ON MODEL-BASED DIFFEOMORPHIC SHAPE EVOLUTION AND DIFFEOMORPHIC SHAPE REGISTRATION

by
Dai-Ni Hsieh

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
July 2021

Abstract

Shape registration is fundamental in many applications. However, the shape registration problem is usually ill posed unless further information is provided. In this dissertation, we examine a scenario when one of the two shapes to be registered is assumed to have evolved from the other shape according to a known model. The shape registration problem is then formulated as a variational problem subject to the dynamics of the shape evolution model. We provide sufficient conditions on models so that diffeomorphic shape evolution and diffeomorphic shape registration are guaranteed theoretically. In addition, we illustrate this model-based registration by applications of piecewise-rigid motion and biological atrophy. Numerical experiments of the two applications are presented with a GPU-accelerated implementation.

Primary reader and advisor: Laurent Younes
Secondary reader: Nicolas Charon

Acknowledgements

I would like to thank my advisor Professor Laurent Younes. He was always patient and willing to give more time for discussion. I was very lucky to have him as my advisor. I will miss our weekly meetings, including discussions of typesetting style.

I would also like to thank my undocumented co-advisors Professor Nicolas Charon and Dr. Sylvain Arguillère. I was spoiled to have the three advisors.

I would like to thank my dissertation committee members Professor Michael Miller and Professor Fadil Santosa, my Graduate Board Oral Examination committee members Professor Yannis Kevrekidis and Professor Misha Kazhdan, and my Candidacy Examination committee member Professor Gregory Eyink.

And thanks to my friends, with hair or with fur.

Last but not least, I am sincerely grateful to my family, who is always there for me.

To my mother, Li-Lin Wu

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Shape registration aims to establish a sensible point-to-point correspondence between two shapes. For example, let us denote a ball by $B \subset \mathbb{R}^3$ and denote a cube by $C \subset \mathbb{R}^3$. A registration from B to C, loosely speaking, is a one-to-one function $\xi : B \to \mathbb{R}^3$ such that $\xi(B) \approx C$, and ideally $\xi(B) = C$. Clearly, there are infinitely many registrations from a ball to a cube, hence we may further restrict our selection criterion, e.g., requiring a higher regularity on ξ. On the other hand, it will be in vain if we try to find a continuous function that registers a sphere to a torus. In addition to the interplay between the class of shapes and the class of registrations, what the meaning of a sensible registration is, what we can say from mathematical and theoretical perspectives, and how to compute such a sensible registration are all active research areas.

The applications of shape registration are vast. For example, image registration can be used to track temporal changes of tissues and organs by registering longitudinal medical images of the same patient [81, 32] and can also be used to compare and measure the difference between an individual medical image and a population-based atlas [10, 97]; curve registration occurs in the registration of vascular structures, e.g., coronary arteries, to assist surgical procedures [13, 84]; surface registration appears in analyzing human facial data [91]. The list is by no means exhaustive. Here we restrict our attention of shape registration to a function between two shapes of interest. However, we note that within the computer graphics community shape registration refers to a much wider spectrum of problems (see [98] and [61, Part III]). For example, it includes building a 3D computer model of a physical object from multiple data point clouds scanned from different viewpoints around the object. In other words, it encompasses partial shape registration, a more complicated and hard-to-define problem. We will focus our discussion on the narrow meaning of whole shape

registration.

Various approaches in the literature have been proposed to compute a shape registration. One approach relies on manually labelled landmarks, or reproducible significances [18, 43, 56]. The labelled landmarks on two shapes to be registered provide a sparse point-to-point correspondence based on which a smooth registration is then extrapolated to a whole-shape point-to-point correspondence. Another approach is based on the idea of small distortion of some attribute, e.g., as rigid as possible [53, 95] and volume preserving [85, 50, 39]. Still another approach registers two shapes by gradually morphing one shape into the other shape [30, 19, 28, 66, 11]. Specifically for surfaces homeomorphic to a sphere, conformal spherical parametrizations also have been used in different methods for surface registration [7, 48, 88]. We now pay full attention to the morphing approach in order to attain a more physically sensible registration.

