ξ. A Fréchet mean between $m_1, \ldots, m_n$ for the metamorphosis metric should therefore minimize, with respect to $\bar{m}$ and $\xi_1, \ldots, \xi_n$,

$$\sum_{k=1}^n \int_0^1 \|\xi_k(t,\cdot)\|_H^2\, dt + \frac{1}{\sigma^2}\sum_{k=1}^n \int_\Omega \frac{d_M(\bar{m}(\tau), m_k(g_k(1,\tau)))^2}{c_{g_k}(\tau)}\, d\tau$$

with $\partial_t g_k(t,\tau) = \xi_k(t, g_k(t,\tau))$. One can use an alternating minimization scheme to solve this problem since, with fixed $\bar{m}$, $\xi_1, \ldots, \xi_n$ are solutions of independent "ordinary" metamorphosis problems on M, and for fixed $\xi_1, \ldots, \xi_n$, the average $\bar{m}(\tau)$ can be obtained, for each τ, as a weighted Fréchet average minimizing

$$\sum_{k=1}^n d_M(\bar{m}(\tau), m_k(g_k(1,\tau)))^2 / c_{g_k}(\tau).$$

Some modifications to this formulation are still needed in the case of shape spaces acted upon by diffeomorphisms, because the geodesic distance in this case is generally not computable exactly, but must be approximated through algorithms such as LDDMM. For example, on spaces of surfaces, one can replace $d_M(\bar{m}(\tau), m_k(g_k(1,\tau)))^2$ by the minimizer of

$$\int_0^1 \|v_k^\tau\|_V^2\, dt + D(\varphi_{01}^{v_k^\tau} \cdot \bar{m}, m_k \circ g_k(1,\tau))$$

for some discrepancy measure D, such as those described in Sect. 9.7.3. The minimization then needs to be done with respect to $\xi_1, \ldots, \xi_n$, $v_1, \ldots, v_n$ and $\bar{m}$. When $\xi_1, \ldots, \xi_n$ is fixed, this is the same problem as the one considered in (13.1) and below, and can be solved separately for each τ.

With fixed $\bar{m}$, $v_1, \ldots, v_n$ the problem splits into n independent problems, each of them requiring the minimization of a function taking the form

$$\int_0^1 \|\xi_k\|_H^2\, dt + \int_\Omega \frac{\Phi_k(\tau, m_k \circ g_k(1,\tau))}{c_{g_k}(\tau)}\, d\tau,$$

with

$$\Phi(\tau, \tilde{m}) = \int_0^1 \|v_k^\tau\|_V^2\, dt + D(\varphi_{01}^{v_k^\tau} \cdot \bar{m}, \tilde{m}).$$

The gradient of this objective function with respect to ξ_k can be obtained using the formulas developed in Sect. C.5 for the differentiation of solutions of ordinary differential equations (we skip the details). One can obtain a simpler method

by disregarding the weights c_g that come from the metamorphosis metric, and just minimize

$$\sum_{k=1}^{n} \int_0^1 \|\xi_k\|_H^2 \, dt + \sum_{k=1}^{n} \int_0^1 \int_\Omega \|v_k^\tau\|_V^2 \, d\tau \, dt + \sum_{k=1}^{n} \int_\Omega D(\varphi_{01}^{v_k^\tau} \cdot \bar{m}, m_k \circ g_k(1,\tau)) \, d\tau,$$

leading to a formulation similar to that developed in [95].

When dealing with sparse observations, i.e., when the kth trajectory is observed at a small number of time points $\tau_{k,1}, \ldots, \tau_{k,j_k}$, one can simply replace integrals over Ω by discrete sums, so that the last two terms in the previous expression become

$$\sum_{k=1}^{n} \sum_{i=1}^{j_k} \int_0^1 \|v_k^i\|_V^2 \, dt + \sum_{k=1}^{n} \sum_{i=1}^{j_k} D(\varphi_{01}^{v_k^i} \cdot \bar{m}, m_k \circ g_k^i).$$

Some regularization must then also be added to this objective function to ensure that $\bar{m}$ is smooth as a function of τ. One can, for example, ensure that $\tau \mapsto \bar{m}(\tau)$ is a geodesic on M, or a Riemannian spline as described in the previous section.

13.3.3 Reference-Centered Representations of Time Series

We now focus on methods that place the observed trajectories in a single coordinate system, allowing for the use of a statistical methods designed for linear spaces. This was not done in the previous discussion, which addressed the computation of an average curve.

We first point out that the reference-based representation discussed in Sect. 13.1 is still an option here, in the sense that, given a reference $\bar{m}_0 \in M$, one can still consider a representation of a family of curves $m_1(\cdot), \ldots, m_n(\cdot)$ as $v_1(\cdot), \ldots, v_n(\cdot)$, where $v_1, \ldots, v_n$ are curves in $T_{\bar{m}_0} M$ such that $m_k(\tau) = \mathrm{Exp}_{\bar{m}_0}(v_k(\tau))$ for all τ. This approach, or its registration counterpart in which one computes a collection of diffeomorphisms φ_k^τ for $k = 1, \ldots, n$ such that $\varphi_k^\tau \cdot \bar{m}_0 = m_k$, is probably the most commonly used in applications.

However, when using this approach, it is difficult to untangle the part of $v_k(\cdot)$ that describes the evolution within the trajectory from that describing the translation from the reference to that trajectory. Because of this, several methods have been designed that move trajectories as a whole rather than each point individually. More precisely, assume that each trajectory, m_k, has a representation with respect to its own reference, or baseline, $\bar{m}_k$ in the form

$$m_k(\tau) = \mathrm{Exp}_{\bar{m}_k}(v_k(\tau)),$$

with $v_k(\tau) \in T_{\bar{m}_k} M$ for all τ. (If one uses, for example, geodesic regression, then $v_k(\tau) = \tau v_k(1)$.) Given a global reference, $\bar{m}_0$, one builds a reference-centered rep-

resentation by "translating" each $v_k(\tau)$ from $T_{\bar{m}_k}M$ to $T_{\bar{m}_0}M$. Notice that, in Euclidean spaces, this operation is trivial, because $v_k(\tau) = m_k(\tau) - \bar{m}_k$ and its translation is just itself!

In Riemannian manifolds, the natural operation for translating tangent vectors is parallel transport, which is described in Sect. B.6.5. This operation must be done along a curve in M connecting the original basis point of the vector that needs to be translated to its target, and the result depends on the chosen curve. When no such curve is specified, it is also natural to choose a minimizing geodesic.

Let, therefore, γ_k be a geodesic such that $\gamma_k(1) = \bar{m}_k$ and $\gamma_k(0) = \bar{m}_0$. The representation of the trajectory m_k in $T_{\bar{m}}M$ is then given by $w_k(\cdot)$ such that $w_k(\tau)$ is the parallel transport of $v_k(\tau)$ along γ_k. After applying this to all curves, we indeed end up with a description of the dataset given by $w_1, \ldots, w_n$, which are all curves in $T_{\bar{m}_0}M$. Parallel transport was introduced for the analysis of manifold data in [163, 176] and for groups of diffeomorphisms and the associated shape spaces in [309], followed by [235, 303, 312].

Using this construction, one therefore represents each trajectory in the form

$$m_k(\tau) = \mathrm{Exp}_{\gamma_k(1)}(\mathcal{T}_{\gamma_k,0,1} w_k(\tau)),$$

where $\mathcal{T}_{\gamma,0,\tau}$ denotes the parallel transport along γ from time 0 to τ. One can also use an alternative approach, proposed in [94] (to which we refer for more details), in which the construction is done in the reverse order. In addition to $w_1, \ldots, w_n$ in $T_{\bar{m}}M$, this approach also requires an average curve $\bar{m}(\cdot)$ with $\bar{m}(0) = \bar{m}_0$, and the observed trajectories are represented in the form

$$m_k(\tau) = \mathrm{Exp}_{\bar{m}(\tau)}(\mathcal{T}_{\bar{m},0,\tau}(w_k(\tau))) .$$

We conclude this chapter with a description of the parallel transport equations in the diffeomorphism groups and their image via Riemannian submersions. Let, as usual, V be an admissible Hilbert space, and consider the right-invariant metric on Diff defined by $\|\delta\varphi\|_\varphi = \|\delta\varphi \circ \varphi^{-1}\|_V$. One can check (after a lengthy application of Eq. (B.9)) that the Levi-Civita connection on this space is given by

$$\nabla_X Y(\varphi) = \left(\frac{1}{2}(\mathbb{K}\,\mathrm{ad}_v^* \mathbb{L} w + \mathbb{K}\,\mathrm{ad}_w^* \mathbb{L} v - \mathrm{ad}_v w) + X w\right) \circ \varphi,$$

where $v(\varphi) = X(\varphi) \circ \varphi^{-1}$ and $w(\varphi) = Y(\varphi) \circ \varphi^{-1}$ are functions defined on Diff and taking values in V, and $\mathrm{ad}_v w = dv\, w - dw\, v$. In particular, if φ depends on time with $\partial_t \varphi = v \circ \varphi$ and $Y(t) = w(t) \circ \varphi(t)$ is a vector field along this curve, then

$$\frac{DY}{Dt} = \left(\frac{1}{2}(\mathbb{K}\,\mathrm{ad}_v^* \mathbb{L} w + \mathbb{K}\,\mathrm{ad}_w^* \mathbb{L} v - \mathrm{ad}_v w) + \partial_t w\right) \circ \varphi$$

and parallel transport is equivalent to

$$\partial_t w + \frac{1}{2}(\mathbb{K}\,\mathrm{ad}_v^* \mathbb{L} w + \mathbb{K}\,\mathrm{ad}_w^* \mathbb{L} v - \mathrm{ad}_v w) = 0.$$

Taking $v = w$, one retrieves the geodesic equation (EPDiff) given by $\partial_t v + \mathbb{K}\,\mathrm{ad}_v^* \mathbb{L} v = 0$.

Now consider a shape space M on which Diff acts, such that $\pi(\varphi) = \varphi \cdot m_0$ is a Riemannian submersion (for a fixed $m_0 \in M$). The vertical space at φ is the space $V_m \circ \varphi$, where $V_m = \{v : v \cdot m = 0\}$, and the horizontal space is $H_m \circ \varphi$ with $H_m = V_m^\perp$. The horizontal lift of $\xi \in T_m M$ is the unique vector $v^\xi \in H_m$ such that $v^\xi \cdot m = \xi$ (cf. Sect. 11.5).

If $m(\cdot)$ is a curve on M and $\eta_0 \in T_{m(0)} M$, its parallel transport $\eta(\cdot)$ along m is characterized by

$$\left(\partial_t v^\eta + \frac{1}{2}(\mathbb{K}\,\mathrm{ad}_{v^\xi}^* \mathbb{L} v^\eta + \mathbb{K}\,\mathrm{ad}_{v^\eta}^* \mathbb{L} v^\xi - \mathrm{ad}_{v^\xi} v^\eta)\right) \cdot m(t) = 0$$

at all times, with $\xi = \partial_t m$ (this results from Eq. (B.15)). Assume, to simplify the discussion, that M is an open subset of a Banach space Q (otherwise, consider the following computation as valid in a local chart). Writing $v^\eta \cdot m = \eta$, we have $\partial_t \eta = (\partial_t v^\eta) \cdot m + dA_{v^\eta}(m)\xi$, where we have denoted by A_w the mapping $m \mapsto w \cdot m$. Using this, we obtain the parallel transport equation along m

$$\partial_t \eta - dA_{v^\eta}(m)\xi + \left(\frac{1}{2}(\mathbb{K}\,\mathrm{ad}_{v^\xi}^* \mathbb{L} v^\eta + \mathbb{K}\,\mathrm{ad}_{v^\eta}^* \mathbb{L} v^\xi - \mathrm{ad}_{v^\xi} v^\eta)\right) \cdot m(t) = 0\,. \quad (13.6)$$

On shape spaces of point sets, i.e., $m = (x_1, \ldots, x_N)$, the infinitesimal action is just $v \cdot m = (v(x_1), \ldots, v(x_n))$ and the horizontal lift is such that

$$\mathbb{L} v^\xi = \sum_{k=1}^N \alpha_k^\xi \delta_{x_k},$$

where $(\alpha_1^\xi, \ldots, \alpha_N^\xi)$ are obtained by solving the equations $\sum_{j=1}^N K(x_k, x_i)\alpha_j^\xi = \xi_k$. Moreover, we have

$$dA_{v^\eta}\xi = \sum_{k=1}^N (\partial_1 K(x_k, x_i)\xi_k)\alpha_j^\eta.$$

This makes all terms in (13.6) explicit.

The situation is not as simple on spaces of images, in which $\varphi \cdot m = m \circ \varphi^{-1}$ and $v \cdot m = -\nabla m^T v$. One has, in this case, $dA_{v^\eta}\xi = -\nabla \xi^T v$, which is simple, but the horizontal lift of ξ consists of minimizing $\|v\|_V^2$ subject to $\xi = -\nabla m^T v$. While this problem has a unique minimizer, the characterization of this minimizer using Lagrange multipliers requires finding a Banach space W such that $h \mapsto -\nabla m^T h$, from V to W, is bounded and has closed range (see Theorem D.4). This problem is, to our knowledge, still open in the general case.

Appendix A
Elements from Functional Analysis

This first chapter of the appendix includes results in functional analysis that are needed in the main part of the book, focusing mostly on Hilbert spaces, and providing proofs when these are simple enough. Many important results of the theory are left aside, and the reader is referred to the many treatises on the subject, including [45, 74, 246, 306] and other references listed in this chapter, for a comprehensive account.

A.1 Definitions and Notation

Definition A.1 A set H is a (real) Hilbert space if:

(i) H is a vector space on $\mathbb{R}$.
(ii) H has an inner product denoted $(h, h') \mapsto \langle h\ ,\ h' \rangle_H$, for $h, h' \in H$. This inner product is a symmetric positive definite bilinear form. The associated norm is denoted $\|h\|_H = \sqrt{\langle h\ ,\ h \rangle_H}$.
(iii) H is a complete space with respect to the topology associated to the norm.

If condition (ii) is weakened to the fact that $\|\cdot\|_H$ is a norm (not necessarily induced by an inner product), one says that H is a Banach space.

Convergent sequences in the norm topology are sequences h_n for which there exists an $h \in H$ such that $\|h - h_n\|_H \to 0$. Property (iii) means that if a sequence $(h_n, n \geq 0)$ in H is a Cauchy sequence, i.e., it collapses in the sense that, for every positive ε there exists an $n_0 > 0$ such that $\|h_n - h_{n_0}\|_H \leq \varepsilon$ for $n \geq n_0$, then it necessarily has a limit: there exists an $h \in H$ such that $\|h_n - h\|_H \to 0$ as n tends to infinity.

If H satisfies (i) and (ii), it is called a pre-Hilbert space. On pre-Hilbert spaces, the Schwarz inequality holds:

Proposition A.2 (Schwarz inequality) *If H is pre-Hilbert, and $h, h' \in H$, then*

$$\langle h\ ,\ h' \rangle_H \leq \|h\|_H\, \|h'\|_H .$$

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

The first consequence of this property is:

Proposition A.3 *The inner product on H is continuous in the norm topology.*

Proof The inner product is a function $H \times H \to \mathbb{R}$. Letting $\varphi(h, h') = \langle h\ ,\ h' \rangle_H$, we have, by the Schwarz inequality, and introducing a sequence (h_n) which converges to h

$$\left|\varphi(h, h') - \varphi(h_n, h')\right| \leq \|h - h_n\|_H \, \|h'\|_H \to 0,$$

which proves the continuity with respect to the first coordinate, and also with respect to the second coordinate by symmetry. □

Working with a complete normed vector space is essential when dealing with infinite sums of elements: if (h_n) is a sequence in H, and if $\|h_{n+1} + \cdots + h_{n+k}\|$ can be made arbitrarily small for large n and any $k > 0$, then completeness implies that the series $\sum_{n=0}^{\infty} h_n$ has a limit in h. In particular, absolutely converging series in H converge:

$$\sum_{n \geq 0} \|h_n\|_H < \infty \Rightarrow \sum_{n \geq 0} h_n \text{ converges.} \tag{A.1}$$

(In fact, (A.1) is equivalent to (iii) in normed spaces.)

We add a fourth condition to Definition A.1.

(iv) H is separable in the norm topology: there exists a countable subset $\mathcal{S}$ in H such that, for any $h \in H$ and any $\varepsilon > 0$, there exists an $h' \in \mathcal{S}$ such that $\left\|h - h'\right\|_H \leq \varepsilon$.

In the following, Hilbert spaces will always be separable without further mention.

An isometry between two Hilbert spaces H and H' is an invertible linear map $\varphi : H \to H'$ such that, for all $h, h' \in H$, $\langle \varphi(h)\ ,\ \varphi(h') \rangle_{H'} = \langle h\ ,\ h' \rangle_H$.

A Hilbert subspace of H is a subspace H' of H (i.e., a non-empty subset, invariant under linear combination) which is closed in the norm topology. Closedness implies that H' is itself a Hilbert space, because Cauchy sequences in H' are also Cauchy sequences in H, hence converge in H, hence in H' because H' is closed. The next proposition shows that every finite-dimensional subspace is a Hilbert subspace.

Proposition A.4 *If K is a finite-dimensional subspace of H, then K is closed in H.*

Proof Let $e_1, \ldots, e_p$ be a basis of K. Let (h_n) be a sequence in K which converges to some $h \in H$; we need to show that $h \in K$. We have, for some coefficients (a_{kn}), $h_n = \sum_{k=1}^{p} a_{kn} e_k$ and for all $l = 1, \ldots, p$:

$$\langle h_n\ ,\ e_l \rangle_H = \sum_{k=1}^{p} \langle e_k\ ,\ e_l \rangle_H a_{kn}.$$

Let a_n be the vector in $\mathbb{R}^p$ with coordinates $(a_{kn}, k = 1, \ldots, p)$ and $u_n \in \mathbb{R}^p$ with coordinates $(\langle h_n\ ,\ e_k \rangle_H, k = 1, \ldots, p)$. Let also S be the matrix with coefficients

$s_{kl} = \langle e_k\ ,\ e_l \rangle_H$, so that the previous system may be written: $u_n = Sa_n$. The matrix S is invertible: if b belongs to the null space of S, a quick computation shows that

$$b^T Sb = \left\| \sum_{i=1}^{p} b_i e_i \right\|_H^2 = 0,$$

which is only possible when $b = 0$, because $(e_1, \ldots, e_n)$ is a basis of K. We therefore have $a_n = S^{-1}u_n$. Since u_n converges to u with coordinates $(\langle h\ ,\ e_k \rangle_H, k = 1, \ldots, p)$ (by continuity of the inner product), we obtain the fact that a_n converges to $a = S^{-1}u$. But this implies that $\sum_{k=1}^{p} a_{kn} e_k \to \sum_{k=1}^{p} a_k e_k$ and, since the limit is unique, we have $h = \sum_{k=1}^{p} a_k e_k \in K$. □

A.2 First Examples

A.2.1 Finite-Dimensional Euclidean Spaces

$H = \mathbb{R}^n$ with $\langle h\ ,\ h' \rangle_{\mathbb{R}^n} = h^T h' = \sum_{i=1}^{n} h_i h'_i$ is the standard example of a finite-dimensional Hilbert space.

A.2.2 The ℓ^2 Space of Real Sequences

Let H be the set of real sequences $h = (h_1, h_2, \ldots)$ such that $\sum_{i=1}^{\infty} h_i^2 < \infty$. Then H is a Hilbert space, with dot product

$$\langle h\ ,\ h' \rangle_2 = \sum_{i=1}^{\infty} h_i h'_i.$$

A.2.3 The L^2 Space of Functions

Let k and d be two integers. Let Ω be an open subset of $\mathbb{R}^k$. We define $L^2(\Omega, \mathbb{R}^d)$ as the set of all square integrable functions $h : \Omega \to \mathbb{R}^d$, with inner product

$$\langle h\ ,\ h' \rangle_2 = \int_\Omega h(x) h'(x) dx.$$

Integrals are taken with respect to the Lebesgue measure on Ω, and two functions which coincide everywhere except on a set of null Lebesgue measure are identified. The fact that $L^2(\Omega, \mathbb{R}^d)$ is a Hilbert space is a standard result in integration theory.

A.3 Orthogonal Spaces and Projection

Let O be a subset of H. The orthogonal space to O is defined by

$$O^{\perp} = \left\{h \in H : \forall o \in O, \langle h\,,\, o\rangle_H = 0\right\}.$$

Theorem A.5 *$O^{\perp}$ is a Hilbert subspace of H.*

Proof Stability under linear combination is obvious, and closedness is a consequence of the continuity of the inner product. □

When K is a subspace of H and $h \in H$, one can define the variational problem:

$$(\mathcal{P}_K(h)):\ \text{find } k \in K \text{ such that } \|k-h\|_H = \min\left\{\left\|h-k'\right\|_H : k' \in K\right\}.$$

The following theorem is fundamental:

Theorem A.6 *If K is a closed subspace of H and $h \in H$, $(\mathcal{P}_K(h))$ has a unique solution, k, characterized by the conditions: $k \in K$ and $h-k \in K^{\perp}$.*

Definition A.7 The solution of problem $(\mathcal{P}_K(h))$ in the previous theorem is called the orthogonal projection of h on K and denoted $\pi_K(h)$.

Proposition A.8 *$\pi_K : H \to K$ is a linear, continuous transformation and $\pi_{K^{\perp}} = \mathrm{id} - \pi_K$.*

Proof We prove the theorem and the proposition together. Let

$$d = \inf\left\{\left\|h-k'\right\|_H : k' \in K\right\}.$$

If $k, k' \in K$, a direct computation shows that

$$\left\|h - \frac{k+k'}{2}\right\|_H^2 + \left\|k-k'\right\|_H^2 \Big/ 4 = \left(\|h-k\|_H^2 + \left\|h-k'\right\|_H^2\right)\Big/ 2.$$

The fact that $(k+k')/2 \in K$ implies $\left\|h - \frac{k+k'}{2}\right\|_H^2 \geq d$, so that

$$\left\|k-k'\right\|_H^2 \leq \frac{1}{2}\left(\|h-k\|_H^2 + \left\|h-k'\right\|_H^2\right) - d^2.$$

Now, from the definition of the infimum, one can find a sequence k_n in K such that $\|k_n - h\|_H^2 \leq d^2 + 2^{-n}$ for each n. The previous inequality implies that

$$\|k_n - k_m\|_H^2 \leq \left(2^{-n} + 2^{-m}\right)/2,$$

which implies that k_n is a Cauchy sequence, which therefore converges to a limit k that belongs to K (because K is closed) and is such that $\|k-h\|_H = d$. If the minimum is attained for another k' in K, we have, by the same inequality

$$\left\|k-k'\right\|_H^2 \leq \frac{1}{2}\left(d^2+d^2\right) - d^2 = 0,$$

so that $k = k'$, uniqueness is proved, hence π_K is well-defined.

Let $k = \pi_K(h), k' \in K$ and consider the function $f(t) = \left\|h-k-tk'\right\|_H^2$, which is by construction minimal for $t=0$. We have $f(t) = \|h-k\|_H^2 - 2t\left\langle h-k\,,\,k'\right\rangle_H + t^2\left\|k'\right\|_H^2$ and this can be minimal at 0 only if $\left\langle h-k\,,\,k'\right\rangle_H = 0$. Since this has to be true for every $k' \in K$, we obtain the fact that $h-k \in K^\perp$. Conversely, if $k \in K$ and $h-k \in K^\perp$, we have for any $k' \in K$

$$\left\|h-k'\right\|_H^2 = \|h-k\|_H^2 + \left\|k-k'\right\|_H^2 \geq \|h-k\|_H^2\,,$$

so that $k = \pi_K(h)$. This proves Theorem A.6.

We now prove Proposition A.8: let $h, h' \in H$ and $\alpha, \alpha' \in \mathbb{R}$. Let $k = \pi_K(h), k' = \pi_K(h')$; we want to show that $\pi_K(\alpha h + \alpha' h') = \alpha k + \alpha' k'$, for which it suffices to prove (since $\alpha k + \alpha' k' \in K$) that $\alpha h + \alpha' h' - \alpha k - \alpha' k' \in K^\perp$. This is true because $\alpha h + \alpha' h' - \alpha k - \alpha' k' = \alpha(h-k) + \alpha'(h'-k')$, $h-k \in K^\perp$, $h'-k' \in K^\perp$, and $K^\perp$ is a vector space. Continuity comes from

$$\|h\|_H^2 = \|h-\pi_K(h)\|_H^2 + \|\pi_K(h)\|_H^2\,,$$

so that $\|\pi_K(h)\|_H \leq \|h\|_H$.

Finally, if $h \in H$ and $k = \pi_K(h)$, then $k' = \pi_{K^\perp}(h)$ is characterized by $k' \in K^\perp$ and $h-k' \in (K^\perp)^\perp$. The first property is certainly true for $h-k$, and for the second, we need to show that $K \subset (K^\perp)^\perp$, which is a direct consequence of the definition of the orthogonal space. □

We have the interesting property:

Corollary A.9 *K is a Hilbert subspace of H if and only if $(K^\perp)^\perp = K$.*

Proof The $\Leftarrow$ implication is a consequence of Theorem A.5. Now assume that K is a Hilbert subspace. The fact that $K \subset (K^\perp)^\perp$ is obviously true for any subset $K \subset H$, so that it suffices to show that every element of $(K^\perp)^\perp$ belongs to K. Assume that $h \in (K^\perp)^\perp$: this implies that $\pi_{K^\perp}(h) = 0$ but since $\pi_{K^\perp}(h) = h - \pi_K(h)$, this implies that $h = \pi_K(h) \in K$. □

If K is finite-dimensional, i.e.,

$$K = \left\{\sum_{i=1}^n \alpha_i f_i, \alpha_1, \ldots, \alpha_n \in \mathbb{R}\right\}$$

where $f_1, \ldots, f_n$ are linearly independent elements of H, then π_K can be calculated in closed form. Indeed, if $h \in H$, its orthogonal projection $\hat{h} = \pi_K(h)$ is characterized by $\hat{h} \in K$ and $\langle \hat{h}\ ,\ f_i \rangle_H = \langle h\ ,\ f_i \rangle_H$ for $h = 1, \ldots, N$. This implies that

$$\hat{h} = \sum_{i=1}^{n} \alpha_i f_i$$

where $\alpha_1, \ldots, \alpha_n$ are obtained by solving the system

$$\sum_{j=1}^{n} \langle f_i\ ,\ f_j \rangle_H \alpha_i = \langle f_i\ ,\ h \rangle_H.$$

A.4 Orthonormal Sequences

A sequence $(e_1, e_2, \ldots)$ in a Hilbert space H is orthonormal if and only if $\langle e_i\ ,\ e_j \rangle_H = 1$ if $i = j$ and 0 otherwise. In such a case, if $\alpha = (\alpha_1, \alpha_2, \ldots) \in \ell^2$ (the space defined is Sect. A.2.2), the series $\sum_{i=1}^{\infty} \alpha_i e_i$ converges in H (its partial sums form a Cauchy sequence) and if h is the limit, then $\alpha_i = \langle h\ ,\ e_i \rangle_H$.

Conversely, if $h \in H$, then the sequence $(\langle h\ ,\ e_1 \rangle_H, \langle h\ ,\ e_2 \rangle_H, \ldots)$ belongs to ℓ^2. Indeed, letting $h_n = \sum_{i=1}^{n} \langle h\ ,\ e_i \rangle_H e_i$, one has

$$\langle h_n\ ,\ h \rangle_H = \sum_{i=1}^{n} \langle h\ ,\ e_i \rangle_H^2 = \|h_n\|_H^2\ .$$

On the other hand, one has, by Schwarz's inequality $\langle h_n\ ,\ h \rangle_H \leq \|h_n\|_H\ \|h\|_H$, which implies that $\|h_n\|_H \leq \|h\|_H$: therefore

$$\sum_{i=1}^{\infty} \langle h\ ,\ e_i \rangle^2 < \infty.$$

Denoting by $K = \mathrm{Hilb}(e_1, e_2, \ldots)$ the smallest Hilbert subspace of H containing this sequence, one has the identity

$$K = \left\{ \sum_{n=1}^{\infty} \alpha_k e_k : (\alpha_1, \alpha_2, \ldots) \in l^2 \right\}. \tag{A.2}$$

The proof is left to the reader. As a consequence of this, we see that $h \mapsto (\langle h\ ,\ e_1 \rangle_H, \langle h\ ,\ e_2 \rangle_H, \ldots)$ is an isometry between $\mathrm{Hilb}(e_1, e_2, \ldots)$ and ℓ^2. Moreover, we have, for $h \in H$

$$\pi_K(h) = \sum_{i=1}^{\infty} \langle h\,,\ e_i \rangle_H e_i$$

(because $h - \pi_K(h)$ is orthogonal to every e_i).

An orthonormal set $(e_1, e_2, \ldots)$ is *complete* in H if $H = \mathrm{Hilb}(e_1, e_2, \ldots)$. In this case, we see that H is itself isometric to ℓ^2, and an interesting point is that (in the separable case) this fact is always true.

Theorem A.10 *Every (separable) Hilbert space has a complete orthonormal sequence.*

A complete orthonormal sequence in H is also called an orthonormal basis of H.

Proof The proof relies on the important Schmidt orthonormalization procedure. Let $f_1, f_2, \ldots$ be a dense sequence in H. We let $e_1 = f_{k_1} / \left\| f_{k_1} \right\|_H$ where f_{k_1} is the first non-vanishing element in the sequence.

Assume that, for $n \geq 1$, an orthonormal sequence $(e_1, \ldots, e_n)$ has been constructed with a sequence $k_1, \ldots, k_n$ such that $e_i \in V_i := \mathrm{span}(f_1, \ldots, f_{k_i})$ for each i. First assume that $f_k \in \mathrm{span}(e_1, \ldots, e_n)$ for all $k > k_n$: then $H = V_n$ is finite-dimensional. Indeed, H is equal to the closure of $(f_1, f_2, \ldots)$, which is included in $\mathrm{span}(e_1, \ldots, e_n)$, which is closed as a finite-dimensional vector subspace of H.

Otherwise, let k_{n+1} be the smallest $k > k_n$ such that $f_{k_{n+1}} \notin V_n$. Then, we may set

$$e_{n+1} = \lambda \left(f_{k_{n+1}} - \pi_{V_n}(f_{k_{n+1}}) \right),$$

where λ is selected so that e_{n+1} has unit norm, which is always possible.

So there are two cases: either the previous construction stops at some point, and H is finite-dimensional and the theorem is true, or the process carries on indefinitely, yielding an orthonormal sequence $(e_1, e_2, \ldots)$ and an increasing sequence of integers $(k_1, k_2, \ldots)$ such that $\mathrm{span}(e_1, \ldots, e_n) = \mathrm{span}(f_1, \ldots, f_{k_n})$. In this case because $\mathrm{Hilb}(e_1, e_2, \ldots)$ contains (f_n), which is dense in H, this space must be equal to H, which shows that the orthonormal sequence is complete. □

A.5 Nested Subspaces

Consider a sequence $(H_n, n \geq 1)$ of Hilbert subspaces of H such that $H_1 \subset H_2 \subset \cdots$. Let $H_\infty = \overline{\bigcup_{n=1}^{\infty} H_n}$ be the smallest Hilbert space that contains all H_n. Then we have the following result.

Proposition A.11 *Let $h \in H$. With the notation above, one has*

$$\lim_{n \to \infty} \pi_{H_n}(h) = \pi_{H_\infty}(H).$$

Proof Let $h_n = \pi_{H_n}(h)$ (including for $n = \infty$). Then $\pi_{H_n}(h_{n+1}) = h_n$ and $(h_1, h_2 - h_1, h_3 - h_2, \ldots)$ forms an orthogonal sequence. This implies that (letting $h_0 = 0$)

$$\sum_{k=1}^{n} \|h_k - h_{k-1}\|_H^2 = \left\| \sum_{k=1}^{n} (h_k - h_{k-1}) \right\|_H^2 = \|h_n\|_H^2.$$

Because $\|h_n\|_H \leq \|h\|_H$, we have $\sum_{k=1}^{n} \|h_k - h_{k-1}\|_H^2 < \infty$, and since, for $m < n$

$$\|h_n - h_m\|_H^2 = \sum_{k=m+1}^{n} \|h_k - h_{k-1}\|_H^2$$

we find that (h_n) is a Cauchy sequence, and therefore has a limit h' that necessarily belongs to H_∞. Moreover, for any $\tilde{h} \in \bigcup_{n \geq 1} H_n$, we have $\tilde{h} \in H_m$ for large enough m and $\langle h - h_n \,,\, \tilde{h} \rangle_H = 0$ for $n \geq m$. Passing to the limit $n \to \infty$, we find that $h - h'$ is perpendicular to $\bigcup_{n \geq 1} H_n$, and therefore also to H_∞, which shows that $h' = h_\infty$.

A.6 The Riesz Representation Theorem

The (topological) dual space of a normed vector space H is the space containing all bounded linear functionals $\varphi : H \to \mathbb{R}$. It is denoted H^*, and we will use the notation, for $\varphi \in H^*$ and $h \in H$:

$$\varphi(h) = (\varphi \mid h). \tag{A.3}$$

Thus, parentheses indicate linear forms and angles indicate inner products. Recall that φ being bounded means that, for some constant c,

$$(\varphi \mid h) \leq c \|h\|_H$$

for all $h \in H$, which is also equivalent to φ being continuous.

H being a normed space, H^* also has a normed space structure defined by:

$$\|\varphi\|_{H^*} = \sup \{(\varphi \mid h) : h \in H, \|h\|_H = 1\}.$$

When H is Hilbert, the function $\varphi_h : h' \mapsto \langle h \,,\, h' \rangle_H$ belongs to H^*, and by the Schwarz inequality $\|\varphi_h\|_{H^*} = \|h\|_H$. The Riesz representation theorem states that there exists no other bounded linear form on H.

Theorem A.12 (Riesz) *Let H be a Hilbert space. If $\varphi \in H^*$, there exists a unique $h \in H$ such that $\varphi = \varphi_h$.*

Proof For uniqueness, it suffices to prove that $\varphi_h = 0 \Rightarrow h = 0$, which is obvious since $\varphi_h(h) = \|h\|_H^2$. To prove existence, we introduce the orthogonal of the null space of φ. So, let $K = \left\{h' \in H, \left(\varphi \,\middle|\, h'\right) = 0\right\}$; K is a closed linear subspace of H (because φ is linear and continuous). If $K^\perp = \{0\}$, then $K = (K^\perp)^\perp = H$ and $\varphi = 0 = \varphi_h$ for $h = 0$, which proves the theorem in this case. So, assume that $K^\perp \neq 0$ and let $h_0 \in K^\perp$, $h_0 \neq 0$, which implies $(\varphi \mid h_0) \neq 0$. For all $h \in H$, we have

$$h - \frac{(\varphi \mid h)}{(\varphi \mid h_0)} h_0 \in K$$

so that (taking the inner product with h_0),

$$\left\langle h_0 \,,\, h \right\rangle_H - \frac{(\varphi \mid h)}{(\varphi \mid h_0)} \|h_0\|_H^2 = 0,$$

which is exactly $\varphi = \varphi_{\alpha h_0}$ with $\alpha = (\varphi \mid h_0)/\|h_0\|_H^2$. □

Riesz's Theorem and the fact that $\|\varphi_h\|_{H^*} = \|h\|_H$ imply that the linear map $\mathbb{L}_H : h \mapsto \varphi_h$ is an isometry between H and H^* (which is therefore also a Hilbert space). We will refer to it as the duality operator of H, and generally denote its inverse by $\mathbb{K}_H = \mathbb{L}_H^{-1} : H^* \to H$.

A.7 Embeddings

Assume that H and H_0 are two Banach spaces. An embedding of H in H_0 is a bounded, one-to-one, linear mapping from H to H_0, i.e., a mapping $\mathfrak{i} : H \to H_0$ such that, for all $h \in H$,

$$\|\mathfrak{i}(h)\|_{H_0} \leq C \, \|h\|_H \,. \tag{A.4}$$

The smallest C for which this is true is the operator norm of $\mathfrak{i}$, denoted $\|\mathfrak{i}\|_{op(H,H_0)}$. If such an embedding exists, we will say that H is (continuously) embedded in H_0.

One says that the embedding is *compact* when the set $\{\mathfrak{i}(h), \|h\|_H \leq 1\}$ is compact in H_0. In the separable case (to which we restrict ourselves), this means that for any bounded sequence $(h_n, n > 0)$ in H, there exists a subsequence of $(\mathfrak{i}(h_n))$ that converges in H_0. One says that the embedding is *dense* if $\mathfrak{i}(H)$ is dense in H_0.

In all the applications we will be interested in, H and H_0 will be function spaces, and we will have a set inclusion $H \subset H_0$. For example H may be a set of smooth functions and H_0 a set of less smooth functions (see the examples of embeddings below). Then, one says that H is embedded (resp. compactly embedded) in H_0 if the canonical inclusion map: $\mathfrak{i} : H \to H_0$ is continuous (resp. compact).

If φ is a bounded linear form on H_0, and $\mathfrak{i} : H \to H_0$ is an embedding, then one can define the form $\mathfrak{i}^*(\varphi)$ on H by $(\mathfrak{i}^*(\varphi) \mid h) = (\varphi \mid \mathfrak{i}(h))$, and $\mathfrak{i}^*(\varphi)$ is bounded on H. Indeed, we have, for all $h \in H$:

$$\left|\left(\mathfrak{i}^*(\varphi) \,\middle|\, h\right)\right| = |(\varphi \mid \mathfrak{i}(h))| \leq \|\varphi\|_{H_0^*} \|\mathfrak{i}(h)\|_{H_0} \leq \|\mathfrak{i}\|_{op(H,H_0)} \|\varphi\|_{H_0^*} \|h\|_H ,$$

where the first inequality comes from the continuity of φ and the last one from (A.4). This proves the boundedness of $\mathfrak{i}^*(\varphi)$ as a linear form on H together with the inequality:

$$\left\|\mathfrak{i}^*(\varphi)\right\|_{H^*} \leq \|\mathfrak{i}\|_{op(H,H_0)} \|\varphi\|_{H_0^*} .$$

This in fact proves the first statement of the theorem:

Theorem A.13 *Let $\mathfrak{i} : H \to H_0$ be a Banach space embedding. Then $\mathfrak{i}^* : H_0^* \to H^*$ is a bounded linear map.*

If $\mathfrak{i}$ is dense, then $\mathfrak{i}^*$ is also an embedding.

If H is a Hilbert space, and $\mathfrak{i}$ is dense, then $\mathfrak{i}^*$ is dense too.

Proof Assume that $\mathfrak{i}$ is dense. To prove that it is an embedding, it remains to show that it is one-to-one. So, assume that $\mathfrak{i}^*(\varphi) = 0$. This implies that $(\varphi \mid \mathfrak{i}(h)) = 0$ for all $h \in H$ so that φ vanishes on $\mathfrak{i}(H)$, which implies that $\varphi = 0$ because $\mathfrak{i}(H)$ is dense in H_0 and φ is continuous.

Still with $\mathfrak{i}$ dense, we now assume that H is a Hilbert space and prove that $\mathfrak{i}^*(H_0)$ is dense in H^*. Because H^* is also a Hilbert space, it suffices to show that no non-zero element of H^* is orthogonal to $\mathfrak{i}^*(H_0^*)$, which, by Theorem A.12, is equivalent to showing that no non-zero vector $h \in H$ is such that $(\mathfrak{i}^*(\varphi) \mid h) = 0$ for all $\varphi \in H_0^*$. But since $(\mathfrak{i}^*(\varphi) \mid h) = (\varphi \mid \mathfrak{i}(h))$, this is only possible when $\mathfrak{i}(h) = 0$, which yields $h = 0$ because $\mathfrak{i}$ is one-to-one. □

A.8 Examples

A.8.1 Banach Spaces of Continuous Functions

Let Ω be an open subset of $\mathbb{R}^d$. The space of continuous functions on Ω with at least p continuous derivatives is denoted $C^p(\Omega, \mathbb{R})$. The elements of $C^p(\Omega, \mathbb{R})$ that are bounded, together with their derivatives up to the pth order, form the set $C_b^p(\Omega, \mathbb{R})$, which is a Banach space for the norm

$$\|f\|_{p,\infty} = \max_{|J| \leq p} \|\partial_J f\|_\infty, \tag{A.5}$$

where $J = (j_1, \ldots, j_d)$ are d-tuples of non-negative integers, $|J| = j_1 + \cdots j_k$ and $\partial_J = \partial_1^{j_1} \ldots \partial_d^{j_d}$. We obviously have, almost by definition, the fact that $C_b^p(\Omega, \mathbb{R})$ is embedded in $C_b^q(\Omega, \mathbb{R})$ as soon as $p \leq q$.

Within the set $C_b^p(\Omega, \mathbb{R})$, we will distinguish functions that vanish at infinity according to the following definition.

Definition A.14 Let $\Omega \subset \mathbb{R}^d$ be open. A function $f \in C^p(\Omega, \mathbb{R})$ tends to 0 at infinity at the pth order if, for all $\varepsilon > 0$, there exists a compact subset $K \subset \Omega$ such that

$$\max_h \sup\{|h(x)|, x \in \Omega \setminus K\} < \varepsilon,$$

where h varies over the set of partial derivatives of f of order less than or equal to p.

We let $C_0^p(\Omega, \mathbb{R})$ denote the subset of $C^p(\Omega, \mathbb{R})$ containing functions that tend to 0 at infinity at the pth order. It forms a closed subset of $C_b^p(\Omega, \mathbb{R})$ and is therefore also Banach. If $f \in C_0^p(\Omega, \mathbb{R})$, it can be extended in a unique way to a function $\tilde{f} \in C_0^p(\mathbb{R}^d, \mathbb{R})$ such that $\tilde{f}(x) = 0$ for $x \in \Omega^c$. It is sometimes convenient to identify f and $\tilde{f}$ and therefore consider elements of $C_0^p(\Omega, \mathbb{R})$ as defined over all $\mathbb{R}^d$. Notice that one can also show that $C_0^p(\Omega, \mathbb{R})$ is the completion of the set of C^p functions with compact support included in Ω within $C_b^p(\Omega, \mathbb{R})$.

The following simple proposition will be useful.

Proposition A.15 *All functions in $C_0^0(\Omega, \mathbb{R})$ are uniformly continuous on Ω.*

Proof Let $f \in C_0^0(\Omega, \mathbb{R})$. Fix $\varepsilon > 0$ and K compact in Ω such that $|f(x)| < \varepsilon/2$ for $x \in \Omega \setminus K$. Let $c = \text{dist}(K, \Omega^c)/2$, which is a positive number. Let $K' = \{x \in \Omega, \text{dist}(x, K) \leq c\}$, which is also a compact subset of Ω. By compactness, there exists $\delta \in (0, c)$ such that $x, y \in K'$ and $|x - y| < \delta$ imply that $|f(x) - f(y)| < \varepsilon$.

Take $x, y \in \Omega$ with $|x - y| < \delta$. If neither x or y are elements of K, then $|f(x) - f(y)| \leq |f(x)| + |f(y)| < \varepsilon$. If at least one of them, say x, belongs to K, then $|x - y| < \delta$ and $\delta < c$ implies that $y \in K'$ and $|f(x) - f(y)| < \varepsilon$ also. This proves that f is uniformly continuous. □

If Ω is bounded, so that $\bar{\Omega}$ is compact, $C^p(\bar{\Omega}, \mathbb{R})$ will denote the set of functions on Ω which are p times differentiable on Ω, each partial derivative being extended to a continuous function on $\bar{\Omega}$; $C^p(\bar{\Omega}, \mathbb{R})$ has a Banach space structure when equipped with $\| \cdot \|_{p,\infty}$. Note that $C_0^p(\Omega, \mathbb{R})$ can be considered as a subset of $C^p(\bar{\Omega}, \mathbb{R})$.

For bounded Ω, relatively compact sets in $C^0(\bar{\Omega}, \mathbb{R})$ are exactly described by Ascoli's theorem: they are bounded subsets $M \subset C^0(\bar{\Omega}, \mathbb{R})$ (for the supremum norm) which are uniformly continuous, meaning that, for any $x \in \bar{\Omega}$, for any $\varepsilon > 0$ there exists an $\eta > 0$ such that

$$\sup_{\substack{y \in \Omega \\ |x-y|<\eta}} \sup_{h \in M} |h(x) - h(y)| < \varepsilon .$$

Compact sets in $C^p(\bar{\Omega}, \mathbb{R})$ are bounded subsets of $C^p(\bar{\Omega}, \mathbb{R})$ over which all the pth partial derivatives are uniformly continuous.

When Ω is not bounded, Ascoli's theorem implies that uniformly continuous and bounded (for the p first derivatives) subsets of $C_b^p(\Omega, \mathbb{R}^d)$ are relatively compact with respect to uniform convergence over compact sets.

Compact subsets of $C_0^p(\Omega, \mathbb{R})$ can be obtained via Ascoli's theorem as follows. Let Ω^* be the one point compactification Ω (obtained by adding a new point to it, called ∞, and new open subsets taking the form $\{\infty\} \cup K^c$, where K varies over all compact subsets of Ω). If $f : \Omega \to \mathbb{R}$ is any function, let f^* be its extension to Ω^* such that $f^*(\infty) = 0$. Then $C_0^p(\Omega, \mathbb{R})$ can also be seen as the set of functions $f \in C^p(\Omega, \mathbb{R})$ such that $(\partial_J f)^* \in C^0(\Omega^*, \mathbb{R})$ for all partial derivatives of order p or less. Then relatively compact subsets of $C_0^0(\Omega, \mathbb{R})$ are sets that are bounded and uniformly continuous on Ω and such that for any $\varepsilon > 0$ there exists a compact subset $K \subset \Omega$ such that

$$\sup_{y \in \Omega \setminus K} \sup_{h \in M} |h(y)| < \varepsilon .$$

The relatively compact subsets of $C_0^p(\Omega, \mathbb{R})$ are such that this condition is true for all partial derivatives of order p or less.

The spaces $C^p(\Omega, \mathbb{R}^d)$ are spaces of functions $f : \Omega \to \mathbb{R}^d$ with all coordinates in $C^p(\Omega, \mathbb{R})$. The extension of the previous discussion to these spaces is straightforward by considering coordinates.

A.8.2 Hilbert Sobolev Spaces

We now define the space $H^1(\Omega, \mathbb{R})$ of functions with square integrable generalized derivatives. A function u belongs to this set if and only if $u \in L^2(\Omega, \mathbb{R})$, and for each $i = 1, \ldots, k$, there exists a function $u_i \in L^2(\Omega, \mathbb{R})$ such that for any C^∞ function φ with compact support in Ω, one has

$$\int_\Omega u(x) \partial_i \varphi(x) dx = - \int_\Omega u_i(x) \varphi(x) dx.$$

The function u_i is called the (generalized) directional derivative of u with respect to the ith variable, and is denoted $\partial_i u$. The integration by parts formula shows that this derivative coincides with the standard partial derivative when the latter exists.

$H^1(\Omega, \mathbb{R}^d)$ has a Hilbert space structure with the inner product:

$$\langle u\ ,\ v \rangle_{H^1} = \langle u\ ,\ v \rangle_2 + \sum_{i=1}^{k} \langle \partial_i u\ ,\ \partial_i v \rangle_2.$$

The space $H^m(\Omega, \mathbb{R})$ can be defined by induction as the set of functions $f \in H^1(\Omega, \mathbb{R})$ with all partial derivatives belonging to $H^{m-1}(\Omega, \mathbb{R})$. Partial derivatives of increasing order are defined by induction, and one can define an inner product

on H^m as the sum of the L^2 inner products of all partial derivatives up to order m, for which H^m is Hilbert.

These spaces can be embedded in classical spaces of functions. We will in particular use a special case of Morrey's theorem, stated below. This theorem requires additional assumptions on the regularity of $\partial\Omega$, that we will not detail here, referring the reader to [3] for a complete statement.

Theorem A.16 *Assume that $m - d/2 > 0$ and that Ω has a regular boundary. Then, for any $j \geq 0$, $H^{j+m}(\Omega, \mathbb{R})$ is embedded in $C_b^j(\Omega, \mathbb{R})$.*

If Ω is bounded, then $H^{j+m}(\Omega, \mathbb{R})$ is compactly embedded in $C^j(\bar{\Omega}, \mathbb{R})$ and, if $\theta \in]0, m - d/2]$, and $u \in H^{j+m}(\Omega, \mathbb{R})$, then every partial derivative of $u \in H^{j+m}(\Omega, \mathbb{R})$ of order j has Hölder regularity θ: for all $x, y \in \Omega$

$$|\partial_J u(x) - \partial_J u(y)| \leq C \|u\|_{H^{m+j}} |x - y|^\theta$$

if $|J| = j$.

As a final definition, we let $H_0^m(\Omega, \mathbb{R})$ be the completion in $H^m(\Omega, \mathbb{R})$ of the set of C^∞ functions with compact support in Ω: u belongs to $H_0^m(\Omega, \mathbb{R})$ if and only if $u \in H^m(\Omega, \mathbb{R})$ and there exists a sequence of functions u_n, C^∞ with compact support in Ω, such that $\|u - u_n\|_{H^m}$ tends to 0. A direct application of Theorem A.16 shows that, if $m - d/2 > 0$, then, for any $j \geq 0$, $H_0^{j+m}(\Omega, \mathbb{R})$ is embedded in $C_0^j(\Omega, \mathbb{R})$.

General Sobolev spaces ($W^{p,k}$) are defined similarly to the spaces H^p, using L^k norms of derivatives instead of L^2 norms. The interested reader can refer to classical textbooks (e.g., [3, 45, 306]) on functional analysis for more detail.

A.8.3 The Duality Paradox

The Riesz representation theorem allows one to identify a Hilbert space H and its dual H^*. However, when H is densely embedded in another Hilbert space H_0, every continuous linear form on H_0 is also continuous on H, and H_0^* is densely embedded in H^*. We therefore have the sequence of embeddings

$$H \to H_0 \simeq H_0^* \to H^*$$

but this sequence loops since $H^* \simeq H$. This indicates that H_0 is also embedded in H. This is arguably a strange result. For example, let $H = H^1(\Omega, \mathbb{R})$, and $H_0 = L^2(\Omega, \mathbb{R})$: the embedding of H in H_0 is canonical, and clear from their definition (it is dense because C^∞ functions form a dense subset for both of them); but the converse inclusion does not seem natural, since there are more constraints in belonging to H than to H_0. To understand this reversed embedding, we must think in terms of linear forms.

If $u \in L^2(\Omega, \mathbb{R})$, we may consider the linear form φ_u defined by

$$(\varphi_u \mid v) = \langle u\,,\ v\rangle_2 = \int_\Omega u(x)v(x)dx.$$

When $v \in H^1(\Omega, \mathbb{R})$, we have

$$(\varphi_u \mid v) \leq \|u\|_2 \|v\|_2 \leq \|u\|_2 \|v\|_{H^1},$$

so that φ_u is continuous when seen as a linear form on $H^1(\Omega, \mathbb{R})$. The Riesz representation theorem implies that there exists a $\tilde{u} \in H^1(\Omega, \mathbb{R})$ such that, for all $v \in H^1(\Omega, \mathbb{R})$, $\langle u\,,\ v\rangle_2 = \langle \tilde{u}\,,\ v\rangle_{H^1}$: the relation $u \mapsto \tilde{u}$ provides the embedding of $L^2(\Omega, \mathbb{R})$ into $H^1(\Omega, \mathbb{R})$. Let us be more specific and take $\Omega =]0, 1[$. The relation states that, for any $v \in H^1(\Omega, \mathbb{R})$,

$$\int_0^1 \partial\tilde{u}(t)\,\partial v(t)dt + \int_0^1 \tilde{u}(t)v(t)dt = \int_0^1 \tilde{u}(t)v(t)dt.$$

In order to integrate the first integral by parts, let us make the simplifying assumption that $\tilde{u}$ has two derivatives, and obtain

$$\partial u(1)v(1) - \partial u(0)v(0) - \int_0^1 \partial^2\tilde{u}(t)v(t) + \int_0^1 \tilde{u}(t)v(t)dt = \int_0^1 \tilde{u}(t)v(t)dt. \tag{A.6}$$

Such an identity can be true for every v in $H^1(\Omega, \mathbb{R})$ if and only if $\partial u(0) = \partial u(1) = 0$ and $-\partial^2\tilde{u} + \tilde{u} = u$: $\tilde{u}$ is thus a solution of a second-order differential equation, with first-order boundary conditions, and the embedding of $L^2((0, 1), \mathbb{R})$ into $H^1((0, 1), \mathbb{R})$ just shows that a unique solution exists, at least in the generalized sense of Eq. (A.6).

As seen in these examples, even if, from an abstract point of view, we have an identification between two Hilbert spaces, the associated embeddings are of a very different nature, the first one corresponding to a set inclusion (it is canonical), the second to the solution of a differential equation in one dimension, and to a partial differential equation in the general case.

A.9 Weak Convergence

Let us start with the definition:

Definition A.17 When V is a Banach space, a sequence (v_n) in V is said to weakly converge to some $v \in V$ if and only if, for all bounded linear forms $\alpha \in V^*$, one has $(\alpha \mid v_n) \to (\alpha \mid v)$ as n tends to infinity.

Recall that $h_n \to h$ "strongly" if $\|h_n - h\|_V \to 0$. The following proposition describes how convergence is related to inclusion.

Proposition A.18 *Assume that V and W are Banach spaces and that $W \subset V$ with a continuous inclusion. If a sequence (w_n) in W converges weakly (in W) to some $w \in W$, then w_n also converges weakly to w in V.*

This just says that if $(\alpha \mid w_n) \to (\alpha \mid w)$ for all bounded linear functionals on W, then the convergence holds for all bounded linear functionals on V, which is in fact obvious because the restriction to W of any bounded linear functional on V is *a fortiori* bounded on W.

In the case of a Hilbert space, the Riesz representation theorem immediately provides the following proposition:

Proposition A.19 *Let V be a Hilbert space. A sequence v_n in V converges weakly to an element $v \in V$ if and only if, for all $w \in V$,*

$$\lim_{n\to\infty} \langle w\,,\, v_n\rangle_V = \langle w\,,\, v\rangle_V.$$

Moreover, if v_n converges weakly to v, then $\|v\|_V \leq \liminf \|v_n\|_V$.

The last statement comes from the inequality: $\langle v_n\,,\, v\rangle_V \leq \|v_n\|_V \, \|v\|_V$, which provides at the limit

$$\|v\|_V^2 \leq \|v\|_V \liminf_{n\to\infty} \|v_n\|_V \,.$$

Finally, the following result is about the weak compactness of bounded sets [306, 315].

Theorem A.20 *If V is a Hilbert space and (v_n) is a bounded sequence in V (there exists a constant C such that $\|v_n\| \leq C$ for all n), then one can extract a subsequence from v_n which weakly converges to some $v \in V$.*

A.10 The Fourier Transform

We provide here a brief overview of Fourier transforms with their definition and basic properties. We let $H = L^2(\mathbb{R}^d, \mathbb{C})$. For a function $f \in H$, one defines

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\iota\pi\xi^T x} dx.$$

This is well defined as soon as f is absolutely integrable, i.e.,

$$\int_{\mathbb{R}^d} |f(x)|\, dx < \infty,$$

but can be extended (based on isometry properties, see below) to the whole space H.

We have the following inversion formula:

Theorem A.21 *Assume that f and $\hat{f}$ are both absolutely integrable. Then, we have*

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\iota\pi x^T \xi} d\xi.$$

The inversion formula illustrates the bijective nature of the Fourier transform. Another important property is that, if both f and $\hat{f}$ are in $L^1 \cap L^2$, then $\|\hat{f}\|_2 = \|f\|_2$. This implies that there is a unique isometric extension of the Fourier transform to L^2. In particular, we have, for this extension:

$$\langle\!\langle \hat{f}, \hat{g} \rangle\!\rangle_2 = \langle\!\langle f, g \rangle\!\rangle_2, \tag{A.7}$$

with the notation

$$\langle\!\langle f, g \rangle\!\rangle_2 = \int_{\mathbb{R}^d} f \bar{g}\, dx$$

for complex-valued functions ($\bar{g}$ being the complex conjugate of g).

The Fourier transform has some other useful properties. In the following proposition, x and ξ are dummy variables, $x_0, \xi_0 \in \mathbb{R}^d$ and $\alpha_1, \alpha_2, a \in \mathbb{R}$.

Proposition A.22

(1) $\mathcal{F}(f(x - x_0))(\xi) = e^{-2\iota\pi\xi^T x_0} \hat{f}(\xi)$;

(2) $\mathcal{F}\left(e^{2\iota\pi x^T \xi_0} f(x)\right)(\xi) = \hat{f}(\xi - \xi_0)$;

(3) $\mathcal{F}(f(ax))(\xi) = \dfrac{1}{|a|^d} \hat{f}\left(\dfrac{\xi}{a}\right)$;

(4) $\mathcal{F}(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 \hat{f}_1 + \alpha_2 \hat{f}_2$;

(5) $\mathcal{F}\left(e^{-\pi x^T x}\right)(\xi) = e^{-\pi \xi^T \xi}$;

(6) $\mathcal{F}(\partial_i f)(\xi) = (2\iota\pi\xi_i)\hat{f}$.

The transform also interacts nicely with convolutions.

Proposition A.23 *If h and f are absolutely integrable, then*

$$\mathcal{F}(h * f) = \hat{h}\hat{f}$$

with

$$(h * f)(x) = \int_{\mathbb{R}^d} h(x - y) f(y) dy.$$

Appendix B
Elements from Differential Geometry

B.1 Introduction

This chapter describes a few notions from differential geometry that are relevant to this book. One of the primary goals of differential geometry is to provide mathematical tools that make possible the use of infinitesimal calculus for sets that are more general than Euclidean spaces. In this setting, we will not go further than the notion of tangent spaces and Riemannian metrics, leaving aside some fundamental features (e.g., curvature), for which we refer the reader to general textbooks on the subject, such as [1, 43, 87, 142, 162, 198].

B.2 Differential Manifolds

B.2.1 Definition

A differential manifold is a topological space within which points can be described by coordinate systems, which must satisfy some compatibility conditions allowing for the definition of intrinsic differential operations. We start with the definition:

Definition B.1 Let M be a topological Hausdorff space. A d-dimensional local chart on M is a pair (U, Φ) where U is an open subset of $\mathbb{R}^d$ and Φ a homeomorphism between U and some open subset of M.

Two d-dimensional local charts, (U_1, Φ_1) and (U_2, Φ_2), are C^∞-compatible if either $\Phi_1(U_1)$ and $\Phi_2(U_2)$ do not overlap, or the function $\Phi_1^{-1} \circ \Phi_2$ is a C^∞-diffeomorphism between $U_2 \cap \Phi_2^{-1}(\Phi_1(U_1))$ and $\Phi_1^{-1}(\Phi_2(U_2)) \cap U_1$.

A d-dimensional atlas on M is a family of pairwise compatible local charts $((U_i, \Phi_i), i \in I)$ such that $M = \bigcup_I \Phi_i(U_i)$. Two atlases on M are equivalent if their union is also an atlas, i.e., if every local chart of the first one is compatible with every local chart of the second one.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

A Hausdorff space with a d-dimensional atlas is called a d-dimensional (C^∞) differential manifold.

One defines C^k differential manifolds simply by replacing C^∞ by C^k in the definition above. We take $k = \infty$ here only for convenience, all definitions and statements below generalizing easily to the C^k case, for large enough k.

If M is a manifold, a local chart on M will always be assumed to be compatible with the atlas on M. If M and N are two manifolds, their product $M \times N$ is also a manifold; if (U, Φ) is a chart on M and (V, Ψ) a chart on N, then $(U \times V, (\Phi, \Psi))$ is a chart on $M \times N$, and one shows easily that one can form an atlas for $M \times N$ using such cross products between two atlases of M and N.

When a local chart (U, Φ) is given, the coordinate functions $x_1, \ldots, x_d$ are defined by $\Phi^{-1}(p) = (x_1(p), \ldots, x_d(p))$ for $p \in U$. Formally, x_i is a function from $\Phi(U)$ to $\mathbb{R}$. However, when a point p is given, one generally refers to $x_i = x_i(p) \in \mathbb{R}$ as the ith coordinate of p in the chart (U, Φ).

According to these definitions, $\mathbb{R}^d$ is a differential manifold, and so are open sets in $\mathbb{R}^d$. Another example is given by the d-dimensional sphere, S^d, defined as the set of points $p \in \mathbb{R}^{d+1}$ with $|p| = 1$. The sphere can be equipped with an atlas with $2(d + 1)$ charts, (U_i, Φ_i), $i = 1, \ldots, 2(d + 1)$, letting

$$U_1 = \{p = (p_2, \ldots, p_{d+1}) \in \mathbb{R}^d : p_2^2 + \cdots + p_{d+1}^2 < 1\}$$

and

$$\Phi_1(p_2, \ldots, p_{d+1}) = \left(\sqrt{1 - p_2^2 - \cdots - p_{d+1}^2}, p_2, \ldots, p_{d+1}\right),$$

$$\Phi_2(p_2, \ldots, p_{d+1}) = \left(-\sqrt{1 - p_2^2 - \cdots - p_{d+1}^2}, p_2, \ldots, p_{d+1}\right),$$

and so on with each of the coordinates $p_2, \ldots, p_{d+1}$ expressed as functions of the others.

We now consider functions on manifolds.

Definition B.2 Let $k \leq \infty$. A function $\psi : M \to \mathbb{R}$ is C^k if, for every local chart (U, Φ) on M, the function $\psi \circ \Phi : U \subset \mathbb{R}^d \to \mathbb{R}$ is C^k in the usual sense. The function $\psi \circ \Phi$ is called the interpretation of ψ in (U, Φ).

From the compatibility condition, if this property is true for an atlas, it is true for all charts compatible with it. The set of C^k functions on M is denoted $C^k(M, \mathbb{R})$, or just $C^k(M)$. If U is open in M, the set $C^k(U)$ contains functions defined on U which can be interpreted as C^k functions of the coordinates for all local charts of M that are contained in U. The first examples of C^∞ functions are the coordinates: if (U, Φ) is a chart, the ith coordinate $(x_i(p),\ p \in U)$ belongs to $C^\infty(U)$, since, when interpreted in (U, Φ), it reduces to $(x_1, \ldots, x_d) \mapsto x_i$.

If M is only a C^l manifold with $l < \infty$, one can only speak of C^k functions for $k \leq l$, if one wants this property to be independent of the local chart representation.

B.2.2 Vector Fields, Tangent Spaces

In this section we fix a differential manifold, denoted M, of dimension d. We define vector fields and tangent vectors via their actions on functions, and later provide alternative interpretations.

Definition B.3 A vector field on M is a function $X : C^\infty(M) \to C^\infty(M)$ such that: $\forall \alpha, \beta \in \mathbb{R}, \forall \varphi, \psi \in C^\infty(M)$:

$$X(\alpha.\varphi + \beta.\psi) = \alpha.X(\varphi) + \beta.X(\psi)\,,$$

$$X(\varphi\psi) = X(\varphi)\psi + \varphi X(\psi)\,.$$

The set of vector fields on M is denoted by $\mathcal{X}(M)$.

Definition B.4 If $p \in M$, a tangent vector to M at p is a function $\xi : C^\infty(M) \to \mathbb{R}$ such that: $\forall \alpha, \beta \in \mathbb{R}, \forall \varphi, \psi \in C^\infty(M)$

$$\xi(\alpha.\varphi + \beta.\psi) = \alpha.\xi(\varphi) + \beta.\xi(\psi)\,,$$

$$\xi(\varphi\psi) = \xi(\varphi)\psi(p) + \varphi(p)\xi(\psi)\,.$$

The set of tangent vectors to M at p is denoted T_pM.

So vector fields assign C^∞ functions to C^∞ functions and tangent vectors assign real numbers to C^∞ functions.

One can go from vector fields to tangent vectors and vice versa as follows. If $X \in \mathcal{X}(M)$ is a vector field on M, and if $p \in M$, define $X_p : C^\infty(M) \to \mathbb{R}$ by

$$X_p(\varphi) = (X(\varphi))(p)$$

to obtain a tangent vector at p. Conversely, if a collection $(X_p \in T_pM,\ p \in M)$ is given, one can define $(X(\varphi))(p) = X_p(\varphi)$ for $\varphi \in C^\infty(M)$ and $p \in M$; X will be a vector field on M if and only if, for all $\varphi \in C^\infty(M)$ the function $p \mapsto X_p(\varphi)$ is C^∞. Finally, one can show [142] that, for all $\xi \in T_pM$, there exists a vector field X such that $\xi = X_p$.

The linear nature of Definitions B.3 and B.4 is clarified in the next proposition:

Proposition B.5 *For all $p \in M$, the tangent space T_pM is a d-dimensional vector space.*

Proof Let $C = (U, \Phi)$ be a local chart with $p \in \Phi(U)$. Define $x^0 = \Phi^{-1}(p)$, $x^0 \in U$. If $\varphi \in C^\infty(M)$, then, by definition,

$$\begin{aligned} \varphi_C :\quad & U \subset \mathbb{R}^d \to \mathbb{R} \\ & x \mapsto \varphi \circ \Phi(x) \end{aligned}$$

is C^∞. Define, for $i = 1, \ldots, d$,

$$\partial_{i,p}(\varphi) := \partial_i \varphi_C(x_0).$$

It is easy to check that $\partial_{i,p}$ satisfies the conditions given in Definition B.4, so that $\partial_{i,p} \in T_p M$. We show that every $\xi \in T_p M$ may be uniquely written in the form

$$\xi = \sum_{i=1}^{d} \lambda_i \partial_{i,p}. \tag{B.1}$$

Indeed,

$$\varphi_C(x) = \varphi_C(x^0) + \sum_{i=1}^{d} (x_i - x_i^0)\psi_i^*(x)$$

with $\psi_i^*(x) = \displaystyle\int_0^1 \partial_i \varphi_C(x^0 + t(x - x^0))dt$. Thus, if $p' \in U$,

$$\varphi(p') = \varphi_0 + \sum_{i=1}^{d} (x_i(p') - x_i(p))\psi_i(p')$$

with $\psi_i(p') = \psi_i^*(\Phi^{-1}(p'))$, and $\varphi_0 = \varphi(p)$. If $\xi \in T_p M$ and f is constant, we have $\xi(f) = f\xi(1) = f\xi(1^2) = 2f\xi(1)$ so that $\xi(f) = 0$. Thus, for all $\xi \in T_p M$,

$$\xi(\varphi) = 0 + \sum_{i=1}^{d} \psi_i(p)\xi(x_i)\,.$$

But $\psi_i(p) = (\partial_{i,p})(\varphi)$, which yields (B.1) with $\lambda_i = \xi(x_i)$. □

This result also implies that, *within a local chart*, a vector field can always be interpreted in the form

$$X = \sum_{i=1}^{d} \varphi_i \partial_i,$$

with $\varphi_i \in C^\infty(M)$ and $[\partial_i]_p = \partial_{i,p}$.

There is another standard definition of tangent vectors on M, in relation to differentiable curves on M. This starts with the following definitions.

Definition B.6 Let $t \mapsto \mu(t) \in M$ be a continuous curve, $\mu : [0, T] \to M$. One says that this curve is C^∞ if, for any local chart $C = (U, \Phi)$, the curve $\mu^C : s \mapsto \Phi^{-1} \circ \mu(s)$, defined on $\{t \in [0, T] : \mu(t) \in U\}$, is C^∞.

Let $p \in M$. One says that two C^∞ curves, μ and ν, starting at p (i.e., $\mu(0) = \nu(0) = p$) have the same tangent at p if, and only if, for all charts $C = (U, \Phi)$, the curves μ^C and ν^C have identical derivatives at $t = 0$.

Proposition B.7 *The tangential identity at p is an equivalence relation. The tangent space to M at p can be identified with the set of equivalence classes for this relation.*

Proof We sketch the argument. If a curve μ is given, with $\mu(0) = p$, define, for $\varphi \in C^\infty(M)$,

$$\xi^\mu(\varphi) = \partial(\varphi \circ \mu)(0).$$

One can check that $\xi^\mu \in T_pM$, and that $\xi^\mu = \xi^\nu$ if μ and ν have the same tangent at p.

Conversely, if $\xi \in T_pM$ there exists a curve μ such that $\xi = \xi^\mu$, and the equivalence class of μ is uniquely specified by ξ. To show this, consider a chart (U, Φ). We must have

$$\xi^\mu(x_i) = \xi(x_i) = \partial(x_i \circ \mu)(0),$$

which indeed shows that the tangent at p is uniquely defined. To determine μ, start from a line segment in U, passing through $\Phi^{-1}(p)$, with direction given by $(\xi(x_i), i = 1, \ldots, d)$ and apply Φ to it to obtain a curve on M. □

When $X \in \mathcal{X}(M)$ is given, one can consider the differential equation

$$\partial_t \mu(t) = X_{\mu(t)}.$$

Such a differential equation always admits a unique solution given an initial condition $\mu(0) = p$, at least for t small enough. This can be proved by translating the problem in a local chart and applying the results of Appendix C.

Finally, we can define the differential of a scalar-valued function, followed by that of a function between manifolds.

Definition B.8 If $\varphi \in C^\infty(M)$, one defines a linear form on T_pM by $\xi \mapsto \xi(\varphi)$. It will be denoted $d\varphi(p)$, and called the differential of φ at p.

Definition B.9 Let M and M' be two differential manifolds. A mapping $\Phi : M \to M'$ has class C^∞ if and only if, for all $\varphi \in C^\infty(M')$, one has $\varphi \circ \Phi \in C^\infty(M)$.

Definition B.10 If $p \in M$ and $p' = \Phi(p)$, define the tangent map of Φ at p,

$$d\Phi(p) : T_pM \to T_{p'}M',$$

by: for all $\xi \in T_pM$, $\varphi \in C^\infty(M')$,

$$(d\Phi(p)\xi)(\varphi) = \xi(\varphi \circ \Phi).$$

The tangent map $d\Phi(p)$ is also called the differential of Φ at p.

The chain rule is true for tangent mappings: if $\Phi : M \to M'$ and $\Psi : M' \to M''$ are differentiable, then $\Psi \circ \Phi : M \to M''$ is also differentiable and

$$d(\Psi \circ \Phi)(p) = d\Psi(\Phi(p)) \circ d\Phi(p).$$

This almost directly follows from the definition.

The set $TM = \{(p, \xi) : p \in M, \xi \in T_pM\}$ is also a manifold called the tangent bundle of M. Local charts on TM can be derived from local charts (U, Φ) on M as $(U \times \mathbb{R}^d, \tilde{\Phi})$ with

$$\tilde{\Phi}(x, \xi) = (\Phi(x), d\Phi(x)\xi),$$

where $d\Phi(x)$ simply maps $(\xi_1, \ldots, \xi_d)$ to $\sum_{i=1}^d \xi_i \partial_i$.

One similarly defines the frame bundle over M as the set

$$FM = \{(p, \xi_1, \ldots, \xi_d) : p \in M, (\xi_1, \ldots, \xi_d) \text{ basis of } T_pM\},$$

which is also a manifold, associating with (U, Φ) the chart $(U \times \mathrm{GL}_d(\mathbb{R}), \tilde{\Phi})$ with

$$\tilde{\Phi}(x, A) = (\Phi(x), d\Phi(x)A)$$

(where $\mathrm{GL}_d(\mathbb{R})$ is the set of invertible d by d matrices with real coefficients, an open subset of $\mathbb{R}^{d^2}$).

B.3 Submanifolds

One efficient way to build manifolds is to characterize them as submanifolds of simple manifolds like $\mathbb{R}^d$. If M is a manifold, a submanifold of M is a subset of M that is itself a manifold, but also inherits the manifold structure of M.

Definition B.11 Let M be a d-dimensional differential manifold. We say that P is a d'-dimensional submanifold of M (with $d' \leq d$) if, for all $m_0 \in P$, there exists a local chart (U, Φ) of M such that $m_0 \in \Phi(U)$, with local coordinates $(x_1, \ldots, x_d)$, such that

$$U \cap \Phi^{-1}(P) = \{m \in M : x_i = 0, i = 1, \ldots, d - d'\}.$$

The next theorem is one of the main tools for defining manifolds:

Theorem B.12 *Let M be a d-dimensional differential manifold, Φ a differentiable map from M to $\mathbb{R}^k$. Let $a \in \Phi(M)$ and*

$$P = \Phi^{-1}(a) = \{p \in M : \Phi(p) = a\}.$$

If for all $p \in P$ the linear map $d\Phi(p) : T_pM \to \mathbb{R}^k$ has full rank k, then P is a submanifold of M, with dimension $d' = d - k$.

Here is a quick justification. Working in a local chart if needed, one can assume without loss of generality that M is an open subset of $\mathbb{R}^d$. Moreover, if $d\Phi(p)$

has rank k, then one can find k of its columns that are linearly independent, and these column will remain independent in a neighborhood of p. Applying the implicit function theorem, the identity $\Phi(p') = a$ can then be solved in a neighborhood of p by expressing k coordinates (say $x_1 \ldots, x_k$) as functions $f_1, \ldots f_k$ of the $d - k$ remaining ones, so that P is, in a neighborhood of P, equivalent to the set provided by the k equations $x_j - f_j(x_{k+1}, \ldots, x_d) = 0$, $j = 1, \ldots, k$. This shows that P is a $(d - k)$-dimensional submanifold of M. (The local chart in Definition B.11 is $\Phi(x_1, \ldots, x_n) = (x_1 - f_1(x_{k+1}, \ldots, x_d), \ldots, x_k - f_k(x_{k+1}, \ldots, x_d), x_{k+1}, \ldots, x_d)$.)

This result can be applied, for example, to the sphere S^d, defined by $x_1^2 + \cdots + x_{d+1}^2 = 1$, which is a submanifold of $\mathbb{R}^{d+1}$ of dimension d.

If $P \subset M$ is a submanifold of M defined as in Theorem B.12, the tangent space to P at p can be identified with the null space of $d\Phi(p)$ in T_pM:

$$T_pP = \{\xi \in T_pM, d\Phi(p)\xi = 0\}.$$

Another way to define submanifolds is via embeddings, as outlined below.

Definition B.13 Let M and P be two differential manifolds. An embedding of M into P is a C^∞ map $\Phi : M \to P$ such that:

(i) For all $p \in M$, the tangent map, $d\Phi(p)$, is one-to-one, from T_pM to $T_{\Phi(p)}P$.
(ii) Φ is a homeomorphism between M and $\Phi(M)$ (this last set being considered with the topology induced by P).

The second condition means that Φ is one-to-one, and, for all open subsets U in M, there exists an open subset V in P such that $\Phi(U) = V \cap \Phi(M)$. Maps that satisfy (i) (but not necessarily (ii)) are called *immersions*. We then have:

Proposition B.14 *If $\Phi : M \to P$ is an embedding, then $\Phi(M)$ is a submanifold of P, with same dimension as M.*

B.4 Lie Groups

B.4.1 Definitions

A group is a set G with a composition rule $(g, h) \mapsto gh$ which is associative, has an identity element (denoted id_G, or id if there is no risk of confusion) and such that every element in G has an inverse in G. A Lie group is both a group and a differential manifold, such that the operations $(g, h) \mapsto gh$ and $g \mapsto g^{-1}$, respectively from $G \times G$ to G and from G to G, are C^∞.

B.4.2 The Lie Algebra of a Lie Group

If G is a Lie group, $g \in G$ and $\varphi \in C^\infty(G)$, one defines $\varphi \cdot g \in C^\infty(G)$ by $(\varphi \cdot g)(g') = \varphi(g'g)$. A vector field on G is right-invariant if, for all $g \in G$ and $\varphi \in C^\infty(G)$, one has $X(\varphi \cdot g) = X(\varphi) \cdot g$. Denoting by R_g the right translation on G (defined by $R_g(g') = g'g$), right invariance is equivalent to the identity $X_{g'g} = dR_g X_{g'}$ being true for all $g, g' \in G$. The set of right-invariant vector fields is called the Lie algebra of the group G, and denoted $\mathfrak{g}$.

Because $(X(\varphi \cdot g))(\mathrm{id}) = (X(\varphi) \cdot g)(\mathrm{id}) = (X(\varphi))(g)$ whenever $X \in \mathfrak{g}$, an element X of $\mathfrak{g}$ is entirely specified by the values of $X(\varphi)(\mathrm{id})$ for $\varphi \in C^\infty(G)$. This implies that the Lie algebra $\mathfrak{g}$ may be identified with the tangent space to G at id, $T_{\mathrm{id}}G$. If $\xi \in T_{\mathrm{id}}G$, its associated right-invariant vector field is

$$X^\xi : g \mapsto dR_g(\mathrm{id})\xi \in T_gG.$$

The operation that provides the structure of an algebra on $\mathfrak{g}$ is called the Lie bracket. Recall that a vector field on a manifold M is a function $X : C^\infty(M) \to C^\infty(M)$ which satisfies the conditions of Definition B.3. When X and Y are two vector fields, it is possible to combine them and compute $(XY)(\varphi) = X(Y(\varphi))$; XY also transforms C^∞ functions into C^∞ functions, but will not satisfy the conditions of Definition B.3, essentially because it involves second derivatives. However, it is easy to check that the second derivatives cancel in the difference $XY - YX$, which is a vector field on M, denoted $[X, Y]$, and called the bracket of X and Y. A few important properties of Lie brackets are listed (without proof) in the next proposition.

Proposition B.15

i. $[X, Y] = -[Y, X]$.
ii. $[[X, Y], Z] = [X, [Y, Z]]$.
iii. $[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0$.
iv. *If* $\Phi \in C^\infty(M, N)$, $d\Phi[X, Y] = [d\Phi X, d\Phi Y]$.

Here, $d\Phi X$ is the vector field $\varphi \mapsto X(\varphi \circ \Phi)$ such that $(d\Phi X)_p = d\Phi(p)X_p$. This last property is important for Lie groups, because when it is applied with $\Phi = R_g : G \mapsto G$, and $X, Y \in \mathfrak{g}$, it yields

$$dR_g[X, Y] = [dR_g X, dR_g Y] = [X, Y],$$

so that $[X, Y] \in \mathfrak{g}$. The Lie algebra of G is therefore closed under the Lie bracket operation (which is the reason for the term 'Lie algebra'). Because of the identification of $\mathfrak{g}$ with $T_{\mathrm{id}}G$, the bracket notation is also used for tangent vectors at the identity, letting $[\xi, \eta] = [X^\xi, X^\eta]_{\mathrm{id}}$.

There is alternative, equivalent, definition of the Lie bracket on $\mathfrak{g}$. For $g \in G$, one can define the group isomorphism $I_g : h \mapsto ghg^{-1}$. It is differentiable, and the

differential of I_g at $h = \mathrm{id}$ is denoted $\mathrm{Ad}_g : T_{\mathrm{id}}G \to T_{\mathrm{id}}G$. We therefore have, for $\eta \in T_{\mathrm{id}}G$,

$$\mathrm{Ad}_g(\eta) = dI_g(\mathrm{id})\eta.$$

The transformation $\mathrm{Ad} : g \mapsto \mathrm{Ad}_g$ maps G into the group of linear transformations of $T_{\mathrm{id}}G$ and is called the adjoint representation of G.

Now consider the map $U_\eta : g \mapsto \mathrm{Ad}_g(\eta)$, which is defined on G and takes values in $T_{\mathrm{id}}G$. One can show [142] that

$$dU_\eta(\mathrm{id})\xi = [\xi, \eta]. \tag{B.2}$$

The notation $\mathrm{ad}_\xi \eta = [\xi, \eta]$ is commonly used to represent the Lie bracket. The transformation $\xi \mapsto \mathrm{ad}_\xi$ now maps $T_{\mathrm{id}}G$ (or, after identification, the Lie algebra $\mathfrak{g}$) to the set of linear transformations of $T_{\mathrm{id}}G$ and is called the adjoint representation of $\mathfrak{g}$.

When a vector field $X \in \mathfrak{g}$ is given, the solution of the associated differential equation

$$\partial_t \mu(t) = X_{\mu(t)} \tag{B.3}$$

with initial condition $\mu(0) = \mathrm{id}$ always exists, not only for small time, but for arbitrary times. The small time existence comes from the general theory of ordinary differential equations, and the existence for arbitrary time comes from the fact that, wherever it is defined, $\mu(t)$ satisfies the semi-group property $\mu(t + s) = \mu(t)\mu(s)$: this implies that if $\mu(t)$ is defined on some interval $[0, T]$, one can always extend it to $[0, 2T]$ by letting $\mu(t + T) = \mu(t)\mu(T)$ if $t > 0$. The semi-group property can be proved to be true as follows: if $X = X^\xi$, for $\xi \in T_{\mathrm{id}}G$, the ordinary differential equation can be written

$$\partial_t \mu(t) = dR_{\mu(t)}(\mathrm{id})\xi.$$

Consider now $\nu : s \mapsto \mu(t + s)$. It is a solution of the same equation with initial condition $\nu(0) = \mu(t)$. If $\tilde{\nu}(s) = \mu(s)\mu(t)$, we have

$$\begin{aligned}\partial_s \tilde{\nu}(s) &= dR_{\mu(t)}(\mu(s))(\partial_s \mu(s)) = dR_{\mu(t)}(\mu(s))\, dR_{\mu(t)}(\mathrm{id})\xi \\ &= d(R_{\mu(t)} \circ R_{\mu(s)})(\mathrm{id})\xi \\ &= dR_{\nu(s)}(\mathrm{id})\xi.\end{aligned}$$

Thus, ν and $\tilde{\nu}$ satisfy the same differential equation, with the same value, $\mu(t)$, at $s = 0$. They therefore coincide, which is the semi-group property.

The solution of (B.3) with initial condition id is called the exponential map on G, and is denoted $\exp(tX)$ or $\exp(t\xi)$ if $X = X^\xi$. The semi-group property becomes $\exp((t + s)X) = \exp(tX)\exp(sX)$. Using the exponential map, Eq. (B.2) can be written as

$$\partial_t \left[\partial_s \exp(t\xi)\exp(s\eta)\exp(-t\xi)_{|s=0}\right]_{|t=0} = [\xi, \eta]\,.$$

We finally quote one last important property of the exponential map [87, 142]:

Theorem B.16 *There exists a neighborhood V of 0 in $\mathfrak{g}$ and a neighborhood U of* id *in G such that* exp *is a diffeomorphism between V and U.*

B.4.3 Finite-Dimensional Transformation Groups

Finite-dimensional transformation groups, and in particular matrix groups, are fundamental examples of Lie groups. They also provide important transformations in the analysis of shapes. Denote by $\mathcal{M}_n(\mathbb{R})$ the n^2-dimensional space of real n by n matrices. For $i, j \in \{1, \ldots, n\}$, denote by ∂_{ij} the matrix with (i, j) coefficient equal to 1, and all others to 0. Let $\mathrm{Id}_{\mathbb{R}^n}$ denote the identity matrix of size n.

Linear Groups

$\mathrm{GL}_n(\mathbb{R})$ is the group of invertible matrices in $\mathcal{M}_n(\mathbb{R})$. It is open in $\mathcal{M}_n(\mathbb{R})$ and therefore is a submanifold of this space, of same dimension, n^2. The Lie algebra of $\mathrm{GL}_n(\mathbb{R})$ is equal to $\mathcal{M}_n(\mathbb{R})$, and is generated by all $(\partial_{ij}, i, j = 1, \ldots, n)$.

If $\xi \in \mathcal{M}_n(\mathbb{R})$, the associated right-invariant vector field is $X^\xi : g \mapsto \xi g$. The adjoint map is $\eta \mapsto g\eta g^{-1}$, and the Lie bracket is $[\xi, \eta] = \xi\eta - \eta\xi$. Finally, the exponential is the usual matrix exponential:

$$\exp(\xi) = \sum_{k=0}^{\infty} \frac{\xi^k}{k!}.$$

Special Linear Group

$\mathrm{SL}_n(\mathbb{R})$ is the subgroup of $\mathrm{GL}_n(\mathbb{R})$ containing all matrices with determinant 1. The determinant is a C^∞ function. Its derivative at $g \in \mathrm{GL}_n(\mathbb{R})$ is a linear map from $\mathcal{M}_n(\mathbb{R})$ to $\mathbb{R}$, given by $d(\det(g))\,\xi = \det(g)\,\mathrm{trace}(g^{-1}\xi)$. Since this differential has rank one, Theorem B.12 implies that $\mathrm{SL}_n(\mathbb{R})$ is a submanifold of $\mathrm{GL}_n(\mathbb{R})$, of dimension $n^2 - 1$. The Lie algebra of $\mathrm{SL}_n(\mathbb{R})$ is defined by $d(\det(\mathrm{Id}_{\mathbb{R}^n}))\,\xi = 0$, and therefore consists of matrices with vanishing trace.

Rotations

$\mathrm{O}_n(\mathbb{R})$ is the group of matrices g such that $g^T g = \mathrm{Id}_{\mathbb{R}^n}$. $\mathrm{SO}_n(\mathbb{R})$ is the subgroup of $\mathrm{O}_n(\mathbb{R})$ containing all matrices of determinant 1. The map $\Phi : g \mapsto g^T g$ is C^∞, and its differential is

$$d\Phi(g)\xi = g^T\xi + \xi^T g.$$

The null space of $d\Phi(g)$ therefore contains matrices $\xi = g\eta$ such that η is skew symmetric, and has dimension $n(n-1)/2$. Thus, again by Theorem B.12, $\mathrm{O}_n(\mathbb{R})$ and $\mathrm{SO}_n(\mathbb{R})$ are submanifolds of $\mathcal{M}_n(\mathbb{R})$, of dimension $n(n-1)/2$.

The Lie algebra of $\mathrm{O}_n(\mathbb{R})$ is the space of skew-symmetric matrices of size n.

Similitudes

$Sim_n(\mathbb{R})$ is the group of similitudes. It is composed of matrices g such that $g^T g = \lambda \mathrm{Id}$, for some $\lambda > 0$ in $\mathbb{R}$. In fact, one must have $\lambda = \det(g)^{2/n}$, so that $Sim_n(\mathbb{R})$ is the set of invertible matrices for which $\Phi(g) = 0$, with

$$\begin{aligned} \Phi : \quad & \mathrm{GL}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R}) \\ & g \mapsto g^T g - \det(g)^{2/n} \mathrm{Id}_{\mathbb{R}^n}. \end{aligned}$$

One can check that Φ has constant rank and that $Sim_n(\mathbb{R})$ is a submanifold of $\mathcal{M}_n(\mathbb{R})$ of dimension $1 + n(n-1)/2$.

The Lie algebra of $Sim_n(\mathbb{R})$ contains all matrices of the form $\alpha \mathrm{Id}_{\mathbb{R}^n} + \xi$, with ξ skew-symmetric.

Affine Groups

Groups of affine transformations are obtained by combining the previous linear transformations with translations. Let G be one of the linear groups we have discussed. Associating a translation to a vector in $\mathbb{R}^n$, we can represent an affine transformation with linear part in G by a pair (g, a) with $g \in G$ and $a \in \mathbb{R}^n$. The set of such pairs is denoted $G \ltimes \mathbb{R}^n$. This notation indicates that we have a semi-direct product: if (g, a) and (g', a') belong to $G \ltimes \mathbb{R}^n$, their product must be defined (in order to be consistent with the composition of maps) by

$$(g, a)(g', a') = (gg', ga' + a)$$

(and not $(gg', a + a')$, which would correspond to a direct product).

One can also find alternative notations for some affine groups, like $\mathrm{GA}_n(\mathbb{R})$ for $\mathrm{GL}_n(\mathbb{R}) \ltimes \mathbb{R}^n$, or $SA_n(\mathbb{R})$ for $\mathrm{SL}_n(\mathbb{R}) \ltimes \mathbb{R}^n$.

Affine groups of dimension n can also be represented as subgroups of $\mathrm{GL}_{n+1}(\mathbb{R})$: to $(g, a) \in G \times \mathbb{R}^n$, one can associate

$$\Phi(g, a) = \begin{pmatrix} g & a \\ 0 & 1 \end{pmatrix}.$$

It is easy to check that this is a bijection and a group isomorphism, in the sense that

$$\Phi(g, a)\Phi(g', a') = \Phi((g, a)(g', a')).$$

This allows one to identify the Lie algebra of $G \ltimes \mathbb{R}^n$ with the set of matrices of the form

$$\begin{pmatrix} A & a \\ 0 & 0 \end{pmatrix}$$

where A belongs to the Lie algebra of the subgroup of $\mathrm{GL}_n(\mathbb{R})$ on which the affine group is built, and $a \in \mathbb{R}^n$.

Projective Group

Definition B.17 The set of real lines passing through the origin in an $(n+1)$-dimensional vector space is the n-dimensional projective space, denoted by $\mathbb{P}^n(\mathbb{R})$.

An element of $\mathbb{P}^n(\mathbb{R})$ is therefore a collection of lines $m = \mathbb{R}\,x = \{\lambda x, \lambda \in \mathbb{R}\}$, x being a non-vanishing vector in $\mathbb{R}^{d+1}$. There is no loss of generality in assuming that x has norm 1, and since x and $-x$ provide the same point m, $\mathbb{P}^n(\mathbb{R})$ can be seen as the sphere S^n in which antipodal points are identified.

The set $\mathbb{P}^n(\mathbb{R})$ is a differential manifold of dimension n. It can be equipped with the quotient topology of the space $\mathbb{R}^{n+1}$ under the equivalence relation of being collinear with the origin (we skip the details). One can define the local chart (U_i, Φ_i) by

$$U_i = \{x = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \in \mathbb{R}^n\} \tag{B.4}$$

and

$$\begin{aligned} \Phi_i : \quad & U_i \to \mathbb{R}^n \\ & (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \mapsto \mathbb{R}(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_{n+1}). \end{aligned} \tag{B.5}$$

One can easily check that the family $((U_i, \Phi_i),\ i = 1, \ldots, d)$ forms an atlas of $\mathbb{P}^n(\mathbb{R})$.

Definition B.18 Because linear maps transform lines into lines, one can associate to each $g \in \mathrm{GL}_{n+1}(\mathbb{R})$ an induced transformation of $\mathbb{P}^n(\mathbb{R})$, still denoted g, defined by

$$g(\mathbb{R}\,x) = \mathbb{R}(gx).$$

The set of all such transformations $g : \mathbb{P}^n(\mathbb{R}) \to \mathbb{P}^n(\mathbb{R})$ is a group, called the n-dimensional projective group, and denoted $\mathrm{PGL}_n(\mathbb{R})$. This group is a Lie group (we skip the proof). It can be identified to the quotient of $\mathrm{GL}_{n+1}(\mathbb{R})$ with respect to the relation: $g \sim h$ if there exists a $\lambda \neq 0$ such that $g = \lambda h$. This group has dimension $(n+1)^2 - 1$.

The Lie algebra of $\mathrm{PGL}_n(\mathbb{R})$ is the Lie algebra of $\mathrm{GL}_{n+1}(\mathbb{R})$ in which two matrices are identified if they differ by a matrix of the form $\alpha \mathrm{Id}_{\mathbb{R}^{n+1}}$. It can also be identified with the set of $(n+1)$ by $(n+1)$ matrices with zero trace.

A local chart for this group in a neighborhood of the identity is formed by the set U of matrices $g = (g_{ij}) \in \mathrm{GL}_{n+1}(\mathbb{R})$ such that $g_{n+1,n+1} = 1$, with $\Phi(g) = \mathbb{R}g$.

B.5 Group Actions

B.5.1 Definitions

One says that a group G acts (on the left) on a set M if there exists a mapping Φ from $G \times M$ to M which associates to a pair (g, p) the result of the action of g on p with the properties that $\Phi(g, \Phi(h, p)) = \Phi(gh, p)$ and $\Phi(\mathrm{id}, p) = m$. The map Φ is a right action if the first property is replaced by $\Phi(g, \Phi(h, p)) = \Phi(hg, p)$. Left actions are usually denoted $(g, p) \mapsto g \cdot p$, and right actions $(g, p) \mapsto p \cdot g$, and the associativity property becomes $g \cdot (h \cdot p) = (gh) \cdot p$ in the first case and $(p \cdot h) \cdot g = p \cdot (hg)$ in the second case.

The orbit, or coset, of $p \in M$ under a left action is the set $G \cdot p = \{g \cdot p, g \in G\}$. Orbits either coincide or are disjoint, and they form a partition of M. We let $M/G = \{G \cdot p, p \in M\}$. A similar definition holds for right actions.

The action of G is transitive if there exists only one orbit, i.e., for all $p, p' \in M$, there exists a $g \in G$ such that $g \cdot p = p'$.

The isotropy subgroup of a point $p \in M$ is the collection of elements $g \in G$ such that $g \cdot p = p$. It is denoted $\mathrm{Iso}_p(G)$, and forms a subgroup of G, (i.e., it is closed under group products and under group inversion). The isotropy subgroup of M is the intersection of all $\mathrm{Iso}_p(G)$, and is denoted $\mathrm{Iso}_M(G)$.

When G is a Lie group and M is a manifold, one implicitly assumes that, in addition, the map $(g, p) \mapsto g \cdot p$ is C^∞.

B.5.2 Homogeneous Spaces

If H is a subgroup of G, the map $(h, g) \mapsto gh$ defines a right action of H on G. The coset space G/H is the set of orbits $\{[g]_H = gH, g \in G\}$ for this action. When G is a Lie group and H a closed subgroup of G, G/H is called a homogeneous space. The differential structure of G can be transferred to G/H to provide this set with the structure of a differential manifold. To achieve this, one expresses the Lie algebra of G as a direct sum $\mathfrak{g} = \mathfrak{h} + \mathfrak{q}$, where $\mathfrak{h}$ is the Lie algebra of H, and shows that the exponential map of G restricted to $\mathfrak{q}$ provides, after projection on G/H, a local chart of G/H in a neighborhood of $[\mathrm{id}]_H \in G/H$. We refer to [142] for a complete construction.

The group G acts on the left on G/H through $g \cdot (g'H) = (gg')H$. This action is transitive and H is the isotropy subgroup of $[\mathrm{id}]_H$. Conversely, the following is true.