There is a trend in the morphing approach to incorporate a prior of the application of interest into the class of feasible registrations. Some generic priors in the literature include local affine deformations [6, 80], local radial and angular deformation functions [45], local deformation modules [47, 46], and linear operator constraints [4]. There are also application-targeted priors. For example, Hogea et al. [51] modeled the tumor growth dynamics and the surrounding tissue deformation for image registration; Sundar et al. [94] estimated myocardial displacements constrained by prior knowledge of cardiac mechanics; Werner et al. [99] estimated respiratory lung motion based on physiology of breathing; Ratnanather et al. [82] measured the thickness of cerebral cortex assuming that cortical columns are orthogonal to cortical layers. From a theoretical viewpoint, theorems for generic priors can be developed once and for all, while theorems for application-targeted priors need to be customized. From a practical viewpoint, however, application-targeted priors usually lead to more satisfying registration results. To achieve a balance between generic and application-targeted priors, our goal is a *form* of priors that can be adapted to a wide range of applications.

In this dissertation, we will examine the shape registration problem when we have prior knowledge of possible registrations. To be more precise, we assume that one of the two shapes to be registered has evolved from the other shape through a very general and abstract model. We will provide sufficient conditions on the abstract model for diffeomorphic shape evolution and diffeomorphic shape registration. The versatility of this approach will be demonstrated by two seemingly unrelated shape registration problems, whose registration results will be assured to be diffeomorphic by our theorems. In the following chapters,

we will first introduce our proposed framework, interpretable deformation vector fields, in Chapter 2, then we will go through examples, theorems, and computation in Chapters 3 to 6. We documented our computation as detailed as possible to promote greater transferability. Most of the meshes in this dissertation were generated with the help of `distmesh` [79].

Chapter 2

Prelude

In this chapter, we establish necessary terminology and our proposed framework. Some notation conventions that suffice to facilitate our discussion are compiled in Section 2.1. Further notation will be introduced when needed. After we review the Bochner integral in Section 2.2, we then define the notions of shapes, deformations, and motions in Section 2.3. Section 2.4 presents some basics of reproducing kernel Hilbert spaces, which are essential in the framework of large deformation diffeomorphic metric mapping (LDDMM) covered in Section 2.5. With all the background in place, we state our proposed framework, interpretable deformation vector fields, motivated by LDDMM in Section 2.6.

2.1 Notation

We differentiate different types of equal sign. The asymmetric sign ":=" indicates equal by definition. As its asymmetry suggests, $a := b$ means that a is defined by b, while $a =: b$ defines b by a. The asymmetry offers an advantage over other symmetric alternatives, like $\triangleq$, $\doteq$, etc. We will use ":=" only for the first occurrence of a notion and use "=" in its later appearances. The sign "$\approx$" means approximately equal in a loose sense. Occasionally, we will use "$f \equiv g$" to stress the equality of two functions when the context includes a discussion of the equality $f(x) = g(x)$ at some point x. In particular, $f \equiv 0$ means that f is a constant zero function. We will simply write $f = g$ if the context is clear.

The prime symbol will *always* be used to denote another object of the same "kind." For example, if $f \in C^1([a, b])$, then f' denotes another function in $C^1([a, b])$; we will use $\dot{f}$ to denote the derivative of a univariate function. By the same token, if B is a Banach space, then B' denotes another Banach space; we will use B^* to denote the topological dual.

For a function $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ which depends on time and space, the notation

$f(t)$ means the function $f(t) : \mathbb{R}^d \to \mathbb{R}^d$ defined by $f(t)(x) := f(t,x)$. We will also write $\dot{f}(t,x) := \partial_t f(t,x)$ and $Df(t,x) := \partial_x f(t,x)$, hence $\dot{f}(t)$ and $Df(t)$ should be understood by combining the two conventions. We will constantly switch between these two viewpoints either regarding f as $(t,x) \mapsto f(t,x)$ or $t \mapsto f(t)$.