Proposition B.19 *Let G be a group acting transitively on the left on a set M. Fix $p \in M$ and let $H = \mathrm{Iso}_p(G)$ be the isotropy subgroup of p. The map*

$$\begin{aligned} \Phi : \ & G/H \to M \\ & [g]_H \mapsto g \cdot p \end{aligned}$$

is a bijection which commutes with the actions of G on G/H and of G on M.

Proof First notice that this mapping is well-defined: if $[g]_H = [g']_H$, then $g^{-1}g' \in H$ so that $g^{-1} \cdot (g' \cdot p) = p$, which implies $g' \cdot p = g \cdot p$. It is onto because the action is transitive, and one-to-one because $g \cdot p = g' \cdot p$ if and only if $[g]_H = [g']_H$. The fact that Φ commutes with the actions means that $\Phi(g \cdot [g']_H) = g \cdot \Phi([g'_H])$, which is obvious. □

When G is a Lie group and M a differential manifold, one shows [142] that, in addition, Φ is differentiable; this provides an identification between M and a homogeneous space.

B.5.3 Infinitesimal Actions

Many Lie group concepts can be interpreted infinitesimally as essentially equivalent concepts on the group Lie algebra. This applies, in particular, to group actions.

We consider a left action, and focus on the mapping

$$\begin{aligned} \Phi_p : \ & G \to M \\ & g \mapsto g \cdot p, \end{aligned}$$

where G is a Lie group acting on a manifold M. Let ξ be an element of $\mathfrak{g}$, the Lie algebra of G (i.e., a tangent vector to G at the identity id). For all $p \in M$, we let

$$\xi \cdot p = d\Phi_p(\mathrm{id})\xi.$$

Define $\rho(\xi) : p \in M \mapsto \xi \cdot p$, which is a vector field on M, so that ρ itself is a mapping $\rho : \mathfrak{g} \to \mathcal{X}(M)$, called the *infinitesimal action of G on M*.

The map ρ being linear, its range, $\rho(\mathfrak{g})$, forms a linear subspace of $\mathcal{X}(M)$. Its dimension is finite and must be smaller than or equal to the dimension of G. Its elements are called infinitesimal generators.

For the dimension of $\rho(\mathfrak{g})$ to be strictly smaller than the dimension of G, there must exist a non-vanishing ξ such that $\rho(\xi) = 0$. It is easy to see that this implies that $\exp(t\xi) \cdot p = p$ for all $p \in M$, which means that the isotropy group of G is non-trivial. Conversely, one can show that any element ξ of the Lie algebra of $\mathrm{Iso}_M(G)$ is such that $\rho(\xi) = 0$.

B.6 Riemannian Manifolds

B.6.1 Introduction

In this section, M is a differential manifold of dimension d.

Definition B.20 A Riemannian structure on M is the definition of a C^∞ inner product between vector fields:

$$(X, Y) \in \mathcal{X}(M) \times \mathcal{X}(M) \mapsto \langle X\ ,\ Y\rangle \in C^\infty(M)$$

such that $\langle X\ ,\ Y\rangle = \langle Y\ ,\ X\rangle$, $\langle X\ ,\ X\rangle \geq 0$, $\langle X\ ,\ X\rangle = 0$ if and only if $X = 0$, and for all $\varphi, \psi \in C^\infty(M)$

$$\langle \varphi X + \psi X'\ ,\ Y\rangle = \varphi\langle X\ ,\ Y\rangle + \psi\langle X'\ ,\ Y\rangle.$$

The value of $\langle X\ ,\ Y\rangle$ at $p \in M$ will be denoted $\langle X\ ,\ Y\rangle_p$, and it can be shown that it depends only on X_p and Y_p. An equivalent construction is to assume that, for all $p \in M$, an inner product denoted $\langle \cdot\ ,\ \cdot\rangle_p$ is given on T_pM, which is such that, if X and Y are vector fields, the function $p \mapsto \langle X_p\ ,\ Y_p\rangle_p$ is C^∞. We shall use the notation

$$|\xi|_p = \sqrt{\langle \xi\ ,\ \xi\rangle_p}\,.$$

In a local chart, $C = (U, \Phi)$, with coordinates $(x_1, \ldots, x_d)$, a tangent vector at $p \in U$ can be written as a linear combination of the basis vectors $\partial_{i,p}$, $i = 1, \ldots, d$. From elementary linear algebra, there exists a positive definite symmetric matrix S_p, the coefficients of which being C^∞ functions of p, such that, if $\xi = \sum \lambda_i \partial_{i,p}$, $\eta = \sum \mu_i \partial_{i,p}$, then,

$$\langle \xi\ ,\ \eta\rangle_p = \lambda^T S_p \mu\,.$$

(Indices in summations always go from 1 to d, here and below.)

The Riemannian structure permits us, among other things, to measure lengths of displacements on the manifold. If $\mu : [0, T] \to M$ is continuous and piecewise differentiable, its length is defined by

$$L(\mu) = \int_0^T |\dot{\mu}|_{\mu(t)}\, dt.$$

In other terms, one defines infinitesimal length elements from norms on tangent spaces to M. Similarly, the energy of μ is defined by

$$E(\mu) = \frac{1}{2}\int_0^T |\dot{\mu}|^2_{\mu(t)}\, dt.$$

The extremal curves of the energy are called geodesics (one says that a curve is an extremal of a given variational problem if any first-order local perturbation of the curve has only second-order effects on the functional). In a chart where $\mu(t) = (y^{(1)}(t), \dots, y^{(n)}(t))$, and where $S(y) = (s^{(ij)}(y))$ is the matrix associated to the inner product, we have

$$|\dot{\mu}|^2_{\mu(t)} = \sum_{ij} s^{(ij)}(y(t))\dot{y}^{(i)}\dot{y}^{(j)}.$$

Making a local variation $y^{(i)} \mapsto y^{(i)} + h^{(i)}$, we find that extremals of the energy are characterized by: for all i

$$\int_0^T \left(\sum_{lj} \dot{h}^{(l)} s^{(lj)}(y(t))\dot{y}^{(j)} + \sum_{i,j,l} \dot{y}^{(i)}\dot{y}^{(j)}\partial_l s^{(ij)} h^{(l)} \right) dt = 0,$$

which yields, after an integration by parts

$$-2\int_0^T \Big(\sum_{l,j} h^{(l)} \partial_t \left(s^{(lj)}(y(t))\dot{y}^{(j)} \right) + \sum_{i,j,l} \dot{y}^{(i)}\ \dot{y}^{(j)}\partial_l s^{(ij)} h^{(l)} \Big) dt = 0.$$

This relation being true for every h, we can conclude that extremals must satisfy, for all $l = 1, \dots, d$,

$$-2\sum_j s^{(lj)}(y)\ddot{y}^{(j)} - 2\sum_{ij} \partial_i s^{(lj)}\dot{y}^{(i)}\dot{y}^{(j)} + \sum_{i,j} \dot{y}^{(i)}\dot{y}^{(j)}\partial_l s^{(ij)} = 0.$$

Let $\tilde{s}^{(ij)}$ denote the coefficients of S^{-1}. The previous identities give (with a symmetrized second term)

$$-2\ddot{y}^{(k)} = \sum_{ij} \dot{y}^{(i)}\dot{y}^{(j)} \sum_l \tilde{s}^{(kl)}(\partial_i s^{(lj)} + \partial_j s^{(li)} - \partial_l s^{(ij)}).$$

Defining

$$\Gamma^k_{ij} = \frac{1}{2}\sum_l \tilde{s}^{(kl)}(\partial_i s^{(lj)} + \partial_j s^{(li)} - \partial_l s^{(ij)}), \tag{B.6}$$

this is

$$\ddot{y}^{(k)} + \sum_{i,j} \Gamma^k_{ij}\dot{y}^{(i)}\dot{y}^{(j)} = 0. \tag{B.7}$$

The coefficients Γ^k_{ij} only depend on the Riemannian metric. They are called the Christoffel symbols of the manifold at a given point. Therefore, geodesics (expressed in a local chart) are solutions of a second-order differential equation. This implies

that they are uniquely specified by their value at time, say, $t = 0$ and their derivative $\dot{\mu}$ at $t = 0$. Using this property, one defines the Riemannian exponential at $p \in M$ in the direction $v \in T_pM$ by

$$\mathrm{Exp}_p(tv) = \mu(t), \tag{B.8}$$

where $\mu(\cdot)$ is the geodesic with $\mu(0) = p$ and $\dot{\mu}(0) = v$. (The capital "E" and the index p differentiate this notation from the Lie group exponential that was defined previously.) Such a geodesic exists, as a solution of a differential equation, at least for small times, so that the exponential is well-defined at least for small enough t. If this exponential exists at all p's for all times, M is said to be a (geodesically) complete manifold. The Riemannian exponential can be restricted to some neighborhood of 0 in T_pM to form a local chart of the manifold.

B.6.2 Geodesic Distance

When M is a Riemannian manifold, one defines the distance between two points p and p' in M to be the length of the shortest path that links them, setting

$$d(p, p') = \inf\{L(\mu) : \mu : [0, 1] \to M, \\ \mu \text{ continuous, piecewise differentiable, } \mu(0) = p, \mu(1) = p'\}.$$

The following theorem is standard, and may be proved as an exercise or read, for example, in [87]:

Theorem B.21 *The function d which is defined above is a distance on M. Moreover,*

$$d(m, m') = \inf\{\sqrt{2E(\mu)} : \mu : [0, 1] \to M, \\ \mu \textit{ continuous, piecewise differentiable, } \mu(0) = p, \mu(1) = p'\}.$$

B.6.3 Lie Groups with a Right-Invariant Metric

On Lie groups, Riemannian structures can be coupled with invariance constraints. As seen in Chap. 7 when considering groups of diffeomorphisms, the suitable way of "moving" within a group is by iterating small steps through the composition rule. For a curve $g(\cdot)$ on the group, the length of a portion between $g(t)$ and $g(t + \varepsilon)$ should measure the increment $g(t + \varepsilon)g(t)^{-1}$. Fix t and let $u(\varepsilon) = g(t + \varepsilon)g(t)^{-1}$: one has $u(0) = 0$ and $g(t + \varepsilon) = u(\varepsilon)g(t)$. If there is a Riemannian structure on G, the length of the displacement from $g(t)$ to $g(t + \varepsilon)$ is

$$\int_t^{t+\varepsilon} |\dot{g}_s|_{g(s)}\, ds \simeq \varepsilon\, |\partial_\varepsilon g(t + \varepsilon)|_{g(t)}\,,$$

where the last derivative is taken at $\varepsilon = 0$. The right-invariance constraint says that this length should in fact only be measured by the increment on the group and therefore be a function of $u(\varepsilon)$. But $u(\cdot)$ is itself a curve on G, between id and $g(t+\varepsilon)g(t)^{-1}$, and its length is, at first order, given by $\varepsilon\,|\dot{u}(0)|_{\mathrm{id}}$. Thus, with the invariance constraint, we should take

$$|\dot{g}(t)|_{g(t)} = |\dot{u}(0)|_{\mathrm{id}}\,.$$

Introduce the right translation in the Lie group:

$$\begin{aligned} R_g : \quad & G \to G \\ & h \mapsto hg \end{aligned}$$

so that $g(t+\varepsilon) = R_{g(t)}(u(\varepsilon))$. We have, by the chain rule,

$$\dot{g}(t) = dR_{g(t)}(\mathrm{id})\dot{u}(0).$$

This leads to the following definition:

Definition B.22 A Riemannian metric on a Lie group G is said to be right-invariant if and only if, for all $u \in \mathfrak{g} = T_{\mathrm{id}}G$, for all $g \in G$

$$\left|dR_g(\mathrm{id})u\right|_g = |u|_{\mathrm{id}}\,.$$

Thus, the metric on any T_gG may be obtained from the metric on $\mathfrak{g}$ by right translation.

The conservation of momentum (see Sect. 10.5) is in fact true in any Lie group with a right-invariant metric. The proof is very similar to the one we have given for diffeomorphisms [18, 19, 188].

B.6.4 Covariant Derivatives

We briefly describe here how a Riemannian structure leads naturally to the notion of a directional derivative of a vector field relative to another. This can provide an introduction to, and a motivation for, the more abstract theory of affine connections [87]. This will allow us to interpret a geodesic as a curve with vanishing acceleration, similar to straight lines in Euclidean spaces.

Let us work on the simpler case in which M is a submanifold of $\mathbb{R}^N$. In this case, the tangent spaces to M can be considered as affine subspaces of $\mathbb{R}^N$, which inherit its standard dot product, which is the Riemannian metric on M. A curve μ on M is also a curve in $\mathbb{R}^N$, and its energy is given by

$$E(\mu) = \frac{1}{2}\int_0^1 |\dot{\mu}|^2\,dt,$$

where the norm here is simply the Euclidean norm in $\mathbb{R}^N$. To compute the geodesics, we need the extremals of E, subject to the constraint that the curves must remain on M.

For such an extremal, μ, and a small perturbation h such that $h(t) \in T_{\mu(t)}M$ for all t, we have

$$E(\mu + h) \simeq E(\mu) + \int_0^1 \langle \dot{\mu}\,,\, \dot{h} \rangle dt + o(h) = E(\mu) - \int_0^1 \langle \ddot{\mu}\,,\, h \rangle dt + o(h).$$

Thus, μ is an extremal if and only if

$$\langle \ddot{\mu}\,,\, h \rangle = 0$$

for all $h \in T_{\mu(t)}M$, which is equivalent to, for all t

$$\Pi_{\mu(t)}\,(\ddot{\mu}) = 0,$$

where Π_p is the orthogonal projection of $\mathbb{R}^N$ on T_pM, with $p \in M$. This provides another characterization of geodesics (which does not require the introduction of local coordinates), in the particular case of a submanifold of $\mathbb{R}^N$.

Still restricting ourselves to this case, let us fix a curve μ on M, and define, for all vector fields Y on M, the derivative of Y along μ by

$$\frac{DY}{Dt}_{|_{\mu(t)}} = \Pi_{\mu(t)} \left(\partial_t Y_{\mu(t)} \right).$$

This is a vector field along μ, and the previous computation shows that a geodesic is characterized by the equation

$$\frac{D\dot{\mu}}{Dt} = 0.$$

One can show (we skip the details) that the expression in a local chart $C = (U, \Phi)$ of the derivative along μ of a vector field $Y = \sum_{i=1}^d \eta_i \partial_i$ is given by

$$\sum_{i=1}^d \rho_i \partial_{i,\mu(t)}$$

with

$$\rho_i = \partial(\eta_i \circ \mu(t)) + \sum_{j,k=1}^d \Gamma_{jk}^i(\mu(t))\eta_j(\mu(t))\eta_k(\mu(t)).$$

The Γ_{jk}^i's are the same Christoffel symbols as defined in (B.6). Introducing $\lambda_1, \ldots, \lambda_d$ such that

$$\dot{\mu}(t) = \sum_{i=1}^{d} \lambda_i(t)\partial_{x_i}(\mu(t))$$

one can write

$$\rho_i = \sum_{j=1}^{d} \lambda_j(t)\partial_{x_j}\eta_i + \sum_{j,k=1}^{d} \Gamma^i_{jk}(\mu(t))\eta_j(\mu(t))\eta_k(\mu(t)).$$

This expression is intrinsic: it does not depend on the ambient space $\mathbb{R}^N$, but on quantities that are computed on the manifold. So, assuming now that M is a general Riemannian manifold, we can define, in a chart $C = (U, \Phi)$, and for two vector fields $X = \sum_{i=1}^{d} \xi_i \partial_i$ and $Y = \sum_{i=1}^{d} \eta_i \partial_i$, a third vector field, called the *covariant derivative* of Y along X, by

$$(\nabla_X Y)_p = \sum_{i=1}^{d} \rho_i \partial_{i,p}$$

with

$$\rho_i = \sum_{j=1}^{d} \xi_j \partial_j \eta_i + \sum_{j,k=1}^{d} \Gamma^i_{jk}\eta_j\xi_k.$$

(Note that $(\nabla_X Y)_p$ only depends on the coordinates of X at p.)

From our definition of the Christoffel symbols in the general case, we see that curves of minimal energy still satisfy

$$\frac{D\dot{\mu}}{Dt} := \nabla_{\dot{\mu}}\,\dot{\mu} = 0.$$

The equation $\nabla_{\dot{\mu}}\,\dot{\mu}$ is therefore called the geodesic equation for curves on M. Curves that satisfy it are called geodesics, even when they are not energy-minimizing (although they are always locally so, see [87]).

Covariant derivatives can be defined in more general contexts than on Riemannian manifolds [87, 142]. The one we have defined above is adapted to the Riemannian metric and called the *Levi-Civita connection*. It satisfies the two characteristic properties:

$$\begin{aligned}
&(1)\ Z\langle X\,,\ Y\rangle = \langle \nabla_Z X\,,\ Y\rangle + \langle X\,,\ \nabla_Z Y\rangle \\
&(2)\ \nabla_X Y - \nabla_Y X = [X, Y] \text{ with } [X, Y] = XY - YX,
\end{aligned}$$

and is explicitly defined by the identity

$$\langle \nabla_X Y\ ,\ Z\rangle = \frac{1}{2}\Big(X\langle Y\ ,\ Z\rangle + Y\langle X\ ,\ Z\rangle - Z\langle Y\ ,\ X\rangle \\ - \langle [X, Z]\ ,\ Y\rangle - \langle [Y, Z]\ ,\ X\rangle + \langle [X, Y]\ ,\ Z\rangle\Big). \tag{B.9}$$

B.6.5 Parallel Transport

Parallel transport (or translation) is the displacement of a vector along a curve with vanishing covariant derivative. It is the generalization of translation in Euclidean spaces. Given a curve $t \mapsto \mu(t)$ on a Riemannian manifold M, a time-dependent tangent vector $t \mapsto X(t)$ with $X(t) \in T_{\mu(t)}M$ is said to be parallel on μ if its derivative along μ vanishes, i.e.,

$$\nabla_{\dot\mu} X = \frac{DX}{Dt} = 0.$$

So, by definition, a geodesic is a curve with derivative moving parallel to itself. Parallel transport derives from a first-order differential equation (for X) along the curve, which, in a chart, is given by:

$$\partial_t(\rho_k \circ \mu) + \sum_{i,j=1}^{d} \Gamma_{ij}^k\, \rho_i \circ \mu\, \dot\mu_j = 0$$

with $X = \sum_{k=1}^{d} \rho_k \partial_{x_k}$. This first-order linear system of equations in $\rho \circ \mu$ can be integrated, with a unique solution, as soon as an initial condition $\rho \circ \mu(0)$ is given. This leads to the following definition.

Definition B.23 Let M be a Riemannian manifold, $p \in M$ and $\xi_0 \in T_pM$. Let μ be a curve on M with $\mu(0) = p$. The parallel transport of ξ_0 along μ is the time-dependent vector $\xi(t) \in T_{\mu(t)}M$ such that $\xi(0) = \xi_0$ and

$$\nabla_{\dot\mu}\xi = 0.$$

It is important to remember that parallel transport is only defined *along a curve*. If $p, \tilde p \in M$, $\xi \in T_pM$, and μ and $\tilde\mu$ are two curves linking p to $\tilde p$, the results of the parallel transport of ξ along μ and $\tilde\mu$ are generally distinct.

B.6.6 A Hamiltonian Formulation

Geodesics in a chart have a Hamiltonian formulation that is sometimes convenient. Define the function (Hamiltonian)

$$H(p,a) = \frac{1}{2} a^T S(p)^{-1} a, \tag{B.10}$$

where S is the metric, $p \in M$ and a is a d-dimensional vector. One can prove that a curve $t \mapsto m(t)$ on M is a geodesic if and only if it satisfies the system

$$\begin{cases} \partial_t m = S(\mu(t))^{-1} a(t) = \partial_a H(m(t), a(t)), \\ \\ \partial_t a = -\partial_p H(m(t), a(t)). \end{cases}$$

(This is an immediate consequence of the Pontryagin maximum principle that will be discussed in Appendix D.) Introducing the coefficients $\tilde{s}^{(ij)}$ of S^{-1}, this is

$$\begin{cases} \partial_t m_i = \displaystyle\sum_{j=1}^{d} \tilde{s}^{(ij)}(m) a_j, \\ \\ \partial_t a_i = -\dfrac{1}{2} \displaystyle\sum_{k,l=1}^{d} a_k a_l \partial_i (\tilde{s}^{(kl)})(m). \end{cases}$$

Note that the energy of the curve is (at time t)

$$\frac{1}{2} \langle \dot{m} \, , \, \dot{m} \rangle_m = \frac{1}{2} \dot{m}^T S(m) \dot{m} = \frac{1}{2} a^T \dot{m}.$$

The vector $a(t)$ is *cotangent* to the motion (because it acts as a linear form on tangent vectors) and must be considered as an element of $T_p M^*$. It is the *momentum* of the motion.

The covariant derivative can also be written in terms of a. If $a = SX$, $b = SY$ and if we let

$$\nabla_a^* b = S(\nabla_X Y),$$

then, the kth coordinate of $\nabla_a^* b$ is

$$\begin{aligned} (\nabla_a^* b)_k = & \sum_{j=1}^{d} \partial_j (b_k) X_j \\ & + \frac{1}{2} \left(\sum_{i,j=1}^{d} \partial_k \tilde{s}^{(ij)} a_i a_j + \sum_{i,j,l,q=1}^{d} s_{kl} \partial_j s^{(lq)} (b_q X_j - a_q Y_j) \right). \end{aligned}$$

This expression also provides the equation for parallel translation of b along a curve m with $S\dot{m} = a$, namely

$$\partial_t b_k + \frac{1}{2}\left(\sum_{i,j=1}^{d} \partial_k \tilde{s}^{(ij)} a_i b_j + \sum_{j,l,q=1}^{d} s_{kl}\partial_j s^{(lq)}(b_q X_j - a_q Y_j)\right) = 0.$$

B.6.7 Riemannian Submersions

All submanifolds of $\mathbb{R}^d$ (or of a Riemannian manifold) are Riemannian manifolds when they inherit the inner product of the larger space. This provides a large number of examples of such manifolds.

Another way to define metrics on manifolds is through Riemannian submersions. A differentiable mapping $\pi : M \to B$ (where B and M are differential manifolds of respective dimensions $k \leq d$) is a submersion if π is onto and $d\pi(p)$ has full rank, d, for all $p \in M$.

Assume that M and B are Riemannian and that $\pi : M \to B$ is a submersion. Then the set $F_b = \pi^{-1}(b)$ is a submanifold of M for any $b \in B$, with dimension $d - k$. For $p \in M$, we let $\mathcal{V}_p = T_p F_{\pi(p)}$ (vertical space at p) and $\mathcal{H}_p = \mathcal{V}_p^{\perp}$ (horizontal space at p), where the orthogonality is taken with respect to the metric on M. Then, $d\pi(p)v = 0$ for any $v \in \mathcal{V}_p$ (because π is constant on $F_{\pi(p)}$, so that $\mathcal{V}_p = \text{Null}(d\pi(p))$. This shows that $d\pi(p)$ restricted to $\mathcal{H}_p$ is an isomorphism between $\mathcal{H}_p$ and $T_{\pi(p)}M$, i.e., it is one-to-one and onto. One says that π is a *Riemannian submersion* if and only if, for all $p \in M$, this isomorphism is, in addition, an isometry, i.e.,

$$|d\pi(p)h|_{\pi(p)} = |h|_p \tag{B.11}$$

for all $h \in \mathcal{H}_p$.

Given a submersion $\pi : M \to B$ and a Riemannian metric on M, one can define a Riemannian metric on B so that π is a Riemannian submersion if and only if the horizontal spaces are isometric via the submersion, i.e.,

$$h_1 \in \mathcal{H}_{p_1}, h_2 \in \mathcal{H}_{p_2}, d\pi(p_1)h_1 = d\pi(p_2)h_2 \Longrightarrow |h_1|_{p_1} = |h_2|_{p_2}. \tag{B.12}$$

Notice that, for $b \in B$, $\xi \in T_b B$, and $b = \pi(p)$, there is a unique $h \in \mathcal{H}_p$ such that $d\pi(p)h = \xi$, which is called the *horizontal lift* of ξ to $T_p M$. Equation (B.12) expresses the condition that the norms of all horizontal lifts of a given tangent vector to M must coincide.

Another way to express the relationship between the metrics is via

$$|u|_{\pi(p)} = \min\left\{|h|_p, d\pi(p)h = u\right\} \tag{B.13}$$

for $u \in T_{\pi(p)}B$, which derives from the orthogonality of $\mathcal{V}_p$ and $\mathcal{H}_p$. To be well defined, this formula should only depend on $\pi(p)$. This requirement is equivalent to (B.12).

An important property of Riemannian submersions is that geodesics on M that start horizontal remain so over all times.

Proposition B.24 *Let $\pi : M \to B$ be a Riemannian submersion. Assume that $\mu : [0, 1] \to M$ is a geodesic such that $\dot{\mu}(0) \in \mathcal{H}_{\mu(0)}$. Then $\dot{\mu}(t) \in \mathcal{H}_{\mu(t)}$ for all $t \in [0, 1]$.*

Proof We start with a lemma that defines horizontal lifts of curves on B.

Lemma B.25 *If μ^* is a curve on B and $p \in M$ such that $\pi(p) = \mu(0)$, there exists a unique curve $\hat{\mu}$, called the horizontal lift of μ^* to M starting at p, such that $\pi(\hat{\mu}) = \mu^*$, $\hat{\mu}(0) = p$ and $\partial_t \hat{\mu}(t)$ horizontal for all t.*

To prove this lemma, we note that the lift can be defined (locally) by taking local charts around p in M and around $\mu^*(0)$ in B such that π in this chart maps coordinates $x_1, \ldots, x_d$ to $x_1, \ldots, x_k$. If the metric of M in this chart takes the form

$$S = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

(where A is k by k), a horizontal vector takes the form $h = (h_1, -C^{-1}B^T h_1)^T$ in the chart, as can be checked by expressing the condition that $h^T S v = 0$ for any vertical vector $v = (0, v_2)^T$. Letting ξ^* denote the expression of μ^* in the chart, one can define the lift as $\hat{\xi} = (\xi^*, \eta)$ where η satisfies $\dot{\eta} = -C^{-1}B^T\dot{\xi}^*$, which characterizes η uniquely once $\eta(0)$ is specified. This specifies the lift in a neighborhood of $\mu(0)$, and it now suffices to extend it as needed by iterating the construction.

If X^* is a vector field on B, there exists a unique horizontal vector field X on M such that $d\pi(p)X(p) = X^*(\pi(p))$ for all $p \in M$. Let ∇ and ∇^* denote the covariant derivatives on M and on B. One then has the following lemma.

Lemma B.26 *If X and Y are horizontal,*

$$d\pi\, \nabla_X Y = \nabla^*_{X^*} Y^*. \tag{B.14}$$

Moreover, $\nabla_X X$ is also horizontal.

To show the first statement of this lemma, it suffices to notice that the left-hand side satisfies the properties of a covariant derivative, including those that characterize the Levi-Civita connection on B, and we skip the details. For the second one, letting V be a vertical vector field on M, one needs to show that $\langle \nabla_X X\,,\ V\rangle = 0$. One has $0 = X\langle X\,,\ V\rangle = \langle \nabla_X X\,,\ V\rangle + \langle X\,,\ \nabla_X V\rangle$. Moreover $\langle X\,,\ \nabla_X V\rangle = \langle X\,,\ \nabla_V X\rangle + \langle X\,,\ [X, V]\rangle$, and the conclusion comes from the fact that both terms on the right-hand side vanish: the first one because $0 = V\langle X\,,\ X\rangle = 2\langle \nabla_V X\,,\ X\rangle$, because the metric is constant along vertical directions; the second one because $d\pi[X, V] = [d\pi X, d\pi V] = 0$, so that $[X, V]$ is vertical.

Going back to the statement of Proposition B.24, let μ^* denote the geodesic in B such that $\dot{\mu}^*(0) = d\pi(\mu(0))\dot{\mu}(0)$ and $\hat{\mu}$ its horizontal lift at $\mu(0)$. Equation (B.14) implies that $D\hat{\mu}/Dt$ is horizontal, but, since it is also vertical by the second part of the lemma, we see that $D\hat{\mu}/Dt = 0$ so that $\hat{\mu}$ is a geodesic on M. We have $\mu(0) = \hat{\mu}(0)$ (by construction) and $\partial\mu(0) = \partial\hat{\mu}(0)$ (because they are both horizontal and project on $\partial\mu^*(0)$). From the uniqueness of geodesics given initial position and derivative, we find that $\mu = \hat{\mu}$, so that μ is horizontal. □

While proving Proposition B.24, we have also proved the following important result.

Proposition B.27 *Let $\pi : M \to B$ be a Riemannian submersion. If μ^* is a geodesic on B, then all its horizontal lifts are geodesics on M.*

The relationship between parallel transport on M and on B is less direct. If γ^* is a curve on B and $\xi_0^* \in T_{\gamma^*(0)}B$, the parallel transport of ξ_0^* along γ^* is defined by

$$\frac{D^*\xi^*}{Dt} = 0$$

with the notation $D^*\xi^*/Dt = \nabla^*_{\dot{\gamma}^*}\xi^*$ for a vector field along γ^*. If p is such that $\pi(p) = \gamma^*(0)$ and $\gamma(t)$ is the unique horizontal curve such that $\gamma(0) = p$ and $\pi(\gamma(t)) = \gamma^*(t)$ for all t, then, letting ξ denote the horizontal lift of ξ^*,

$$\frac{D^*\xi^*}{Dt} = d\pi(\gamma(t))\frac{D\xi}{Dt}\,.$$

Parallel transport on B can therefore be described in terms of the submersion and the metric on M via

$$d\pi(\gamma(t))\frac{D}{Dt}(\ell_{\gamma(t)}(\xi^*(t))) = 0, \tag{B.15}$$

where $\ell_p : T_{\pi(p)}B \to \mathcal{H}_p$ denotes the horizontal lift.

More details on Riemannian submersions can be found in [225]. We now review some examples of Riemannian manifolds built according to this principle.

Unit Spheres and Projective Spaces

Take $B = \mathbb{P}^d(\mathbb{R})$, which, we recall, is the set of all real lines passing through 0 in $\mathbb{R}^{d+1}$. Letting $M = S^d \subset \mathbb{R}^{d+1}$, B can also be considered as the set of all pairs $\{p, -p\}$ for $p \in M$ with the surjection $\pi : p \mapsto \{-p, p\}$. It is easy to show that π is smooth: consider the local charts (U_i, Φ_i) described for the projective space in Eqs. (B.4) and (B.5). Consider the charts (U_i, Φ_i^+) and (U_i, Φ_i^-) on S^d, with

$$\Phi^+(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) = (x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_{n+1})/\lambda,$$

where $\lambda = \sqrt{x_1^2 + \cdots + x_{i-1}^2 + 1 + x_{i+1}^2 + \cdots + x_{n+1}^2}$ and $\Phi^- = -\Phi^+$. Then the mapping π interpreted in coordinates between (U_i, Φ_i) and $(U_i, \Phi_i^\pm)$ is the identity, proving that it is smooth and has a full rank.

Trivially, the vertical fiber above $b = \{-p, p\} \in B$ is b itself (considered as a subset of S^d), and the tangent planes at these points are clearly isometric via the submersion π. One can therefore define a Riemannian metric on $T_b B$ by duplicating the one on S^d at one of the antipodal points that maps onto b.

Stiefel and Grassmann Manifolds

We now let $M = St(n, k)$ be the set of all orthonormal families of size k in $\mathbb{R}^n$ and $B = Gr(n, k)$ the space of all subspaces of dimension k in $\mathbb{R}^n$. This example generalizes the previous one because $S^d = St(d + 1, 1)$ and $\mathbb{P}^d = Gr(d + 1, 1)$.

Both M and B are manifolds: M can be identified with the subset of $\mathcal{M}_{n,k}(\mathbb{R})$ comprising $n \times k$ matrices A that satisfy the equation $A^T A = \mathrm{Id}_{\mathbb{R}^k}$. The mapping $\Phi : \mathcal{M}_{n,k}(\mathbb{R}) \to \mathcal{M}_k(\mathbb{R})$ defined by $\Phi(A) = A^T A$ is such that $d\Phi(A)H$, which is equal to $A^T H + H^T A$, is a symmetric matrix. Conversely, given any symmetric matrix, S, one has $d\Phi(A)(AS/2) = S$ if $A^T A = \mathrm{Id}_{\mathbb{R}^k}$, which shows that $d\Phi(A)$ is a surjection onto the space of symmetric matrices. This shows that Φ has constant rank, given by $k(k + 1)/2$, on M, which is therefore a submanifold of $\mathcal{M}_{n,k}(\mathbb{R})$ of dimension $nk - k(k + 1)/2$. M also inherits the Riemannian structure of this space, with $|H|_A^2 = \mathrm{trace}(H^T H)$.

To define local charts on $B = \mathrm{Gr}(n, k)$, first notice that k linearly independent vectors in $\mathbb{R}^n$ can be represented as an n by k matrix $A = (x_{ij})$ such that, for at least one tuple $J = (r_1, \ldots, r_k)$ with $1 \leq r_1 < \cdots < r_k \leq n$, the submatrix A_J with entries $(x_{r_i j}, i, j = 1, \ldots, k)$ is invertible. A change of basis in the subspace generated by the columns of A can be obtained by multiplying A on the right by a k by k invertible matrix. In particular, the matrix $\tilde{A} = A\, A_J^{-1}$ generates the same subspace and satisfies $\tilde{A}_J = \mathrm{Id}_{\mathbb{R}^k}$ (and is the only matrix with both these properties).

For a k-tuple J as above, let $U = \mathcal{M}_{n-k,k}$, the set of all $(n - k) \times k$ matrices and $\Phi_J : U \to \mathrm{Gr}(n, k)$ be such that $\Phi_J(A) = \mathrm{Range}(\hat{A})$, where $\hat{A}$ is such that $\hat{A}_J = \mathrm{Id}_{\mathbb{R}^k}$ and $\hat{A}_{J^c} = A$, where the latter expression denotes the $(n - k) \times k$ matrix obtained by removing from $\hat{A}$ the rows corresponding to indices in J. The family (U, Φ_J) provide an atlas on B, which therefore has dimension $k(n - k)$. Note that the transformation $\Phi_{J,R}$ in which the condition $\hat{A}_J = \mathrm{Id}_{\mathbb{R}^k}$ is replaced by $\hat{A}_J = R$ can also be chosen as a local chart for any $R \in \mathrm{GL}_k(\mathbb{R})$.

Consider the mapping $\pi : M \to B$ defined by $\pi(A) = \mathrm{Range}(A)$. Then π is a submersion. To see this, first notice that π is the restriction to M of the mapping $\hat{\pi} : \mathcal{M}_{n,k} \to B$, also defined by $\hat{\pi}(A) = \mathrm{Range}(A)$. Then $\hat{\pi}$ interpreted between the charts $\hat{\pi}^{-1}(\Phi_J(U))$ and U is simply the transformation $A \mapsto (A\, A_J^{-1})_{J^c} = A_{J^c} A_J^{-1}$, which is smooth. This implies that $\hat{\pi}$ and therefore π is smooth. One has (with the same chart interpretation)

$$d\hat{\pi}(A)H = H_{J^c} A_J^{-1} - A_{J^c} A_J^{-1} H_J A_J^{-1},$$

which vanishes when $H_{J^c} = A_{J^c} A_J^{-1} H_J$ so that $\mathrm{Null}(d\hat{\pi}(A))$ has dimension $k^2 = \dim(\mathcal{M}_{nk}) - \dim(B)$, which shows that $d\hat{\pi}$ has full rank. Note that this condition can also be written as $H = A\tilde{H}$ for some $\tilde{H} \in \mathcal{M}_k(\mathbb{R})$ (namely $\tilde{H} = A_J^{-1} H_J$). The fact that these matrices belong to the null space of $d\pi(A)$ results from the

observation that $\pi(AR) = \pi(R)$ for any $R \in \mathrm{GL}_k(\mathbb{R})$. If $A \in M$ (so that $A^T A = \mathrm{Id}$), matrices H in the intersection of this null space with the tangent space to M at A must also satisfy $A^T H + H^T A = 0$. Writing $H = A\tilde{H}$, we see that the condition requires that $\tilde{H}$ is skew-symmetric, so that $\mathrm{Null}(d\pi(A))$ has dimension $k(k-1)/2 = \dim(M) - \dim(B)$ and π is also a submersion.

The vertical fiber that contains A is the set of matrices AR for some $R \in O_k(\mathbb{R})$ and the vertical space $\mathcal{V}_A$ is the set of matrices $H \in T_A M$ such that $H = A\tilde{H}$ for $\tilde{H}$ skew-symmetric (which implies $A^T H + H^T A = 0$). Therefore, the horizontal space contains matrices H such that $A^T H + H^T A = 0$ and $\mathrm{trace}(H^T A\tilde{H}) = 0$ for all skew-symmetric matrices H. The latter condition implies that $A^T H$ is symmetric, which, combined with the former, implies that $A^T H = 0$.

The left action of $O_k(\mathbb{R})$ on M induces transformations $\psi_R : A \mapsto AR$. They are such that $d\psi_R(A)H = HR$ and are therefore isometries between tangent spaces along vertical fibers: $|d\psi_R(A)H|^2_{AR} = \mathrm{trace}(HRR^T H^T) = \mathrm{trace}(HH^T) = |H|^2_A$. Since these transformations map vertical spaces onto vertical spaces, they also conserve horizontality. Using the fact that $\pi \circ \psi_R = \pi$, we see that $d\pi \circ \psi_R$ and $d\pi$ are related by an isometry ($d\psi_R$), which shows that π is a Riemannian submersion.

We can therefore transport the metric of M onto B to provide the Grassmann manifold with a Riemannian structure. For $q \in B$ and $\beta \in T_q B$, it is defined by

$$|\beta|^2_q = \mathrm{trace}(HH^T),$$

where $H \in \mathcal{M}_{n,k}(\mathbb{R})$ is such that $A^T H = 0$ and $d\pi(A)H = \beta$ for some $A \in \pi^{-1}(q)$. In the local chart considered above, in which β is represented as an $(n-k)$ by k matrix, this identity corresponds to $H_{J^c} - A_{J^c} A_J^{-1} H_J = \beta A_J$ with $A_J^T H_J + A_{J^c}^T H_{J^c} = 0$, which uniquely defines H with

$$\begin{cases} H_{J^c} = \left(\mathrm{Id}_{\mathbb{R}^{n-k}} + A_{J^c} A_J^{-1} A_J^{-T} A_{J^c}^T\right)^{-1} \beta A_J \\ H_J = -A_J^{-T} A_{J^c}^T \left(\mathrm{Id}_{\mathbb{R}^{n-k}} + A_{J^c} A_J^{-1} A_J^{-T} A_{J^c}^T\right)^{-1} \beta A_J \end{cases}$$

so that

$$\begin{aligned} |\beta|^2_q &= \mathrm{trace}(H_J^T H_J + H_{J^c}^T H_{J^c}) \\ &= \mathrm{trace}\left(A_J^T \beta^T \left(\mathrm{Id}_{\mathbb{R}^{n-k}} + A_{J^c} A_J^{-1} A_J^{-T} A_{J^c}^T\right)^{-1} \beta A_J\right). \end{aligned}$$

B.7 Differential Forms

B.7.1 Multilinear Forms

A k-linear form on $\mathbb{R}^d$ is a function $\alpha : (\mathbb{R}^d)^k \to \mathbb{R}$ which is linear with respect to each of its variables. We will use the notation

$$\alpha(u_1, \ldots, u_k) = (\alpha \mid u_1, \ldots, u_k),$$

which is consistent with our notation for $k = 1$.

A k-linear form is skew-symmetric (or alternating) if $(\alpha \mid u_1, \ldots, u_k) = 0$ whenever $u_i = u_j$ for some $i \neq j$. Equivalently, such a form satisfies

$$(\alpha \mid u_1, \ldots, u_k) = \varepsilon(\sigma)\big(\alpha \;\big|\; u_{\sigma(1)}, \ldots, u_{\sigma(k)}\big)$$

for $\sigma \in \mathfrak{S}_k$, the group of permutations of $\{1, \ldots, k\}$ and $\varepsilon(\sigma) \in \{-1, 1\}$ its signature, associated with the parity of the decomposition of the permutation as a sequence of inversions.

The k-linear skew-symmetric forms define a vector space, denoted by Λ_k (or $\Lambda_k(\mathbb{R}^d)$). A one-form can always be expressed as $(\eta_a \mid u) := a^T u$ for some $a \in \mathbb{R}^d$, and we retrieve the fact that $\Lambda_1(\mathbb{R}^d) \simeq \mathbb{R}^d$.

Given two one-forms α_1 and α_2, the product $(u_1, u_2) \mapsto (\alpha_1 \mid u_1)(\alpha_2 \mid u_2)$ is bilinear, but not alternating. It can be modified by defining

$$(\alpha_1 \wedge \alpha_2 \mid u_1, u_2) = (\alpha_1 \mid u_1)(\alpha_2 \mid u_2) - (\alpha_1 \mid u_2)(\alpha_2 \mid u_1),$$

yielding an alternating two-form called the *wedge product* of α_1 and α_2. More generally, one defines the wedge product between k 1-forms as the k-form given by

$$(\alpha_1 \wedge \cdots \wedge \alpha_k \mid u_1, \ldots, u_k) = \sum_{\sigma \in \mathfrak{S}_k} \varepsilon(\sigma)\big(\alpha_1 \;\big|\; u_{\sigma(1)}\big) \cdots \big(\alpha_k \;\big|\; u_{\sigma(k)}\big).$$

One also defines the product between a k-form α and an l-form β so that the above expression becomes associative, i.e.,

$$(\alpha_1 \wedge \cdots \wedge \alpha_k) \wedge (\beta_1 \wedge \cdots \wedge \beta_l) = \alpha_1 \wedge \cdots \wedge \alpha_k \wedge \beta_1 \wedge \cdots \wedge \beta_l,$$

where all forms in this identity are one-forms. This results in the definition

$$(\alpha \wedge \beta \mid u_1, \ldots, u_{k+l}) = \frac{1}{k!\,l!} \sum_{\sigma \in \mathfrak{S}_{k+l}} \varepsilon(\sigma)\big(\alpha \;\big|\; u_{\sigma(1)}, \ldots, u_{\sigma(k)}\big)\big(\beta \;\big|\; u_{\sigma(k+1)}, \ldots, u_{\sigma(k+l)}\big).$$

The sum in the previous equation is redundant, in the sense that the summand only depend on the *sets* $\{\sigma(1), \ldots, \sigma(k)\}$ and $\{\sigma(k+1), \ldots, \sigma(k+l)\}$, and not on the orders within these sets. This explains the normalization by $k!\,l!$.

If $e_1, \ldots, e_d$ is a basis of $\mathbb{R}^d$, one defines the dual basis $e_1^*, \ldots, e_d^* \in \Lambda_1(\mathbb{R}^d)$ by $\big(e_i^* \;\big|\; e_j\big) = 1$ if $i = j$ and 0 otherwise, so that $e_i^*(u)$ is the ith coefficient of $u \in \mathbb{R}^d$ in its decomposition over the basis $e_1, \ldots, e_d$, yielding

$$u = \sum_{j=1}^{d} \big(e_j^* \,\big|\, u\big) e_j.$$

Using this expression and after some algebra, one sees that any k-form α can be decomposed as

$$\alpha = \sum_{1 \le i_1 < \cdots < i_k \le d} \big(\alpha \,\big|\, e_{i_1}, \ldots, e_{i_k}\big) e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*.$$

Using the fact that

$$\big(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^* \;\big|\; e_{j_1}, \ldots, e_{j_k}\big) = 1$$

if $(i_1, \ldots, i_k) = (j_1, \ldots, j_k)$ and 0 otherwise (assuming that the two sequences are increasing), one proves that such a sum vanishes if and only if all its coefficients vanish, so that $\Lambda_k(\mathbb{R}^d)$ is $\binom{d}{k}$-dimensional, with $(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*, 1 \le i_1 < \cdots < i_k \le d)$ as a basis.

In particular, $\Lambda_d(\mathbb{R}^d)$ has dimension one so that there is, up to a multiplicative constant, only one non-vanishing skew-symmetric d-linear functional. If $(e_1, \ldots, e_d)$ is orthonormal, it is clear from the definition of the determinant that

$$\big(e_1^* \wedge \cdots \wedge e_d^* \;\big|\; u_1, \ldots, u_d\big) = \det(u_1, \ldots, u_d),$$

so that, for any d-form

$$(\alpha \mid u_1, \ldots, u_d) = (\alpha \mid e_1, \ldots, e_d) \det(u_1, \ldots, u_d).$$

In fact, for any $k \le d$, the skew-symmetry property implies that, for any alternating k-linear form α and any $k \times k$ matrix $A = (a_{ij})$,

$$(\alpha \mid v_1, \ldots, v_k) = \det(A)(\alpha \mid u_1, \ldots, u_k) \tag{B.16}$$

when $v_i = \sum_{j=1}^{k} a_{ij} u_j$, $i = 1, \ldots, k$.

For $k < d$, one defines the cross-product between k vectors $u_1, \ldots, u_k$ as the skew-symmetric $(d-k)$-linear functional

$$(u_1 \times \cdots \times u_k \mid v_1, \ldots, v_{d-k}) = \det(u_1, \ldots, u_k, v_1, \ldots, v_{d-k}).$$

In particular, if $e_1, \ldots, e_d$ is orthonormal, then

$$e_{i_1} \times \cdots \times e_{i_k} = \varepsilon\, e_{i_{k+1}}^* \wedge \cdots \wedge e_{i_d}^*,$$

where $\{i_{k+1}, \ldots, i_d\} = \{1, \ldots, d\} \setminus \{i_1, \ldots, i_k\}$ and

$$\varepsilon = \det(e_{i_1}, \ldots, e_{i_d}) = (-1)^{i_1 + \cdots + i_k - k(k+1)/2}.$$

If $k = d - 1$, one usually identifies $u_1 \times \cdots \times u_{d-1}$ with the unique vector w such that

$$(u_1 \times \cdots \times u_{d-1} \mid v) = w^T v$$

and w is also denoted $u_1 \times \cdots \times u_{d-1}$ in this case. (This provides the usual definition of the cross-product in $\mathbb{R}^3$, for example.)

One defines an inner product on Λ_k by first selecting an orthonormal basis $(e_1, \ldots, e_d)$ of $\mathbb{R}^d$ and deciding that $(e_{j_1} \times \cdots \times e_{j_{d-k}}, 1 \leq j_1 < \cdots < j_{d-k} \leq d)$ is an orthonormal family in Λ_k. This inner product will be denoted $\langle \cdot, \cdot \rangle_{\Lambda_k}$ and does not depend on the choice made for the orthonormal basis of $\mathbb{R}^d$.

B.7.2 Differential Forms

A differential k-form on $\mathbb{R}^d$ is a function $x \mapsto \alpha(x)$ such that, for all x, $\alpha(x)$ is a k-linear skew symmetric form, i.e., α is a function from $\mathbb{R}^d$ to $\Lambda_k(\mathbb{R}^d)$. For example, 0-forms are ordinary functions. The space of differential k-forms is denoted Ω_k (or $\Omega_k(\mathbb{R}^d)$). The set of m-times continuously differentiable k-forms is denoted Ω_k^m. The wedge product between k-forms immediately extends to differential k-forms via

$$(\alpha \wedge \beta)(x) = \alpha(x) \wedge \beta(x).$$

If M is a d-dimensional manifold, one similarly defines $\Omega_k(M)$ as the set of mappings $\omega : p \in M \to \omega(p)$ where $\omega(p) \in \Lambda_k(T_p M)$. If the mapping $p \mapsto (\omega(p) \mid X_1(p), \ldots, X_k(p))$ has class C^m for any collection of C^m vector fields $X_1, \ldots, X_k$, then one says that ω has class C^m, i.e., $\omega \in \Omega_k^m(M)$. Therefore, $\Omega_k^m(M)$ is the set of skew linear maps

$$\omega : \mathcal{X}(M)^k \to C^m(M)$$

with

$$(\omega \mid X_1, \ldots, X_k) : p \mapsto (\omega(p) \mid X_1(p), \ldots, X_k(p)).$$

If $\Phi : U \to M$ is a local chart, then $(\partial_1, \ldots, \partial_d)$ form a basis of the tangent space. Let $x_1, \ldots, x_d$ denote the coordinate functions in the chart (x_i is the ith coordinate of Φ^{-1}). Then, the differential, dx_i, of the function $p \mapsto x_i$ coincides with the ith element of the dual basis: $dx_j = \partial_j^*$ (we leave the proof of this statement to the reader). Recall that ∂_j and dx_j depend on $p \in \Phi(U)$, and are respectively (local) vector fields and differential 1-forms. Any differential k-form can be expressed in a local chart as

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \left(\omega \mid \partial_{i_1}, \ldots, \partial_{i_k}\right) dx_{i_1} \wedge \cdots \wedge dx_{i_k} \tag{B.17}$$

and ω is C^m as soon as the coefficients in this decomposition are m times continuously differentiable.

If $\omega \in \Omega_k(M)$, its *exterior derivative* $d\omega$ is a $(k+1)$-form defined in a chart as follows. If ω is given by (B.17), then

$$\begin{aligned} d\omega &= \sum_{1\le i_1<\cdots<i_k\le n} d\big(\big(\omega \mid \partial_{i_1},\ldots,\partial_{i_k}\big)\big)\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k} \\ &= \sum_{1\le i_1<\cdots<i_k\le n}\ \sum_{j=1}^{d} \partial_j\big(\omega \mid \partial_{i_1},\ldots,\partial_{i_k}\big)\, dx_j\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}. \end{aligned}$$

Note that only terms $j \notin \{i_1,\ldots,i_k\}$ are in the right-hand term, because the wedge product of any form with itself is zero. One can show that this formula does not depend on the local chart, and actually corresponds to the intrinsic definition, applied to vector fields $X_0,\ldots,X_k$ on M:

$$\begin{aligned} (d\omega \mid X_0, X_1,\ldots,X_k) &= \sum_{i=0}^{k}(-1)^i X_i(\omega \mid X_0,\ldots,X_{i-1},X_{i+1},\ldots,X_k) \\ &+ \sum_{0\le i<j\le k}(-1)^{i+j}\big(\omega \mid [X_i,X_j],X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_{j-1},X_{j+1},\ldots,X_k\big), \end{aligned}$$

where $[X_i,X_j] = X_iX_j - X_jX_i$ is the Lie bracket between X_i and X_j. If α is a k-form and β is an l-form, one shows from the definition that

$$d(\alpha\wedge\beta) = d\alpha\wedge\beta + (-1)^k\alpha\wedge d\beta$$

and that, for any k-form ω, one has $dd\omega = 0$.

If $g : M \to M'$ is a smooth mapping, where M and M' are manifolds, and $\omega \in \Omega_k(M')$, one defines the pullback, $g^*\omega \in \Omega_k(M)$, by

$$g^*\omega(X_1,\ldots,X_k) = \omega(dg\,X_1,\ldots dg\,X_k)\circ g.$$

The exterior derivative commutes with pullbacks: $d(g^*\omega) = g^*d\omega$. One also has, given two differential forms α and β: $g^*(\alpha\wedge\beta) = (g^*\alpha)\wedge(g^*\beta)$.

B.7.3 Integration of Differential Forms

Boundaries

If M is a manifold and P is a submanifold of M, we can define $\bar{P}$ as the topological closure of P in M (the set of limits of sequences in P that converge in M). The set

$\partial P = \bar{P} \setminus P$ is called the boundary of P. If P is closed, so that $\partial P = \emptyset$, P is called a manifold without boundary.

For example, an open line segment $P = \{p + t(q - p), t \in (0, 1)\}$ in $M = \mathbb{R}^2$ is such that $\partial P = \{p, q\}$. The unit circle $S^1 \subset \mathbb{R}^2$ is such that $\partial S^1 = \emptyset$, while the unit open disc, $P = B(0, 1)$, is such that $\partial P = S^1$. Note that this definition of a boundary is different from the usual topological boundary $\bar{P} \setminus \mathring{P}$ (since submanifolds have empty interior as soon as their dimension is smaller than that of the ambient manifold).

Let $k = \dim(P)$ and $d = \dim(M)$. A point $p \in \partial P$ is called a regular boundary point if there exists a local chart $\Phi : U \to M$, where U is an open neighborhood of 0 in $\mathbb{R}^d$, with $p = \Phi(0)$ and

$$\begin{aligned} P \cap \Phi(U) &= \Phi\left(\{x \in U : x_k > 0, x_{k+1} = \cdots = x_d = 0\}\right), \\ \partial P \cap \Phi(U) &= \Phi\left(\{x \in U : x_k = 0, x_{k+1} = \cdots = x_d = 0\}\right). \end{aligned} \tag{B.18}$$

This says that in a small neighborhood of p, M can be parametrized so that P is a half subspace in that chart and ∂P is the boundary of this subspace. For example, the interior of a triangle is a submanifold of $\mathbb{R}^2$ (since it is an open subset) and its boundary is the triangle itself. The vertices of the triangle are not regular boundary points, however.

Denote by $\partial' P$ the regular boundary of P (the sets of regular boundary points). Then $\partial' P$ is a submanifold of M, with dimension $k - 1$, using the charts that were just described to provide an atlas. We will say that P is a submanifold of M with regular boundary if $\partial' P = \partial P$. One can extend tangent spaces on P to regular boundary points by simply letting $T_p P = \text{span}(\partial_1, \ldots, \partial_k) \subset T_p M$ if $p \in \partial' P$ and (U, Φ) is as above. Obviously $T_p \partial' P = \text{span}(\partial_1, \ldots, \partial_{k-1})$.

Note that one can define an abstract notion of "manifold with boundary" [175] by letting local charts be open subsets of closed half spaces. Using this terminology P is a manifold without boundary, and $P \cup \partial' P$ is a manifold with boundary. We will however not need this concept since all the manifolds we work with can be easily represented as submanifolds of Euclidean spaces.

Orientation

Let M be a d-dimensional manifold. One says that an atlas $\mathcal{A} = ((U_i, \Phi_i), i \in \mathcal{I})$ on M is oriented if and only if the changes of coordinates $\Phi_j^{-1} \circ \Phi_i$ have positive Jacobian determinant whenever they are defined over non-empty sets (i.e., whenever the charts overlap). Two atlases $\mathcal{A}$ and $\mathcal{A}'$ define the same orientation if and only if their union is oriented. The manifold is orientable if an oriented atlas exists. If M is orientable, one can fix an orientation of M by specifying an oriented atlas, say $\mathcal{A}_0$, and one then says that M is oriented. In this case, any new local chart (U, Φ) is said to be positively oriented if $\mathcal{A}_0 \cap \{(U, \Phi)\}$ is oriented (i.e., (U, Φ) has the same orientation as the charts that compose the atlas) and negatively oriented otherwise.

An orientation of M immediately provides an orientation of $T_p M$ for $p \in M$, simply by letting $(\partial_1, \ldots, \partial_d)$ be a positively oriented basis of $T_p M$ as soon as (U, Φ) is a positively oriented local chart at p. An orientation of M can therefore be

seen as a continuous mapping $\varepsilon : FM \to \{-1, 1\}$, where FM is the frame bundle of M, associating 1 or -1 to a basis of T_pM depending on whether this basis is positively oriented or not, and M is orientable if such a continuous mapping exists.

Assume now that P is an oriented submanifold of M. One can then provide $\partial' P$ with a natural orientation, which will always be implicitly assumed in the rest of this discussion. If (U, Φ) is a boundary chart around $p \in \partial' P$, we will say that a basis $(\xi_1, \ldots, \xi_{k-1})$ of $T_p(\partial' P)$ is positively oriented if and only if $(\xi_1, \ldots, \xi_{k-1}, \partial_k)$ is positively oriented in P. This definition is independent of the chosen chart: assume a change of chart, say, $(x_1, \ldots, x_k) \to (y_1, \ldots, y_k)$. Let ∂'_k denote the element of the basis of T_pP associated with the new chart (replacing ∂_k). Then

$$\partial'_k = (\partial_1 y_k)\partial_1 + \cdots + (\partial_k y_k)\partial_k.$$

Because $\partial_1, \ldots, \partial_{k-1} \in T_p\partial' P$, one can also write

$$\partial'_k = \alpha_1\xi_1 + \cdots + \alpha_{k-1}\xi_{k-1} + (\partial_k y_k)\partial_k$$

for some coefficients $\alpha_1, \ldots, \alpha_{k-1}$. This implies that the determinant of the family $(\xi_1, \ldots, \xi_{k-1}, \partial'_k)$ with respect to $(\xi_1, \ldots, \xi_{k-1}, \partial_k)$ is equal to $\partial_k y_k$. However, we have $y_k(p) = x_k(p) = 0$, and we know that $x_k > 0$ implies $y_k > 0$. This is only possible if $\partial_k y_k > 0$ so that the two charts give the same orientation on $\partial' P$.

Taking, for example, $M = \mathbb{R}^2$, P an open subset of $\mathbb{R}^2$ such that ∂P is a smooth Jordan curve, parametrized as $\gamma(t)$, the orientation of the boundary should be such that $\det(\dot\gamma(t), N) > 0$ for the inward normal N. This is consistent with our definition of a positive orientation of a closed curve in Definition 1.14.

Note that using the last coordinate, x_k, as the one defining the boundary differs from the choice often made in the literature (which is x_1), and will result in a $(-1)^d$ factor in Stokes's formula below. This choice however ensures that the inward normal is always equal to the cross product between basis vectors in positively oriented charts of the boundary, which is why we preferred it.

Integration of d-Forms on M

We assume from now on that M is oriented and any local chart on M will be assumed to be positively oriented unless specified otherwise.

Let ω be a d-form on M. Let S be a chartable set in M, i.e., an open set in M that can be completely covered by a chart, so that there exists a (positively oriented) local chart (U, Φ) such that $S = \Phi(U)$. We say that ω is integrable on S if

$$\int_S |\omega| := \int_U |(\omega \mid \partial_1, \ldots, \partial_d)| \circ \Phi \, dx_1 \ldots dx_d < \infty$$

and one then defines the integral of ω on S by

$$\int_S \omega = \int_U (\omega \mid \partial_1, \ldots, \partial_d) \circ \Phi \, dx_1 \ldots dx_d.$$

These definitions are parametrization-independent. Assume that $(\tilde{U}, \tilde{\Phi})$ is another positively oriented local chart satisfying $\tilde{\Phi}(\tilde{U}) = S$. Let $\psi = \tilde{\Phi}^{-1} \circ \Phi : U \to \tilde{U}$ (a diffeomorphism). Let $\tilde{\partial}_1, \ldots, \tilde{\partial}_d$ be the coordinate vector fields associated with the new chart. Then, using the identity

$$(\tilde{\partial}_1, \ldots, \tilde{\partial}_d)_p = d\psi(x)(\partial_1, \ldots, \partial_d)_p,$$

we have

$$\begin{aligned}
&\int_{\tilde{U}} \left(\omega \,\middle|\, \tilde{\partial}_1, \ldots, \tilde{\partial}_d\right) \circ \tilde{\Phi}\, d\tilde{x}_1 \ldots d\tilde{x}_d \\
&= \int_{\tilde{U}} \det(d\psi) \circ \psi^{-1}\, (\omega \mid \partial_1, \ldots, \partial_d) \circ \Phi \circ \psi^{-1} d\tilde{x}_1 \ldots d\tilde{x}_d \\
&= \int_U (\omega \mid \partial_1, \ldots, \partial_d) \circ \Phi\, dx_1 \ldots dx_d,
\end{aligned}$$

where the last identity uses the change of variable formula and the fact that $\det(d\psi) > 0$. This shows that our formula for the integral of ω over S did not depend on the chosen chart. The same conclusion holds for the integral of $|\omega|$, which does not require the positivity assumption.

One can define the integral of $|\omega|$ or ω on M by covering M by a locally finite family of chartable sets $(S_i, i \in I)$ and using a partition of the identity, which is a set of smooth functions $(\alpha_i, i \in I)$ satisfying

$$\begin{cases}
\alpha_i(p) \geq 0, & \forall i \in I, \forall p \in M \\
\alpha_i(p) = 0, & \forall i \in I, \forall p \in M \setminus S_i \\
\displaystyle\sum_{i \in I} \alpha_i(p) = 1, & \forall p \in M.
\end{cases}$$

(One says that $(S_i, i \in I)$ is locally finite if, given any $p \in M$, there exists only a finite number of indices i such that $p \in S_i$. Under this assumption, the sum in the third equation only involves a finite number of terms. One can show that locally finite chartable sets and associated partitions of unity always exist on M.)

One then says that ω is integrable over M if

$$\int_M |\omega| := \sum_{i \in I} \int_{S_i} \alpha_i |\omega| < \infty$$

and one defines

$$\int_M \omega = \sum_{i \in I} \int_{S_i} \alpha_i \omega.$$

These expressions do not depend on the family $(S_i, \alpha_i, i \in I)$. Assuming a second family, $(T_j, \beta_j, j \in J)$, one only needs to observe that integrals over chartable sets are obviously additive, so that

$$\sum_{i\in I}\int_{S_i}\alpha_i\omega = \sum_{i\in I}\sum_{j\in J}\int_{S_i\cap T_j}\alpha_i\beta_j\omega = \sum_{j\in J}\int_{T_j}\beta_j\omega.$$

A d-form ω on M is called a volume form if $\omega(p) \not\equiv 0$ for all $p \in M$. If ω is such a form and ω' another d-form, there exists a function $\alpha : M \to \mathbb{R}$ such that $\omega' = \alpha\omega$. If ω' is itself a volume form, then α never vanishes.

Change of Variables

The change of variables formula is remarkably simple when expressed with differential forms. Let $g : M \to M'$ be a C^1 function between two d-dimensional manifolds M and M'. Let ω be a d-form on M'. Then

$$\int_M g^*\omega = \int_{M'}\omega.$$

Canonical Volume form on a Riemannian Manifold

The volume of a d-dimensional parallelogram with edges $u_1, \ldots, u_d \in \mathbb{R}^d$,

$$U = \{\alpha_1 u_1 + \cdots + \alpha_d u_d, 0 \le \alpha_1, \ldots, \alpha_d \le 1\},$$

is equal to $|\det(u_1, \ldots, u_d)|$. If A is the d by d matrix with columns $u_1, \ldots, u_d$, then $A^T A$ is the Gram matrix with coefficients $u_i^T u_j$ and $|\det A| = \sqrt{\det(A^T A)}$. On a Riemannian manifold M, the infinitesimal volume in a local chart is defined by

$$d\sigma_M = \sqrt{\det S(x)}\, dx_1 \ldots dx_d$$

where d is the dimension of M and $s(ij)(x) = \left\langle \partial_{x_i}\,,\, \partial_{x_j}\right\rangle_x$. This directly provides the integral of functions $f : M \to \mathbb{R}$ which are supported by the image of a local chart (U, Φ) using

$$\int_M f d\sigma_M = \int_U f\circ\Phi(x)\sqrt{\det S(x)}\, dx_1 \ldots dx_d$$

and one passes to integrals of arbitrary functions on M using partitions of unity.

If M is oriented, the Riemannian volume form is defined on a positively oriented chart (U, Φ) by

$$\omega_M = \sqrt{\det S(x)} dx_1 \wedge \cdots \wedge dx_d$$

so that one can also write

$$\int_M f d\sigma_M = \int_M f\omega_M.$$

More generally, for any $\xi_1, \ldots, \xi_d \in T_pM$, one has

$$\omega_M(\xi_1, \ldots, \xi_d) = \varepsilon(\xi_1, \ldots, \xi_p)\sqrt{\det\left(\langle \xi_i\ ,\ \xi_j \rangle_p\right)}, \tag{B.19}$$

where $\varepsilon(\xi_1, \ldots, \xi_p)$ is the orientation of the family.

B.7.4 Stokes's Theorem

Stokes's theorem expresses that the integral of the exterior derivative of a differential form over a bounded domain is equal to the integral of this form on the boundary of the domain. This generalizes the fundamental theorem of calculus (where the domain is simply a bounded interval) and implies the divergence theorem and other important results in vector calculus. We state the following theorem without proof (see, e.g., [175]).

Theorem B.28 (Stokes's Theorem) *Let P be an oriented k-dimensional submanifold of a manifold M such that P has a regular boundary. Let ω be a C^1 differential $(k-1)$-form on M. Then*

$$\int_{\partial P} \omega = (-1)^k \int_P d\omega.$$

The requirement that ∂P is regular can be partially lifted to allow for singularities on "negligible subsets" of the boundary (see [175], in which it is shown that negligible sets can be defined as finite unions of images of closed rectangles of dimension $\leq k-2$).

Divergence and the Kelvin–Stokes Theorems on $\mathbb{R}^d$

We now review a few applications of this formula. Let $M = \mathbb{R}^d$ and P an open subset of $\mathbb{R}^d$ such that ∂P is regular and therefore a smooth submanifold of $\mathbb{R}^d$. Let ω be a C^1 $(d-1)$-form on $\mathbb{R}^d$ so that ω can be written as

$$\omega = \sum_{k=1}^{d} (-1)^{k-1} \alpha_k dx_1 \wedge \cdots \wedge dx_{k-1} \wedge dx_{k+1} \wedge \cdots \wedge dx_d, \tag{B.20}$$

where $\alpha_1, \ldots, \alpha_d$ are C^1 functions on $\mathbb{R}^d$ and $x_1, \ldots, x_d$ are the Euclidean coordinates on $\mathbb{R}^d$. In particular, letting $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, so that α is a vector field, we have

$$\omega(\xi_1, \ldots, \xi_{d-1}) = \det(\alpha, \xi_1, \ldots, \xi_{d-1}).$$

Moreover, we have $d\omega = \operatorname{div}(\alpha)\, dx_1 \wedge \cdots \wedge dx_d$.

There exists a unique vector $u = u(\xi_1, \ldots, \xi_{d-1})$ such that, for any vector $\beta \in \mathbb{R}^d$,

$$\det(\xi_1, \ldots, \xi_{d-1}, \beta) = \beta^T u$$

(because the right-hand side is a linear form as a function of β). This vector is the cross product of $\xi_1, \ldots, \xi_{d-1}$, denoted $\xi_1 \times \cdots \times \xi_{d-1}$ if $d > 2$ and $\xi_1^\perp$ if $d = 2$. If $\xi_1, \ldots, \xi_{d-1}$ are linearly dependent, then obviously $\xi_1 \times \cdots \times \xi_{d-1} = 0$. Otherwise the cross-product is perpendicular to $\xi_1, \ldots, \xi_{d-1}$ and completes it into a positively oriented basis of $\mathbb{R}^d$. Moreover, one has

$$|\xi_1 \times \cdots \times \xi_{d-1}| = \sqrt{\det(\mathrm{Gram}(\xi_1, \ldots, \xi_{d-1}))},$$

where $\mathrm{Gram}(\xi_1, \ldots, \xi_{d-1})$ is the matrix with entries $\xi_i^T \xi_j$. To prove this identity, first make a change of positively oriented orthonormal basis in $\mathbb{R}^d$ to reduce to the case in which $\xi_1, \ldots, \xi_{d-1}$ are perpendicular to the last basis vector, say e_d, and write

$$\det(\xi_1, \ldots, \xi_{d-1}, e_d) = e_d^T(\xi_1 \times \cdots \times \xi_{d-1}) = |\xi_1 \times \cdots \times \xi_{d-1}|,$$

so that the norm of the cross product is equal to the determinant of the coefficients $\xi_1, \ldots, \xi_{d-1}$ in the first $d-1$ vectors of the basis, and is therefore also equal to the square root of the Gram matrix.

Now, if $p \in \partial P$, then, in a chart (U, Φ)

$$N(p) = \frac{\partial_1 \times \cdots \times \partial_{d-1}}{|\partial_1 \times \cdots \times \partial_{d-1}|}$$

is the positively oriented unit normal at p to ∂P and

$$\begin{aligned}\omega(\partial_1, \ldots, \partial_{d-1})\, dx_1 \ldots dx_{d-1} &= (-1)^{d-1} \alpha^T(\partial_1 \times \cdots \times \partial_{d-1})\, dx_1 \ldots dx_{d-1} \\ &= (-1)^{d-1} \alpha^T N d\sigma_{\partial P}.\end{aligned}$$

Using this identity, we can write Stokes's theorem in this case as

$$\int_{\partial P} \alpha^T N d\sigma_{\partial P} = -\int_P \mathrm{div}\alpha\, dx_1 \ldots dx_d, \tag{B.21}$$

which is the divergence theorem (N being the inward normal).

Here is another application of Stokes's theorem. Let P be a bounded surface in $\mathbb{R}^3$ with a regular boundary. Let ω be a one-form on $\mathbb{R}^3$, so that

$$\omega = \alpha_1 dx_1 + \alpha_2 dx_2 + \alpha_3 dx_3$$

for some vector field α on $\mathbb{R}^3$. Then $d\omega$ is a two-form with

$$d\omega = \beta_1 dx_2 \wedge dx_3 - \beta_2 dx_1 \wedge dx_3 + \beta_3 dx_1 \wedge dx_2$$

such that $\beta = \mathrm{curl}(\alpha)$ is formally defined by

$$\mathrm{curl}(\alpha) = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} \times \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Moreover, one has (see the discussion above)

$$d\omega(\xi_1, \xi_2) = \det(\beta, \xi_1, \xi_2) = \det(\xi_1, \xi_2, \beta) = \beta^T(\xi_1 \times \xi_2).$$

We therefore have

$$\int_P d\omega = \int_P \beta^T N d\sigma_P.$$

Here N is the unit normal to P defined by $N = (\partial_1 \times \partial_2)/|\partial_1 \times \partial_2|$ in any positively oriented chart of P. Let T be the positively oriented unit tangent on ∂P. Then we have

$$\int_{\partial P} \omega = \int_{\partial P} \alpha^T T d\sigma_{\partial P}$$

and we conclude from Stokes's theorem (with $d = 2$) that

$$\int_P (\mathrm{curl}(\alpha))^T N d\sigma_P = \int_{\partial P} \alpha^T T d\sigma_{\partial P}.$$

This is the Kelvin–Stokes theorem.

The Divergence Theorem on a Riemannian Manifold

If M is a d-dimensional Riemannian manifold, and ω_M its volume form, one can associate to any vector field v the $(d-1)$-form $I_v\omega_M$

$$(\xi_1, \ldots, \xi_{d-1}) \mapsto \omega_M(v, \xi_1, \ldots, \xi_{d-1})$$

(I_v is the insertion operator that transforms any k form into a $k-1$ form by replacing the first vector with v.) The exterior derivative $d(I_v\omega_M)$ is then a d-form, necessarily proportional to ω_M. The proportionality coefficient is called the divergence of v, $\mathrm{div}_M v$, therefore defined by

$$d(I_v\omega_M) = (\mathrm{div}_M v)\, \omega_M.$$

The divergence theorem on M is obtained by applying Stokes's theorem to $I_v\omega_M$: If P is an open subset of M with a regular boundary, then

$$\int_{\partial P} I_v\omega_M = (-1)^d \int_P (\mathrm{div}_M v)\, d\sigma_M.$$

In a local chart (U, Φ) such that ∂P corresponds to $x_d = 0$ and P to $x_d > 0$, $T_p\partial P$ is generated by $\partial_{x_1}, \ldots, \partial_{x_{d-1}}$, so that

$$\int_{\Phi(U)\cap\partial P} I_v\omega_M = \int_{U\cap\{x_d=0\}} \omega_M(v, \partial_1, \ldots, \partial_{d-1})dx_1 \ldots dx_{d-1}.$$

Defining $\partial_1 \times \cdots \times \partial_{d-1}$ by

$$\omega_M(\partial_1, \ldots, \partial_{d-1}, w) = \left\langle \partial_1 \times \cdots \times \partial_{d-1}\,,\; w \right\rangle_p$$

one defines the inward normal to ∂P by

$$\nu = \frac{\partial_1 \times \cdots \times \partial_{d-1}}{|\partial_1 \times \cdots \times \partial_{d-1}|_p}.$$

Noticing that (using (B.19))

$$|\partial_1 \times \cdots \times \partial_{d-1}|_p = \omega_M(\partial_1, \ldots, \partial_{d-1}, \nu) = \omega_{\partial P}(\partial_1, \ldots, \partial_{d-1}),$$

we can conclude that

$$\int_{\partial P} (I_v\omega_M) = (-1)^{d-1} \int_{\partial P} \left\langle v\,,\; \nu \right\rangle_p d\sigma_{\partial P},$$

yielding

$$\int_{\partial P} \left\langle v\,,\; \nu \right\rangle_p d\sigma_{\partial P} = -\int_P (\mathrm{div}_M v)\, d\sigma_M. \tag{B.22}$$

Manifold Evolution

Let ω be a C^1 k-form in $\mathbb{R}^d$ and M a k-dimensional oriented manifold, with a regular boundary ∂M. Let $m : (t, p) \mapsto m(t, p)$ be a C^1 mapping from $[0, 1] \times M$ to $\mathbb{R}^d$, and let $m_t : p \mapsto m(t, p)$. Define

$$F(t) = \int_M m_t^*\omega.$$

When m_t is an embedding, one has (letting $M_t = m_t(M)$)

$$F(t) = \int_{M_t} \omega$$

and we are studying the evolution of the integral of a fixed form over a moving submanifold of $\mathbb{R}^d$.

Theorem B.29 *One has*

$$\partial_t F(t) = (-1)^k \int_{\partial M} m_t^* I_{\partial_t m} \omega + \int_M m_t^* I_{\partial_t m} d\omega.$$

When the m_t's are embeddings, this reduces to

$$\partial_t F(t) = (-1)^k \int_{\partial M_t} I_{\partial_t m} \omega + \int_{M_t} I_{\partial_t m} d\omega.$$

Proof Consider the manifold $\hat{M} = (t, t+\varepsilon) \times M$. One can define local charts $(\hat{U}, \hat{\Phi})$ on $\hat{M}$ from local charts (U, Φ) on M by $\hat{U} = (t, t+\varepsilon) \times U$ and $\hat{\Phi}(u, x) = (u, \Phi(x))$. We orient $\hat{M}$ so that $(\hat{U}, \hat{\Phi})$ is positive as soon as (U, Φ) is.

Then $\partial\hat{M}$ has a regular boundary given by the union of the sets $\{t\} \times M$, $\{t+\varepsilon\} \times M$ and $(t, t+\varepsilon) \times \partial M$. The singular part is $(\{t\} \times \partial M) \cup (\{t+\varepsilon\} \times \partial M)$ and can be neglected when applying Stokes's theorem. If (U, Φ) is a positive chart on M, one can take the boundary charts $(\tilde{U}, \tilde{\Phi})$ on $\{t\} \times M$ and $(\tilde{U}, \tilde{\Phi}')$ on $\{t+\varepsilon\} \times M$ by letting $\tilde{U} = U \times (-\delta, \delta)$, $\tilde{\Phi}(x, u) = (t+u, \Phi(x))$ and $\tilde{\Phi}'(x, u) = (t+\varepsilon-u, \Phi(x))$. Note that these two charts have reversed orientations, which depend on the parity of k: $(-1)^k$ for the former and $(-1)^{k+1}$ for the latter. On $(t, t+\varepsilon) \times \partial M$, starting from a positively oriented boundary chart (U, Φ) of ∂M, the chart $(\hat{U}, \hat{\Phi})$ is also a boundary chart of $\hat{M}$ and is positively oriented.

Stokes's theorem applied to $\hat{M}$ implies that

$$\int_{\partial\hat{M}} m^*\omega = (-1)^{k+1} \int_{\hat{M}} d(m^*\omega) = (-1)^{k+1} \int_{\hat{M}} m^* d\omega.$$

We have

$$\int_{\partial\hat{M}} m^*\omega = (-1)^{k+1} F(t+\varepsilon) + (-1)^k F(t) + \int_t^{t+\varepsilon} \int_{\partial M} I_{\partial_t} m^*\omega$$

and

$$\int_{\hat{M}} m^* d\omega = \int_t^{t+\varepsilon} \int_M I_{\partial_t} m^* d\omega = \int_t^{t+\varepsilon} \int_M m_t^* I_{\partial_t m} d\omega,$$

so that

$$F(t+\varepsilon) - F(t) = (-1)^k \int_t^{t+\varepsilon} \int_{\partial M} I_{\partial_t} m^*\omega + \int_t^{t+\varepsilon} \int_M m_t^* I_{\partial_t m} d\omega,$$

and the conclusion follows. □

Consider the case $k = d - 1$ and ω given by (B.20). Assume that the m_t's are embeddings. Then, we have

$$F(t) = (-1)^{d-1} \int_{M_t} \alpha^T N \, d\sigma_{M_t},$$

where N is the unit normal to M_t. We also have

$$d\omega(\partial_t m, \partial_1, \ldots, \partial_{d-1}) = (-1)^{d-1} \mathrm{div}\alpha \, (\partial_t m)^T (\partial_1 \times \cdots \times \partial_{d-1})$$

and

$$\int_{M_t} I_{\partial_t m} d\omega = \int_{M_t} \mathrm{div}\alpha \, (\partial_t m)^T N d\sigma_{M_t}.$$

Moreover,

$$I_{\partial_t m}\omega(\xi_1, \ldots, \xi_{d-2}) = \det(\alpha, \partial_t m, \xi_1, \ldots, \xi_{d-2}).$$

If $d = 2$ we have

$$\int_{\partial M_t} I_{\partial_t m}\omega = \sum_{p \in \partial M_t} \delta(p) \det(\alpha(p), \partial_t m),$$

where $\delta(p) = 1$ if the oriented arc that contains p has p as a starting point, and $\delta(p) = -1$ if p is an end point.

For $d > 2$, let $n(p) \in T_p M_t$ be the inward-pointing unit normal to ∂M_t ($p \in \partial M_t$). If $\xi_1, \ldots, \xi_{d-2}$ is a positively oriented basis of $T_{\partial M_t}$, then $(n, N, \xi_1, \ldots, \xi_{d-2})$ is a positive basis of $\mathbb{R}^d$ and

$$\begin{aligned} &\det(\alpha, \partial_t m, \xi_1, \ldots, \xi_{d-2}) \\ &\quad = ((\alpha^T n)(\partial_t m^T N) - (\alpha^T N)(\partial_t m^T n)) \det(n, N, \xi_1, \ldots, \xi_{d-2}). \end{aligned}$$

Note that

$$\det(n, N, \xi_1, \ldots, \xi_{d-2}) = \omega_{\partial M}(\xi_1, \ldots, \xi_{d-2})$$

(the volume form on ∂M) as soon as $(\xi_1, \ldots, \xi_{d-2})$ is positively oriented. Therefore

$$\int_{\partial M_t} I_{\partial_t m}\omega = \int_{\partial M_t} (\alpha^T n \, \partial_t m^T N - \alpha^T N \, \partial_t m^T n) d\sigma_{\partial M_t}.$$

Collecting all terms in Theorem B.29, we obtain the following corollary.

Corollary B.30 *Assume that $m(t, \cdot)$ is C^1 from $[0, 1] \times M$ to $\mathbb{R}^d$, where M is a $(d - 1)$-dimensional manifold with regular boundary, such that $m(t, \cdot)$ is an embedding for all t. Let α be a C^1 vector field on $\mathbb{R}^d$. Then, if $d > 2$,*

$$\partial_t \int_{M_t} \alpha^T N d\sigma_{M_t} = \int_{\partial M_t} ((\alpha^T n)\,(\partial_t m^T N) - (\alpha^T N)\,(\partial_t m^T n)) d\sigma_{\partial M_t} \qquad \text{(B.23)}$$
$$+ \int_{M_t} \mathrm{div}\alpha\,(\partial_t m)^T N d\sigma_{M_t},$$

where $n(t,\cdot)$ is the inward normal to ∂M_t and $N(t,\cdot)$ is the positively oriented normal to M_t. If $d = 2$, the boundary term reduces to

$$\sum_{p\in\partial M_t} \delta(p)\,\det(\alpha(p), \partial_t m),$$

where $\delta(p) = 1$ if the oriented arc that contains p has p as a starting point, and $\delta(p) = -1$ if p is an end point.

Appendix C
Ordinary Differential Equations

In the first part of this chapter, we review basic existence theorems and properties of ordinary differential equations (ODEs) on Banach spaces. The presentation will follow those provided in standard textbooks [69], although we will work with slightly relaxed regularity conditions. The second part will provide a partial overview of some of the most important numerical methods designed to solve ODEs.

We start with some definitions about differentials in infinite dimensions.

C.1 Calculus in Banach Space

A domain Ω in a Banach space $\mathbb{B}$ is open if, for every point $x \in \Omega$, there exists an $\varepsilon(x) > 0$ such that the open ball $B(x, \varepsilon(x))$ is included in Ω, with

$$B(x, \varepsilon(x)) = \{y \in \mathbb{B} : \|y - x\|_{\mathbb{B}} < \varepsilon(x)\}.$$

A set F is closed in $\mathbb{B}$ if $\mathbb{B} - F$ is open. The closure of a set A (denoted $\text{clos}(A)$ or $\bar{A}$) is the smallest closed set that contains it and its interior ($\text{int}(A)$ or $\mathring{A}$) is the largest open set included in it. Its boundary is $\partial A = \bar{A} \setminus \mathring{A}$.

A function f defined on a subset Ω' of a Banach set $\mathbb{B}'$ with values in $\mathbb{B}$ is continuous if $\|f(x_n) - f(x)\|_{\mathbb{B}} \to 0$ whenever (x_n) is a sequence in Ω' such that $\|x_n - x\|_{\mathbb{B}'} \to 0$ for some $x \in \Omega'$. The function f is Lipschitz if there exists a constant c such that

$$\|f(x) - f(y)\|_{\mathbb{B}} \leq c\|x - y\|_{\mathbb{B}'}$$

for all $x, y \in \Omega'$. The smallest c for which this is true is called the Lipschitz constant of f and denoted $\text{Lip}(f)$.

A function $f : \mathbb{R} \to \mathbb{B}$ is differentiable at $t \in \mathbb{R}$ if there exists an element of $\mathbb{B}$, denoted $\partial f(t)$ or $\dot{f}(t)$, such that

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

$$\lim_{\varepsilon\to 0}\left\|\frac{1}{\varepsilon}(f(t+\varepsilon)-f(t))-\partial f(t)\right\|_{\mathbb{B}}=0.$$

If $\mathbb{B}$ and $\mathbb{B}'$ are Banach spaces and $f:\mathbb{B}'\to\mathbb{B}$, we say that f is *Gâteaux differentiable* at $x\in\mathbb{B}'$ if, for all $h\in\mathbb{B}'$, $t\mapsto f(x+th)$ is differentiable at $t=0$ and there exists a bounded linear transformation, denoted $df(x):\mathbb{B}'\to\mathbb{B}$, such that, for all $h\in\mathbb{B}'$:

$$\partial_t f(x+th)(0)=df(x)h.$$

Then, $df(x)$ is called the Gâteaux derivative of f at x.

We say that f is differentiable in the Fréchet sense, or *Fréchet differentiable* at $x\in\mathbb{B}'$, if there exists a bounded linear transformation $df(x):\mathbb{B}'\to\mathbb{B}$ such that, for any $h\in\mathbb{B}'$,

$$\lim_{\|h\|_{\mathbb{B}'}\to 0}\frac{1}{\|h\|_{\mathbb{B}'}}\|f(x+h)-f(h)-df(x)h\|_{\mathbb{B}}=0;$$

$df(x)$ is called the Fréchet derivative of f at x.

There is no ambiguity in the notation, because it is clear that if f is Fréchet differentiable, then it is also Gâteaux differentiable, and the two definitions of $df(x)$ coincide. When writing that f is differentiable without qualifier, we will always mean Fréchet differentiable.

If f is differentiable over $\Omega'\subset\mathbb{B}'$, then df maps Ω' to the space $\mathcal{L}(\mathbb{B}',\mathbb{B})$ of bounded linear functionals from $\mathbb{B}'$ to $\mathbb{B}$, which is a Banach space for the operator norm

$$\|A\|_{op(\mathbb{B}',\mathbb{B})}=\sup\left\{\|Ah\|_{\mathbb{B}}:h\in\mathbb{B}',\|h\|_{\mathbb{B}'}=1\right\}.$$

(We will write $\|\cdot\|_{op(\mathbb{B})}$ when $\mathbb{B}=\mathbb{B}'$.) The second derivative can therefore be defined, and, when it exists, maps Ω' to $\mathcal{L}(\mathbb{B}',\mathcal{L}(\mathbb{B}',\mathbb{B}))$. This space is usually identified with $\mathcal{L}^2(\mathbb{B}',\mathbb{B})$, the Banach space of bilinear forms from $\mathbb{B}'\times\mathbb{B}'$ to $\mathbb{B}$, and one uses the notation, for the second derivative

$$d^2f(x)(h_1,h_2)=(d^2f(x)h_1)h_2.$$

Higher order derivatives are defined similarly, so that $d^pf(x)\in\mathcal{L}^p(\mathbb{B}',\mathbb{B})$ is p-linear. The space of p-time continuously differentiable maps from Ω' to $\mathbb{B}$ will be denoted $C^p(\Omega',\mathbb{B})$ and differentials of order p or less of functions in that space are symmetric linear forms. Similar to finite dimensions, we let $C_b^p(\Omega',\mathbb{B})$ denote the set of functions $f\in C^p(\Omega',\mathbb{B})$ with bounded derivatives up to order p. This space is a Banach space for

$$\|f\|_{p,\infty}=\max_{0\le k\le p}\|d^kf\|_\infty.$$

We will also denote by $C_0^p(\Omega',\mathbb{B})$ the Banach space of functions $f\in C^p(\Omega',\mathbb{B})$ that are restrictions of functions $\tilde{f}\in C_b^p(\mathbb{B}',\mathbb{B})$ that vanish on $(\Omega')^c\cup\{\infty\}$. This space coincides with the space of functions $f\in C_b^p(\Omega',\mathbb{B})$ such that for any sequence

$x_n \in \Omega'$ such that $\|x_n\|_{\mathbb{B}'} \to \infty$ or $\|x_n - x\|_{\mathbb{B}'} \to 0$ with $x \notin \Omega'$, one has $d^k f(x_n) \to 0$ for all $k = 0, \ldots, p$.

The following standard theorems are used at several points in this book. Their proofs can be found in many textbooks, such as [175].

Theorem C.1 (Banach fixed-point theorem) *Let $\mathbb{B}$ be a Banach space. Let $U \subset \mathbb{B}$ be a closed set and Φ be a contraction of U, i.e., $\Phi : U \to U$ is such that there exists a constant $c \in [0, 1)$ satisfying, for all $x, y \in U$,*

$$\|\Phi(x) - \Phi(y)\|_{\mathbb{B}} \leq c \|x - y\|_{\mathbb{B}} .$$

Then, Φ has a unique fixed point in U, i.e., there exists a unique $x_0 \in U$ such that $\Phi(x_0) = x_0$.

Theorem C.2 (Inverse mapping theorem) *Assume that Ω is an open subset of the Banach space $\mathbb{B}$ and that $\Psi : \Omega \to \mathbb{B}$ is continuously differentiable, such that $d\psi(x_0)$ is invertible for some $x_0 \in \Omega$. Then there exists an open neighborhood $\tilde{\Omega} \subset \Omega$ of x_0 such that the restriction of Ψ to $\tilde{\Omega}$ is a diffeomorphism between $\tilde{\Omega}$ and $\Psi(\tilde{\Omega})$ with*

$$d(\Psi^{-1})(y) = (d\Psi(\Psi^{-1}(y))^{-1}$$

for all $y \in \Psi(\tilde{\Omega})$.

Theorem C.3 (Implicit mapping theorem) *Let Ω and Ω' be open subsets of Banach spaces $\mathbb{B}$ and $\mathbb{B}'$ and $\Phi : \Omega \times \Omega' \to \mathbb{B}'$ be continuously differentiable. Assume that $(x_0, y_0) \in \Omega \times \Omega'$ is such that $\Phi(x_0, y_0) = 0$ and $d_2\Phi(x_0, y_0)$ is invertible. Then there exist open neighborhoods $\tilde{\Omega} \subset \Omega$ of x_0 and $\tilde{\Omega}' \subset \Omega'$ of y_0 and a continuously differentiable function $F : \tilde{\Omega}' \to \tilde{\Omega}$ such that, for all $y \in \tilde{\Omega}'$, $x = F(y)$ is the unique solution in $\tilde{\Omega}$ of the equation $\Phi(\tilde{x}, y) = 0$ and*

$$dF(y) = -d_2\Phi(x, y)^{-1} d_1\Phi(x, y).$$

(Here, d_1 and d_2 denote the differentials with respect to the first and second variables.)

C.2 The Bochner Integral

The integral of a Banach-valued function $f : [a, b] \to \mathbb{B}$ can be defined, as in finite dimensions, by the limit of the integral of sums of simple functions (i.e., constant over a finite number of measurable subsets of $[a, b]$)). More precisely, a function $g : [a, b] \to \mathbb{B}$ is simple if and only if it is measurable and takes a finite number of different values, and one then defines

$$\int_a^b g(t)\, dt = \sum_{x \in \text{Range}(g)} x |g^{-1}(x)|,$$

where $|g^{-1}(x)|$ is the Lebesgue measure of the set of $t \in [a, b]$ such that $g(t) = x$, the sum being, by assumption, over a finite set.

One then says that f is Bochner integrable (see [38, 83, 84, 202, 306]) if and only if there exists a sequence f_n of simple functions such that

$$\lim_{n\to\infty} \int_a^b \|f(t) - f_n(t)\|_{\mathbb{B}}\, dt = 0$$

and the Bochner integral of f is then defined by

$$\int_a^b f(t)\, dt = \lim_{n\to\infty} \int_a^b f_n(t)\, dt,$$

the right-hand term being well defined and independent of the choice of the approximating sequence of simple functions. A Bochner integrable function is automatically measurable.

The following result characterizes Bochner integrable functions and greatly simplifies their analysis.

Theorem C.4 (Bochner) *A measurable function $f : [a, b] \to \mathbb{B}$ is Bochner integrable if and only if $\|f\|_{\mathbb{B}} : [a, b] \to \mathbb{R}$ is Lebesgue integrable. Moreover, one has*

$$\left\| \int_a^b f(t)\, dt \right\|_{\mathbb{B}} \leq \int_a^b \|f(t)\|_{\mathbb{B}}\, dt.$$

The space of Bochner integrable functions from an interval $[a, b]$ to $\mathbb{B}$ will be denoted $L^1([a, b], \mathbb{B})$. Functions that coincide on sets of Lebesgue measure 0 are identified in this space.

The dominated convergence theorem is true for Bochner integrals, and states that, if f_n is a sequence of measurable function in $\mathbb{B}$ that converges strongly to a limit f, and if there exists a Lebesgue-integrable function $g : [a, b] \to \mathbb{R}$ such that $\|f_n(t)\|_{\mathbb{B}} \leq g(t)$ for all $t \in [a, b]$, then

$$\lim_{n\to\infty} \int_a^b f_n(t)\, dt = \int_a^b f(t)\, dt.$$

Lebesgue's theorem also holds, so that, if f is Bochner integrable, then

$$\lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} f(u)\, du = f(t)$$

almost everywhere on $[a, b]$. Moreover, if f is continuously differentiable,

$$f(t) - f(a) = \int_a^t \partial f(t)\, dt.$$

The Bochner integral commutes with bounded linear operators, in the sense that, if $A : \mathbb{B} \to \mathbb{B}'$ is such an operator between Banach spaces $\mathbb{B}$ and $\mathbb{B}'$ and if $f : [a, b] \to \mathbb{B}$ is Bochner integrable, then so is Af with

$$\int_a^b Af(t)\,dt = A\int_a^b f(t)\,dt.$$

We will use this integral in the proof of the next proposition, which provides conditions for Gâteaux differentiability to imply Fréchet differentiability.

Proposition C.5 *Let $f : \Omega' \subset \mathbb{B}' \to \mathbb{B}$ be Gâteaux differentiable at all points $x \in \Omega'$, which is an open subset of the Banach space $\mathbb{B}'$. Assume that df is continuous on Ω' in the sense that, if*

$$\omega(x, \delta, df) = \sup\left\{\|(df(y) - df(x))h\|_{\mathbb{B}},\ \|h\|_{\mathbb{B}'} = 1,\ y \in \Omega',\ \|x - y\|_{\mathbb{B}'} \le \delta\right\},$$

then, for all $x \in \Omega'$,

$$\lim_{\delta \to 0} \omega(x, \delta, df) \to 0.$$

Then f is also Fréchet differentiable.

(In other terms, if f is continuously Gâteaux differentiable, then it is Fréchet differentiable.)

Proof Note that, for $x, h \in \Omega'$, such that the segment $[x, x + h]$ is included in Ω', the transformation $t \mapsto \partial_t f(x + th)(t) = df(x + th)\,h$ is integrable (because df is continuous) and

$$f(x + h) - f(x) = \int_0^1 \partial_t f(x + th)dt.$$

This implies

$$\begin{aligned} f(x + h) - f(x) - df(x)h &= \int_0^1 (df(x + th) - df(x))\,h\,dt \\ &\le \omega(x, \|h\|_{\mathbb{B}'}, df)\,\|h\|_{\mathbb{B}'} \\ &= o(\|h\|_{\mathbb{B}'}). \end{aligned}$$

□

C.3 A Class of Ordinary Differential Equations

Let $\mathbb{B}$ be a Banach space. We define the Banach space $\mathbb{L}(\mathbb{B}, \mathbb{B})$ of Lipschitz functions from $\mathbb{B}$ to itself, equipped with the norm

$$\|v\|_{\mathbb{L}(\mathbb{B},\mathbb{B})} = \|v(x_*)\|_{\mathbb{B}} + \mathrm{Lip}(v),$$

where $x_* \in \Omega$ is fixed and $\mathrm{Lip}(v)$ is the Lipschitz constant of v. For an open set $\Omega \subset \mathbb{B}$, we let $\mathbb{L}_{(0)}(\Omega, \mathbb{B})$ be the set of restrictions to Ω of functions $\tilde{v} \in \mathbb{L}(\mathbb{B}, \mathbb{B})$ that vanish outside of Ω, with $\|v\|_{\mathbb{L}_{(0)}(\Omega,\mathbb{B})} = \|\tilde{v}\|_{\mathbb{L}(\mathbb{B},\mathbb{B})}$.