We will use C to denote a generic constant and C_a to show its dependency on a. The value of such constants may change from equation to equation. It may happen that the constant depends on a, while a depends on b, and b further depends on d. In this case, we will choose to write C_a, C_b, or C_d according to which one is the most pertinent to the context of discussion. For symbols other than C, we will sometimes use $\delta(\varepsilon)$ to indicate the dependency at its first occurrence and write δ in the discussion for simplicity. We will always use subscripts for the dependency of generic constants to avoid ambiguous expressions like $C(a+b)$.

The notation $\mathscr{L}(X, X')$ will denote the vector space of continuous linear operators from a topological vector space X to another topological vector space X'. Weak convergence of a sequence $(x_n)_{n=1}^{\infty}$ in a Banach space will be denoted by $x_n \rightharpoonup x$. We will use the notation $(\mu \mid v)$ rather than $\langle \mu, v \rangle$ or $\mu(v)$ to denote the natural pairing of $\mu \in B^*$ and $v \in B$ between a Banach space B and its topological dual B^*. The notation $\langle \cdot, \cdot \rangle$ will always refer to an inner product. The notation $|\cdot|$ denotes a norm on a finite-dimensional vector space; without further specification, it is the Euclidean norm if the underlying space is $\mathbb{R}^d$. For a metric space (X, d_X), we will denote the open ball centered at c with radius r by $B(c,r) := \{x \in X : d_X(x,c) < r\}$ and denote the closed ball by $\overline{B}(c,r) := \{x \in X : d_X(x,c) \le r\}$. The bar of $\overline{B}$ should be interpreted as a mnemonic notation, not the closure $\overline{B(c,r)}$. The metric for the ball will be clear from the context.

Let $\Omega \subset \mathbb{R}^d$ be a nonempty open set. We will denote by $\mathscr{D}(\Omega)$ the topological space of C^∞ functions with compact support in Ω. The topology on $\mathscr{D}(\Omega)$ is chosen so that the dual $\mathscr{D}^*(\Omega)$ is the space of distributions on Ω. We will denote the L^p spaces on an interval by $L^p([a,b])$, even though there is no difference between $L^p([a,b])$ and $L^p\big((a,b)\big)$. Given a Banach space B, we will denote by $L^p([a,b], B)$ the space of strongly measurable functions such that $\int_a^b \|f(t)\|_B^p \, dt < \infty$. Strongly measurable functions will be briefly reviewed in the next section (Section 2.2).

For an integer $p \ge 1$, we let $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ denote the space of p-times continuously differentiable vector fields v such that the j-th derivative $D^j v$ tends to 0 at infinity for every $j \le p$. The space $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ is a Banach space equipped with the norm $\|v\|_{p,\infty} := \sum_{j=0}^p \sup_{x \in \mathbb{R}^d} |D^j v(x)|$, where $|\cdot|$ denotes the operator norm of a multilinear function

on a product of finite-dimensional vector spaces equipped with the Euclidean norm. Let $id : \mathbb{R}^d \to \mathbb{R}^d$ be the identity function, i.e., $id(x) := x$. We denote by $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ the set of C^p-diffeomorphisms on $\mathbb{R}^d$ that tend to identity at infinity. Thus every element $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$ can be written as $\xi = id + v$, where $v \in C^p_0(\mathbb{R}^d, \mathbb{R}^d)$; that is, $\mathrm{Diff}^p_{id}(\mathbb{R}^d) \subset id + C^p_0(\mathbb{R}^d, \mathbb{R}^d)$. The affine Banach space $id + C^p_0(\mathbb{R}^d, \mathbb{R}^d)$ and the subset $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ are both equipped with the induced metric $d_{p,\infty}(\xi, \eta) := \|\xi - \eta\|_{p,\infty}$. We note that $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ is an open subset of $id + C^p_0(\mathbb{R}^d, \mathbb{R}^d)$.