We will also consider the spaces $C^p_{(0)}(\Omega, \mathbb{B})$ of restrictions to Ω of functions $\tilde{v} \in C^p(\mathbb{B}, \mathbb{B})$ with bounded derivatives of order 1 to p (without assuming that the function $\tilde{v}$ itself is bounded) that vanish on Ω^c. This space is Banach with norm

$$\|v\|_{C^p_{(0)}(\Omega,\mathbb{B})} = \|\tilde{v}\|_{C^p_{(0)}(\mathbb{B},\mathbb{B})} = \|\tilde{v}(x_*)\|_{\mathbb{B}} + \|d\tilde{v}\|_{p-1,\infty}.$$

We let $\mathbb{V} = L^1([0, T], \mathbb{L}_{(0)}(\Omega, \mathbb{B}))$ and $\mathbb{V}^{(p)} = L^1([0, T], C^p_{(0)}(\Omega, \mathbb{B}))$ denote the spaces of Bochner integrable functions with values in $\mathbb{L}_{(0)}(\Omega, \mathbb{B})$ and $C^p_{(0)}(\Omega, \mathbb{B})$. Given $v \in \mathbb{V}$, we will say that a continuous function $y : [0, T] \to \Omega$ is a solution of the ordinary differential equation $\partial_t y = v(t, y)$ with initial condition $y(t) = x$ if it satisfies

$$y(s) = x + \int_t^s v(u, y(u))du.$$

The right-hand term is well defined as a Bochner integral because

$$\|v(u, y(u))\|_{\mathbb{B}} \leq \|v(u, x_*)\|_{\mathbb{B}} + \mathrm{Lip}(v(u))\|y(u) - x_*\|_{\mathbb{B}} \tag{C.1}$$

is integrable when y is continuous (hence bounded over $[0, T]$). We make the notation abuse of writing $v(t, x) = v(t)(x)$.

C.3.1 Existence and Uniqueness of Solutions

Theorem C.6 *Let $v \in \mathbb{V}$. For all $x \in \Omega$ and $t \in [0, T]$, there exists a unique solution on $[0, T]$ of the ordinary differential equation $\partial_t y = v(t, y)$ with initial condition $y(t) = x$.*

Proof The proof slightly deviates from the standard Picard–Lindelöf theorem, which assumes that $\mathrm{Lip}(v(t, \cdot))$ is uniformly bounded [69, 252]. We have here an integrability condition instead, so that the statement is slightly more general, although the proof, which follows [93, 277], is very similar.

We first prove the result for $\Omega = \mathbb{B}$. Fix $x \in \mathbb{B}$, $t \in [0, T]$ and $\delta > 0$. Let $I = I(t, \delta)$ denote the interval $[0, T] \cap [t - \delta, t + \delta]$. If φ is a continuous function from I to $\mathbb{B}$ such that $\varphi(t) = x$, we define the transformation $\Gamma(\varphi) : I \to \mathbb{B}$ by

$$\Gamma(\varphi)(s) = x + \int_t^s v(u, \varphi(u))du.$$

The function $s \mapsto \Gamma(\varphi)(s)$ is continuous and is such that $\Gamma(\varphi)(t) = x$. The set of continuous functions from the compact interval I to $\mathbb{B}$, equipped with the supremum norm, is a Banach space, and we now show that for δ small enough, Γ satisfies

$$\left\|\Gamma(\varphi) - \Gamma(\varphi')\right\|_\infty \leq \gamma \left\|\varphi - \varphi'\right\|_\infty$$

with $\gamma < 1$. The fixed-point theorem will then imply that there is a unique function φ such that $\Gamma(\varphi) = \varphi$, and this is the definition of a solution of the ordinary differential equation on I.

Since $\Gamma(\varphi)(s) - \Gamma(\varphi')(s) = \int_t^s (v(u, \varphi(u)) - v(u, \varphi'(u)))du$ we have

$$\|\Gamma(\varphi) - \Gamma(\varphi')\|_\infty \leq \left(\int_I \mathrm{Lip}(v(u, \cdot))du\right) \|\varphi - \varphi'\|_\infty.$$

But $\int_I \mathrm{Lip}(v(u, \cdot))du$ can be made arbitrarily small by reducing δ so that existence and uniqueness over I is proved.

We also note that δ can be taken independent of t. This is because the function $\alpha : s \mapsto \int_0^s \mathrm{Lip}(v(u, \cdot))du$ is continuous, hence uniformly continuous on the interval $[0, T]$, so that there exists a constant $\eta > 0$ such that $\left|s - s'\right| < \eta$ implies that $\left|\alpha(s) - \alpha(s')\right| < 1/2$, and it suffices to take $\delta < \eta/2$. From this, we can conclude that a unique solution of the ordinary differential equation exists over all $[0, T]$, because it is now possible, starting from the interval $I(t, \delta)$, to extend the solution from both sides, by jumps of $\delta/2$ at least, until the boundaries are reached.

This proves the result for $\Omega = \mathbb{B}$ and we now consider arbitrary open sets Ω. By extending $v(t)$ by 0 on Ω^c, the value of $\mathrm{Lip}_T(v(t, \cdot))$ remains unchanged and we can apply the result over $\mathbb{B}$ to ensure existence and uniqueness of the solution with a given initial condition. So, it only remains to show that solutions such that $y(t) \in \Omega$ for some t belong to Ω at all times. This is true because if there exists an s such that $y(s) = x' \notin \Omega$, then the function $\tilde{y}(u) = x'$ for all u is a solution of the equation, since $v(u, x') = 0$ for all u. Uniqueness implies $\tilde{y} = y$ at all times, which is impossible. □

Important special cases of this theorem are linear equations. They correspond to taking $v(t, x) = A(t)x + b(t)$ in Theorem C.6. Noting that

$$\|v\|_{\mathbb{V}} \leq \int_0^T \left(\|A(t)\|_{op(\mathbb{B})}\|x_*\|_{\mathbb{B}} + \|b(t)\|_{\mathbb{B}}\right) dt + \int_0^T \|A(t)\|_{op(\mathbb{B})}\, dt,$$

we have the following corollary.

Corollary C.7 *Assume that, for $t \in [0, T]$, $A(t) : \mathbb{B} \to \mathbb{B}$ is a bounded linear operator, $b(t) \in \mathbb{B}$, and that they are integrable and satisfy*

$$\int_0^T (\|A(t)\|_{op(\mathbb{B})} + \|b(t)\|_{\mathbb{B}})\, dt < \infty.$$

Then the equation

$$\partial_t y = A(t)y + b(t)$$

has a unique solution over $[0, T]$ *for a given initial condition.*

C.3.2 Flow Associated to an ODE

Definition C.8 Let $v \in \mathbb{V}$. We denote by $\varphi_{st}^v(x)$ the solution at time t of the equation $\partial_t y = v(t, y)$ with initial condition $y(s) = x$. The function $(t, x) \mapsto \varphi_{st}^v(x)$ is called the flow associated to v starting at s. It is defined on $[0, T] \times \Omega$ and takes values in Ω.

From the definition, we have the following property.

Proposition C.9 *If* $v \in \mathbb{V}$ *and* $s, r, t \in [0, T]$*, then*

$$\varphi_{st}^v = \varphi_{rt}^v \circ \varphi_{sr}^v.$$

In particular, $\varphi_{st}^v \circ \varphi_{ts}^v = \mathrm{id}$ *and* φ_{st}^v *is invertible for all* s *and* t.

Proof If $x \in \Omega$, $\varphi_{st}^v(x)$ is the value at time t of the unique solution of $\partial_t y = v(t, y)$, which is equal to x at time s. It is equal to $x' := \varphi_{sr}^v(x)$ at time r, and thus also equal to $\varphi_{rt}^v(x')$, which is the statement of the proposition. □

Theorem C.10 *If* $v \in \mathbb{V}$*, the associated flow,* φ_{st}^v*, is at all times* s *and* t *a homeomorphism of* Ω *(it is continuous, invertible, with a continuous inverse).*

The proof of this result uses Gronwall's lemma, which we first state and prove.

Theorem C.11 (Gronwall's lemma) *Consider two non-negative functions* $\alpha(s)$ *and* $u(s)$*, defined for* $s \in I$ *where* I *is an interval in* $\mathbb{R}$ *containing 0. Assume that* u *is bounded, and that, for some integrable function* c*, and for all* $t \in I$*,*

$$u(t) \leq c(t) + \left| \int_0^t \alpha(s)u(s)\, ds \right|. \tag{C.2}$$

Then, for all $t \in I$*,*

$$u(t) \leq c(t) + \left| \int_0^t c(s)\alpha(s)e^{\left|\int_0^s \alpha(u)\, du\right|}\, ds \right|. \tag{C.3}$$

When c *is a constant, this upper bound becomes*

$$u(t) \leq ce^{\left|\int_0^t \alpha(s)\, ds\right|}. \tag{C.4}$$

Proof To address simultaneously the cases $t > 0$ and $t < 0$, we let $\varepsilon = 1$ in the first case and $\varepsilon = -1$ in the second case. Inequality (C.2) now becomes:

$$u(t) \leq c(t) + \varepsilon \int_0^t \alpha(s)u(s)ds.$$

Iterating this inequality once yields

$$u(t) \leq c(t) + \varepsilon \int_0^t c(s)\alpha(s)ds + \varepsilon^2 \int_0^t \int_0^{s_1} \alpha(s_1)\alpha(s_2)u(s_2)ds\,,$$

and iterating further, we get

$$\begin{aligned} u(t) \leq c(t) &+ \varepsilon \int_0^t c(s)\alpha(s)ds + \cdots \\ &+ \varepsilon^n \int_{0\leq \varepsilon s_1 \leq \cdots \leq \varepsilon s_n \leq \varepsilon t} c(s_n)\alpha(s_1)\ldots\alpha(s_n)ds_1\ldots ds_n \\ &+ \varepsilon^{n+1} \int_{0\leq \varepsilon s_1 \leq \cdots \leq \varepsilon s_n \leq \varepsilon t} \alpha(s_1)\ldots\alpha(s_{n+1})u(s_{n+1})ds_1\ldots ds_n ds_{n+1}\,. \end{aligned}$$

Consider the integral

$$I_n = \int_{0\leq \varepsilon s_1 \leq \cdots \leq \varepsilon s_n} \alpha(s_1)\ldots\alpha(s_{n-1})ds_1\ldots ds_{n-1}.$$

By symmetry of the integration indices, we have

$$\begin{aligned} I_n &= \frac{1}{(n-1)!} \int_{[0,s_n]^{n-1}} \alpha(s_1)\ldots\alpha(s_{n-1})ds_1\ldots ds_{n-1} \\ &= \frac{1}{(n-1)!} \left(\int_0^{s_n} \alpha(s)ds\right)^{n-1}. \end{aligned}$$

Therefore, using the fact that u is bounded, we have

$$\begin{aligned} u(t) \leq c(t) + \sum_{k=1}^{n} \frac{\varepsilon^k}{(k-1)!} \int_0^t c(s_n)\alpha(s_n) \left(\int_0^{s_n} \alpha(s)ds\right)^{k-1} ds_n \\ + \frac{\varepsilon^{n+1}\sup(u)}{(n+1)!} \left(\int_0^t \alpha(s)ds\right)^{n+1}, \end{aligned}$$

and passing to the limit yields (C.3). If c is constant, the previous upper bound can be written

$$u(t) \leq c \sum_{k=0}^{n} \frac{\varepsilon^k}{k!} \left(\int_0^t \alpha(s) ds \right)^k + \frac{\varepsilon^{n+1} \sup(u)}{(n+1)!} \left(\int_0^t \alpha(s) ds \right)^{n+1},$$

which yields (C.4). □

We now pass to the proof of Theorem C.10.

Proof (Theorem C.10) It suffices to show that φ_{st}^v is continuous, because Proposition C.9 implies that $(\varphi_{st}^v)^{-1} = \varphi_{ts}^v$. Take $x, y \in \Omega$. We have

$$\begin{aligned} \left\| \varphi_{st}^v(x) - \varphi_{st}^v(y) \right\|_{\mathbb{B}} &= \left\| x - y + \int_s^t \left(v(s, \varphi_{sr}^v(x)) - v(s, \varphi_{sr}^v(y)) \right) dr \right\|_{\mathbb{B}} \\ &\leq \|x - y\|_{\mathbb{B}} + \int_s^t \mathrm{Lip}(v(t, \cdot)) \left\| \varphi_{sr}^v(x) - \varphi_{sr}^v(y) \right\|_{\mathbb{B}} ds. \end{aligned}$$

We apply Gronwall's lemma, equation (C.4), with $c = \|x - y\|_{\mathbb{B}}$, $\alpha(r) = \mathrm{Lip}(v(r, \cdot))$ and $u(r) = \left\| \varphi_{0r}^v(x) - \varphi_{0r}^v(y) \right\|_{\mathbb{B}}$, which is bounded because φ_{sr}^v is continuous in r. This yields

$$\left\| \varphi_{sr}^v(x) - \varphi_{sr}^v(y) \right\|_{\mathbb{B}} \leq \|x - y\|_{\mathbb{B}} \exp \left(\int_s^t \mathrm{Lip}(v(r, .)) dr \right), \tag{C.5}$$

which shows that φ_{st}^v is continuous on Ω, and even Lipschitz, with a Lipschitz constant smaller than $\exp(\|v\|_{\mathbb{V}})$. □

Remark Using the fact that (C.1) implies that, if $\partial_t y = v(t, y)$ and $y(0) = x$, then

$$\|y(t) - x_*\|_{\mathbb{B}} \leq \|x - x_*\|_{\mathbb{B}} + \int_0^t \left(\|v(u, x_*)\|_{\mathbb{B}} + \mathrm{Lip}(v(u)) \|y(u) - x_*\|_{\mathbb{B}} \right) du,$$

Gronwall's lemma implies that

$$\|y(t) - x_*\|_{\mathbb{B}} \leq \left(\|x - x_*\|_{\mathbb{B}} + \int_0^t \|v(u, x_*)\|_{\mathbb{B}} \, dt \right) \exp \left(\int_0^t \mathrm{Lip}(v(u) \, du \right),$$

which, in turn, implies the simpler expression

$$\|y(t) - x_*\|_{\mathbb{B}} \leq \left(\|v\|_{\mathbb{V}} + \|x - x_*\|_{\mathbb{B}} \right) \exp \left(\|v\|_{\mathbb{V}} \right). \tag{C.6}$$

C.4 Variation in the Initial Condition

We now discuss the differentiability of the flow associated with an ODE. Because we consider variations along a given trajectory, we will only need to require that the vector field is regular in a neighborhood of this trajectory. We first formalize this concept.

Definition C.12 Let v be a time-dependent vector field on Ω and $y : [0, T] \to \Omega$ a continuous curve. Let $\delta > 0$ be small enough so that $y(t) + x \in \Omega$ for all $t \in [0, T]$ and $x \in B(0, \delta)$. Let $v^y(t, x) = v(t, y(t) + x)$, $x \in B(0, \delta)$. We then say that v is Lipschitz in a δ-tube along y if $v^y \in L^1([0, T], \mathbb{L}(B(0, \delta), \mathbb{B}))$. Similarly, v is C_b^p in this tube if $v^y \in L^1([0, T], C_b^p(B(0, \delta), \mathbb{B}))$.

(As usual, $B(x, \delta)$ denotes the open ball with center x and radius δ in $\mathbb{B}$.)

We then have the following proposition.

Proposition C.13 *Assume that v is Lipschitz in a δ_0-tube along a continuous curve $y_0 : [0, T] \to \Omega$ which furthermore satisfies (with a fixed $s \in [0, T]$)*

$$y_0(t) = x_0 + \int_s^t v(s, y_0(s))\, ds.$$

Then there exists a $\delta > 0$ and a flow φ_{st}, $t \in [0, T]$, of homeomorphisms of Ω such that, for $\|x - x_0\| < \delta$, $y(t) = \varphi_{st}(x)$ is the unique solution of the equation $\partial_t y = v(t, y)$ with condition at time s: $y(s) = x$. Moreover, φ_{st} is such that $\|\varphi_{st}(x) - y_0(t)\|_{\mathbb{B}} < \delta_0$ at all times.

Proof Let $\chi : [0, +\infty) \to [0, 1]$ be a smooth function such that $\chi(\tau) = 1$ for $\tau \in [0, 1]$ and $\chi(\tau) = 0$ for $\tau \geq 2$. For $\delta > 0$, let $v_\delta(t, x) = \chi(\|x - y_0(t)\|_{\mathbb{B}}/\delta)v(t, x)$, with $\delta < \delta_0/2$. Then $v_\delta \in \mathbb{V} = L^1([0, T], \mathbb{L}_{(0)}(\Omega, \mathbb{B}))$, and Theorem C.6 and Proposition C.9 imply that it generates a flow $\varphi_{st}^{v_\delta}$ of homeomorphisms of Ω.

Clearly, $\varphi_{st}^{v_\delta}(x)$ is solution of $\partial_t y = v(t, y)$ as long as $\|\varphi_{st}^{v_\delta}(x) - y_0(t)\| \leq \delta$, which also ensures that $y_0(t) = \varphi_{st}^{v_\delta}(x_0)$. Using Eq. (C.5), we see that there exists a $\delta' > 0$ such that $\|x - x_0\| \leq \delta'$ implies $\|\varphi_{st}^{v_\delta}(x) - y_0(t)\|_{\mathbb{B}} \leq \delta$ for $t \in [0, T]$, so that $\varphi_{st}^{v_\delta}(x)$ is the solution of $\partial_t y = v(t, y)$ with $y(s) = x$. This completes the proof of the proposition. □

Let us now assume that v is C_b^1 in a δ_0-tube along one of its solutions, y_0, such that $y_0(s) = x_0$. Denote by $\mathbb{W}$ the space $L^1([0, T], C_b^1(B(0, \delta_0), \mathbb{B}))$. Let φ_{st} be the flow provided by Proposition C.13, such that $\varphi_{st}(x)$ is a solution of $\partial_t y = v(t, y)$ if $x \in B(x_0, \delta)$. We now study the differentiability of φ_{st} in that ball. In the following, $d\varphi, dv$, etc. will always refer to derivatives over the "space" variable $x \in \Omega$. Because $\varphi_{st}(x)$ remains in the δ_0-tube around y_0 for $\|x - x_0\|_{\mathbb{B}} < \delta$, we have, following the proof of Theorem C.10

$$\|\varphi_{st}(x) - \varphi_{st}(x')\|_{\mathbb{B}} \leq \|x - x'\|_{\mathbb{B}} + \int_s^t \|v^{y_0}(r)\|_{1,\infty} \|\varphi_{sr}(x) - \varphi_{sr}(x')\|_{\mathbb{B}}\, dr,$$

with

$$\begin{aligned} v^{y_0} : \quad & [0, T] \times B(0, \delta_0) \to \mathbb{B} \\ & (t, z) \mapsto v(t, y_0(t) + z)\,. \end{aligned}$$

Using Gronwall's lemma, we get, for $x, x' \in B(x_0, \delta)$,

$$\|\varphi_{st}(x) - \varphi_{st}(x')\|_{\mathbb{B}} \leq \|x - x'\|_{\mathbb{B}} \exp(\|v^{y_0}\|_{\mathbb{W}}).$$

Because we have

$$\varphi_{st}(x) = x + \int_s^t v(t, \varphi_{sr}(x))dr,$$

a natural candidate for $W_h(t, x) := d\varphi_{st}(x)h$ (where $h \in \mathbb{B}$) should satisfy

$$W_h(t, x) = h + \int_s^t dv(r, \varphi_{sr}(x))W_h(r, x)\, dr,$$

which is well defined as soon as $x \in B(x_0, \delta)$, which implies that $\varphi_{sr}(x)$ belongs to the δ_0-tube along y_0, within which $v(r, \cdot)$ is differentiable. This defines W_h as a solution of a linear equation for which Corollary C.7 applies as soon as $x \in B(x_0, \delta)$. We now prove that $\varphi_{st}(x + h) - \varphi_{st}(x) - W_h(t, x) = o(\|h\|_{\mathbb{B}})$, writing

$$\begin{aligned}
&\varphi_{st}(x + h) - \varphi_{st}(x) - W_h(t, x) = \\
&\int_s^t (v(r, \varphi_{sr}(x + h)) - v(r, \varphi_{sr}(x)) - dv(r, \varphi_{sr}(x))(\varphi_{sr}(x + h) - \varphi_{sr}(x)))\, dr \\
&+ \int_s^t dv(r, \varphi_{sr}(x))(\varphi_{sr}(x + h) - \varphi_{sr}(x) - W_h(r, x))\, dr
\end{aligned}$$

so that

$$\begin{aligned}
&\|\varphi_{st}(x + h) - \varphi_{st}(x) - W_h(t, x)\|_{\mathbb{B}} \\
\leq &\int_s^t \int_0^1 \|dv(r, \varphi_{sr}(x + \varepsilon h)) - dv(r, \varphi_{sr}(x))\|_{op(\mathbb{B})} \|\varphi_{sr}(x + h) - \varphi_{sr}(x))\|_{\mathbb{B}}\, d\varepsilon\, dr \\
&+ \int_s^t \|dv(r, \varphi_{sr}(x))\|_{op(\mathbb{B})} \|\varphi_{sr}(x + h) - \varphi_{sr}(x) - W_h(r, x)\|_{\mathbb{B}}\, dr \\
\leq &\|h\|_{\mathbb{B}} \exp(\|v^{y_0}\|_{\mathbb{W}}) \int_0^1 \int_0^T \|dv(r, \varphi_{sr}(x + \varepsilon h)) - dv(r, \varphi_{sr}(x))\|_{op(\mathbb{B})}\, d\varepsilon\, dr \\
&+ \int_s^t \|dv(r, \varphi_{sr}(x))\|_{op(\mathbb{B})} \|\varphi_{sr}(x + h) - \varphi_{sr}(x) - W_h(r, x)\|_{\mathbb{B}}\, dr.
\end{aligned}$$

The continuity of dv implies that $\|dv(r, \varphi_{sr}(x + \varepsilon h)) - dv(r, \varphi_{sr}(x))\|_{op(\mathbb{B})}$ tends to 0 with $\|h\|_{\mathbb{B}}$ for all $\varepsilon \in [0, 1]$. This norm being moreover bounded by $2\|dv(r)\|_\infty$, which is integrable by assumption, the dominated convergence theorem implies that

$$\lim_{h \to 0} \int_0^1 \int_0^T \|dv(r, \varphi_{sr}(x + \varepsilon h)) - dv(r, \varphi_{sr}(x))\|_{op(\mathbb{B})}\, d\varepsilon\, dr = 0.$$

The fact that $\varphi_{st}(x + h) - \varphi_{st}(x) - W_h(t, x) = o(\|h\|_{\mathbb{B}})$ is then a consequence of Gronwall's lemma. Gronwall's lemma applied again to

$$\|W_h(t,x)\|_{\mathbb{B}} = \|h\|_{\mathbb{B}} + \int_s^t \|dv(r,\varphi_{sr}(x))\|_{\mathrm{op}(\mathbb{B})}\|W_h(r,x)\|_{\mathbb{B}}\,dr$$

shows that $W(t,x): h \mapsto W_h(t,x) \in \mathcal{L}(\mathbb{B},\mathbb{B})$ with

$$\|W(t,x)\|_{\mathrm{op}(\mathbb{B})} \le \exp(\|v^{y_0}\|_{\mathbb{W}}).$$

Finally, writing

$$\begin{aligned} W_h(t,x) - W_h(t,x') &= \int_s^t (dv(r,\varphi_{sr}(x)) - dv(r,\varphi_{sr}(x')))W_h(r,x)\,dr \\ &\quad + \int_s^t dv(r,\varphi_{sr}(x'))(W_h(r,x) - W_h(r,x'))\,dr, \end{aligned}$$

so that

$$\begin{aligned} &\|W_h(t,x) - W_h(t,x')\|_{\mathbb{B}} \\ &\quad \le \|h\|_{\mathbb{B}} \exp(C(v)) \int_s^t \|dv(r,\varphi_{sr}(x)) - dv(r,\varphi_{sr}(x'))\|_{\mathrm{op}(\mathbb{B})}\,dr \\ &\quad + \int_s^t \|dv(r,\varphi_{sr}(x'))\|_{\mathrm{op}(\mathbb{B})}\|W_h(r,x) - W_h(r,x')\|_{\mathbb{B}}\,dr, \end{aligned}$$

and applying Gronwall's lemma and the continuity of dv ensures that W is continuous in x. We have therefore proved the following theorem.

Theorem C.14 *Assume that $y_0 : [0,T] \to \mathbb{B}$ is a continuous curve, that v is C_b^1 in a δ_0-tube along y_0, and that y_0 is a solution of $\partial_t y = v(t,y)$ with $y_0(x) = x_0$. Then there exists a $\delta > 0$ such that the flow φ_{st} associated with the ODE is well defined and continuously differentiable for $x \in B(x_0,\delta)$, and $d\varphi_{st}^v(x)h$ is the solution at time t of the equation*

$$\partial_t W_h = dv(t,\varphi_{st}^v(x))W_h \tag{C.7}$$

with $W_h(s) = h$.

The following global version of the theorem is then immediate.

Theorem C.15 *Assume that $v \in L^1([0,T], C_{(0)}^1(\Omega,\mathbb{B}))$. Then φ_{st}^v is a diffeomorphism of Ω and $d\varphi_{st}^v(x)h$ is the solution at time t of (C.7) with $W_h(s) = h$.*

We can alternatively restate this theorem by writing that $d\varphi_{st}^v(x)$ if the solution of the equation

$$\partial_t W = dv(t,\varphi_{st}^v(x))W \tag{C.8}$$

with $W(0) = \mathrm{Id}$, which is now a linear ODE over $\mathcal{L}(\mathbb{B},\mathbb{B})$.

C.5 Variation with Respect to a Parameter

We now assume that v depends on a parameter λ. We introduce for this a set U, an open subset of a Banach space $\mathbb{U}$, and consider a mapping $\lambda \mapsto v^\lambda$ defined on U and returning a time-dependent vector field on Ω.

We fix $\lambda_0 \in U$ and consider a continuous curve $y_0 : [0, T] \to \Omega$ such that, defining $w^\lambda(t, x) = v^\lambda(t, y_0(t) + x)$, the mapping $\lambda \mapsto w^\lambda$ is continuously differentiable from U to $\mathbb{W} = L^1([0, T], C_b^1(B(0, \delta_0), \mathbb{B}))$ for some $\delta_0 > 0$. We also assume that

$$y_0(t) = x_0 + \int_0^T v^{\lambda_0}(t, y_0(t))\, dt$$

i.e., y_0 is a solution of $\partial_t y = v^{\lambda_0}(y)$. In the following computation, we will use the letter d to denote a derivative with respect to the x variable ($x \in \mathbb{B}$), and ∂_λ for derivatives with respect to λ. Restricting the domain U to a small neighborhood of λ_0 if needed, we will furthermore assume that U is convex and that $\|w^\lambda\|_{\mathbb{W}} \le C_0$ for some constant C_0 and all $\lambda \in U$.

If $\delta < \delta_0/2$ and $\chi : [0, +\infty) \to [0, 1]$ is a smooth function such that $\chi(\tau) = 1$ on $[0, 1]$ and $\chi(\tau) = 0$ on $[2, +\infty)$, we can associate to each λ the Lipschitz vector field $\hat{v}^\lambda(t, x) = \chi(\|x - y_0(t)\|_{\mathbb{B}}/\delta)v^\lambda(t, x)$ whose solution y^λ starting at x_0 provides a solution of $\partial_t y = v^\lambda(t, y)$ as long as it remains at distance less that δ from y_0. Using Gronwall's lemma, we can show that this remains true over the whole interval $[0, T]$. Indeed,

$$\begin{aligned} y^\lambda(t) - y_0(t) = &\int_0^t \left(\hat{v}^\lambda(s, y^\lambda(s)) - \hat{v}^{\lambda_0}(s, y^\lambda(s))\right) ds \\ &+ \int_0^t \left(\hat{v}^{\lambda_0}(s, y^\lambda(s)) - \hat{v}^{\lambda_0}(s, y_0(s))\right) ds. \end{aligned}$$

We have

$$\hat{v}^\lambda(t, x) - \hat{v}^{\lambda_0}(t, x) = \chi\left(\frac{\|x - y_0(t)\|_{\mathbb{B}}}{\delta}\right)(v^\lambda(t, x) - v^{\lambda_0}(t, x))$$

so that the left-hand side vanishes if $x - y_0(t) \notin B(0, \delta_0)$ and is such that

$$\begin{aligned} \|\hat{v}^\lambda(t, x) - \hat{v}^{\lambda_0}(t, x)\|_{\mathbb{B}} &\le \chi\left(\frac{\|x - y_0(t)\|_{\mathbb{B}}}{\delta}\right)\|v^\lambda(t, x) - v^{\lambda_0}(t, x)\| \\ &\le \left(\int_0^1 \sup_{x \in B(y_0(t), \delta_0)} \|\partial_\lambda v^{\lambda_0 + \alpha(\lambda - \lambda_0)}(t, x)\|_{\mathrm{op}(\mathbb{U}, \mathbb{B})} d\alpha\right)\|\lambda - \lambda_0\|_{\mathbb{U}}, \end{aligned}$$

so that

$$\int_0^t \|\hat{v}^\lambda(s, y^\lambda(s)) - \hat{v}^{\lambda_0}(s, y^\lambda(s)\|_{\mathbb{B}}\, ds \le \left(\int_0^1 \|w^{\lambda_0+\alpha(\lambda-\lambda_0)}\|_{\mathbb{W}}\, d\alpha\right) \|\lambda - \lambda_0\|_{\mathbb{U}} \le C_0\|\lambda - \lambda_0\|_{\mathbb{U}}.$$

We also have

$$\|\hat{v}^{\lambda_0}(s, y^\lambda(s)) - \hat{v}^{\lambda_0}(s, y_0(s))\|_{\mathbb{B}} \le \sup_{x\in B(0,\delta_0)} \|dw^{\lambda_0}(s, x)\|_{\mathrm{op}(\mathbb{B})} \|y^\lambda(s) - y_0(s)\|_{\mathbb{B}}.$$

Combining these inequalities and applying Gronwall's lemma implies that

$$\|y^\lambda(s) - y_0(s)\|_{\mathbb{B}} \le (C_0 e^{C_0})\|\lambda - \lambda_0\|_{\mathbb{U}}\,. \tag{C.9}$$

We can therefore find some $\delta > 0$ such that $\|y^\lambda(s) - y_0(s)\|_{\mathbb{B}} < \delta_0$ at all times $s \in [0, T]$ for all λ such that $\|\lambda - \lambda_0\|_{\mathbb{U}} < \delta$. Hence, y^λ is a solution of $\partial_t y = v^\lambda(y)$ for $\lambda \in B(\lambda_0, \delta)$. To simplify the discussion below, and without loss of generality, we will assume that $U = B(\lambda_0, \delta)$ in the following.

We now study the differentiability of y^λ with respect to λ. Formally differentiating the identity

$$y^\lambda(t) = x_0 + \int_0^t v^\lambda(s, y^\lambda(s))\, ds,$$

we get, for small enough $\eta \in \mathbb{U}$,

$$\partial_\lambda y^\lambda(t)\eta = \int_0^t (\partial_\lambda v^\lambda \eta)(s, y^\lambda(s))\, ds + \int_0^t dv^\lambda(s, y^\lambda(s))(\partial_\lambda y^\lambda(s)\eta)\, ds.$$

Given this, we introduce the solution W_η of the equation

$$W_\eta(t) = \int_0^t (\partial_\lambda v^\lambda \eta)(s, y^\lambda(s))\, ds + \int_0^t dv^\lambda(s, y^\lambda(s)) W_\eta(s)\, ds$$

and show that $W_\eta(t)$ is the differential of $y^\lambda(t)$ with respect to λ in the direction η.

Using the same argument as the one leading to (C.9), we have, for small enough η,

$$\|y^{\lambda+\eta}(t) - y^\lambda(t)\|_{\mathbb{B}} \le C(\lambda)\|\eta\|_{\mathbb{U}}. \tag{C.10}$$

Consider now the difference

$$
\begin{aligned}
&y^{\lambda+\eta}(t) - y^{\lambda}(t) - W_\eta(t) = \\
&\int_0^t \left(v^{\lambda+\eta}(s, y^{\lambda+\eta}(s)) - v^{\lambda}(s, y^{\lambda+\eta}(s)) - (\partial_\lambda v^\lambda \eta)(s, y^{\lambda+\eta}(s))\right) ds \\
&+ \int_0^t \left(v^{\lambda}(s, y^{\lambda+\eta}(s)) - v^{\lambda}(s, y^{\lambda}(s)) - dv^{\lambda}(s, y^{\lambda}(s))(y^{\lambda+\eta}(s) - y^{\lambda}(s))\right) ds \\
&+ \int_0^t ((\partial_\lambda v^\lambda \eta)(s, y^{\lambda+\eta}(s)) - (\partial_\lambda v^\lambda \eta)(s, y^{\lambda}(s)))ds \\
&+ \int_0^t dv^{\lambda}(s, x^{\lambda}(s))(y^{\lambda+\eta}(s) - y^{\lambda}(s) - W_\eta(s))ds
\end{aligned}
\tag{C.11}
$$

and the upper-bound

$$
\begin{aligned}
&\|y^{\lambda+\eta}(t) - y^{\lambda}(t) - W_\eta(t)\|_{\mathbb{B}} \leq \\
&\int_0^t \int_0^1 \left\|(\partial_\lambda v^{\lambda+\varepsilon\eta}\eta)(s) - (\partial_\lambda v^\lambda \eta)(s)\right\|_\infty d\varepsilon\, ds \\
&+ \int_0^t \int_0^1 \|dv^{\lambda}(s, y^{\lambda+\varepsilon\eta}(s)) - dv^{\lambda}(s, y^{\lambda}(s))\|_{op(\mathbb{B})} \|y^{\lambda+\eta}(s) - y^{\lambda}(s))\|_{\mathbb{B}}\, d\varepsilon\, ds \\
&+ \int_0^t \|d(\partial_\lambda v^\lambda \eta)(s)\|_\infty \|y^{\lambda+\eta}(s) - y^{\lambda}(s)\|_{\mathbb{B}}\, ds \\
&+ \int_0^t \|dv^{\lambda}(s)\|_\infty \|y^{\lambda+\eta}(s) - y^{\lambda}(s) - W_\eta(s)\|_{\mathbb{B}}\, ds.
\end{aligned}
$$

If we prove that the first three terms in the right-hand side are $o(\|\eta\|_{\mathbb{U}})$, then Gronwall's lemma will ensure that the left-hand side is also $o(\|\eta\|_{\mathbb{U}})$.

The first term is bounded by

$$
\|\eta\|_{\mathbb{U}} \max_{\varepsilon\in[0,1]} \|\partial_\lambda w^{\lambda+\varepsilon\eta} - \partial_\lambda w^\lambda\|_{\mathbb{W}},
$$

which is $o(\|\eta\|_{\mathbb{U}})$, because $\lambda \mapsto w^\lambda$ is C^1. The second one is bounded by

$$
\begin{aligned}
&\left(\max_{t\in[0,T]} \|y^{\lambda+\eta}(t) - y^{\lambda}(t)\|_{\mathbb{B}}\right) \\
&\qquad \int_0^T \int_0^1 \|dv^{\lambda}(s, y^{\lambda+\varepsilon\eta}(s)) - dv^{\lambda}(s, y^{\lambda}(s))\|_{op(\mathbb{B})}\, d\varepsilon\, ds \\
&\qquad \leq C_0' \|\eta\|_{\mathbb{U}} \max_{\varepsilon\in[0,1]} \|w^{\lambda+\varepsilon\eta} - w^\lambda\|_{\mathbb{W}}
\end{aligned}
$$

for some constant C_0', so that this term is also $o(\|\eta\|_U)$.

Finally, the third term is less than

$$\|\partial_\lambda v^\lambda\|_{op(\mathbb{V}^{(1)},\mathbb{V}^{(1)})}\,\|\eta\|_{\mathbb{U}}\max_{t\in[0,T]}\|y^{\lambda+\eta}(t)-y^\lambda(t)\|_{\mathbb{B}}=o(\|\eta\|_{\mathbb{U}}).$$

This proves that $W_\eta=\partial_\lambda y^\lambda\eta$.

Theorem C.16 *Assume that $\lambda\mapsto v^\lambda$ is such that, for some $\lambda_0\in U$, and with $w^\lambda(t,x)=v^\lambda(t,y_0(t)+x)$, the mapping $\lambda\mapsto w^\lambda$ is continuously differentiable from U to $\mathbb{W}=L^1([0,T],C_b^1(B(0,\delta_0),\mathbb{B}))$ for some $\delta_0>0$, where y_0 satisfies*

$$y_0(t)=x_0+\int_0^T v^{\lambda_0}(t,y_0(t))\,dt\,.$$

Then, there exists a neighborhood of λ_0 such that $\partial_t y=v^\lambda(t,y)$ has a unique solution y^λ satisfying $y^\lambda(0)=x_0$, and $\lambda\mapsto y^\lambda$ is differentiable. Moreover, for $\eta\in\mathbb{U}$, $\partial_\lambda y^\lambda(t)\eta$ is the solution of the equation

$$\partial_t W_\eta=dv^\lambda(t,y^\lambda(t))W_\eta+(\partial_\lambda v^\lambda\eta)(t,y^\lambda(t)) \tag{C.12}$$

with $W_\eta(0)=0$.

This directly implies the following global version.

Theorem C.17 *Assume that $\lambda\mapsto v^\lambda$ is C^1 from U to $\mathbb{V}^{(1)}$, where U is an open subset of a Banach space $\mathbb{U}$. Fix $x_0\in\mathbb{B}$. Then the solution y^λ of the equation $\partial_t y=v^\lambda(t,y)$ with $y^\lambda(0)=x_0$ is differentiable in λ and for $\eta\in\mathbb{U}$, $\partial_\lambda y^\lambda(t)\eta$ is the solution of* (C.12) *with $W_\eta(0)=0$.*

The solutions of (C.8) and (C.12) are related because the former is the homogeneous part of the latter. This implies that one can deduce the solutions of (C.12) from those of (C.8) using a variation of the constant. To simplify the notation, we fix λ and write v for v^λ. Let also $M(t)$ denote $(\partial_\lambda v^\lambda\eta)(t,y^\lambda(t))$. Take $W_\eta(t)=d\varphi_{0t}^v(x_0)h(t)$, with h to be determined. Then one must have

$$M(t)=d\varphi_{0t}^v(x_0)\dot h(t),$$

so that (given that $W_\eta(0)=0$)

$$h(t)=\int_0^t(d\varphi_{0s}^v(x_0))^{-1}M(s)\,ds=\int_0^t d\varphi_{s0}^v(y^\lambda(s))M(s)\,ds,$$

yielding (returning to the full notation)

$$(\partial_\lambda y^\lambda(t))\eta=\int_0^t d\varphi_{st}^{v^\lambda}(y^\lambda(s))(\partial_\lambda v^\lambda\eta)(s,y^\lambda(s))\,ds. \tag{C.13}$$

As a second connection between Theorems C.15 and C.17, notice that x acts as a parameter in Eq. (C.8). This implies that higher derivatives of φ^v_{st} will follow an ODE provided by Theorem C.17. Assume, for example, that $v \in \mathbb{V}^{(2)} = L^1([0,T], C^2_{(0)}(\Omega, \mathbb{B}))$. Then W in Eq. (C.8) is the solution of a linear ODE with associated vector field in $L^1([0,T], C^\infty_0(\mathbb{B}, \mathbb{B}))$ (because it is linear) which is continuously differentiable in x. This implies that $d^2\varphi_{st}(x)$ is the solution of an ODE over $\mathcal{L}^2(\mathbb{B}, \mathbb{B})$ (symmetric bilinear forms) obtained by formal differentiation of (C.8). Letting $W^{(0)} = \varphi^v_{st}(x)$, $W^{(1)} = d\varphi^v_{st}(x)$, etc., we have

$$\partial_t W^{(2)}(h_1, h_2) = d^2 v(t, W^{(0)})(W^{(1)}(h_1), W^{(1)}(h_2)) + dv(t, W^{(0)})W^{(2)}(h_1, h_2).$$

If one assumes higher derivatives, i.e., $v \in \mathbb{V}^{(p)}$ for $p \geq 1$, then the argument can be iterated to the pth order, leading to the following result.

Theorem C.18 *Assume that $v \in \mathbb{V}^{(p)} = L^1([0,T], C^p_{(0)}(\Omega, \mathbb{B}))$. Then $W^{(0)}(x) = \varphi^v_{st} \in C^p(\Omega, \mathbb{B})$ and, for $k = 1, \ldots, p$, $W^{(k)}(t,x) = d^k\varphi^v_{st}$ is the solution of the linear ODE*

$$\partial_t W^{(k)} = B^{(k)}(t) + dv(t, W^{(0)})W^{(k)}(t), \tag{C.14}$$

where the symmetric k-linear form $B^{(k)} \in \mathcal{L}^k(\mathbb{B}, \mathbb{B})$ depends on v and on $W^{(0)}, \ldots, W^{(k-1)}$ and satisfies

$$\begin{aligned} B^{(l+1)}(h_1, \ldots, h_{l+1}) = {} & d(B^{(l)}(h_1, \ldots, h_l))h_{l+1} + \\ & d^2 v(t, W^{(0)})(W^{(l)}(h_1, \ldots, h_l), W^{(1)}h_{l+1}). \end{aligned} \tag{C.15}$$

The initial conditions are $W^{(1)}(s) = \mathrm{Id}$ and $W^{(k)}(s) = 0$, $s \geq 2$.

The symmetry of $B^{(k)}$ results from the fact that the other two terms in (C.14) are symmetric. Equation (C.14) can also be written in the form

$$\partial_t W^{(k)} = d^k(v(t, W^{(0)})) \tag{C.16}$$

with $W^{(l+1)} = dW^{(l)}$.

C.6 Numerical Integration of ODEs

We end this chapter with a few remarks on how ODEs can be solved numerically (in finite dimensions!). Of course, it cannot be our intention, in these few paragraphs, to include all the information that can be found in specialized textbooks (e.g., [234, 270, 271]), so we refer the reader to these references for a comprehensive presentation.

C.6.1 Explicit Schemes

Consider the differential equation $\partial_t y = v(t, y)$ where $y \in \mathbb{R}^n$. Denote by $\varphi_{st}(z)$ the solution at time t of this equation with initial condition $y(s) = z$. Explicit one-step numerical schemes for this ODE implement iterations of the form

$$\begin{cases} z_{k+1} = \eta(t_k, z_k, h_k) \\ t_{k+1} = t_k + h_k \end{cases}$$

in which z_k is intended to provide a numerical approximation of $\varphi_{0t_k}(z_0)$. If $h_k = h$ is fixed, this is a fixed step method, and is h_k is optimized at each step, this is an adaptive step method.

The simplest such algorithm is the Euler method, for which

$$\eta(t, z, h) = z + hv(t, z).$$

This algorithm has the advantage of requiring only one evaluation of v per iteration, but it also has a limited accuracy, as measured by the difference

$$\delta(t, z, h) = \frac{1}{h}(\varphi_{t,t+h}(z) - \eta(t, z, h)).$$

For the Euler method, we have, making a second-order Taylor expansion of y around $y(t)$:

$$\delta(t, z, h) = \frac{h}{2}\ddot{y}(t) + o(h)$$

with $\ddot{y} = \partial_t v(t, y) + dv(t, y)v(t, y)$ (therefore requiring that v is differentiable in time and space). A method is said to be accurate at order p if $\delta(t, z, h) = O(h^p)$. So the Euler method has order 1.

Slightly more complex is the modified Euler, or Euler midpoint method in which η is defined by the iteration

$$\begin{cases} u_1 = v(t, z), \\ \eta(t, z, h) = z + hv(t, z + \dfrac{h}{2}u_1). \end{cases}$$

Using a second-order expansion, one proves that the Euler midpoint method has order 2, assuming this time two derivatives in time and space for v. A fourth-order method is provided by the Runge–Kutta iteration, defined by

$$\begin{cases} u_1 = v(t, z), \\ u_2 = v(t + \dfrac{h}{2}, z + \dfrac{h}{2}u_1), \\ u_3 = v(t + \dfrac{h}{2}, z + \dfrac{h}{2}u_2), \\ u_4 = v(t + h, z + hu_3), \\ \eta(t, z, h) = z + \dfrac{h}{6}(u_1 + 2u_2 + 2u_3 + u_4). \end{cases}$$

More generally, one defines Runge-Kutta methods with s stages using iterations of the form

$$\begin{cases} u_i = v\Big(t + c_i h, z + h \displaystyle\sum_{j=1}^{i-1} a_{ij} u_j\Big), \ i = 1, \ldots, s, \\ \eta(t, z, h) = z + h \displaystyle\sum_{i=1}^{s} b_i u_i \end{cases}$$

with $c_i = \sum_{j=1}^{i-1} a_{ij}$. The method is therefore specified by the lower triangular s by s matrix (with null diagonal) (a_{ij}) and by the vector (b_i). Such an iteration requires s evaluations of the function v, but can reach higher orders of accuracy. Note that high orders of accuracy can only be achieved provided v has at least the same number of derivatives in time and space.

A lot can be gained in the accuracy vs. efficiency trade-off by using adaptive step size. This can be implemented by comparing two similar Runge–Kutta methods (using different coefficients in the expansion) and comparing their differences to some tolerance level. If the difference is larger than the tolerance, the step is reduced; otherwise, it is enlarged (see [88, 234]).

C.6.2 Implicit Schemes

The above methods are called explicit because z_{k+1} is explicitly given as a function of the current value z_k. Implicit methods relax this constraint, and require solving a nonlinear equation to evaluate $y(t_{k+1})$. In spite of this added complexity, such methods may exhibit improved performances, because they are generally more stable and allow for the use of larger step sizes h. The simplest example is the Euler implicit method, which iterates

$$z_{k+1} = z_k + h_k v(t_{k+1}, z_{k+1}).$$

To see why such a method can be interesting, consider the simple autonomous and linear case in which $v(t, y) = Ay$. The solution of the ODE is $y(t) = e^{tA} y(0)$. When A only has negative eigenvalues, the solution converges to 0 as t tends to infinity (one says that the resulting system is stable). An Euler discretization, say

with constant step h, gives $z_k = (\mathrm{Id} + hA)^k z_0$. Such an iteration will converge to 0 only if the eigenvalues of $(\mathrm{Id} + hA)$ are smaller than 1 in modulus, requiring h to be smaller than $2/|\lambda_{max}|$, where λ_{max} is the eigenvalue of A that has the largest absolute value. The implicit iteration yields $z_k = (\mathrm{Id} - hA)^{-k} z_0$, which always converges to 0.

More general one-step implicit schemes use iterations for which η is given by

$$\theta(t, z, \eta(t, z, h), h) = 0$$

for some function θ such that $\theta(t, z, z', h) - z'$ tends to 0 as h tends to 0. Since an analytical solution (expressing z' as a function of t, z and h) for

$$\theta(t, z, z', h) = 0$$

is almost never available, one typically uses fixed-point iterations. The sequence

$$\zeta_{q+1} = \zeta_q - \theta(t, z, \zeta_q, h)$$

will converge to the required z' for small enough h (ensuring that $\zeta \mapsto (\theta(t, z, \zeta, h) - \zeta)$ is contractive). Doing this may clearly require several evaluations of η (which may in turn require several evaluations of v), so that the gain obtained from being able to use larger steps may be lost if evaluating v is costly.

A natural situation in which implicit methods are needed is when one wants to solve an equation backward in time, in a way which is consistent with the forward solution. Let us restrict ourselves to autonomous equations, so that v does not depend on t, assuming the same property at the discretization level (θ does not depend on t). The uniqueness theorem for ODEs implies that if y is a solution of $\partial_t y = v(y)$, then the solution of $\partial_t \tilde{y} = -v(\tilde{y})$ initialized with $\tilde{y}(0) = y(T)$ is $\tilde{y}(t) = y(T - t)$. This property is useful if one needs to recover the initial condition of an ODE given its state at time T. If one separately discretizes the forward and the backward equations, this property is generally not satisfied exactly at the discrete level. One needs to make sure that the backward iteration is consistent with the forward one. This implies that, if the forward iteration is solved using (assuming constant steps)

$$\theta(z_k, z_{k+1}, h) = 0$$

then the backward scheme should use

$$\tilde{\theta}(\tilde{z}_k, \tilde{z}_{k+1}, h) = 0$$

with

$$\tilde{\theta}(\tilde{z}_k, \tilde{z}_{k+1}, h) = \theta(\tilde{z}_{k+1}, \tilde{z}_k, h)$$

and it is clear that θ and $\tilde{\theta}$ cannot be explicit together. Another interesting notion related to reversibility in time is the adjoint scheme

$$\theta^*(z, z', h) = \theta(z', z, -h),$$

which is such that $\tilde{\theta}(z, z', h) = \theta^*(z, z', -h)$. The adjoint scheme is also a forward scheme. For example, the adjoint of the Euler explicit scheme is the Euler implicit scheme.

One says that the scheme is symmetric if $\theta^* = \theta$. Symmetric schemes are interesting because they can be used to solve both the backward and the forward equation in a consistent way. One simple way to build a symmetric scheme from a non-symmetric one is to combine the latter with its adjoint, i.e., define z' from z by first finding z'' such that $\theta^*(z, z'', h/2) = 0$, then solving $\theta(z'', z', h/2) = 0$. Starting with the Euler method, this yields the implicit midpoint method, which is defined by

$$\begin{cases} y' = y_k + \dfrac{h}{2} v(y') \\ y_{k+1} = y_k + \dfrac{h}{2} v(y') \,. \end{cases}$$

Another important class of problems in which implicit methods are useful are those which require the solution of Hamiltonian systems, such as the one described in Sect. B.6.6, or such as the EPDiff equation, which is extensively discussed in this book. Such systems are split over two variables, traditionally denoted p (momentum, or co-state) and q (state), and involve a real-valued function $H(p, q)$ called the Hamiltonian.

The resulting Hamiltonian system is the ODE

$$\begin{cases} \partial_t q = \partial_p H(p, q), \\ \partial_t p = -\partial_q H(p, q). \end{cases}$$

An important property of Hamiltonian systems is that their flow is "symplectic". To define this notion, introduce the matrix

$$J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}.$$

With this notation, we can rewrite the Hamiltonian system in the form

$$\partial_t \begin{pmatrix} p \\ q \end{pmatrix} = J \nabla H(p, q)$$

or, letting Φ denote the flow associated to this equation:

$$\partial_t \Phi = J \, \nabla H \circ \Phi.$$

The differential of Φ satisfies

$$\partial_t d\Phi = J\,(d^2 H \circ \Phi) d\Phi,$$

from which one easily deduces that $\partial_t (d\Phi^T J d\Phi) = 0$, which implies (because $\Phi =$ id at $t = 0$)

$$d\Phi^T J d\Phi = J.$$

Transformations Φ that satisfy this equation are called symplectic. In fact, such a property characterizes Hamiltonian systems, and a natural requirement is that it should be shared by discrete systems, too. This leads to the definition of symplectic integrators, which are integrators for which the function $z \mapsto \eta(z, h)$ is symplectic for all h (there is no time variable because the Hamiltonian system is autonomous).

The simplest examples of symplectic integrators are the symplectic Euler methods. There are two of them, defined by

$$\begin{cases} p_{n+1} = p_n - h\partial_q H(p_{n+1}, q_n) \\ q_{n+1} = q_n + h\partial_p H(p_{n+1}, q_n) \end{cases} \text{and} \begin{cases} p_{n+1} = p_n - h\partial_q H(p_n, q_{n+1}) \\ q_{n+1} = q_n + h\partial_p H(p_n, q_{n+1}) \end{cases}.$$

Note that the latter equation is the dual of the former. When composed, they form the Störmer–Verlet schemes (there are two of them, depending on which of the symplectic Euler schemes is applied first). Symplectic Euler is a first-order method, whereas Störmer–Verlet has order two. Another symplectic method of order two is the implicit midpoint, defined by

$$\begin{cases} p_{n+1} = p_n - h\partial_q H(p'_n, q'_n) \\ q_{n+1} = q_n + h\partial_p H(p'_n, q'_n) \end{cases}$$

with $p'_n = (p_n + p_{n+1})/2$ and $q'_n = (q_n + q_{n+1})/2$. The proof of these statements and many more details can be found in [98, 191].

Appendix D
Introduction to Optimization and Optimal Control Theory

D.1 Unconstrained Optimization Problems

We consider a function $F : \Omega \to \mathbb{R}$ where Ω is an open subset of a Banach space $\mathbb{B}$, and we first discuss the unconstrained optimization problem of finding

$$x^* = \operatorname*{argmin}_{x \in \Omega} F(x). \tag{D.1}$$

The following theorem gives the usual necessary conditions for optimality. They are the same as those that hold in finite dimensions, with which we assume that the reader is familiar.

Theorem D.1 *Assume that F is differentiable over Ω, and that x^* satisfies* (D.1)*; then $dF(x^*) = 0$.*

When $dF : x \mapsto dF(x)$, which maps Ω to $\mathbb{B}^*$, is differentiable, one defines the second derivative, d^2F, which is such that $d^2F(x) \in \mathcal{L}(\mathbb{B}, \mathbb{B}^*)$, the space of continuous linear transformations from $\mathbb{B}$ to $\mathbb{B}^*$ and say that F is twice differentiable. If F is twice differentiable on Ω, we have

$$\begin{aligned} &F(x + th + th') - F(x + th) - F(x + th') + F(x) \\ &= \int_0^t \big(dF(x + th + u'h') - dF(x + th) \;\big|\; h'\big) du' \\ &= \int_0^t \int_0^t \big(d^2F(x + uh + uh')h \;\big|\; h'\big) du du'. \end{aligned}$$

If $x \mapsto d^2F(x)$ is continuous, i.e., $F \in C^2(\Omega)$, we have

$$\lim_{t \to 0} \Big(F(x + th + th') - F(x + th) - F(x + th') + F(x)\Big) = \big(d^2F(x)h \;\big|\; h'\big).$$

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

Since the left-hand side is symmetric in h and h', this proves that

$$\left(d^2F(x)h \,\middle|\, h'\right) = \left(d^2F(x)h' \,\middle|\, h\right)$$

for any $h, h' \in \mathbb{B}$. This generalizes the usual theorem on symmetry of second partial derivatives in finite dimensions. In this case, $d^2F(x)$ can be identified with a symmetric bilinear form on $\mathbb{B}$,

$$HF(x)(h, h') := \left(d^2F(x)h \,\middle|\, h'\right)$$

called the Hessian of F. With this definition, we can state the generalization of the sufficiency conditions on a local minimum in finite dimensions.

Theorem D.2 *Assume that $F \in C^2(\Omega)$. If $x^* \in \Omega$ is such that $dF(x^*) = 0$ and $HF(x^*)$ is positive definite, then x^* is a local minimum of F.*

(One says that $HF(x^*)$ is positive definite if $HF(x^*)(h, h) \geq 0$ for all h, with equality if and only if $h = 0$.)

D.2 Problems with Equality Constraints

The extension of the Lagrange-multiplier theorem to infinite dimensions is not as straightforward as the extension of the optimality conditions in the unconstrained case. Let us first recall how Lagrange multipliers arise in finite dimensions. Consider the problem of finding

$$x^* = \operatorname*{argmin}_{x \in \Omega} F(x) \tag{D.2}$$

subject to $\Gamma(x) = 0$, where Ω is an open subset of $\mathbb{R}^d$, and $F : \Omega \to \mathbb{R}$ and $\Gamma : \Omega \to \mathbb{R}^k$, $k \leq d$, are continuously differentiable. The following theorem holds.

Theorem D.3 *If x^* is a solution of (D.2), there exists $\rho \in \{0, 1\}$ and $\lambda \in \mathbb{R}^k$, with $(\rho, \lambda) \neq (0, 0)$, such that x^* is a stationary point of the Lagrangian*

$$L(x) = \rho F(x) + \lambda^T \Gamma(x)$$

(i.e., $dL(x^*) = 0$.)
If $d\Gamma(x^*)$ has rank k, one must have $\rho = 1$.

Proof First assume that $d\Gamma(x^*)$ has rank $q < k$. Then there exists a vector $\lambda \neq 0 \in \mathbb{R}^k$ such that $d\Gamma(x^*)^T\lambda = 0$. This is because, for a matrix A, one has $\dim(\text{Range}(A)) = \dim(\text{Range}(A^T))$ and

$$\dim(\text{Range}(A^T)) + \dim(\text{Null}(A^T)) = k,$$

so that

$$\dim(\text{Null}(A^T)) = k - q > 0\,.$$

This proves the theorem with $\rho = 0$. Conversely, if the theorem holds with $\rho = 0$, then $d\Gamma(x^*)$ cannot have rank k.

Now assume that $d\Gamma(x^*)$ has rank k (which is its maximal rank since $k \leq d$). We note that it suffices to show that, whenever $v \in \mathbb{R}^k$ satisfies $d\Gamma(x^*)v = 0$, then $dF(x^*)v = 0$. If this is true, then a standard result in linear algebra (the proof of which being left to the reader) states that one must have $dF(x^*)^T + d\Gamma(x^*)^T\lambda = 0$ for some $\lambda \in \mathbb{R}^k$, which is exactly the statement of the theorem with $\rho = 1$ (note that $d\Gamma(x^*)$ is a k by d matrix and $dF(x^*)$ a 1 by d matrix).

So let us prove that

$$d\Gamma(x^*)v = 0 \Rightarrow dF(x^*)v = 0.$$

We note that one can, without loss of generality, assume that $\Gamma(x) = x_2 - Q(x_1)$ where x_2 is composed with k coordinates from x and x_1 contains the remaining $d - k$. Indeed, since $d\Gamma(x^*)$ has rank k, the implicit function theorem implies that the equation $\Gamma(x) = 0$ can be solved (in some neighborhood of x^*) by expressing k coordinates of x as functions of the remaining ones. Defining the linear projections $\pi_1(x) = x_1$, $\pi_2(x) = x_2$, we will therefore assume that $\Gamma = \pi_2 - Q \circ \pi_1$. In such a case,

$$d\Gamma(x^*)v = \pi_2(v) - dQ(x_1^*)\pi_1(v),$$

which vanishes if and only if $\pi_2(v) = dQ(x_1^*)\pi_1(v)$.

Fix such a v. Consider the function $t \mapsto x(t)$ with $x_1(t) = x_1^* + t\pi_1(v)$ and $x_2(t) = Q(x_1^* + t\pi_1(v))$. Then $\Gamma(x(t)) = 0$ for all t in a neighborhood of $t = 0$ and the function $f(t) = F(x(t))$ is differentiable and minimal at $t = 0$, so that

$$\dot{f}(0) = dF(x^*)\dot{x}(0) = 0.$$

But we have $\pi_1(\dot{x}(0)) = \pi_1(v)$ and

$$\pi_2(\dot{x}(0)) = dQ(x_1^*)\pi_1(v) = \pi_2(v),$$

which implies that $\dot{x}(0) = v$. We have therefore proved that $dF(x^*)v = 0$, which completes the proof of the theorem when $d\Gamma(x^*)$ has full rank. □

Another way to interpret the previous theorem is that the row vectors of the matrix

$$\begin{pmatrix} dF(x^*) \\ d\Gamma(x^*) \end{pmatrix}$$

must be linearly dependent if x^* is a minimizer. The possibility that $\lambda = 0$ is important to keep in mind. It is easy to provide examples for which this occurs. One can, for example, transform any minimization problem such as (D.2) into the problem

$$x^* = \operatorname*{argmin}_{x \in V} F(x) \tag{D.3}$$

subject to $|\Gamma(x)|^2 = 0$. Both problems clearly have the same solutions, but, letting $\tilde{\Gamma}(x) = |\Gamma(x)|^2$, one has $d\tilde{\Gamma}(x) = 2\Gamma(x)^T d\Gamma(x)$, which vanishes for any x that satisfies the constraints. This means that no solution x^* will satisfy $dF(x^*) + \lambda d\tilde{\Gamma}(x^*)$ unless $dF(x^*) = 0$, i.e., x^* is already stationary for F, which is not true in general. To take a less artificial example, let $d = 3$, $k = 2$ with $F(x, y, z) = x/2 + y$ and $\Gamma(x, y, z) = (x^2 - y^2, y - z^2)$. Noting that $\Gamma = 0$ implies that $y = |x|$, one checks that $F(x, y, z) = |x| + x/2 \geq 0$ and vanishes only when $x = y = 0$, in which case $z = 0$. At this point, $dF(0) = (1/2, 1, 0)$,

$$d\Gamma(0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and there is no $\lambda \in \mathbb{R}^2$ such that $dF(0) + \lambda^T d\Gamma(0) = 0$.

Theorem D.3 can be generalized to infinite dimensions, but some precautions must be taken. Two aspects of the proof cannot be immediately transcribed within a Banach space setting. The first one is the separation of the variable x into two parts x_1 and x_2, but this can be fixed by using a quotient space argument instead. The second one is the argument about ranks and null spaces that was applied in the first part of the proof, which now requires the additional assumption, always true in finite dimension, that the range of $d\Gamma(x^*)$ is closed. The closed range assumption means that, if a sequence v_n is such that $w_n = d\Gamma(x^*)v_n$ converges to a limit w, then there exists a v such that $w = d\Gamma(x^*)v$. When this is true, we have the following theorem (see [179]).

Theorem D.4 *Let $\mathbb{B}$ and $\tilde{\mathbb{B}}$ be Banach spaces, and $F : \mathbb{B} \to \mathbb{R}$, $\Gamma : \mathbb{B} \to \tilde{\mathbb{B}}$ be Fréchet differentiable. Assume that x^* is a solution of* (D.2)*, and that* $\text{Range}(d\Gamma(x^*))$ *is a closed subspace of $\tilde{\mathbb{B}}$. Then there exists $\rho \in \{0, 1\}$ and $\lambda \in \tilde{\mathbb{B}}^*$, with $(\rho, \lambda) \neq (0, 0)$, such that x^* is a stationary point of the Lagrangian*

$$L(x) = \rho F(x) + (\lambda \mid \Gamma(x)).$$

If $d\Gamma(x^*)$ is onto, i.e., $\text{Range}(d\Gamma(x^*)) = \tilde{\mathbb{B}}$*, then one has $\rho = 1$.*

D.3 An Optimal Control Problem

When discussing control systems, one usually introduces two types of variables. The first ones, called state variables, contain quantities that must be driven to a desired behavior, while the second ones, called control variables, can be specified by the

user in order to influence the state variables. Both variables evolve with time. The relationship between the state, q, and the control, u, is typically represented by an evolution equation of the form

$$\partial_t q = f(q, u). \tag{D.4}$$

We will let $\mathcal{U}$ denote the set of controls, and Ω the set of states. The former will be assumed to form a Banach space, and the latter an open subset of a Banach space $\mathcal{Q}$. We will also make assumptions ensuring that the state evolution equation is a well-posed ODE in q for fixed u.

We will discuss the problem of finding a control that drives the state variable from an initial position, say q_0 at time $t = 0$, to (or near) a target state, q_1, at a fixed time, $t = 1$, with a minimal cost according to a cost function that will be specified. This is not the only kind of optimal control problems that one can consider. For example, one can relax the fixed time constraint and find an optimal way to reach q_1 at some time $t > 0$, where this time is also part of the optimization process. One can also generalize the unique target setting and rather search for a control that will lead the state to (or near) a target set of states, $\mathcal{Q}_1 \subset \Omega$, any element $q_1 \in \mathcal{Q}_1$ being an acceptable end point. A particular case of this situation will be considered later.

Restricting to the unique target setting, we will consider two types of problems. Both problems assume the definition of a cost function, defined on $\Omega \times \mathcal{U}$, denoted $g : (q, u) \mapsto g(q, u)$. The first optimal control (OC) problem that we consider attempts to drive q to q_1 exactly, and is formulated as follows. Find

$$(q^*, u^*) = \operatorname*{argmin}_{q,u} \int_0^1 g(q(t), u(t))dt \tag{D.5}$$

subject to $q(0) = q_0$, $q(1) = q_1$, $\partial_t q = f(q, u)$.

The second problem, which is both theoretically and numerically easier to address, replaces the constraint $q(1) = q_1$ by an endpoint cost $G(q(1))$, where G, defined on Ω, will penalize large differences between $q(1)$ and q_1. This yields the formulation

$$(q^*, u^*) = \operatorname*{argmin}_{q,u} \left(\int_0^1 g(q(t), u(t))dt + G(q(1)) \right) \tag{D.6}$$

subject to $q(0) = q_0$, $\partial_t q = f(q, u)$. For both problems, the control is assumed to be a bounded measurable function from $[0, 1]$ to $\mathcal{U}$, i.e., $u \in \mathbb{U} := L^\infty([0, 1], \mathcal{U})$.

We now formalize our assumptions. For $u \in \mathcal{U}$, define

$$\begin{aligned} f^u : \quad & \Omega \to \mathcal{Q} \\ & q \mapsto f(q, u). \end{aligned}$$

We assume that f^u takes values in $\mathbb{B}^{(1)} := C^1_{(0)}(\Omega, \mathcal{Q})$ and that the mapping $u \mapsto f^u$ is continuously differentiable from $\mathcal{U}$ to $\mathbb{B}^{(1)}$. We will furthermore assume that

$$\max(\|f^u\|_{\mathbb{B}^{(1)}}, \|\partial_u f^u\|_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})}) \le C(\|u\|_{\mathcal{U}}), \tag{D.7}$$

where C is a continuous function. It is easy to check that these assumptions imply that $f \in C^1(\Omega \times \mathcal{U}, Q)$, with $\partial_2 f(q,u) = \partial_u f^u(q)$, and that the cross derivative exists with $\partial_1\partial_2 f(q,u) = d\partial_u f^u(q) = \partial_u df^u(q) = \partial_2\partial_1 f(q,u)$. We then have the following proposition, in which we let $\mathbb{V}^{(1)} := L^1([0,1], \mathbb{B}^{(1)})$.

Proposition D.5 *With the above notation and assumptions, let $\mathscr{V}^u(t)(q) = \mathscr{V}^u(t,q) = f(q,u(t))$. Then, for $u \in \mathbb{U}$, $\mathscr{V}^u$ takes its values in $\mathbb{V}^{(1)}$ and $u \mapsto \mathcal{V}^u$ is continuously differentiable from $\mathbb{U}$ to $\mathbb{V}^{(1)}$, with*

$$\partial_u \mathscr{V}^u \eta : t \mapsto \partial_u f^{u(t)} \eta(t).$$

(Note that u, $\mathscr{V}^u$ etc. are defined up to modifications over sets of measure 0.)

Proof We have

$$\|\mathcal{V}^u\|_{\mathbb{V}^{(1)}} = \int_0^1 \|f^{u(t)}\|_{\mathbb{B}^{(1)}}\, dt,$$

which is finite if $u(\cdot) \in \mathbb{U}$, from (D.7). The same equation ensures that

$$\begin{aligned}\int_0^1 \|\partial_u \mathscr{V}^u \eta(t)\|_{\mathbb{B}^{(1)}}\, dt &\le \int_0^1 \|\partial_u f^{u(t)}\|_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})} \|\eta(t)\|_{\mathcal{U}}\, dt \\ &\le \|\eta\|_{\mathbb{U}} \int_0^1 C(\|u(t)\|_{\mathcal{U}})\, dt\end{aligned}$$

and the last integral is finite because C is continuous and u is bounded. This shows that $\partial_u \mathcal{V}^u \in \mathcal{L}(\mathbb{U}, \mathbb{V}^{(1)})$. The continuity of $\partial_u \mathscr{V}^u$ with respect to u can be assessed by writing

$$\begin{aligned}&\int_0^1 \|\partial_u f^{u(t)}\eta(t) - \partial_u f^{u'(t)}\eta(t)\|_{\mathbb{B}^{(1)}}\, dt \\ &\qquad \le \int_0^1 \|\partial_u f^{u(t)} - \partial_u f^{u'(t)}\|_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})} \|\eta(t)\|_{\mathcal{U}}\, dt \\ &\qquad \le \|\eta\|_{\mathbb{U}} \int_0^1 \|\partial_u f^{u(t)} - \partial_u f^{u'(t)}\|_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})}\, dt\end{aligned}$$

so that

$$\|\partial_u \mathcal{V}^u - \partial_u \mathcal{V}^u\|_{\mathrm{op}(\mathbb{U},\mathbb{V}^{(1)})} \le \int_0^1 \|\partial_u f^{u(t)} - \partial_u f^{u'(t)}\|_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})}\, dt,$$

which tends to 0 as u' tends to u (using the continuity of $\partial_u f^u$ and the dominated convergence theorem).

Finally, we write

$$f^{u(t)+\eta(t)} - f^{u(t)} = \int_0^1 \partial_u f^{u(t)+\varepsilon\eta(t)} \eta(t)\, d\varepsilon$$

and

$$\int_0^1 \|f^{u(t)+\eta(t)} - f^{u(t)} - \partial_u f^{u(t)}\eta(t)\|_{\mathbb{B}^{(1)}} dt \leq$$
$$\|\eta\|_{\mathbb{U}} \int_0^1 \int_0^1 \|\partial_u f^{u(t)+\varepsilon\eta(t)} - \partial_u f^{u(t)}\|_{op(\mathcal{U},\mathbb{B}^{(1)})}\, d\varepsilon dt = o(\|\eta\|_{\mathbb{U}}),$$

using the fact that $\partial_u f^u$ is continuous. This concludes the proof. □

If we now apply Theorem C.17 to the solution q^u of $\partial_t q = f(q,u)$ with $q^u(0) = q_0^u$, we find that

$$\partial_u q^u(t) = \int_0^t d\varphi_{st}^u(q^u(s))\partial_2 f(q^u(s), u(s))\eta(s) ds,$$

where φ^u is the flow associated with $\partial_t q = \mathcal{V}^u(t,q)$.

Remark D.6 One can prove a similar result with $\mathbb{U} = L^2([0,1],\mathcal{U})$ instead of L^∞ if we replace (D.7) by

$$\begin{aligned} \|f^u\|_{\mathbb{B}^{(1)}} &\leq C\|u\|_{\mathcal{U}} \\ \|\partial_u f^u\|_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})} &\leq C \end{aligned} \tag{D.8}$$

for some constant C. The proof follows the same lines, with the Cauchy–Schwartz inequality replacing direct bounding of integrals using supremum norms. To prove that quantities such as

$$\int_0^1 \|\partial_u f^{u(t)} - \partial_u f^{u'(t)}\|^2_{\mathrm{op}(\mathcal{U},\mathbb{B}^{(1)})}\, dt$$

tend to 0 (as u' tends to u in L^2), it suffices to use the fact that, if g is a bounded function, continuous at 0 with $g(0) = 0$, then

$$\int_0^1 g(\eta(t))\, dt \to 0$$

as $\eta \to 0$ in L^2, which can easily be proved as a consequence of Chebyshev's inequality.

D.3.1 Soft Endpoint Condition

Assume that g is C^1 as a function of q and u, and G as a function of q. Then, noting that (D.6) can be written as

$$u^* = \operatorname*{argmin}_u F(u)$$

with $F(u) = \int_0^1 g(q^u(t), u(t))dt + G(q^u(1))$, we find that an optimal solution must satisfy $(dF(u) \mid \eta) = 0$ for all $\eta \in L^\infty([0,1], \mathcal{U})$ with

$$\begin{aligned}
&(dF(u) \mid \eta) \\
&= \int_0^1 \Big(\partial_1 g(q^u(t), u(t)) \Big| \int_0^t d\varphi_{st}^u(q^u(s)) \partial_2 f(q^u(s), u(s)) \eta(s) ds \Big) dt \\
&\quad + \int_0^1 \big(\partial_2 g(q^u(t), u(t)) \mid \eta(t) \big) \big) \, dt \\
&\quad + \Big(dG(q^u(1)) \Big| \int_0^1 d\varphi_{s1}^u(q^u(s)) \partial_2 f(q^u(s), u(s)) \eta(s) ds \Big) \\
&= \int_0^1 \Big(\int_s^1 \partial_2 f(q^u(s), u(s))^* d\varphi_{st}^u(q^u(s))^* \partial_1 g(q^u(t), u(t)) dt \Big| \eta(s) \Big) ds \\
&\quad + \int_0^1 \big(\partial_2 g(q^u(t), u(t)) \mid \eta(t) \big) \big) \, dt \\
&\quad + \int_0^1 \Big(\partial_2 f(q^u(s), u(s))^* d\varphi_{s1}^u(q^u(s))^* dG(q^u(1)) \Big| \eta(s) \Big) ds \\
&= - \int_0^1 \big(\partial_2 f(q^u(s), u(s))^* p(s) \mid \eta(s) \big) ds + \int_0^1 \big(\partial_2 g(q^u(t), u(t)) \mid \eta(t) \big) \big) \, dt,
\end{aligned}$$

where we have set

$$\begin{aligned}
p(s) &= - \int_s^1 d\varphi_{st}^u(q^u(s))^* \partial_1 g(q^u(t), u(t)) \, dt - d\varphi_{s1}^u(q^u(s))^* dG(q^u(1)) \\
&= -d\varphi_{0s}^u(q_0)^{-*} \int_s^1 d\varphi_{0t}^u(q_0)^* \partial_1 g(q^u(t), u(t)) \, dt \\
&\qquad -d\varphi_{0s}^u(q_0)^{-*} d\varphi_{01}^u(q_0)^* dG(q^u(1)).
\end{aligned} \tag{D.9}$$

Here we have used the identity $d\varphi_{0t}^u = d\varphi_{st}^u \circ \varphi_{0s}^u \, d\varphi_{0s}^u$ (and the "$-*$" exponent represents the inverse of the conjugate operator). We therefore obtain the fact that

$$dF(u)(t) = -\partial_2 f(q^u(t), u(t))^* p(t) + \partial_2 g(q^u(t), u(t)).$$

We now reinterpret this results in a form that will lead us to Pontryagin's maximum principle. The function $p : [0, 1] \to \mathcal{Q}^*$ in (D.9) is continuous and differentiable a.e. From (C.8), we have

$$\begin{aligned}\partial_t (d\varphi^u_{0t}(q_0)^{-*}) &= -d\varphi^u_{0t}(q_0)^{-*}\, \partial_t (d\varphi^u_{0t}(q_0)^*)\, d\varphi^u_{0t}(q_0)^{-*} \\ &= -\partial_1 f(q^u(t), u(t))^*\, d\varphi^u_{0t}(q_0)^{-*}.\end{aligned}$$

Using this identity, one can easily check that, for an optimal control,

$$\partial_t p(t) = -\partial_1 f(q^u(t), u(t))^* p(t) + \partial_1 g(q^u(t), u(t))$$

with $p(1) = -dG(q^u(1))$.

Introduce the family of *Hamiltonian functions* $H_u : \mathcal{Q}^* \times \Omega \to \mathbb{R}, u \in \mathcal{U}$, defined by

$$H_u(p, q) = (p \mid f(q, u)) - g(q, u)\,. \tag{D.10}$$

We have $\partial_1 H_u = f(q, u)$ and $\partial_2 H_u = \partial_1 f^* p - \partial_1 g$. Moreover, we have $\partial_u H_u = \partial_2 f^* p(t) - \partial_2 g$. We can therefore summarize the results we have obtained in the following theorem.

Theorem D.7 *If $(q(\cdot), u(\cdot))$ is a solution of (D.6), there exists a time-dependent linear form $p(\cdot) \in \mathcal{Q}^*$, called the co-state, such that the following conditions are satisfied.*

$$\begin{cases} \partial_t q(t) = \partial_1 H_{u(t)}(p(t), q(t)) \\ \partial_t p(t) = -\partial_2 H_{u(t)}(p(t), q(t)) \\ \partial_u H_{u(t)} = 0 \end{cases} \tag{D.11}$$

with the boundary condition $p(1) = -dG(q(1))$.

A stronger result is actually true. For a control to be optimal, one must have, at all times:

$$u(t) = \operatorname*{argmax}_u (H_u(p(t), q(t))),$$

which implies the last condition in (D.11). The resulting statement is *Pontryagin's maximum principle*, and we have proved, via Theorem D.7, a weak form of this principle. Proving the strong form would require more work [5, 148, 183, 297], which will not be needed in our case, because, in the applications we will consider, H_u will always be strictly convex in u, so that $\partial_u H_u = 0$ will in fact be equivalent to u maximizing H_u.

Note that we have also proved that, for any u (not necessary optimal), one has

$$dF(u)(t) = -\partial_2 f(q^u(t), u(t))^* p(t) + \partial_2 g(q^u(t), u(t)), \tag{D.12}$$

where p and q satisfy $q(0) = q_0$, $p(1) = -dG(q(1))$ and

$$\begin{cases} \partial_t q(t) = \partial_1 H_{u(t)}(p(t), q(t)) \\ \partial_t p(t) = -\partial_2 H_{u(t)}(p(t), q(t)). \end{cases} \tag{D.13}$$

D.3.2 Fixed Endpoint Condition

Problem D.5 is a constrained minimization problem, requiring us to minimize

$$F_0(u) = \int_0^1 g(q^u(t), u(t))\, dt$$

subject to the constraint $\Gamma(u) = q^u(1) - q_1 = 0$. We have

$$d\Gamma(u)\eta = d\varphi_{01}^u(q_0) \int_0^1 d\varphi_{0t}^u(q_0)^{-1} \partial_2 f(q^u(t), u(t)\eta(t)\, dt\,. \tag{D.14}$$

To be able to apply Theorem D.4, we need to ensure that $d\Gamma(u)$ has closed range. When this holds, then we can state that, if u is an optimal control, there exists a $\lambda \in \mathcal{Q}^*$ such that

$$\rho d F_0(u) + d\Gamma(u)^* \lambda = 0 \tag{D.15}$$

with $\rho \in \{0, 1\}$ and $(\rho, \lambda) \neq (0, 0)$. We can then make exactly the same computation as before, replacing g by ρg and $dG(q^u(1))$ by λ. This leads to the following theorem.

Theorem D.8 *Assume that $(q(\cdot), u(\cdot))$ is a solution of (D.5), and that $d\Gamma(u)$ has closed range, where $\Gamma(u) = q^u(1) - q_1$. Then there exists a time-dependent linear form $p(\cdot) \in \mathcal{Q}^*$, called the co-state, such that the following conditions are satisfied.*

$$\begin{cases} \partial_t q(t) = \partial_1 H_{u(t)}(p(t), q(t)) \\ \partial_t p(t) = -\partial_2 H_{u(t)}(p(t), q(t)) \\ \partial_u H_{u(t)} = 0, \end{cases} \tag{D.16}$$

where

$$H_u(p, q) = (p \mid f(q, u)) - \rho g(q, u)\,, \tag{D.17}$$

with $\rho \in \{0, 1\}$ and $(\rho, p(1)) \neq (0, 0)$.

Here again, the stronger result $u(t) = \operatorname{argmax} H_u(p, q)$ is true. Solutions for which $\rho = 1$ are called *normal solutions* of the optimal control problem, and those for which $\rho = 0$ are referred to as *abnormal solutions*. The case excluded from this theorem, for which $d\Gamma(u)$ does not have closed range, or, more generally, for which (D.15) has no non-trivial solution, has been referred to in [17] as an *elusive solution*.

Propositions D.9 and D.11 provides two cases for which the former result applies.

Proposition D.9 *Theorem D.8 holds if $\mathcal{Q}$ has finite dimension.*

(This is obvious, because the range of $d\Gamma(u)$ is a linear subspace of $\mathcal{Q}$.)

Definition D.10 For a state-control pair (q, u) satisfying $\partial_t q = f(q, u)$, define the linearized control system

$$\partial_t h(t) = \partial_1 f(q(t), u(t))h(t) + \partial_2 f(q(t), u(t))\eta(t). \tag{D.18}$$

One says that (q, u) is a regular trajectory if the system (D.18) is controllable, i.e., for any h_1, there exists a pair (h, η) satisfying (D.18) such that $h(0) = 0$ and $h(1) = h_1$.

We then have the following.

Proposition D.11 ([179]) *If an optimal trajectory (q, u) is regular, then $d\Gamma(u)$ is onto and Theorem D.8 holds with $\rho = 1$.*

Proof The solution of (D.18) is given by

$$h(t) = d\varphi^u_{0t}(q_0) \int_0^t d\varphi^u_{0s}(q_0)^{-1} \partial_2 f(q^u(s), u(s))\eta(s)\, ds\,.$$

Since $h(1) = d\Gamma(u)\eta$, (q, u) being regular is indeed equivalent to $d\Gamma(u)$ being onto.

An important special case is when $\mathcal{U} = \mathcal{Q}$ and $f(q, u) = u$ (so that $u = \dot{q}$), which provides the standard formulation of a problem in the Calculus of Variations. Then any (q, u) is regular because one can always take $\eta(t) = h_1$, which gives $h(t) = th_1$ in (D.18), since $\partial_1 f = 0$. The same statement clearly applies to the more general setting in which $\partial_2 f$ is invertible.

To conclude this section, we mention that Pontryagin's maximum principle holds under more general assumptions than the ones we have considered here, with a proof significantly more complicated. The interested reader can refer to textbooks such as [183] or [5].

D.3.3 Control from the Initial Condition

We now consider the problem of minimizing $G(q(1))$ with respect to q_0, subject to the condition $\partial_t q = f(t, q)$, $q(0) = q_0$, or more directly

$$F(q_0) = G(\varphi_{01}(q_0)),$$

where φ is the flow associated with $\partial_t q = f(t,q)$. This can be addressed with an approach similar (and simpler) than the optimal control problem studied in Sect. D.3.1. We have

$$\begin{aligned}(dF(q_0) \mid h) &= (dG(q(1)) \mid d\varphi_{01}(q_0)h) \\ &= \left(d\varphi_{01}(q_0)^* dG(q(1)) \,\middle|\, h\right) \\ &= -(p(0) \mid h)\end{aligned}$$

with

$$p(t) = -d\varphi_{0t}^{-*}(q_0) d\varphi_{01}^*(q_0) dG(q(1)).$$

Using a computation made in Sect. D.3.1, we have

$$\partial_t p(t) = -\partial_2 f(t, q(t))^* p(t)$$

with $p(1) = -dG(q(1))$. We therefore have the following proposition.

Proposition D.12 *The differential of $F(q_0) = G(q(1))$, where q is the solution of $\partial_t q = f(t,q)$ with $q(0) = q_0$ is given by $dF(q_0) = -p(0)$ where $p(t)$ is the solution of $\partial_t p(t) = -\partial_2 f(t, q(t))^* p(t)$ with $p(1) = -dG(q(1))$.*

We could have saved on minus signs by replacing p by $-p$, but we wanted to maintain a similarity with the statement of the maximum principle. Indeed, another way to state the proposition is that the differential of F is computed by solving

$$\begin{cases} \partial_t q = \partial_1 H_t(p,q) \\ \partial_t p = -\partial_2 H_t(p,q), \end{cases}$$

where $H_t(p,q) = (p \mid f(t,q))$, with boundary conditions $q(0) = q_0$ and $p(1) = -dG(q(1))$.

D.4 Time Discretization and Discrete Maximum Principle

We now discuss how the time variable in (D.6) can be discretized, and the associated variational problem. Fixing a positive integer, n, we let $h = 1/n$ and $t_k = kh$ for $k = 0, \ldots, n$. We assume that the control is also a discrete function, $u(t) = u_k$ on $[t_k, t_{k+1})$ for $k = 0, \ldots, n-1$, so that

$$\int_0^1 g(q(t), u(t))dt = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} g(q(t), u_k)dt.$$

Choose an explicit approximation scheme for $\partial_t q = f(t,u)$ in the form $q_{k+1} = \eta(q_k, u_k)$. For example, $\eta(q,u) = q + hf(q,u)$, which corresponds to a simple Euler scheme. Finally, approximate

$$\int_{t_k}^{t_{k+1}} g(q(t), u_k)dt \simeq \gamma(q_k, q_{k+1}, u_k),$$

using, for example, the trapezoidal rule

$$\gamma(q, q', u) = \frac{h}{2}(g(q,u) + g(q',u)).$$

This results in the discrete-time optimal control problem minimizing, letting $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $\boldsymbol{u} = (u_0, \ldots, u_{n-1})$,

$$F(\boldsymbol{q}, \boldsymbol{u}) = \sum_{k=0}^{n-1} \gamma(q_k, q_{k+1}, u_k) + G(q_n)$$

subject to $q_{k+1} = \eta(q_k, u_k)$ (q_0 being given). Since there is a finite number of constraints, whose differential, forming a triangular system, has full range, minimizers must be stationary solutions of

$$\Lambda(\boldsymbol{q}, \boldsymbol{u}) = \sum_{k=0}^{n-1} \gamma(q_k, q_{k+1}, u_k) + G(q_n) + \sum_{k=0}^{n-1} (\lambda_k \mid q_{k+1} - \eta(q_k, u_k)),$$

where $\lambda_0, \ldots, \lambda_{n-1} \in \mathcal{Q}^*$ are Lagrange multipliers. This provides the equations

$$\begin{cases} q_{k+1} = \eta(q_k, u_k), \quad k = 0, \ldots, n-1 \\ \lambda_{n-1} = -dG(q_n) - \partial_2 \gamma(q_{n-1}, q_n, u_{n-1}) \\ \lambda_{k-1} = \partial_1 \eta(q_k, u_k)^* \lambda_k - \partial_1 \gamma(q_k, q_{k+1}, u_k) - \partial_2 \gamma(q_{k-1}, q_k, u_{k-1}), \\ \qquad k = 0, \ldots, n-1 \\ \partial_3 \gamma(q_k, q_{k+1}, u_k) - \partial_2 \eta(q_k, u_k)^* \lambda_k = 0, \quad k = 0, \ldots, n-1. \end{cases} \tag{D.19}$$

We retrieve the Pontryagin maximum principle in a discrete form (λ_k being the co-state). In particular, if $\tilde{F}(\boldsymbol{u})$ is the objective function when $\boldsymbol{q}$ is considered as a function of $\boldsymbol{u}$, then

$$\partial_k \tilde{F}(\boldsymbol{u}) = \partial_3 \gamma(q_k, q_{k+1}, u_k) - \partial_2 \eta(q_k, u_k)^* \lambda_k$$

for $k = 0, \ldots, n-1$, where q and λ satisfy the first three equations of (D.19).

For example, for the simplest choices of η and γ, namely $\eta(q,u) = q + hf(q,u)$ and $\gamma(q, q', u) = hg(q,u)$, (D.19) becomes

$$\begin{cases} q_{k+1} = q_k + hf(q_k, u_k), \quad k = 0, \dots, n-1 \\ \lambda_{n-1} = -dG(q_n) \\ \lambda_{k-1} = \lambda_k + h\partial_1 f(q_k, u_k)^* \lambda_k - h\partial_1 g(q_k, u_k), \quad k = 0, \dots, n-1 \\ \partial_2 g(q_k, u_k) - \partial_2 f(q_k, u_k)^* \lambda_k = 0, \quad k = 0, \dots, n-1. \end{cases} \tag{D.20}$$

D.5 Optimization Algorithms

Similar to our discussion of numerical algorithms for ordinary differential equations (ODEs), we give in this chapter a very limited and very partial account of numerical optimization methods, limiting ourselves to the few concepts and algorithms that are used in this book, and inviting the reader to consult one of many available textbooks (e.g., [40, 112, 221]) for more information.

D.5.1 Directions of Descent and Line Search

Since the problems we consider in this book are nonlinear, and very often non-convex, we will consider the general problem of finding a (local) minimizer of a function $x \mapsto E(x)$, defined on an $\mathbb{R}^n$ and taking values in $\mathbb{R}$ without making any specific assumption on E except that it is sufficiently differentiable. We will discuss iterative methods that update a current value of x by first finding a good direction, $h \in \mathbb{R}^n$, then replacing x by $x + \varepsilon h$ for a suitably chosen ε. We will assume, in the following, that E has enough derivatives for the computations to make sense, without repeating the exact assumptions every time.

The minimal requirement for h is that it must be a direction of descent, i.e., there must exist an $\varepsilon_0 > 0$ such that $E(x + \varepsilon h) < E(x)$ for $\varepsilon \in (0, \varepsilon_0)$. Once h is given (and most of our discussion will be on how to find a good h) determining ε is a one-dimensional operation which is usually referred to as a line search. If we expand $E(x + \varepsilon h)$ to first order in ε, we get

$$E(x + \varepsilon h) = E(x) + \varepsilon(dE(x) \mid h) + o(\varepsilon)$$

and we see that a sufficient condition for h to be a direction of descent is that $(dE(x) \mid h) < 0$. We will only consider directions h that satisfy this condition, and therefore always assume that it holds when speaking of directions of descent.

For such an h, we find that, for any $\gamma \in (0, 1)$, the expression

$$c(\varepsilon, x, h) = \frac{1}{\varepsilon}\big(E(x + \varepsilon h) - E(x) - \gamma\varepsilon(dE(x) \mid h)\big)$$

converges to a non-vanishing negative number when $\varepsilon \to 0$. Given this, one devises a simple line search procedure as follows. Fix two constants $\alpha, \gamma \in (0, 1)$. Let x and a descent direction h be given. Start with some reasonably large value of ε, say, $\bar{\varepsilon}$, and replace ε by $\alpha\varepsilon$ iteratively, as long as $c(\varepsilon, x, h) \geq 0$. This is the backtracking line search technique, and $c(\varepsilon, x, h) < 0$ is called the Armijo rule.

The choice of $\bar{\varepsilon}$ is important because choosing it too large would require trying too many values of ε before finding a satisfactory one, and taking it too small may prevent the algorithm from making large steps. One possibility is to start a minimization procedure with some choice of $\bar{\varepsilon}_0$, and choose at step k of the procedure $\bar{\varepsilon}_k$ to be some multiple of $\bar{\varepsilon}_{k-1}$ by some factor larger than one. Let's summarize this in the following algorithm.

Algorithm 5 *(Generic minimization with backtracking line search) Start with an initial choice for x_0, $\bar{\varepsilon}_0$. Choose positive constants $\alpha, \gamma < 1$ and $\beta > 1$. Let x_k, $\bar{\varepsilon}_k$ be their current values at step k and obtain their values at the next step as follows.*

1. *Compute a good direction of descent h_k.*
2. *Set $\varepsilon_k = \bar{\varepsilon}_k$. While*

$$E(x_k + \varepsilon_k h_k) - E(x_k) - \gamma\varepsilon_k (dE(x_k) \mid h_k) > 0$$

   replace ε_k by $\alpha\varepsilon_k$.
3. *Set $x_{k+1} = x_k + \varepsilon_k h_k$ and $\bar{\varepsilon}_{k+1} = \beta\varepsilon_k$.*

One typically stops the algorithm if the decrease $(E(x_k + \varepsilon_k h_k) - E(x_k))/\varepsilon$ is smaller than a given threshold. Other, more elaborate line search methods can be devised, including a full optimization of the function $\varepsilon \mapsto E(x + \varepsilon h)$. One additional condition which is often imposed in the line search is the Wolfe condition, which ensures (taking $0 < \rho < 1$) that

$$(dE(x_{k+1}) \mid h_k) \geq \rho\, (dE(x_k) \mid h_k),$$

or the strong Wolfe condition, which is

$$|(dE(x_{k+1}) \mid h_k)| \leq \rho\, |(dE(x_k) \mid h_k)|,$$

which force the steps to be larger as long as the slope of E along the direction of descent remains significantly negative. But one must not forget that the line search needs to be repeated many times during the procedure and should therefore not induce too much computation (enforcing the Wolfe condition may be impractical if the computation of the gradient is too costly).

The key part of the algorithm is, however, the choice of the direction of descent, which is now addressed, starting with the simplest one.

D.5.2 Gradient Descent

A direction of descent at x being characterized by $(dE(x) \mid h) < 0$, a natural requirement is to try to find h such that this expression is as negative as possible. Of course, this requirement does not make sense unless some normalization is imposed on h, and this is based, for gradient descent, on the specification of an inner product that may depend on x.

More precisely, assume that, for all $x \in \mathbb{R}^n$, an inner product denoted $\langle \cdot\,,\,\cdot \rangle_x$ is selected (or, in other words, that a Riemannian metric is chosen on $\mathbb{R}^n$; smoothness of the dot product as a function of x is not a requirement, but it is needed, for example, in error estimation formulas like (D.22) below). In finite dimensions, this is equivalent to associating to each x a symmetric, definite positive matrix $A(x)$ and to setting

$$\langle u\,,\,v \rangle_x = u^T A(x) v.$$

Notice that the choice of the metric can be inspired by infinite-dimensional representations of the problem, but it is important to make sure that it remains positive once discretized.

The gradient of E at x for this metric, denoted $\nabla^A E(x)$, is defined by

$$\forall h \in \mathbb{R}^d,\ (dE(x) \mid h) = \langle \nabla^A E(x)\,,\,h \rangle_x.$$

If we denote by $\nabla E(x)$ the column vector representation of $dE(x)$, i.e., the gradient for the usual dot product (with $A(x) = \mathrm{Id}$), this definition implies

$$\nabla^A E(x) = A(x)^{-1} \nabla E(x). \tag{D.21}$$

A gradient descent procedure for the Riemannian metric associated to A selects, at a given point x, the direction of descent h as a minimizer of $(dE(x) \mid h)$ subject to the constraint $\langle h\,,\,A(x)h \rangle_x = 1$. The solution of this problem can readily be computed as

$$h = -\frac{\nabla^A E(x)}{|\nabla^A E(x)|_x}$$

(unless of course $\nabla^A E(x) = 0$, in which case no direction of descent exists).

Since directions of descent need only be defined up to a multiplicative factor, we may as well take $h = -\nabla^A E(x)$. This can be plugged into step 1 of Algorithm 5, which becomes: *Set* $h_k = -A(x_k)^{-1} \nabla E(x_k)$.

Gradient descent is, like all the methods considered in this discussion, only able to find local minima of the function E (one may even not be able to rule out the situation in which it gets trapped in a saddle point). The speed at which it finds it is linear [221], in the sense that, if x^* is the limit point, and if an exact line search algorithm is run, there exists $\rho \in (0, 1)$ such that, for large enough k,

$$E(x_{k+1}) - E(x_*) \leq \rho\,(E(x_k) - E(x_*)). \tag{D.22}$$

The smallest possible ρ in this formula depends on how "spherical" the function E is around its minimum. More precisely, if c^* is the ratio between the largest and the smallest eigenvalue (or condition number) of $A_*^{-1/2} d^2E(x^*) A_*^{-1/2}$ where $A_* = A(x^*)$, then ρ must be larger than $(c^* - 1)^2/(c^* + 1)^2$, so that c^* should be as close to 1 as possible.