2.2 The Bochner integral

In this section, we cite properties of the Bochner integral without proofs. We adapt the statements to functions defined on the interval $[0, T]$ so that we can apply the results more smoothly; the results still hold for functions defined on a finite measure space. We refer the reader to [103, Sections V.4 and V.5] for statements on general measure spaces and [34, Sections II.1 and II.2] on finite measure spaces. Since the theorems in this section are for reference, some assumptions will be repeated in the statements to make each theorem self-contained. The reader only needs Definition 2.2.1, Definition 2.2.2(i), and Corollary 2.2.4 to understand the notation $L^p([a, b], B)$ and read this chapter. Other properties of the Bochner integral may be consulted before reading the proofs in Chapter 4.

We now consider functions defined on the interval $[0, T]$ with values in a Banach space. For these Banach-space-valued functions, we will extend the definitions of measurability and integrability. Recall that the characteristic function on a set $S \subset X$ is defined by

$$\mathbb{1}_S(x) := \begin{cases} 1, & \text{if } x \in S\,; \\ 0, & \text{if } x \in X \setminus S\,. \end{cases}$$

We start simple.

Definition 2.2.1. Let B be a Banach space. A function $v : [0, T] \to B$ is *simple* if there exist distinct vectors $\{b_i\}_{i=1}^N \subset B$ such that

$$v(t) = \sum_{i=1}^{N} b_i \, \mathbb{1}_{S_i}(t),$$

where $S_1, \ldots, S_N$ are Lebesgue measurable subsets of $[0, T]$.

Definition 2.2.2. Let B be a Banach space.

(i) A function $v : [0, T] \to B$ is *strongly measurable* if there exists a sequence of simple functions $(v_n)_{n=1}^\infty$ such that $\|v_n(t) - v(t)\|_B \to 0$ for almost every $t \in [0, T]$.

(ii) A function $v : [0,T] \to B$ is *weakly measurable* if for all $\mu \in B^*$, the real-valued function $t \mapsto (\mu \mid v(t))$ is Lebesgue measurable.

A strongly-measurable function is weakly measurable. Indeed, let $(v_n)_{n=1}^{\infty}$ be a sequence of simple functions that tends to a strongly-measurable function v. Given $\mu \in B^*$, the strong measurability of v gives

$$(\mu \mid v(t)) = \lim_{n\to\infty} (\mu \mid v_n(t)) \quad \text{for almost every } t \in [0,T].$$

Note that the real-valued function $t \mapsto (\mu \mid v_n(t))$ is simple in the classical sense. Thus the real-valued function $t \mapsto (\mu \mid v(t))$ is Lebesgue measurable since it is the almost everywhere pointwise limit of simple, thus measurable, functions and our measure space is complete. Conversely, a weakly-measurable function may not be strongly measurable; we refer the reader to Example 5 of Section II.1 in [34]. A characterization of strongly-measurable functions is given by Pettis' theorem.

Theorem 2.2.3 (Pettis)**.** *Let B be a Banach space. A function $v : [0,T] \to B$ is strongly measurable if and only if v is weakly measurable and almost separably valued, i.e., there exists $S \subset [0,T]$ of Lebesgue measure zero such that $\{v(t) : t \in [0,T] \setminus S\}$ is separable.*

Hence if B is a separable Banach space, the strong measurability is equivalent to the weak measurability. The proof of Pettis' theorem also shows the following useful result.

Corollary 2.2.4. *If a function $v : [0,T] \to B$ is strongly measurable, then $t \mapsto \|v(t)\|_B$ is Lebesgue measurable.*

Now we define the integrability and integrals of strongly-measurable functions. We will not integrate a weakly-measurable function, whose integral is an element in B^{**}; we refer the interested reader to [34, Section II.3]. The integral of a Banach-space-valued simple function is defined in the same way as the integral of a real-valued simple function:

$$\int_0^T \left(\sum_{i=1}^N b_i \, \mathbb{1}_{S_i}(t) \right) dt := \sum_{i=1}^N b_i \, \lambda(S_i) \; \in \; B,$$

where λ is the Lebesgue measure.