The gradient descent algorithm can also be written in continuous time, namely

$$\partial_t x = -A(x)^{-1} \nabla E(x).$$

This is a convenient formulation, provided that one remembers that it is usually preferable to discretize it using a line search rather than using standard ODE methods (since the goal is not to solve the ODE, but to minimize E).

The time-continuous formulation can easily be extended to Riemannian manifolds. If M is such a manifold, and E is defined on M, the gradient descent algorithm runs

$$\partial_t x = -\nabla^M E(x).$$

When discretized, however, one must remember that additions do not make sense on nonlinear manifolds (even when they are submanifolds of Euclidean spaces). The Riemannian equivalent to moving along a straight line is to use geodesics [2], so that gradient descent should be discretized as

$$x_{k+1} = \mathrm{Exp}_{x_k}(-\varepsilon \nabla^M E(x_k))$$

where Exp is the Riemannian exponential (as defined in Eq. B.8).

D.5.3 Newton and Quasi-Newton Directions

If the condition number of $A_*^{-1/2} d^2E(x^*) A_*^{-1/2}$ is 1, which is equivalent to $A^* = d^2E(x^*)$, then ρ can be taken arbitrarily small in (D.22), and one says that the convergence is *superlinear*. This suggests using $A(x) = d^2E(x)$ in the metric, provided, of course, this matrix is positive (it is nonnegative at the minimum, but not everywhere if the function is not convex). The direction $-d^2E(x)^{-1} \nabla E(x)$ is called the Newton direction. It is optimal up to a second-order approximation, as shown by the following computation. For a given x, we have

$$E(x + h) = E(x) + \nabla E(x)^T h + \frac{1}{2} h^T d^2E(x) h + o(h^2).$$

If we neglect the error, and if $d^2E(x)$ is a positive matrix, the minimum of the second-order approximation is indeed given by

$$\hat{h} = -d^2E(x)^{-1}\nabla E(x).$$

If $d^2E(x)$ is not positive, the Newton direction is not necessarily a direction of descent. It has to be modified, the simplest approach being to add a multiple of the identity matrix to the Hessian, i.e., to use $A(x) = d^2E(x) + \lambda(x)\mathrm{Id}$ for a large enough $\lambda(x)$. Given x, the choice of a suitable λ can be based on ensuring that the Cholesky decomposition of $A(x)$ (i.e., the decomposition $A = LL^*$ where L is lower triangular and L^* the conjugate of the transpose of L [131]) only has real coefficients.

When the computation of the Hessian is too complex, or its inversion too costly, using a Newton direction is impractical. Alternative methods, called quasi-Newton, are available in that case. One such method, called BFGS (based on the initials of its inventors, who independently discovered the method [47, 79, 115, 130]), updates an approximation A_k of the Hessian at step k of the algorithm. Let x_k be the current variable at step k, and $h_k = -A_k^{-1}\nabla E(x_k)$ the associated direction of descent, computed using the current A_k. Then, the new value of x is $x_{k+1} = x_k + s_k$ with $s_k = \varepsilon_k h_k$. The BFGS method defines a matrix A_{k+1} for the next step as follows. Letting $y_k = \nabla E(x_{k+1}) - \nabla E(x_k)$, take

$$A_{k+1} = \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) A_k \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{y_k s_k^T}{y_k^T s_k}.$$

We refer the reader to [221] for a justification of this updating rule. It has the important feature to ensure that A_{k+1} is positive as soon as A_k is, and $s_k^T y_k > 0$. This last condition can be written

$$\nabla E(x_k + \varepsilon_k)^T h_k > \nabla E(x_k)^T h_k$$

and one can always find ε_k such that this is satisfied, unless the minimum of E is $-\infty$. This is because the left-hand side is $\partial_\varepsilon E(x_k + \varepsilon h_k)(\varepsilon_k)$, and

$$\partial_\varepsilon E(x_k + \varepsilon h_k) \leq \nabla E(x_k)^T h_k < 0$$

for all $\varepsilon > 0$ implies that $\lim_{\varepsilon\to\infty} E(x_k + \varepsilon h_k) = -\infty$. The condition $s_k^T y_k > 0$ can therefore be added to the line search procedure. It is automatically satisfied if the Wolfe condition is ensured.

Equally important to the fact that A_k can be made to remain positive is the fact that an inverse to it can be computed iteratively too. If we let $B_k = A_k^{-1}$, then

$$B_{k+1} = B_k - \frac{B_k y_k y_k^T B_k}{y_k^T B_k y_k} + \frac{s_k s_k^T}{y_k^T y_k}, \tag{D.23}$$

which allows for an efficient computation of the direction of descent, $h_k = -B_k\nabla E(x_k)$.

A variation of the method offers an even more efficient update rule, namely

$$B_{k+1} = V_k^T B_k V_k + \rho_k s_k s_k^T, \tag{D.24}$$

with $\rho_k = 1/y_k^T s_k$ and $V_k = \mathrm{Id} - \rho_k s_k s_k^T$.

The BFGS method is not directly applicable for large n (the dimension), however, because the computation and storage of n by n matrices like B_k would become impractical. One possibility is to use the iteration specified by (D.24) over a finite time interval only (say p iterations in the past), resetting the value of B_{k-p-1} to Id. The computation of $B_k h_k$ only requires storing the values of y_{k-j} and s_{k-j} for $j = 1, \ldots, p$, and can be done recursively using (D.24), with a storage and computation cost which is now linear in the dimension. This results in the limited-memory BFGS (L-BFGS) method.

D.5.4 Conjugate Gradient

Nonlinear conjugate gradient methods can be seen as intermediate in complexity and efficiency between basic gradient descent and quasi-Newton methods. They may provide a reasonable choice for very large-scale methods, for which even limited-memory quasi-Newton methods may be too costly.

Since non-linear conjugate gradient derives from the linear one, and since linear conjugate gradient is a method of choice for solving large-scale linear systems, we start with this linear case.

Linear Conjugate Gradient

The goal of the conjugate gradient method is to invert a linear system

$$Ax = b$$

where A is an n by n symmetric, positive definite matrix. This problem is equivalent to minimizing the quadratic function

$$E(x) = \frac{1}{2} x^T A x - b^T x.$$

Conjugate gradient works by generating a set of *conjugate directions*, $h_0, \ldots, h_p, \ldots$ that satisfy $h_i^T h_j = 0$ if $i \neq j$, and generate the sequence

$$x_{k+1} = x_k + \alpha_k h_k \tag{D.25}$$

in which α_k is the optimal choice for the minimization of $\alpha \mapsto E(x_k + \alpha h_k)$, namely

$$\alpha_k = -\frac{(Ax_k - b)^T h_k}{h_k^T A h_k}.$$

It is easy to prove by induction that, if $h_1, \dots, h_k$ are non-vanishing conjugate directions, then $Ax_{k+1} - b$ is orthogonal to $\text{span}(h_0, \dots, h_k)$ and, since this space has dimension $k+1$, this implies that $Ax_n - b = 0$ so that the algorithm converges to a solution of the linear system in at most n steps. This also implies that

$$E(x_{k+1}) = \min \left\{ E(x) : x = x_0 + t_1 h_1 + \cdot + t_k h_k, t_1, \dots, t_k \in \mathbb{R} \right\}$$

because $(Ax_{k+1} - b)^T h_j = \partial_{t_j} E(x_0 + t_1 h_1 + \cdot + t_k h_k)$.

It remains to describe how a good sequence $h_1, \dots, h_n$ can be specified. Gram–Schmidt orthogonalization is a standard process to build a conjugate family of vectors starting from linearly independent vectors $r_0, \dots, r_k$. It consists in setting $h_0 = r_0$ and, for $k \geq 1$,

$$h_k = r_k + \sum_{j=0}^{k-1} \mu_{kj} h_j$$

with

$$\mu_{kj} = -\frac{r_k^T A h_j}{h_j^T A h_j}.$$

The beautiful achievement made in conjugate gradient was to recognize that, if $r_k = Ax_k - b$ with x_k coming from (D.25) (which is algorithmically feasible), then $\mu_{kj} = 0$ except for $j = k-1$. To prove this, one needs to first check, by induction, that

$$\text{span}(h_0, \dots, h_k) = \text{span}(r_0, Ar_0, \dots, A^k r_0),$$

the latter space being called the Krylov space of order k associated to A and r_0. This implies that, for $j \leq k-1$,

$$Ah_j \in \text{span}(r_0, Ar_0, \dots, A^{j+1} r_0) = \text{span}(r_0, \dots, r_{j+1}) \subset \text{span}(r_0, \dots, r_k).$$

Since we know that r_k is perpendicular to this space, we have $r_k^T A h_j = 0$ if $j \leq k-1$.

Given these remarks, the following iteration provides a family of conjugate directions, starting with an initial x_0 and $h_0 = r_0 = Ax_0 - b$ (and letting $\beta_k = \mu_{k,k-1}$):

$$\begin{cases} x_{k+1} = x_k + \alpha_k h_k \text{ with } \alpha_k = -\dfrac{r_k^T h_k}{h_k^T A h_k}, \\ r_{k+1} = r_k + \alpha_k A h_k, \\ h_{k+1} = r_{k+1} + \beta_{k+1} h_k \text{ with } \beta_{k+1} = -\dfrac{r_{k+1}^T A h_k}{h_k^T A h_k}. \end{cases}$$

This is, in essence, the conjugate gradient algorithm. The computation can be made slightly more efficient by noticing that $r_k^T h_k = r_k^T r_k$ and $r_{k+1}^T A h_k = r_{k+1}^T r_{k+1}$ [131, 221].

The rate of convergence of conjugate gradient can be estimated from the eigenvalues of A. If $\lambda_1 \leq \cdots \leq \lambda_n$ are the ordered eigenvalues, and $\rho_k = \lambda_{n-k+1}/\lambda_1$, then,

$$|x_k - x_n| \leq \left(\frac{\rho_k - 1}{\rho_k + 1}\right) |x_0 - x_n| \tag{D.26}$$

(x_n being the final —and correct— state of the algorithm). So, if many eigenvalues of A are much larger than the smallest one, conjugate gradient will converge very slowly. One the other hand, if A has a few large eigenvalues, which then drop to being close to the smallest, then a few iterations will be needed to obtain a good approximation of the solution.

We could have formulated the initial problem in terms of any inner product on $\mathbb{R}^n$. Let M be symmetric, and positive definite, and consider the inner product $\langle x\ ,\ y\rangle_M = x^T My$. Let $\tilde{A}$ be self-adjoint for this product, i.e.,

$$\forall x, y,\ \langle x\ ,\ \tilde{A}y\rangle_M = \langle Ax\ ,\ y\rangle_M \text{ or } M\tilde{A} = \tilde{A}^T M.$$

Then solving $\tilde{A}x = \tilde{b}$ is equivalent to minimizing

$$\tilde{E}(x) = \frac{1}{2}\langle x\ ,\ \tilde{A}x\rangle_M - \langle \tilde{b}\ ,\ x\rangle_M$$

and the same argument we made for $M = \mathrm{Id}$ leads to the algorithm (taking $\tilde{r}_0 = \tilde{A}x_0 - \tilde{b}$)

$$\begin{cases} x_{k+1} = x_k + \alpha_k h_k \text{ with } \alpha_k = -\dfrac{\langle \tilde{r}_k\ ,\ \tilde{r}_k\rangle_M}{\langle h_k\ ,\ \tilde{A}h_k\rangle_M}, \\ \tilde{r}_{k+1} = \tilde{r}_k + \alpha_k \tilde{A}h_k, \\ h_{k+1} = \tilde{r}_{k+1} + \beta_{k+1} h_k \text{ with } \beta_{k+1} = -\dfrac{\langle \tilde{r}_{k+1}\ ,\ \tilde{r}_{k+1}\rangle_M}{\langle h_k\ ,\ \tilde{A}h_k\rangle_M}. \end{cases}$$

Now, we can remark that given any symmetric matrix A, we get a self-adjoint matrix for the M dot product by letting $\tilde{A} = M^{-1}A$ and that the problem $Ax = b$ is equivalent to $\tilde{A}x = \tilde{b}$ with $\tilde{b} = M^{-1}b$, which can be solved using the M dot product. Doing this leads to the iterations (in which we set $\tilde{r}_k = M^{-1}r_k$): start with $r_0 = Ax_0 - b$ and iterate

$$\begin{cases} x_{k+1} = x_k + \alpha_k h_k \text{ with } \alpha_k = -\dfrac{r_k^T M^{-1} r_k}{h_k^T A h_k}, \\ r_{k+1} = r_k + \alpha_k A h_k, \\ h_{k+1} = M^{-1} r_{k+1} + \beta_{k+1} h_k \text{ with } \beta_{k+1} = -\dfrac{r_{k+1}^T M^{-1} r_{k+1}}{h_k^T A h_k}. \end{cases}$$

This is preconditioned conjugate gradient. Its speed of convergence is now governed by the eigenvalues of $M^{-1}A$ (or, equivalently of $M^{-1/2}AM^{-1/2}$), and a lot of efficiency can be gained if most of these eigenvalues get clustered near the smallest one. Of course, for this to be feasible, the equation $M\tilde{r}_k = r_k$ has to be easy to solve. Preconditioning has most of the time to be designed specifically for a given problem, but it may result in a dramatic reduction of the computation time.

Nonlinear Conjugate Gradient

Now, assume that E is a nonlinear function. We can formally apply the conjugate gradient iterations by replacing r_k by $\nabla E(x_k)$, which yields, starting with $h_0 = -\nabla E(x_0)$,

$$\begin{cases} x_{k+1} = x_k + \varepsilon_k h_k, \\ h_{k+1} = \nabla E(x_{k+1}) + \beta_{k+1} h_k \\ \qquad \text{with } \beta_{k+1} = -\dfrac{|\nabla E(x_{k+1})|^2}{|\nabla E(x_k)|^2}. \end{cases}$$

This is the Fletcher–Reeves algorithm. In this algorithm, ε_k should be determined using a line search. This algorithm can significantly accelerate the convergence of gradient descent methods, especially when closing up to a minimum, around which E will be roughly quadratic. However, h_k is not guaranteed to always provide a direction of descent (this is true when the line search is exact, or under the strong Wolfe condition; see [221]). It may sometimes be useful to reinitialize the iterations at regular intervals, setting $h_{k+1} = -\nabla E(x_{k+1})$ (or, equivalently, $\beta_{k+1} = 0$).

Variants of this algorithm use different formulae for β_{k+1}. One of them is the Polak–Ribière algorithm, which sets

$$\tilde{\beta}_{k+1} = \frac{\nabla E(x_{k+1})^T (\nabla E(x_{k+1}) - \nabla E(x_k))}{|\nabla E(x_k)|^2}$$

and $\beta_{k+1} = \max(\tilde{\beta}_{k+1}, 0)$.

Appendix E
Principal Component Analysis

E.1 General Setting

Assume that a random variable X takes values in a finite- or infinite-dimensional Hilbert space H. Denote by $\langle \cdot\,,\,\cdot \rangle_H$ the inner product in this space. Assume that $x_1, \ldots, x_N$ are observed.

The goal of principal component analysis (PCA) is to provide, for each finite $p \leq \dim(H)$, an optimal representation of order p in the form

$$x_k = \bar{x} + \sum_{i=1}^{p} \alpha_{ki} e_i + R_k, \qquad k = 1, \ldots, N,$$

where $(e_1, \ldots, e_p)$ is an orthonormal family in H. The error terms, $R_1, \ldots, R_N$, should be as small as possible. More precisely, PCA minimizes the residual sum of squares

$$S = \sum_{k=1}^{N} \|R_k\|_H^2. \tag{E.1}$$

When $\bar{x}$ and $(e_1, \ldots, e_p)$ are fixed, $\sum_{i=1}^{p} \alpha_{ki} e_i$ must then be the orthogonal projection of $x_k - \bar{x}$ on $\text{span}(e_1, \ldots, e_p)$, which implies that

$$\alpha_{ki} = \langle x_k - \bar{x}\,,\, e_i \rangle_H.$$

Still assuming that $(e_1, \ldots, e_p)$ is fixed, it is easy to prove that the optimal choice for $\bar{x}$ is $\bar{x} = \frac{1}{N} \sum_{k=1}^{N} x_k$. For notational simplicity, we assume that $\bar{x} = 0$, which is equivalent to assuming that all x_k's have been replaced by $x_k - \bar{x}$.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

From these results, $(e_1, \ldots, e_p)$ must minimize

$$S = \sum_{k=1}^{N} \|x_k - \sum_{i=1}^{p} \langle x_k\ ,\ e_i\rangle e_i\|_H^2$$
$$= \sum_{k=1}^{N} \|x_k\|_H^2 - \sum_{i=1}^{p}\sum_{k=1}^{N} \langle x_k\ ,\ e_i\rangle_H^2.$$

For $u, v \in H$, define

$$\langle u\ ,\ v\rangle_T = \frac{1}{N}\sum_{k=1}^{N} \langle x_k\ ,\ u\rangle_H \langle x_k\ ,\ v\rangle_H$$

and $\|u\|_T = \langle u\ ,\ u\rangle_T^{1/2}$ (the index T refers to the fact that this norm is associated to a training set). This provides a new quadratic form on H. The formula above shows that minimizing S is equivalent to maximizing

$$\sum_{i=1}^{p} \|e_i\|_T^2$$

subject to the constraint that $(e_1, \ldots, e_p)$ is orthonormal in H.

If μ is a square integrable probability measure on H, such that $\sigma_\mu^2 = \int_H \|x\|_H^2\, d\mu(x) < \infty$, one can more generally define the *covariance bilinear form*

$$\Gamma_\mu(u, v) = \int_H \langle u\ ,\ x\rangle_H \langle v\ ,\ x\rangle_H\, d\mu(x),$$

which is bounded, with $\Gamma_\mu(u, v) \le \sigma_\mu^2 \|u\|_H\, \|v\|_H$. With this notation,

$$\langle u\ ,\ v\rangle_T = \Gamma_{\hat{\mu}_T}(u, v),$$

where $\hat{\mu}_T = (1/N)\sum_{k=1}^{N} \delta_{x_k}$ is the empirical measure. The generalization of the previous problem then requires us to maximize

$$\sum_{k=1}^{p} \Gamma_\mu(e_k, e_k)$$

over all orthonormal families $(e_1, \ldots, e_p)$ in H.

When μ is square integrable, the associated operator, A_μ, defined by $\langle u\ ,\ A_\mu v\rangle_H = \Gamma_\mu(u, v)$ for all $u, v \in H$, is a *Hilbert–Schmidt operator* [306]. Such an operator can, in particular, be diagonalized in an orthonormal basis of H, i.e., there exists an orthonormal basis, $(f_1, f_2, \ldots)$ of H such that $A_\mu f_i = \lambda_i f_i$ for a non-decreasing sequence of eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge 0$.

We have the following result, which we state without proof:

Theorem E.1 *The optimal* $(e_1, \ldots, e_p)$ *must be such that span*$(e_1, \ldots, e_p) =$ *span*$(f_1, \ldots, f_p)$. *In particular* $f_1, \ldots, f_p$ *provide a solution. They are called the* p *principal components of the dataset* $(x_1, \ldots, x_N)$.

The result can be applied to $\mu = \hat{\mu}_T$. In this case, with this choice, S in (E.1) is such that

$$S = N \sum_{i>p} \lambda_i^2.$$

The interest of discussing PCA associated with a covariance operator for a square integrable measure (in which case it is often called a Karhunen–Loeve (KL) expansion) is that it is often important when discussing infinite-dimensional random processes (such as Gaussian random fields). They moreover quite naturally correspond to asymptotic version of sample-based PCA. Interesting issues, that are part of *functional data analysis* [239], address the design of proper estimation procedures to obtain converging estimators of KL expansions based on finite samples for stochastic processes in infinite-dimensional spaces.

E.2 Computation of the Principal Components

E.2.1 Small Dimension

Assume that $H = \mathbb{R}^d$ has finite dimension, d, and that $x_1, \ldots, x_N \in \mathbb{R}^d$ are column vectors. Let the inner product on H be associated to a matrix Q:

$$\langle u\ ,\ v \rangle_H = u^T Q v.$$

Writing $A_T = A_{\hat{\mu}_T}$, and introducing the covariance matrix of the data

$$\Sigma_T = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^T,$$

one can then identify $A_T = \Sigma_T Q$. Eigenvectors f of A_T are such that $Q^{1/2} f$ are eigenvectors of the symmetric matrix $Q^{1/2} \Sigma_T Q^{1/2}$, which shows that they form an orthogonal system in H, which will be orthonormal if the eigenvectors are normalized so that $f^T Q f = 1$.

E.2.2 Large Dimension

It often happens that the dimension of H is much larger than the number of observations, N. In such a case, the previous approach is quite inefficient (especially when the dimension of H is infinite!) and one proceeds as follows.

The basic observation is that there are at most N principal components $f_1, \dots, f_N$ with non-vanishing eigenvalues, and they must belong to the vector space generated by $x_1, \dots, x_N$, so that, for some α_{ik}, $1 \le i, k \le N$:

$$f_i = \sum_{k=1}^{N} \alpha_{ik} x_k.$$

With this notation, we have $\langle f_i\ ,\ f_j\rangle_H = \sum_{k,l=1}^{N} \alpha_{ik}\alpha_{jl}\langle x_k\ ,\ x_l\rangle_H$ and

$$\langle f_i\ ,\ f_j\rangle_T = \sum_{l=1}^{N} \langle f_i\ ,\ x_l\rangle_H \langle f_j\ ,\ x_l\rangle_H = \sum_{k,k'=1}^{N} \alpha_{ik}\alpha_{jk'} \sum_{l=1}^{N} \langle x_k\ ,\ x_l\rangle_H \langle x_{k'}\ ,\ x_l\rangle_H.$$

Let S be the Gram matrix formed by the inner products $\langle x_k\ ,\ x_l\rangle_H$, for $k, l = 1, \dots, N$. We have $\langle f_i\ ,\ f_j\rangle_H = \alpha_i^T S \alpha_j$ and $\langle f_i\ ,\ f_j\rangle_T = \alpha_i^T S^2 \alpha_j$, which implies that, in this representation, the operator A_T is also given by S. Thus, the previous simultaneous orthogonalization problem can be solved with respect to the α's by diagonalizing S and taking the first eigenvectors, normalized so that $\alpha_i^T S \alpha_i = 1$.

E.3 Statistical Interpretation and Probabilistic PCA

We take here $H = \mathbb{R}^d$ with the standard inner product. The statistical interpretation of linear PCA is quite simple: assume that X is a centered random vector with covariance matrix Σ. Consider the problem that consists in finding a decomposition $X = \sum_{i=1}^{p} \xi_i e_i + R$ where $(\xi_1, \dots, \xi_p)$ forms a p-dimensional centered random vector, $e_1, \dots, e_p$ is an orthonormal system, and R is a random vector, uncorrelated to the ξ_i's and as small as possible, in the sense that $E(|R|^2)$ is minimal. One can see that, in an optimal decomposition, one needs $R^T e_i = 0$ for all i, because one can always write

$$\sum_{i=1}^{p} \xi_i e_i + R = \sum_{i=1}^{p} (\xi_i + R^T e_i) e_i + R - \sum_{i=1}^{p} (R^T e_i) e_i$$

and $|R - \sum_{i=1}^{p} (R^T e_i) e_i| \le |R|$. Also, one can always restrict oneself to uncorrelated $(\xi_1, \dots, \xi_p)$ by a change of basis in $\mathrm{span}(e_1, \dots, e_p)$.

Then, we can write

$$E(|X|^2) = \sum_{i=1}^{p} E(\xi_i^2) + E(|R|^2),$$

with $\xi_i = e_i^T X$. So, to minimize $E(|R|^2)$, one needs to maximize

$$\sum_{i=1}^{p} E((e_i^T X)^2),$$

which is equal to

$$\sum_{i=1}^{p} e_i^T \Sigma e_i.$$

The solution for this problem is given by the first p eigenvectors of Σ. PCA exactly applies this procedure, with Σ replaced by the empirical covariance.

Probabilistic PCA is based on the statistical model in which it is assumed that X can be decomposed as $X = \sum_{i=1}^{p} \lambda_i \xi_i e_i + \sigma R$, where R is a d-dimensional standard Gaussian vector and $\xi = (\xi_1, \ldots, \xi_p)$ a p-dimensional standard Gaussian vector, independent of R. The parameters here are the coordinates of $e_1, \ldots, e_p$, the values of $\lambda_1, \ldots, \lambda_p$ and of σ. Introduce the $d \times p$ matrix W with columns given by $\lambda_1 e_1, \ldots, \lambda_p e_p$ to rewrite this model in the form

$$X = W\xi + \sigma^2 R,$$

where the parameters are W and σ^2, with the constraint that $W^T W$ is diagonal. As a linear combination of independent Gaussian random variables, X is Gaussian with covariance matrix $WW^T + \sigma^2 I$. The log-likelihood of the observation $x_1, \ldots, x_N$ is

$$L(W, \sigma) = -\frac{N}{2}(d \log 2\pi + \log \det(WW^T + \sigma^2 I) + \operatorname{trace}((WW^T + \sigma^2 I)^{-1} \Sigma)),$$

where Σ is the empirical covariance matrix of $x_1, \ldots, x_N$. This function can be maximized explicitly in W and σ. One can show that the solution can be obtained by taking $e_1, \ldots, e_p$ to be the first p eigenvectors of Σ, $\lambda_i = \sqrt{\delta_i^2 - \sigma^2}$, where δ_i^2 is the eigenvalue of Σ associated to e_i, and

$$\sigma^2 = \frac{1}{d-p} \sum_{i=p+1}^{d} \delta_i^2.$$

Appendix F
Dynamic Programming

This chapter describes a few basic dynamic programming algorithms.

F.1 Minimization of a Function on a Tree

We consider the issue of minimizing a function $E : x \mapsto E(x)$, defined for all x of the form $x = (x_s, s \in S)$, in the following context:

- S is a finite set which forms the vertices of an oriented tree. We will represent edges in this tree by relations $s \to t$, for some pairs (s, t) in S. The assumption is that, for any $s, s' \in S$, there exists at most one path between s and s' (such that $s = s_0 \to s_1 \to \cdots \to s_N = s'$). For $s \in S$, we denote the product space by $\mathcal{V}_s$ the set of all children of s, i.e., the set of all t such that $s \to t$.
- For all $s \in S$, x_s belongs to a finite set A_s. We denote the product space $\prod_{s \in S} A_s$ by A.
- The function E takes the form $E(x) = \sum_{s,t \in S, s \to t} E_{st}(x_s, x_t)$.

We will consider the following partial order on S: $s < t$ if there exists a sequence $s = s_0 \to s_1 \to \cdots \to s_p = t$. For all $s \in S$, define

$$E_s^+(x) = \sum_{t \in \mathcal{V}_s} E_{st}(x_s, x_t) + \sum_{t > s, u > s, t \to u} E_{tu}(x_t, x_u).$$

Clearly, $E_s^+(x)$ only depends on x_s, and x_t for $t > s$. One can furthermore prove the relation

$$E_s^+(x) = \sum_{t \in \mathcal{V}_s} \left(E_{st}(x_s, x_t) + E_t^+(x) \right). \tag{F.1}$$

Let $F_s^+(x_s) = \min\{E_s^+(y), y \in A, y_s = x_s\}$. The following equation, which is a consequence of (F.1), essentially describes the dynamic programming algorithm.

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

$$F_s^+(x_s) = \min_{x_t, t \in \mathcal{V}_s} \left[\sum_{t \in \mathcal{V}_s} \left(E_{st}(x_s, x_t) + F_t^+(x_t) \right) \right]. \tag{F.2}$$

This implies that, in order to compute the values of $F^+(x_s)$, it suffices to know $F^+(x_t)$ for all $t \in \mathcal{V}_s$, and all x_t. This yields an algorithm that successively minimizes the functions F_s^+, starting from the leaves of the tree (the minimal elements for our order) to the roots (the maximal elements). If $\mathcal{R}$ is the set of roots, we have

$$\min_{x \in A} E(x) = \sum_{s \in \mathcal{R}} \min_{x_s \in A_s} [F_s^+(x_s)]$$

so that this algorithm directly provides a minimum of E. (Notice that we allow "trees" with multiple roots, sometimes called "forests" in the literature.)

There are practical limitations to this approach. First, the minimization involved in (F.2) should not be too difficult, because it has to be replicated at every s. This means that the product space of all A_t for $t \in \mathcal{V}_s$ must not be too large. Second, the memory load should remain tractable; for a given s, the value of $F_s^+(x_s)$ must be kept in memory for all $x_s \in A_s$, until the values of F_t^+ have been computed for the parent of s. One does not need to keep track of the configuration $x_t, t > s$, that achieves this minimum because the optimal configuration can be reconstituted if all the values of $F_s^+(x_s)$ have been saved. Indeed, if this is done, the optimal configuration can be reconstructed by starting with the roots of the tree, and keeping the best x_s at each step.

One can, however, devise upgraded versions of the algorithm for which the values and solutions of some of the $F^+(x_s)$ can be pruned a priori, based on lower bounds on the best value of E they can lead to.

F.2 Shortest Path on a Graph: a First Solution

The search for a shortest path on a graph is important for the methods discussed in this book, because it can be used to compute geodesics on discretized manifolds.

The set S now forms the vertices of a general oriented graph. A cost $\Gamma(s, t)$ is attributed to each edge $s \to t$; the global cost of a path $\boldsymbol{s} = (s_0 \to \cdots \to s_N)$ is defined by

$$E(\boldsymbol{s}) = \sum_{k=1}^{N} \Gamma(s, t).$$

Given s and t in S, we want to compute the shortest path (or path with lowest cost) $\boldsymbol{s} = (s = s_0 \to \cdots \to s_N = t)$. The variables therefore are $s_1, \ldots, s_{N-1}$, and the integer N. The optimal cost will be denoted $d(s, t)$. To avoid infinite minima, we will assume that there exists at least one path from s to t, for any $s, t \in S$.

We fix t and consider s as an additional variable. We denote by $\mathcal{V}_s$ the set of all $u \in S$ such that $s \to u$. We then have the formula, which is the analog of (F.2) in this case:

$$d(s,t) = \min_{u \in \mathcal{V}_s} [d(u,t) + \Gamma(s,u)]. \tag{F.3}$$

This equation provides an algorithm to compute $d(s,t)$ for all $s \in S$. Define, for all $N \geq 0$,

$$d_N(s,t) = \min \{E(\boldsymbol{s}) : \boldsymbol{s} \text{ path between } s \text{ and } t \text{ with at most } N \text{ points}\}.$$

We have

$$d_{N+1}(s,t) = \min_{u \in \mathcal{V}_s} [d_N(u,t) + \Gamma(s,u)]. \tag{F.4}$$

Let $d_0(s,t) = +\infty$ if $s \neq t$ and 0 if $s = t$. Equation (F.4) can be iterated to provide $d_N(u,t)$ for all u and arbitrary N. We have $d(s,t) = \lim_{N \to \infty} d_N(s,t)$, but it is clear that, if $d_N(s,t) = d_{N+1}(s,t)$ for all s, then $d_N(s,t) = d(s,t)$ and the computation is over.

This provides the distance. To also obtain the shortest path between s and t once all distances $d(u,t)$, $u \in S$, have been computed, it suffices to start the path at s and iteratively choose the neighbor of the current point that is closest to t. More precisely, one lets $s_0 = s$, and, given s_k, take s_{k+1} such that $d(s_{k+1}, t) = \min\{d(u,t), u \in \mathcal{V}_{s_k}\}$.

F.3 Dijkstra's Algorithm

Dijkstra's algorithm provides an alternative, and significantly more efficient, method for the computation of shortest paths. This algorithm provides, for a given $s_0 \in S$, the shortest path between s_0 and any $t \in S$ with a number of operations of order $|S| \log |S|$.

To each step of the algorithm are associated two subsets of S; the first one is the set of all unresolved vertices, and will be denoted C; the second is the set of all vertices t for which $d(s_0,t)$ is known, and will be denoted D. Since $d(s_0,s_0) = 0$, the initial partition is $D = \{s_0\}$ and $C = S \setminus \{s_0\}$. We will also ensure that at each step of the algorithm

$$\max \{d(s_0,t) : t \in D\} \leq \min \{d(s_0,t), t \in C\},$$

which is true at the beginning.

Introduce the function F defined on S by:

$$F(t) = \begin{cases} d(s_0,t) \text{ if } t \in D; \\ \inf \{d(s,u) + \Gamma(u,t) : u \in D, u \to t\} \text{ if } t \in C, \text{ with } \inf(\emptyset) = +\infty. \end{cases}$$

We therefore start with $F(s_0) = 0$, $F(t) = \Gamma(s,t)$ for all $t \in \mathcal{V}_{s_0}$ and $F(t) = +\infty$ otherwise.

The basic observation is that, at each step, the vertex $t \in C$ for which $F(t)$ is minimal is such that $d(s_0,t) = F(t)$; indeed, we can already see that

$$d(s_0,t) = \inf\{d(s_0,u) + \Gamma(u,t) : u \in S\} \leq F(t).$$

Assume $d(s_0,t) < F(t)$. This implies that there exists a $u \in C$ such that $d(s_0,u) + \Gamma(u,t) < F(t)$. Choose an optimal path between s_0 and u and let u' denote the first time that this path leaves D. We have $F(u') = d(s_0,u') \leq d(s_0,u) < F(t)$, which is a contradiction to the fact that $F(t)$ is minimal. Moreover, for all $u \in C$, we have $d(s_0,u) \geq F(t)$: to prove this, it suffices to consider again the exit point from D of an optimal path between s_0 and u to have a contradiction.

One step of Dijkstra's algorithm therefore consists in identifying a minimizer, t, of F in C and moving it from C to D. The function is then updated, and only needs to be modified for points $t' \in \mathcal{V}_t$, for which $F(t')$ must be replaced by

$$\min(F(t'), F(t) + \Gamma(t,t')).$$

The algorithm stops when $C = \emptyset$. Similar to the previous section, the function F can be used to reconstruct optimal paths.

F.4 Shortest Paths in Continuous Spaces

We now provide an efficient algorithm to compute a certain class of geodesic distances in $\mathbb{R}^d$ [257], which can be seen as a numerically consistent version of Dijkstra's algorithm. Although the algorithm generalizes to any dimension, we restrict ourselves here to $d = 2$ in order to simplify the presentation.

Let W be a positive function and consider the Riemannian metric on $\mathbb{R}^2$ given by $|u|_x = W(x)|u|$, for $u \in \mathbb{R}^2$. The length of a path $\gamma(\cdot)$ defined over $[a,b]$ is therefore given by

$$L(\gamma) = \int_a^b |\dot{\gamma}(t)|\, W(\gamma(t))dt.$$

The geodesic distance is then given by the length of the shortest path between two points. Now let $f(x)$ denote the distance of x to a set $S_0 \subset \mathbb{R}^2$. The function f can be shown to satisfy the following *eikonal equation*

$$|\nabla f| = W \tag{F.5}$$

and the algorithm consists in solving (F.5) given that $f = 0$ on S_0.

The principle is to progressively update f starting from S_0. The norm of the gradient is discretized according to the following formulas. Let

$$\begin{aligned}
\partial_1^+ f(i,j) &= (f_{i+1j} - f_{ij})/h, \\
\partial_1^- f(i,j) &= (f_{ij} - f_{i-1j})/h, \\
\partial_2^+ f(i,j) &= (f_{ij+1} - f_{ij})/h, \\
\partial_2^- f(i,j) &= (f_{ij} - f_{ij-1})/h
\end{aligned}$$

and [248]

$$\begin{aligned}
|\nabla f(i,j)| = \Big(&\max\left(\max(\partial_1^- f(i,j), 0), -\min(\partial_1^+ f(i,j), 0)\right)^2 \\
&+ \max\left(\max(\partial_2^- f(i,j), 0), -\min(\partial_2^+ f(i,j), 0)\right)^2 \Big)^{1/2},
\end{aligned}$$

which can also be written

$$|\nabla f(i,j)| = \left((f_{ij} - \min(f_{ij}, U_{ij}))^2 + (f_{ij} - \min(f_{ij}, V_{ij}))^2\right)^{1/2}$$

with $U_{ij} = \min(f_{i-1j}, f_{i+1j})$ and $V_{ij} = \min(f_{ij-1}, f_{ij+1})$.

Before proceeding further, let us try to understand why this is a well-chosen formula, starting with a one-dimensional example, and restricting ourselves to the particular case $W \equiv 1$, which corresponds to the Euclidean distance function. In one dimension, it suffices to consider the case in which the set with respect to which the distance is computed has two points, say 0 and 1. We also focus on the interval [0, 1] in which interesting things happen. The distance function in this case is given by the first plot of Fig. F.1. The following two plots in the same figure provide a discretization of the same function over respectively an odd and an even number of points. For such a function, it is easy to check that

$$|f_i - \min(f_i, f_{i-1}, f_{i+1})| = 1$$

for every point; this partially justifies the choice made for the discretization of the eikonal equation. Approximating the gradient by central differences (for example, by $(f_{i+1} - f_{i-1})/2$) would not satisfy the equation at the central points, since the result would be 0 in the odd case, and ± 0.5 in the even case. Note that the formula

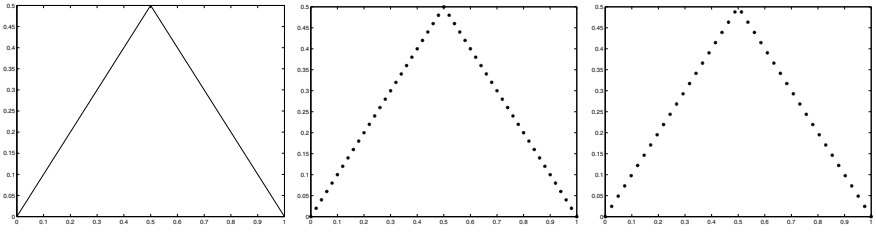

Fig. F.1 A typical distance map in 1D, followed by a discretization over an odd and even number of points

would have worked without including f_i in the minimum, but this is because, in one dimension, the situation in which f_i is smaller that its two neighbors does not happen, unless the distance is zero.

This can happen in two dimensions, however. Let us illustrate the formula in the simplest case when one computes the distance to a single point, say $(0, 0)$; that is, we consider the function $f(x, y) = \sqrt{x^2 + y^2}$. Let this function be discretized over a grid (ih, jh), $i, j \in \mathbb{Z}$, h being the discretization step. When i and j are non-vanishing, a first-order expansion shows that the norm of the previous estimation of the gradient is equal to 1 at first order in h (we skip the computation). If, say, $i = 0$, the provided approximation formula works exactly, since in this case $\min(f_{1j}, f_{0j}, f_{-1j})$ is equal to f_{0j} so that the approximation of the first derivative is zero, while the approximation of the other derivative is exactly 1. But this would not have worked if the term f_{ij} had been omitted in the minimum. The reader is referred to [255] for further justifications of this discretization.

We now solve the discretized eikonal equation,

$$(f_{ij} - \min(f_{ij}, U_{ij}))^2 + (f_{ij} - \min(f_{ij}, V_{ij}))^2 = W_{ij}^2,$$

with respect to f_{ij}. For this, it suffices to discuss the relative position of f_{ij}, U_{ij} and V_{ij}, and solve a second-degree polynomial equation in each case. The first observation is that, unless $W_{ij} = 0$, which is excluded, there is no solution, f_{ij}, which is smaller than both U_{ij} and V_{ij}. Let us consider the case $U_{ij} \leq f_{ij} \leq V_{ij}$. In such a case the equation is

$$(f_{ij} - U_{ij})^2 = W_{ij}^2,$$

which yields $f_{ij} = W_{ij} + U_{ij}$. To satisfy the current assumption $f_{ij} \leq V_{ij}$, we need $W_{ij} \leq V_{ij} - U_{ij}$. We have the same conclusion inverting the roles of U_{ij} and V_{ij}, and both cases can be summarized in the unique statement:

$$f_{ij} = W_{ij} + \min(U_{ij}, V_{ij}) \text{ if } W_{ij} \leq |V_{ij} - U_{ij}|.$$

To have f_{ij} larger that both U_{ij} and V_{ij}, we must solve

$$(f_{ij} - U_{ij})^2 + (f_{ij} - V_{ij})^2 = W_{ij}^2,$$

which yields

$$f_{ij} = \frac{1}{2}\left(U_{ij} + V_{ij} + \sqrt{2W_{ij}^2 - (U_{ij} - V_{ij})^2}\right).$$

One can prove that f_{ij} is larger than $\max(U_{ij}, V_{ij})$ only if $W_{ij} \geq |U_{ij} - V_{ij}|$ which is complementary to the previous case. To summarize

$$f_{ij} = \begin{cases} W_{ij} + \min(U_{ij}, V_{ij}) & \text{if } W_{ij} \leq |U_{ij} - V_{ij}|, \\ \\ \left(\sqrt{2W_{ij}^2 - (U_{ij} - V_{ij})^2} + U_{ij} + V_{ij}\right)/2 & \text{if } W_{ij} \geq |U_{ij} - V_{ij}|. \end{cases} \tag{F.6}$$

This will be called the *update formula*. From this formula, one can organize an iterative algorithm in which one progressively updates the f_{ij}'s one at a time according to this formula until stabilization [255]. Notice that, if, at some point, the values of U_{ij} and V_{ij} are exactly known, then the value of f_{ij} will also be exactly known after the application of the update formula.

The construction introduced in [257] makes it possible to organize the algorithm so that the correct values U_{ij} and V_{ij} are known at the time they are needed. This parallels Dijkstra's algorithm (Sect. F.3) and goes as follows.

Denote by S the complete discretized grid. This set will be divided into two subsets which evolve during the algorithm. These subsets are denoted C and D, the former containing the pixels which have not been updated yet, and the latter those for which the values of f are already known. At the beginning of the algorithm, D contains the points in S_0 (after discretization), f being 0 on D, and C contains all the other points, on which we temporarily set $f = \infty$. The algorithm stops when $C = \emptyset$.

Preliminary step. Compute the value of f for all points m which are neighbors of D according to the update formula (F.6).

Then iterate the main loop until C is empty:

Main loop. Select a point t in C for which f is minimal, and add it to D. Then recompute f for the neighbors of t according to (F.6). (Provided of course these neighbors do not already belong to D.)

Figure F.2 provides an example of a distance function to a plane curve, computed using this algorithm.

To show why the algorithm converges, we first note that if t_n is the point added to D at step n, and f^n the current function f at the same time, the sequence $f^n_{t_n}$ is increasing. Indeed, consider t_{n+1}. If t_{n+1} is not a neighbor of t_n, then the value of f is not updated after step n and $f^{n+1}_{t_{n+1}} = f^n_{t_{n+1}} \geq f^n_{t_n}$ since t_n was the minimum. So, if $f^{n+1}_{t_{n+1}} < f^n_{t_n}$, then t_n and t_{n+1} are neighbors, and the value of f at t_{n+1} must have changed; this implies that the new value at t_n must have affected $U_{t_{n+1}}$ or $V_{t_{n+1}}$ depending on whether t_n and t_{n+1} are horizontal or vertical neighbors. Say it has affected $U_{t_{n+1}}$. The only way this could have happened is when $U_{t_{n+1}} = f^n_{t_n}$. We are therefore in the situation when $f^{n+1}_{t_{n+1}} \leq U_{t_{n+1}}$, which implies $f^{n+1}_{t_{n+1}} = V_{t_{n+1}} + W_{t_{n+1}}$. But this value has not been affected by the previous step, since, for $V_{t_{n+1}}$ to be smaller than $f^n_{t_n}$, it must be equal to the value at a point which is already in D, which has not changed. This is a contradiction to $f^n_{t_{n+1}} < f^n_{t_n}$ and proves our initial statement.

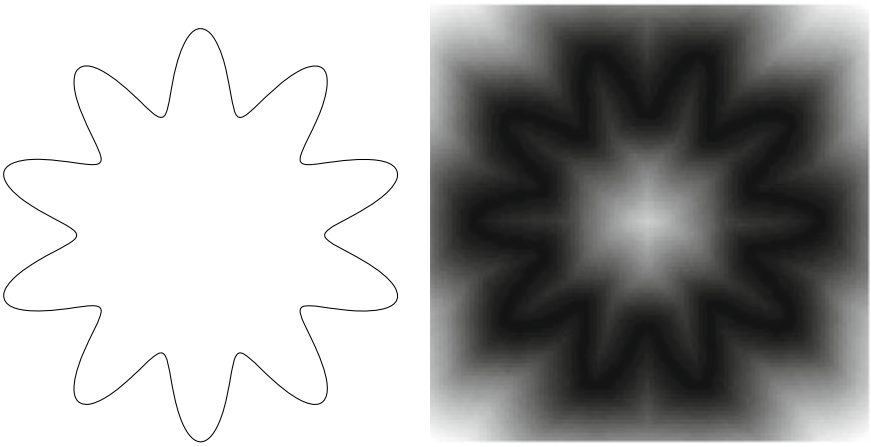

Fig. F.2 A curve and its distance map

Given this, any updated value after step n must be larger than the stored value at f_{t_n}. But this cannot affect the update formula, since this one only depends on the values of f at the neighbors which are smaller than f_{t_n}. So the update formula is satisfied, and remains so for all points added to D, which proves the convergence of the method.

References

1. R. Abraham, J.E. Marsden, *Foundations of Mechanics* (Perseus Publishing, Cambridge, 1994)
2. P.-A. Absil, R. Mahony, R. Sepulchre, *Optimization Algorithms on Matrix Manifolds* (Princeton University, Princeton, 2009)
3. R.A. Adams, *Sobolev Spaces* (Academic, New York, 1975)
4. B. Afsari, Riemannian l^p center of mass: existence, uniqueness, and convexity. Proc. Am. Math. Soc. **139**(2), 655–673 (2011)
5. A.A. Agrachev, Y. Sachkov, *Control Theory from the Geometric Viewpoint*, vol. 87 (Springer Science and Business Media, New York, 2013)
6. D.C. Alexander, J.C. Gee, R. Bajcsy, Strategies for data reorientation during non-rigid warps of diffusion tensor images, in *MICCAI* (1999), pp. 463–472
7. S. Allassonnière, Y. Amit, A. Trouvé, Towards a coherent statistical framework for dense deformable template estimation. J. R. Stat. Soc. B **69**(1), 3–29 (2007)
8. S. Allassonnière, E. Kuhn, Convergent stochastic expectation maximization algorithm with efficient sampling in high dimension. application to deformable template model estimation. Comput. Stat. Data Anal. **91**, 4–19 (2015)
9. L. Alvarez, P.-L. Lions, J.-M. Morel, Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal. **29**, 845–866 (1992)
10. L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, Axioms and fundamental equations of image processing. Arch. Ration. Mech. **123**(3), 199–257 (1993)
11. N. Amenta, M. Bern, Surface reconstruction by Voronoi filtering. Discret. Comput. Geom. **22**, 481–504 (1999)
12. Y. Amit, U. Grenander, M. Piccioni, Structural image restoration through deformable templates. JASA **86**, 376–387 (1989)
13. Y. Amit, P. Piccioni, A non-homogeneous Markov process for the estimation of Gaussian random fields with non-linear observations. Ann. Probab. **19**, 1664–1678 (1991)
14. Y. Amit, U. Grenander, M. Piccioni, Structural image restoration through deformable templates. J. Am. Stat. Assoc. **86**(414), 376–387 (1991)
15. N. Arad, N. Dyn, D. Reisfeld, Y. Yeshurun, Image warping by radial basis functions: application to facial expressions. CVGIP: Graph. Models Image Process. **56**(2), 161–172 (1994)
16. N. Arad, D. Reisfeld, Image warping using few anchor points and radial functions. Comput. Graph. Forum **14**, 35–46 (1995)
17. S. Arguillere, E. Trélat, Sub-Riemannian structures on groups of diffeomorphisms. J. Inst. Math. Jussieu **16**(4), 745–785 (2017)
18. V.I. Arnold, Sur un principe variationnel pour les ecoulements stationnaires des liquides parfaits et ses applications aux problèmes de stanbilité non linéaires. J. Mécanique **5**, 29–43 (1966)

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5

19. V.I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd edn. 1989 (Springer, Berlin, 1978)
20. N. Aronszajn, Theory of reproducing kernels. Trans. Am. Math. Soc. **68**, 337–404 (1950)
21. F. Arrate, J.T. Ratnanather, L. Younes, Diffeomorphic active contours. SIAM J. Imaging Sci. (2010) (to appear)
22. J. Ashburner, A fast diffeomorphic image registration algorithm. Neuroimage **38**(1), 95–113 (2007)
23. J. Ashburner, K.J. Friston, Voxel based morphometry - the methods. Neuroimage **11**(6), 805–821 (2000)
24. J. Ashburner, G.R. Ridgway, Symmetric diffeomorphic modeling of longitudinal structural MRI. Front. Neurosci. **6**, 197 (2013)
25. J. August, A. Tannebaum, S.W. Zucker, On the evolution of the skeleton, in *Proceedings of ICCV* (IEEE Computer Society, 1999), pp. 315–322
26. B. Avants, C.L. Epstein, J.C. Gee, Geodesic image normalization in the space of diffeomorphisms, in 1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy: Geometrical. Statistical and Registration Methods for Modeling Biological Shape Variability 125–135 (2006)
27. R. Bajcsy, C. Broit, Matching of deformed images, in *The 6th International Conference in Pattern Recognition* (1982), pp. 351–353
28. R. Bajcsy, S. Kovacic, Multiresolution elastic matching. Comput. Vis., Graph., Image Process. **46**, 1–21 (1989)
29. Martin Bauer, Martins Bruveris, Stephen Marsland, Peter W. Michor, Constructing reparameterization invariant metrics on spaces of plane curves. Differ. Geom. Appl. **34**, 139–165 (2014)
30. R.K. Beatson, L. Greengard, A short course on fast multipole methods, in *Wavelets, Multilevel Methods and Elliptic PDEs*, ed. by M. Ainsworth, J. Levesley, W.A. Light, M. Marlette (1997), pp. 1–37
31. M.F. Beg, Variational and computational methods for flows of diffeomorphisms in image matching and growth in computational anatomy. Ph.D. thesis, Johns Hopkins University (2003)
32. M.F. Beg, M.I. Miller, A. Trouvé, L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comput. Vis. **61**(2), 139–157 (2005)
33. S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts. IEEE Trans. PAMI **24**(24), 509–522 (2002)
34. J.-D. Benamou, Y. Brénier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math. **84**(3), 375–393 (2000)
35. K.K. Bhatia, P. Aljabar, J.P. Boardman, L. Srinivasan, M. Murgasova, S.J. Counsell, M.A. Rutherford, J.V. Hajnal, A.D. Edwards, D. Rueckert, *Groupwise combined segmentation and registration for atlas construction, in International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Berlin, 2007), pp. 532–540
36. K.K. Bhatia, J.V. Hajnal, B.K. Puri, A.D. Edwards, D. Rueckert, Consistent groupwise non-rigid registration for atlas construction, in *IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004* (IEEE, 2004), pp. 908–911
37. H. Blum, A transformation for extracting new descriptors of shape, in *Proceedings of Symposia Models for the Perception of Speech and Visual Form*, ed. by W. Dunn (MIR, 1967), pp. 362–380
38. S. Bochner, Integration von funktionen, deren werte die elemente eines vectorraumes sind. Fundamenta Mathmaticae **20**, 262–276 (1933)
39. J.-D. Boissonnat, M. Teillaud, *Effective Computational Geometry for Curves and Surfaces* (Springer, Berlin, 2007)
40. J.F. Bonnans, J.-C. Gilbert, C. Lemaréchal, C.A. Sagastizábla, *Numerical Optimization: Theoretical and Practical Aspects*, 2nd edn. (Springer, Berlin, 2009)
41. F.L. Bookstein, Principal warps: thin plate splines and the decomposition of deformations. IEEE Trans. PAMI **11**(6), 567–585 (1989)

42. F.L. Bookstein, *Morphometric Tools for Landmark Data; Geometry and Biology* (Cambridge University, Cambridge, 1991)
43. W. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry* (Academic, New York, 2002). Original edition 1986
44. Y. Brénier, Polar factorization and monotone rearrangement of vector-valued functions. Commun. Pure Appl. Math. **44**, 375–417 (1991)
45. H. Brezis, *Analyse Fonctionnelle, Théorie et Applications* (Masson, Paris) (Springer-Verlag, English translation, 1983)
46. Ethan S. Brown, Tony F. Chan, Xavier Bresson, Completely convex formulation of the chan-vese image segmentation model. Int. J. Comput. Vis. **98**(1), 103–121 (2012)
47. Charles G. Broyden, The convergence of a class of double-rank minimization algorithms: 2. the new algorithm. IMA J. Appl. Math. **6**(3), 222–231 (1970)
48. A.M. Bruckstein, E. Rivlin, I. Weiss, Scale space semi-local invariants. Image Vis. Comput. **15**, 335–344 (1997)
49. L. Caffarelli, The regularity of mappings with a convex potential. J. Am. Math. Soc. **5**, 99–104 (1992)
50. R. Camassa, D.D. Holm, An integrable shallow water equation with peaked solitons. Phys. Rev. Lett. **71**, 1661–1664 (1993)
51. V. Camion, L. Younes, Geodesic interpolating splines, in *EMMCVPR*, Lecture notes in computer science, ed. by M. Figueiredo, J. Zerubia, A.K. Jain, vol. 2134 (Springer, Berlin, 2001)
52. J. Canny, A computational approach to edge detection. IEEE PAMI **6**, 679–698 (1986)
53. F. Cao, *Geometric Curve Evolution and Image Processing* (Springer, Berlin, 2003)
54. F. Cao, J.-L. Lisani, J.-M. Morel, P. Musé, F. Sur, *A Theory of Shape Identification* (Springer, Berlin, 2008)
55. E. Cartan, *La Théorie des Groupes Finis et Continus et la Géométrie Différentielle Traitée par la Méthode du Repère Mobile*, originale Gauthiers-Villars 1937 edn. (Jacques Gabay, 1992)
56. V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, in *Proceedings of the 1995 ECCV* (1995)
57. S.L. Chan, E.O. Purisima, A new tetrahedral tesselation scheme for isosurface generation. Comput. Graph. **22**(1), 83–90 (1998)
58. T.F. Chan, L.A. Vese, Active contours without edges. IEEE Trans. Image Proc. **10**(2), 266–277 (2001)
59. Benjamin Charlier, Nicolas Charon, Alain Trouvé, The fshape framework for the variability analysis of functional shapes. Found. Comput. Math. **17**(2), 287–357 (2017)
60. B. Charlier, J. Feydy, D.W. Jacobs, A. Trouvé, Distortion minimizing geodesic subspaces in shape spaces and computational anatomy, in *VipIMAGE 2017: Proceedings of the VI ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing Porto, Portugal, October 18-20, 2017*, ed. by J.M.R.S. Tavares, R.M.N. Jorge (Springer International Publishing, Cham, 2018), pp. 1135–1144
61. Nicolas Charon, Alain Trouvé, The varifold representation of nonoriented shapes for diffeomorphic registration. SIAM J. Imaging Sci. **6**(4), 2547–2580 (2013)
62. I. Chavel, *Riemannian Geometry: A Modern Introduction*, vol. 98 (Cambridge University, Cambridge, 2006)
63. S.S. Chern, Curves and surfaces in euclidean space, in *Studies in Global Geometry and Analysis*, ed. by S.S. Chern (1967)
64. E.V. Chernyaev, Matching cubes 33: Construction of topologically correct isosurfaces. Technical report, CERN (1995). http://wwwinfo.cern.ch/asdoc/psdir/mc.ps.gz
65. H.-I. Choi, S.-W. Choi, H.-P. Moon, Mathematical theory of medial axis transform. Pac. J. Math. **181**(1), 57–88 (1997)
66. K.-S. Chou, X.-P. Zhu, *The Curve Shortening Problem* (Chapman and Hall, London, 2001)
67. G.E. Christensen, R.D. Rabbitt, M.I. Miller, Deformable templates using large deformation kinematics. IEEE Trans. Image Proc. **5**, (1996)
68. M.K. Chung, K.J. Worsley, T. Paus, C. Cherif, D.L. Collins, J.N. Giedd, J.L. Rapoport, A.C. Evans, A unified statistical approach to deformation-based morphometry. NeuroImage **14**(3), 595–606 (2001)

69. E.A. Coddington, N. Levinson., *Theory of Ordinary Differential Equations* (McGraw-Hill, USA, 1955)
70. L.D. Cohen, On active contours models and balloons. Comput. Vis., Graph. Image Process.: Image Underst. **53**(2), 211–218 (1991)
71. D. Cohen-Steiner, J.-M. Morvan, Second fundamental measure of geometric sets and local approximation of curvature. J. Differ. Geom. **74**, 363–394 (2006)
72. H. Cohn, *Conformal Mapping in Riemann Surfaces* (Dover Publications, New York, 1980)
73. T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models: their training and application. Comp. Vis. Image Underst. **61**(1), 38–59 (1995)
74. R. Courant, D. Hilbert, *Methods of Mathematical Physics, I* (Wiley, New York, 1955). Original edition: 1937
75. T.M. Cover, J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991)
76. J. Damon, Smoothness and geometry of boundaries associated to skeletal structures i: Sufficient conditions for smoothness. Ann. Inst. Fourier **53**, 1941–1985 (2003)
77. J. Damon, Smoothness and geometry of boundaries associated to skeletal structures ii: Geometry in the blum case. Compos. Math. **140**(6), 1657–1674 (2004)
78. J. Damon, Determining the geometry of boundaries of objects from medial data. Int. J. Comput. Vis. **63**(1), 45–64 (2005)
79. F. David, Shanno, Conditioning of quasi-newton methods for function minimization. Math. Comput. **24**(111), 647–656 (1970)
80. M.C. Delfour, J.-P. Zolésio, *Shapes and Geometries* (Analysis, Differential Calculus and Optimization) (SIAM, 2001)
81. Loïc Devilliers, Stéphanie Allassonnière, Alain Trouvé, Xavier Pennec, Template estimation in computational anatomy: Fréchet means top and quotient spaces are not consistent. SIAM J. Imaging Sci. **10**(3), 1139–1169 (2017)
82. F. Dibos, V. Caselles, T. Coll, F. Catté, Automatic contours detection in image processing, in *Proceedings of the First World Congress of Nonlinear Analysts*, ed. by V. Lakshmikantham (de Gruyter, 1996), pp. 1911–1921
83. J. Diestel, J.J. Uhl, *Vector Measures, Mathematical surveys and monographs* (American Mathematical Society, 1977)
84. J. Diestel, *Sequences and Series in Banach Spaces*, vol. 92 (Springer Science and Business Media, 2012)
85. Discrete differential geometry: An applied introduction, Course at SIGGRAPH'06 (2006)
86. M.P. Do Carmo, *Differential Geometry of Curves and Surfaces* (Prentice Hall, USA, 1976)
87. M.P. Do Carmo, *Riemannian Geometry* (Birkaüser, 1992)
88. J.R. Dormand, P.J. Prince, A family of embedded runge-kutta formulae. J. Comput. Appl. Math. **6**(1), 19–26 (1980)
89. T.R. Downie, L. Shepstone, B.W. Silverman, A wavelet approach to deformable templates, in *Image Fusion and Shape Variability Techniques*, ed. by K.V. Mardia, C.A. Gill, I.L. Dryden (Leeds University, 1996), pp. 163–169
90. M. Droske, M. Rumpf, A variational approach to non-rigid morphological registration. SIAM Appl. Math. **64**(2), 668–687 (2004)
91. I.L. Dryden, K.V. Mardia, *Statistical Shape Analysis* (Wiley, New York, 1998)
92. I.L. Dryden, K.V. Mardia, *Statistical Shape Analysis: With Applications in R* (Wiley, New York, 2016)
93. P. Dupuis, U. Grenander, M. Miller, *Variational problems on flows of diffeomorphisms for image matching* (Q. Appl, Math, 1998)
94. Stanley Durrleman, Xavier Pennec, Alain Trouvé, José Braga, Guido Gerig, Nicholas Ayache, Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. Int. J. Comput. Vis. **103**(1), 22–59 (2013)
95. S. Durrleman, M. Prastawa, J.R. Korenberg, S. Joshi, A. Trouvé, G. Gerig, *Topology preserving atlas construction from shape data without correspondence using sparse parameters, in International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Berlin, 2012), pp. 223–230

96. N. Dyn, Interpolation and approximation by radial and related functions, in *Approximation Theory VI*, ed. by C.K. Chui, L.L. Shumaker, J.D. Ward, vol. 1 (Academic, New York, 1989), pp. 211–234
97. G. Dziuk, E. Kuwert, R. Schäzle, Evolution of elastic curves in R^n: existence and computation. SIAM J. Math. Anal. **33**(5), 1228–1245 (2002)
98. E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration*, 2nd edn. (Springer, Berlin, 2006)
99. A. Edelman, T.A. Arias, S.T. Smith, The geometry of algorithms with orthogonality contraints. SIAM J. Matrix Anal. Appl. **20**(2), 303–353 (1998)
100. H. Edelsbrunner, *Geometry and Topology for Mesh Generation, Monographs on Applied and Computational Mathematics* (Cambridge Unversity, Cambridge, 2001)
101. C.L. Epstein, M. Gage, The curve shortening flow, in *Wave Motion*, ed. by A.J. Chorin, A.J. Madja (Springer, Berlin, 1987)
102. L.C. Evans, *Partial Differential Equations* (American Mathematical Society, 1998)
103. L.C. Evans, J. Spruck, Motion of level sets by mean curvature I. J. Differ. Geom. **33**, 635–681 (1991)
104. O. Faugeras, Cartan's moving frame method and its application to the geometry and evolution of curves, in *the euclidean, affine and projective plane, in Applications of Invariance in Computer Vision*, Lecture notes in computer sciences, ed. by J.L. Mundy, A. Zisserman, D. Forsyth, vol. 825 (Springer, Berlin, 1994), pp. 11–46
105. A.C. Faul, G. Goodsell, M.J.D. Powell, A Krylov subspace algorithm for multiquadric interpolation in many dimensions. IMA J. Numer. Anal. **25**, 1–24 (2005)
106. H. Federer, Curvature measures. Trans. Am. Math. Soc. **93**, 418–491 (1959)
107. H. Federer, *Geometric Measure Theory* (Springer, Berlin, 1969)
108. M. Fels, P.J. Olver, Moving coframes I. a practical algorithm. Acta Appl. Math. **51**, 161–213 (1998)
109. M. Fels, P.J. Olver, Moving coframes II. regularization and theoretical foundations. Acta Appl. Math. **55**, 127–208 (1999)
110. J. Fishbaugh, M. Prastawa, G. Gerig, S. Durrleman, Geodesic shape regression in the framework of currents, *Information processing in medical imaging: proceedings of the... conference*, vol. 23 (NIH Public Access, 2013), p. 718
111. M.A. Fishler, R.A. Elschager, The representation and matching of pictorial structures. IEEE Trans. Comput. **C–22**(73), 67–92 (1973)
112. B. Fletcher, *Practical Methods of Optimization*, 2nd edn. (Wiley, New York, 2000)
113. P.T. Fletcher, C. Lu, M. Pizer, S. Joshi, Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans. Med. Imaging 995–1005, (2004)
114. P.T. Fletcher, Geodesic regression and the theory of least squares on riemannian manifolds. Int. J. Comput. Vis. **105**(2), 171–185 (2013)
115. Roger Fletcher, A new approach to variable metric algorithms. Comput. J. **13**(3), 317–322 (1970)
116. C. Foias, D.D. Holm, E.S. Titi, The Navier-Stokes-alpha model of fluid turbulence. Phys. D **152**, 505–519 (2001)
117. M. Gage, An isoperimetric inequality with applications to curve shortening. Duke Math. J. **50**(4), 1225–1229 (1983)
118. M. Gage, R.S. Hamilton, The heat equation shrinking convex plane curves. J. Differ. Geom. **53**, 69–96 (1986)
119. W. Gangbo, R. McCann, The geometry of optimal transportation. Acta Math. **177**, 113–161 (1996)
120. L. Garcin, L. Younes, Geodesic image matching: a wavelet based energy minimization scheme. in *EMMCVPR'05*, pp. 349–364 (2005)
121. L. Garcin, L. Younes, Geodesic matching with free extremities. J. Math. Imaging Vis. **25**(3), 329–340 (2006)
122. François Gay-Balmaz, Darryl D. Holm, David M. Meier, Tudor S. Ratiu, François-Xavier Vialard, Invariant higher-order variational problems. Commun. Math. Phys. **309**(2), 413–458 (2012)

123. J.C. Gee, P.D. Peralta, Continuum models for Bayesian image matching, in *Maximum Entropy and Bayesian Methods*, ed. by K.H. Hanson, R.N. Silver (Kluwer Academic, Dordrecht, 1995)
124. P.J. Giblin, B.B. Kimia, *On the local form and transitions of symmetry sets, and medial axes, and shocks in 2D, in Proceedings of ICCV* (IEEE Computer Society, Greece, 1999), pp. 385–391
125. C.A. Glasbey, K.V. Mardia, A review of image-warping methods. J. Appl. Stat. **25**(2), 155–171 (1998)
126. J. Glaunès. Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique. Ph.D. thesis, University Paris 13, 2005. in French
127. J. Glaunès, A. Qiu, M.I. Miller, L. Younes, Large deformation diffeomorphic curve matching. Int. J. Comput. Vis. **80**(3), 317–336 (2008)
128. J. Glaunès, A. Trouvé, L. Younes, Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching (2004)
129. J. Glaunès, M. Vaillant, M.I. Miller, Landmark matching via large deformation diffeomorphisms on the sphere. J. Math. Imaging Vis. **20**, 179–200 (2004)
130. Donald Goldfarb, A family of variable-metric methods derived by variational means. Math. Comput. **24**(109), 23–26 (1970)
131. G.H. Golub, C.F. Van Loan, *Matrix Computations*, 3rd edn. (Johns Hopkins University, 1996)
132. A.A. Goshtasby, *Image Registration: Principles, Tools and Methods* (Springer Science and Business Media, 2012)
133. M. Grayson, The heat equation shrinks embedded plane curves to round points. J. Differ. Geom. **26**, 285–314 (1987)
134. U. Grenander, Lectures in Pattern Theory. Applied Mathematical Sciences, vol. 33, (1981)
135. U. Grenander, *General Pattern Theory* (Oxford Science Publications, Oxford, 1993)
136. U. Grenander, Y. Chow, D.M. Keenan, *Hands: A Pattern Theoretic Study of Biological Shapes* (Springer, Berlin, 1991)
137. U. Grenander, M.I. Miller, Computational anatomy: an emerging discipline. Q. Appl. Math. LV **I**(4), 617–694 (1998)
138. U. Grenander, M.I. Miller, *Pattern Theory: From Representation to Knowledge* (Oxford University, Oxford, 2007)
139. U. Grenander, M.I. Miller, *Pattern Theory: From Representation to Inference* (Oxford University, Oxford, 2007)
140. N.A. Gumerov, R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions* (Elsevier, 2004)
141. N.A. Gumerov, R. Duraiswami, Fast radial basis function interpolation via preconditioned Krylov iteration. SIAM J. Sci. Comput. **29**(5), 1876–1899 (2008)
142. S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces* (Academic, USA, 1978)
143. P.A. Helm, L. Younes, M.L. Beg, D.B. Ennis, C. Leclercq, O.P. Faris, E. McVeigh, K. Kass, M.I. Miller, R.L. Winslow, Evidence of structural remodeling in the dyssynchronous failing heart. Circ. Res. **98**(1), 125–132 (2006)
144. G. Hermosillo, Variational methods for multimodal matching. Ph.D. thesis, Université de Nice, 2002
145. G. Hermosillo, C. Chefd'Hotel, O. Faugeras, Variational methods for multimodal image matching. Int. J. Comput. Vis. **50**(3), 329–343 (2002)
146. Klaus Hildebrandt, Konrad Polthier, Max Wardetzky, On the convergence of metric and geometric properties of polyhedral surfaces. Geom. Dedicata **123**(1), 89–112 (2006)
147. C. Hilditch, Linear skeleton from square cupboards. Mach. Intell **6**, 403–409 (1969)
148. L.M. Hocking, *Optimal control: an introduction to the theory with applications* (Oxford University, Oxford, 1991)
149. D.D. Holm, *Geometric Mechanics* (Imperial College, 2008)
150. D.D. Holm, J.E. Marsden, T.S. Ratiu, The Euler-Poincaré equations and semidirect products with applications to continuum theories. Adv. Math. **137**, 1–81 (1998)

151. D.R. Holm, J.T. Ratnanather, A. Trouvé, L. Younes, Soliton dynamics in computational anatomy. Neuroimage **23**, S170–S178 (2004)
152. D.R. Holm, A. Trouvé, L. Younes, The euler poincaré theory of metamorphosis. Q. Appl. Math. (2009). (to appear)
153. Yi Hong, Sarang Joshi, Mar Sanchez, Martin Styner, Marc Niethammer, Metamorphic geodesic regression. Med. Image Comput. Comput.-Assist. Interv.-MICCAI **2012**, 197–205 (2012)
154. H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, W. Stuetzle, Surface reconstruction from unorganized points. in *ACM SIGGRAPH 1992* (1992)
155. A. Jain, Particle Methods for Diffeomorphic Registration. Ph.D. thesis, Johns Hopkins University, 2011
156. Aastha Jain, Laurent Younes, A kernel class allowing for fast computations in shape spaces induced by diffeomorphisms. J. Comput. Appl. Math. **245**, 162–181 (2013)
157. S. Jain, D.J. Tward, D.S. Lee, A. Kolasny, T. Brown, J.T. Ratnanather, M.I. Miller, L. Younes, Computational anatomy gateway: leveraging xsede computational resources for shape analysis, in *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment* (ACM, 2014), p. 54
158. S. Joshi. Large Deformation Diffeomorphisms and Gaussian Random Fields for Statistical Characterization of Brain Sub-manifolds. Ph.D. thesis, Sever institute of technology, Washington University, 1997
159. S. Joshi, M. Miller, Landmark matching via large deformation diffeomorphisms. IEEE Trans. Image Process. **9**(8), 1357–1370 (2000)
160. Sarang Joshi, Brad Davis, Matthieu Jomier, Guido Gerig, Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage **23**, S151–S160 (2004)
161. S.H. Joshi, E. Klassen, A. Srivastava, I. Jermyn, A novel representation for Riemannian analysis of elastic curves in r^n (2007)
162. J. Jost, *Riemannian Geometry and Geometric Analysis*, 2nd edn. (Springer, Berlin, 1998)
163. P.E. Jupp, J.T. Kent, Fitting smooth paths to spherical data. Appl. Stat. **36**, 34–46 (1987)
164. H. Karcher, Riemannian center of mass and mollifier smoothing. Commun. Pure Appl. Math. **30**(5), 509–541 (1977)
165. M. Kass, A. Witkin, D. Terzopoulos, *Snakes: active contour models* (Int. J. Comput, Vis, 1988)
166. D.G. Kendall, Shape manifolds, Procrustean metrics and complex projective spaces. Bull. Lond. Math. Soc. **16**, 81–121 (1984)
167. D.G. Kendall, D. Barden, T.K. Carne, H. Le, *Shape and Shape Theory* (Wiley, New York, 1999)
168. W.S. Kendall, Probability, convexity, and harmonic maps with small image i: uniqueness and fine existence. Proc. Lond. Math. Soc. **3**(2), 371–406 (1990)
169. D. Keren, D.B. Cooper, J. Subrahmonia, Describing complicated objects by implicit polynomials. PAMI **16**(1), 38–53 (1994). January
170. R. Kimmel, D. Shaked, N. Kiryati, A.M. Bruckstein, Skeletonization via distance maps and level sets. Comput. Vis. Image Underst. **62**(3), 382–391 (1995)
171. E. Klassen, A. Srivastava, W. Mio, S.H. Joshi, Analysis of planar shapes using geodesic paths on shape spaces. IEEE Trans. Pattern Anal. Mach. Intell. **26**(3), 372–383 (2004)
172. N. Koiso, On the motion of a curve towards elastica. in *Actes de la Table Ronde de Géométrie Différentielle, Luminy 1992*, Sémin. Congr. 1. Soc. Math. (France, 1996)
173. E. Kuhn, S. Allassonnière, S. Durrleman, Convergent stochastic expectation maximization algorithm with efficient sampling in high dimension. application to deformable template model estimation, in *8. International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2015)* (2015)
174. S. Kurtek, T. Needham, Simplifying transforms for general elastic metrics on the space of plane curves (2018). arXiv:1803.10894
175. S. Lang, *Introduction to Differentiable Manifolds* (Interscience, New York, 1962)
176. H. Le, Unrolling shape curves. J. Lond. Math. Soc. **68**, 511–526 (2003)

177. H. Le, Estimation of Riemannian barycentres. Lond. Math. Soc. J. Comput. Math. **7**, 193–200 (2004)
178. W.E. Lorensen, H.E. Cline, Marching cubes: a high resolution 3d surface construction algorithm (1987)
179. D.G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969)
180. Jun Ma, Michael I. Miller, Alain Trouvé, Laurent Younes, Bayesian template estimation in computational anatomy. NeuroImage **42**(1), 252–261 (2008)
181. J .Ma, M.I Miller, L. Younes, A bayesian generative model for surface template estimation. J. Biomed. Imaging **2010**, 16 (2010)
182. L. Machado, F. Silva Leite, K. Krakowski, Higher-order smoothing splines versus least squares problems on riemannian manifolds. J. Dyn. Control. Syst. **16**(1), 121–148 (2010)
183. J. Macki, A. Strauss, *Introduction to Optimal Control Theory* (Springer Science and Business Media, 2012)
184. Ryuji Maehara, The jordan curve theorem via the brouwer fixed point theorem. Am. Math. Mon. **91**(10), 641–643 (1984)
185. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens, Multimodality image registration by maximization of mutual information. Trans. Med. Imaging **16**(2), 187–198 (1997)
186. D. Marr, E. Hildreth, Theory of edge detection. Proc. R. Soc. Lond. Ser. B **207**(1167), 187–217 (1980)
187. J.E. Marsden, *Lectures on Geometric Mechanics* (Cambridge University, Cambridge, 1992)
188. J.E. Marsden, T.S. Ratiu, *Introduction to Mechanics and Symmetry* (Springer, Berlin, 1999)
189. S. Marsland, C.J. Twining, Clamped-plate splines and the optimal flow of bounded diffeomorphisms. in *Statistics of Large Datasets, Proceedings of Leeds Annual Statistical Research Workshop*, pp. 91–95 (2002)
190. R. McCann, Existence and uniqueness of monotone measure-preserving maps. Duke Math. J. **80**(2), 309–323 (1995)
191. R.I. McLachlan, G.R.W. Quispel, Six lectures on geometric integration, in *Foundations of Computational Mathematics*, ed. by R. De Vore, A. Iserles, E. Süli (Cambridge University, Cambridge, 2001), pp. 155–210
192. J. Meinguet, Multivariate interpolation at arbitrary points made simple. J. Appl. Math. Phys. **30**, 292–304 (1979)
193. A. Mennucci, A. Yezzi, Metrics in the space of curves. Technical report (2005). arXiv:math.DG/0412454v2
194. M. Meyer, M. Desbrun, P. Schroeder, A. Barr, Discrete differential geometry operators for triangulated 2-manifolds (2002)
195. Mario Micheli, Joan A. Glaunès, Matrix-valued kernels for shape deformation analysis. Geom. Imaging Comput. **1**(1), 57–139 (2014)
196. I. Michael, Miller, Alain Trouvé, Laurent Younes, Hamiltonian systems and optimal control in computational anatomy: 100 years since d'arcy thompson. Ann. Rev. Biomed. Eng. **17**, 447–509 (2015)
197. I. Michael, Miller, Laurent Younes, Alain Trouvé, Diffeomorphometry and geodesic positioning systems for human anatomy. Technology **2**(01), 36–43 (2014)
198. P.W. Michor, *Topics in Differential Geometry*, Graduate studies in mathematics (Amarican Mathematical Society, 2008)
199. P.W. Michor, D. Mumford, Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. Documenta Mathematica **10**, 217–245 (2005)
200. P.W. Michor, D. Mumford, Riemannian geometries on spaces of plane curves. J. Eur. Math. Soc. **8**, 1–48 (2006)
201. P.W. Michor, D. Mumford, An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. Appl. Comput. Harmon. Anal. **23**(1), 74–113 (2007)
202. J. Mikusiński, *The Bochner Integral* (Springer, Berlin, 1978), pp. 15–22
203. M.I. Miller, S.C. Joshi, G.E. Christensen, Large deformation fluid diffeomorphisms for landmark and image matching, in *Brain Warping*, ed. by A. Toga (Academic, USA, 1999), pp. 115–131

204. M.I. Miller, A. Trouvé, L. Younes, On the metrics and Euler-Lagrange equations of computational anatomy. Ann. Rev. Biomed. Eng. **4**, 375–405 (2002)
205. M.I. Miller, A. Trouvé, L. Younes, Geodesic shooting for computational anatomy. J. Math. Image Vis. **24**(2), 209–228 (2006)
206. M.I. Miller, L. Younes, Group action, diffeomorphism and matching: a general framework. Int. J. Comput. Vis. **41**, 61–84 (2001). Originally published in electronic form in: Proceedings of SCTV 99. http://www.cis.ohio-state.edu/~szhu/SCTV99.html
207. Ha Quang Minh, Some properties of gaussian reproducing kernel hilbert spaces and their implications for function approximation and learning theory. Constr. Approx. **32**(2), 307–338 (2010)
208. J. Modersitzki, *Numerical Methods for Image Registration* (Oxford University, Demand, 2004)
209. C. Montani, R. Scateni, R. Scopigno, A modified look-up table for implicit disambiguation of marching cubes. Vis. Comput. **10**, 353–355 (1994)
210. F. Morgan, *Geometric Measure Theory: A Beginner's Guide* (Academic, USA, 1988)
211. Jean-Marie Morvan, Boris Thibert, Approximation of the normal vector field and the area of a smooth surface. Discret. Comput. Geom. **32**(3), 383–400 (2004)
212. S. Mukhopadhyaya, New methods in the geometry of plane arc. Bull. Calcutta Math. Soc. **1**, 31–37 (1909)
213. D. Mumford, J. Shah, Optimal approximation by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. **42**, 577–685 (1989)
214. D. Mumford, A. Desolneux, *Pattern Theory: The Stochastic Analysis of Real-world Signals* (CRC, 2010)
215. B.K. Natarajan, On generating topologically consistent surfaces from uniform samples. Vis. Comput. **11**(1), 52–62 (1994)
216. Y.A. Neretin, On jordan angles and the triangle inequality in Grassmann manifolds. Geom. Dedicata **86**, 81–92 (2001)
217. G.M. Nielson, On marching cubes. IEEE Trans. Vis. Comput. Graph. **9**(3), 283–297 (2003)
218. G.M. Nielson, B. Hamann, The asymptotic decider: removing the ambiguity in marching cubes, in *IEEE Visualization'91*, ed. by G.M. Nielson, L.J. Rosenblum (1991), pp. 83–91
219. M. Niethammer, Y. Huang, F.-X. Vialard, *Geodesic Regression for Image Time-series* (Springer, Berlin, 2011), pp. 655–662
220. Lyle Noakes, Greg Heinzinger, Brad Paden, Cubic splines on curved spaces. IMA J. Math. Control. Inf. **6**(4), 465–473 (1989)
221. J. Nocedal, S.J. Wright, *Numerical Optimization* (Springer, Berlin, 1999)
222. R.L. Ogniewicz, *Discrete Voronoï Skeletons* (Swiss federal institue of technology, 1992)
223. K.B. Oltham, J. Spanier, *An Atlas of Functions* (Springer, Berlin, 1996)
224. P.J. Olver, *Equivalence, Invariants and Symmetry* (Cambridge University, Cambridge, 1995)
225. B. O'Neill, The fundamental equations of a submersion. Michigan Math. J. **13**, 459–469 (1966)
226. J. O'Rourke, *Computational Geometry in C* (Cambridge University, Cambridge, 2000)
227. S. Osher, R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces* (Springer, Berlin, 2003)
228. R. Osserman, The four of more vertex theorem. Am. Math. Mon. **92**(5), 332–337 (1985)
229. Xavier Pennec, Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. J. Math. Imaging Vis. **25**(1), 127 (2006). Jul
230. A. Pentland, S. Sclaroff, Closed-form solutions for physically-based shape modeling and recognition. IEEE TPAMI **13**(7), 715–729 (1991)
231. M. Piccioni, S. Scarlatti, A. Trouvé, A variational problem arising from speech recognition. SIAM J. Appl. Math. **58**(3), 753–771 (1998)
232. Ulrich Pinkall, Konrad Polthier, Computing discrete minimal surfaces and their conjugates. Exp. Math. **2**(1), 15–36 (1993)
233. F.P. Preparata, M.I. Shamos, *Computational Geometry: An Introduction* (Springer, Berlin, 1985)

234. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*, 2nd edn. (Cambridge University, Cambridge, 1992)
235. A. Qiu, A. Albert, L. Younes, M.I. Miller, Time sequence diffeomorphic metric mapping and parallel transport track time-dependent shape changes. Neuroimage (2008). To appear
236. A. Qiu, L. Younes, L. Wang, J.T. Ratnanather, S.K. Gillepsie, K. Kaplan, J. Csernansky, M.I. Miller, Combining anatomical manifold information via diffeomorphic metric mappings for studying cortical thinning of the cingulate gyrus in schizophrenia. NeuroImage **37**(3), 821–833 (2007)
237. R.D. Rabbitt, J.A. Weiss, G.E. Christensen, M.I. Miller, Mapping of hyperelastic deformable templates using the finite element method (1995)
238. S.T. Rachev, *Probability Metrics* (Wiley, New York, 1991)
239. J.O. Ramsay, B.W. Silverman, *Functional Data Analysis* (Springer, Berlin, 1997)
240. A. Rangarajan, H. Chui, F.L. Bookstein, The Softassign Procrustes Matching Algorithm. in *Information Processing in Medical Imaging* (Springer, Berlin,1997), pp. 29–42
241. A. Rangarajan, E. Mjolsness, S. Pappu, L. Davachi, P.S. Goldman-Rakic, J.S. Duncan, A robust point matching algorithm for autoradiograph alignment, in *Visualization in Biomedical Computing (VBC)*, ed. by K.H. Hohne, R. Kikinis (1996), pp. 277–286
242. M. Reuter, B. Fischl, Avoiding asymmetry-induced bias in longitudinal image processing. Neuroimage **57**(1), 19–21 (2011)
243. M. Reuter, N.J. Schmansky, H. Diana Rosas, B. Fischl, Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage **61**(4), 1402–1418 (2012)
244. F. Richard, Résolution de problèmes hyperélastiques de recalage d'images. C. R. Acad. Sci., serie I **335**, 295–299 (2002)
245. Casey L. Richardson, Laurent Younes, Metamorphosis of images in reproducing kernel hilbert spaces. Adv. Comput. Math. **42**(3), 573–603 (2016). Jun
246. F. Riesz, B.S. Nagy, *Functional Analysis* (Ungar Publication Co., 1955)
247. P. Roussillon, J.A. Glaunès, *Surface matching using normal cycles, in International Conference on Geometric Science of Information* (Springer, Berlin, 2017), pp. 73–80
248. E. Rouy, A. Tourin, A viscosity solutions approach to shape from shading. SIAM J. Numer. Anal. **29**(3), 867–884 (1992)
249. W. Rudin. *Real and Complex Analysis* (Tata McGraw Hill, 1966)
250. G. Sapiro, *Geometric Partial Differential Equations and Image Analysis* (Cambridge University, Cambridge, 2001)
251. I.J. Schoenberg, Metric spaces and completely monotone functions. Ann. Math. **39**(4), 811–841 (1938)
252. B. Schroder, *Mathematical Analysis: A Concise Introduction* (Wiley, New York, 2007)
253. S. Sclaroff, Modal Matching: A Method for Describing, Comparing and Manipulating Signals. Ph.D. thesis, MIT, 1995
254. J. Serra, *Image Analysis and Mathematical Morphology*, vol. 1 and 2 (Academic, 1982), p. 88
255. J. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision and Material Science, Second edition: 1999* (Cambridge University, Cambridge, 1996)
256. J.A. Sethian, A review of recent numerical algorithms for hypersurface moving for curvature dependent speed. J. Differ. Geom. **31**, 131–161 (1989)
257. J.A. Sethian, A fast marching level set method for monotonically advancing fronts. Proc. Natl. Acad. Sci. **93**(4), 1591–1595 (1996)
258. R. Shaback, Error estimates and condition numbers for radial basis function interpolation. Adv. Comput. Math. **3**, 251–264 (1995)
259. J. Shah, H^0 type Riemannian metrics on the space of planar curves. Q. Appl. Math. **66**, 123–137 (2008)
260. E. Sharon, D. Mumford, 2d-shape analysis using conformal mapping (2004)
261. E. Sharon, D. Mumford, 2d-shape analysis using conformal mapping. Int. J. Comput. Vis. **70**(1), 55–75 (2006)

262. K. Siddiqi, B. Kimia, A. Tannenbaum, A. Zucker, Shocks, shapes, and wiggles. Image Vis. Comput. **17**(5–6), 365–373 (1999)
263. K. Siddiqi, B.B. Kimia, C.-W. Shu, Geometric shock-capturing ENO schemes for subpixel interpolation, computation and curve evolution. GMIP **59**(5), 278–301 (1997). September
264. K. Siddiqi, K.J. Tresness, B.B. Kimia, On the anatomy of visual form, in *Aspects of Visual Form Processing, 2nd International Workshop on Visual Form*, ed. by C. Arcelli, L.P. Cordella, G.S. di Baja (World Scientific, Singapore, 1994), pp. 507–521. May–June. IAPR, Workshop held in Capri
265. Nikhil Singh, François-Xavier Vialard, Marc Niethammer, Splines for diffeomorphisms. Med. Image Anal. **25**(1), 56–71 (2015)
266. C. Small, *The Statistical Theory of Shape* (Springer, Berlin, 1996)
267. A. Srivastava, E.P. Klassen, *Functional and Shape Data Analysis* (Springer, Berlin, 2016)
268. Ingo Steinwart, Don Hush, Clint Scovel, An explicit description of the reproducing kernel hilbert spaces of gaussian rbf kernels. IEEE Trans. Inf. Theory **52**(10), 4635–4643 (2006)
269. K. Stephenson, *Introduction to Circle Packing* (Cambridge University, Cambridge, 2005)
270. J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, 2nd edn. (Springer, Berlin, 1992)
271. E. Süli, D.F. Mayers, *An Introduction to Numerical Analysis* (Cambridge University, Cambridge, 2003)
272. G. Sundaramoorthi, A. Yezzi, A. Mennucci, Sobolev active contours, *Variational, Geometric and Level Set Methods, VLSM'05*, Lecture notes in computer science, vol. 3752/2005 (Springer, Berlin, 2005), pp. 109–120
273. G. Taubin, Estimating the tensor of curvatures of a surface from a polyhedral interpolation (1995)
274. H. Tek, B.B. Kimia, Curve evolution, wave propagation, and mathematical morphology, in *Mathematical Morphology and its Applications to Image and Signal Processing, Computational imaging and vision*, ed. by H.J.A.M. Heijmans, J.B.T.M. Roerdink, vol. 12 (Kluwer Academic, Amsterdam, 1998), pp. 115–126
275. J.-P. Thirion, Image matching as a diffusion process: an analogy with Maxwell's demons. Med. Image Anal. **2**(3), 243–260 (1998)
276. D.W. Thompson, *On Growth and Form, Revised edition 1992* (Dover Publications, 1917)
277. A. Trouvé. Habilitation à diriger les recherches. Technical report, Université Paris XI, 1996
278. A. Trouvé, Diffeomorphism groups and pattern matching in image analysis. Int. J. Comput. Vis. **28**(3), 213–221 (1998)
279. A. Trouvé, L. Younes, Diffeomorphic matching in 1d: designing and minimizing matching functionals, in *Proceedings of ECCV 2000*, ed. by D. Vernon (2000)
280. A. Trouvé, L. Younes, On a class of optimal matching problems in 1 dimension. Siam J. Control Opt. **39**(4), 1112–1135 (2001)
281. A. Trouvé, L. Younes, Local geometry of deformable templates. SIAM J. Math. Anal. **37**(1), 17–59 (2005)
282. A. Trouvé, L. Younes, Metamorphoses through lie group action. Found. Comput. Math. 173–198, (2005)
283. Alain Trouvé, Action de groupe de dimension infinie et reconnaissance de formes. C. R. Acad. Sci. Paris Sér. I Math. **321**(8), 1031–1034 (1995)
284. A. Trouvé, F.-X. Vialard, A second-order model for time-dependent data interpolation: Splines on shape spaces. in *MICCAI STIA Workshop* (2010)
285. Daniel Tward, Michael Miller, Alain Trouve, Laurent Younes, Parametric surface diffeomorphometry for low dimensional embeddings of dense segmentations and imagery. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1195–1208 (2017)
286. F. Tony, Chan, Selim Esedoglu, Mila Nikolova, Algorithms for finding global minimizers of image segmentation and denoising models. SIAM J. Appl. Math. **66**(5), 1632–1648 (2006)
287. D.J. Tward, J. Ma, M.I. Miller, L. Younes, Robust diffeomorphic mapping via geodesically controlled active shapes. J. Biomed. Imaging **2013**, 3 (2013)
288. C.J. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, C.J. Taylor, *A unified information-theoretic approach to groupwise non-rigid registration and model building* (Springer, Berlin, 2005), pp. 1–14

289. M. Vaillant, J. Glaunès, Surface matching via currents, in *Proceedings of Information Processing in Medical Imaging (IPMI 2005)*, Lecture notes in computer science, ed. by G.E. Christensen, M. Sonka, vol. 3565 (Springer, Berlin, 2005)
290. M. Vaillant, M.I. Miller, A. Trouvé, L. Younes, Statistics on diffeomorphisms via tangent space representations. Neuroimage **23**(S1), S161–S169 (2004)
291. A. Van Gelder, J. Wilhelms, Topological considerations in isosurface generation. ACM Trans. Graph. **13**(4), 337–375 (1994)
292. Oswald Veblen, Theory on plane curves in non-metrical analysis situs. Trans. Am. Math. Soc. **6**(1), 83–98 (1905)
293. F.-X. Vialard, *Hamiltonian approach to geodesic image matching* (Technical report, ArXiv Mathematics, 2008)
294. F.-X. Vialard, Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework: From the Discontinuous Image Matching Problem to a Stochastic Growth Model. Ph.D. thesis, Ecole Normale Supérieure de Cachan, 2009. http://tel.archives-ouvertes.fr/tel-00400379/fr/
295. F.-X. Vialard, F. Santambrogio, Extension to bv functions of the large deformation diffeomorphisms matching approach. Comptes Rendus Math. **347**(1–2), 27–32 (2009)
296. C. Villani, *Topics in Optimal Transportation* (American Mathematical Society, 2003)
297. T.L. Vincent, W.J. Grantham, *Nonlinear and Optimal Control Systems* (Wiley, New York, 1997)
298. P. Viola, W.M. Wells III, Alignment by maximization of mutual information. Int. J. Comput. Vis. **24**(2), 137–154 (1997)
299. G. Wahba. *Spline Models for Observational Data* (SIAM, 1990)
300. L. Wang, M.F. Beg, J.T. Ratnanather, C. Ceritoglu, L. Younes, J. Morris, J. Csernansky, M.I. Miller, Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. IEEE Trans. Med. Imaging **26**, 462–470 (2006)
301. L.N. Wasserstein, Markov processes over denumerable products of spaces describing large systems of automata. Probl. Inf. Transm. **5**, 47–52 (1969)
302. T. Willmore, *Riemannian Geometry* (Clarendon Press, 1993)
303. Qian Xie, Sebastian Kurtek, Huiling Le, and Anuj Srivastava. Parallel transport of deformations in shape space of elastic surfaces. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 865–872, 2013
304. C. Xu, J. Prince, Gradient vector flow: A new external force for snakes (1997)
305. T. Yanao, W.S. Koon, J.E. Marsden, Intramolecular energy transfer and the driving mechanisms for large-amplitude collective motions of clusters. J. Chem. Phys. **130**(14), 144111 (2009)
306. K. Yosida, *Functional Analysis* (Springer, Berlin, 1970)
307. L. Younes, Computable elastic distances between shapes. SIAM J. Appl. Math **58**(2), 565–586 (1998)
308. L. Younes, Optimal matching between shapes via elastic deformations. Image Vis. Comput. **17**(5), 381–389 (1999)
309. L. Younes, Jacobi fields in groups of diffeomorphisms and applications. Q. Appl. Math. **65**, 113–134 (2007)
310. L. Younes, F. Arrate, M.I. Miller, Evolutions equations in computational anatomy. NeuroImage **45**(1, Supplement 1), S40–S50 (2009). Mathematics in Brain Imaging
311. L. Younes, P. Michor, J. Shah, D. Mumford, A metric on shape spaces with explicit geodesics. Rend. Lincei Math. Appl. **9**, 25–57 (2008)
312. L. Younes, A. Qiu, R.L. Winslow, M.I. Miller, Transport of relational structures in groups of diffeomorphisms. J. Math. Imaging Vision **32**(1), 41–56 (2008)
313. Laurent Younes, A distance for elastic matching in object recognition. C. R. Acad. Sci. Paris Sér. I Math. **322**(2), 197–202 (1996)
314. M. Zähle, Integral and current representation of curvature measures. Arch. Math. **46**, 557–567 (1986)

315. E. Zeidler, *Applied Functional Analysis, Applications to mathematical physics* (Springer, Berlin, 1995)
316. H.-K. Zhao, T. Chan, B. Merriman, S. Osher, A variational level set approach to multiphase motion. J. Comput. Phys. **127**, 179–195 (1996)

Index

L. Younes, *Shapes and Diffeomorphisms*, Applied Mathematical Sciences 171,
https://doi.org/10.1007/978-3-662-58496-5