Definition 2.2.5. Let B be a Banach space. A strongly-measurable function $v : [0,T] \to B$ is called *Bochner integrable* if there exists a sequence of simple functions $(v_n)_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty} \int_0^T \|v_n(t) - v(t)\|_B \, dt = 0.$$

If $v : [0,T] \to B$ is Bochner integrable, the *Bochner integral* over a Lebesgue measurable set $S \subset [0,T]$ is defined by

$$\int_S v(t)\,dt := \lim_{n\to\infty} \int_0^T \mathbb{1}_S(t)\, v_n(t)\,dt \;\in\; B.$$

Remark 2.2.6. To validate the notion of Bochner integral, we need to verify the following two conditions: 1. The real-valued function $t \mapsto \|v_n(t) - v(t)\|_B$ is Lebesgue measurable; 2. The limit $\lim_{n\to\infty} \int_0^T \mathbb{1}_S(t)\, v_n(t)\,dt$ exists and is independent of the sequence $(v_n)_{n=1}^\infty$. The first condition is verified by Corollary 2.2.4, and the verification of the second condition can be found in [103, Section V.5].

The following theorem is an indispensable characterization of strongly-measurable functions to be Bochner integrable.

Theorem 2.2.7. *Let B be a Banach space. A strongly-measurable function $v : [0,T] \to B$ is Bochner integrable if and only if the real-valued function $t \mapsto \|v(t)\|_B$ is Lebesgue integrable.*

The Bochner integral is invariant under bounded linear transformations.

Theorem 2.2.8. *Let B and B' be two Banach spaces and $\mathcal{L} \in \mathscr{L}(B, B')$. If $v : [0,T] \to B$ is Bochner integrable, then $t \mapsto \mathcal{L}(v(t))$ is also Bochner integrable and*

$$\int_S \mathcal{L}(v(t))\,dt = \mathcal{L}\left(\int_S v(t)\,dt\right) \;\in\; B',$$

where $S \subset [0,T]$ is a Lebesgue measurable set.

In particular, if $i : B \hookrightarrow B'$ is a continuous embedding, then the Bochner integrals of $v : [0,T] \to B$ in B and B' agree.

Many classical theorems without nonnegativeness have their analogues in Banach-space-valued functions. For example, we have the modulus inequality, the dominated convergence theorem, the Lebesgue differentiation theorem, and Fubini's theorem. Note that Fatou's lemma and the monotone convergence theorem are not in the list. Below we record the modulus inequality, which is the only theorem in the list we need.

Theorem 2.2.9. *Let B be a Banach space and $S \subset [0,T]$ be a Lebesgue measurable set. If $v : [0,T] \to B$ is Bochner integrable, then*

$$\left\| \int_S v(t)\,dt \right\|_B \le \int_S \|v(t)\|_B\,dt.$$

2.3 Shapes, Deformations, and Motions

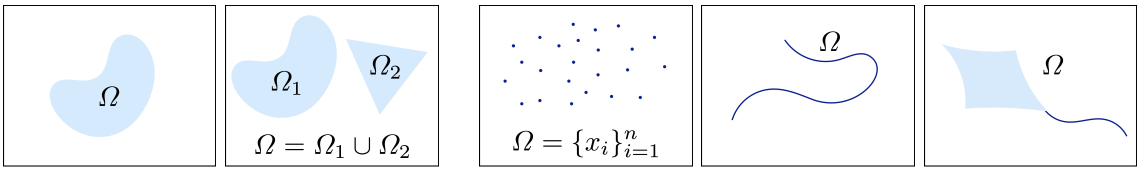
We may define shapes as compact sets in the Euclidean affine space. Although this coordinate-free definition of shapes is more satisfying, it will complicate the analysis of shape deformation as we must cautiously distinguish points from vectors. Hence we technically define shapes as compact sets in $\mathbb{R}^d$, which is identified with a particular rectangular coordinate system, and ensure that the results we obtain are independent of rectangular coordinate systems. The dimension d is 2 or 3 for practical applications, but it can in fact be any positive integer. In this dissertation, we will focus on volumetric shapes, or shapes with nonempty interior (see Figure 2.1(a)). We also rule out volumetric shapes with curves and surfaces attached (see Figure 2.1(b)). Furthermore, in order to compare shapes in a differentiable manner (see (3.5)), and also from a practical point of view, we thus restrict ourselves to the class of building blocks

$$\mathscr{B} := \{\Omega \subset \mathbb{R}^d : \Omega \text{ is homeomorphic to } \overline{B}(0,1) \text{ and } \partial\Omega \text{ is rectifiable}\}$$

and define the class of shapes of interest by

$$\mathscr{S} := \{\Omega \subset \mathbb{R}^d : \Omega = \bigcup_{i=1}^{N} \Omega_i \text{ for some } N \in \mathbb{N}, \text{ where } \Omega_i\text{'s are disjoint and } \Omega_i \in \mathscr{B}\}. \quad (2.1)$$

When we mention a shape hereafter, we mean an element $\Omega \in \mathscr{S}$. We certainly can include more complicated shapes in $\mathscr{S}$, for example, like balls with holes. We choose the class (2.1), which is rich enough to provide concrete examples; more importantly, we are free from the burden of describing shape topology. With our shapes specified, we can direct our full attention to shape deformation.

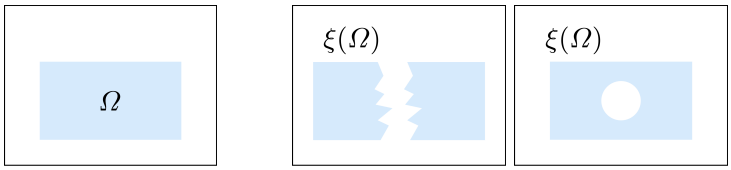


(a) Examples of $\Omega \in \mathscr{S}$. (b) Examples of $\Omega \notin \mathscr{S}$.

Figure 2.1: Illustrations of shapes of interest in $\mathbb{R}^2$.

In the broadest sense, a deformation is a function $\xi : \mathbb{R}^d \to \mathbb{R}^d$, which deforms a shape Ω into a set $\xi(\Omega)$; this definition includes constant functions that annihilate the universe. To be more physical meaningful, we require that a deformation ξ is one-to-one and onto (nuclear reactions are not of our concern) and that $\xi \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ with $\det D\xi(x) \neq 0$ for

all $x \in \mathbb{R}^d$ so that volume elements will not be deformed into lower-dimensional objects. The requirements on ξ are equivalent to saying that ξ is a diffeomorphism, i.e., ξ, $\xi^{-1} \in C^1(\mathbb{R}^d, \mathbb{R}^d)$. Moreover, since our shapes are bounded, there is no loss of generality to only consider diffeomorphisms that tend to identity at infinity, or $\xi \in \mathrm{Diff}^1_{id}(\mathbb{R}^d)$, which turns out to be more convenient when developing our theorems. We will also consider $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$, $p \geq 1$, for higher regularity. Hence we define a Diff^p-deformation by a function $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$. For the class of shapes $\mathscr{S}$ we focus on, note that $\xi(\mathscr{S}) \subset \mathscr{S}$ for all $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$. Therefore, the class of shapes $\mathscr{S}$ and the class of deformations $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ provide an adequate context to discuss shape deformations, i.e., a deformed shape is still a shape. Although diffeomorphisms are widely used to model deformations, we remark that certain deformations such as fractures and punctures cannot be described by continuous functions, let alone by diffeomorphisms (see Figure 2.2).



(a) A shape. (b) Deformed shapes: fracture and puncture.

Figure 2.2: Examples of deformed shapes that cannot be described by diffeomorphisms.

A motion is a one parameter family of deformations. Specifically, a motion is a function $\varphi : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ on a time interval $[0, T]$ and the space $\mathbb{R}^d$. Note that for a fixed time t, the function $\varphi(t) : \mathbb{R}^d \to \mathbb{R}^d$ is a deformation in the general sense. We view time 0 as the beginning of shape changes and impose that $\varphi(0) = id$, the identity function. Given a motion φ and a shape Ω, the function $t \mapsto \varphi(t, \Omega)$ thus plays a movie of shape changes when t goes from 0 to T, while $\varphi(t, \Omega)$ is the snapshot of the deformed shape at time t (see Figure 2.3). Since we do not expect shape teleportation, we require that φ and all its spatial derivatives are continuous in time. Moreover, we would like to be able to talk about velocities $\dot{\varphi}(t, x)$ of motions. We summarize our assumptions in the following definition.

Definition 2.3.1 (Diff^p-motion)**.** We say that a function $\varphi : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a Diff^p-*motion* if φ satisfies:

- $\varphi(0) = id$.
- $\varphi \in C([0, T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$.
- $\varphi : [0, T] \to \mathrm{Diff}^p_{id}(\mathbb{R}^d)$ is differentiable almost everywhere and $\dot{\varphi} \in L^1([0, T], C^p_0(\mathbb{R}^d, \mathbb{R}^d))$.

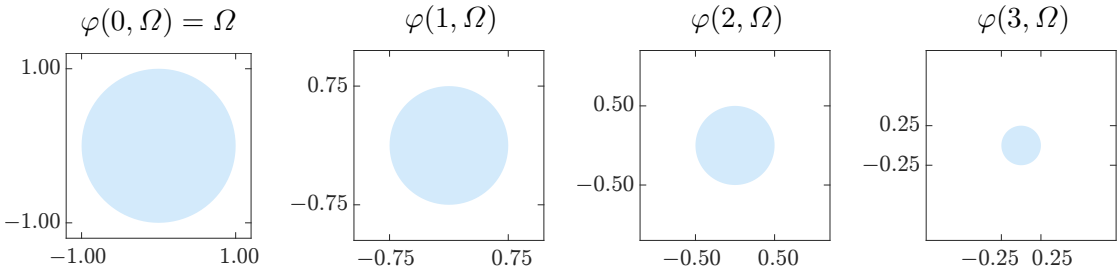


Figure 2.3: Snapshots of the motion $\varphi(t,x) = x - \frac{t}{4}\,x$ deforming the unit disk. This motion is not a Diff^p-motion since $\varphi(t)$ does not tend to identity at infinity.

We remind the reader that $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ is equipped with the metric $d_{p,\infty}(\varphi,\psi) = \|\varphi - \psi\|_{p,\infty}$. For a Diff^p-motion φ, we have $\det D\varphi(t,x) > 0$ for all $(t,x) \in [0,T]\times\mathbb{R}^d$ since $\det D\varphi(0,x) = 1$, $\det D\varphi(t,x) \neq 0$, and $t \mapsto \det D\varphi(t,x)$ is continuous. Sometimes it will be useful to restrict the spatial domain of a motion from $\mathbb{R}^d$ to Ω and consider $\varphi : [0,T]\times\Omega \to \mathbb{R}^d$. In this case, we say that φ is a motion of Ω.

Given a Diff^p-motion φ, we can define its Eulerian velocity field $v(t) := \dot\varphi(t)\circ\varphi^{-1}(t)$ for almost every $t \in [0,T]$ so that

$$\varphi(t,x) = x + \int_0^t \dot\varphi(s,x)\,ds = x + \int_0^t v(s,\varphi(s,x))\,ds.$$

As a partial converse, it can be shown ([104, Theorem 7.11] and Theorem A.2.3) that if $v \in L^1([0,T], C^{p+1}_0(\mathbb{R}^d,\mathbb{R}^d))$, then the initial value problem

$$\varphi(t,x) = x + \int_0^t v(s,\varphi(s,x))\,ds$$

has a unique solution $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$. It follows that $\dot\varphi(t) = v(t)\circ\varphi(t)$ for almost every $t \in [0,T]$ and $\dot\varphi \in L^1([0,T], C^p_0(\mathbb{R}^d,\mathbb{R}^d))$. In other words, the flow of a Eulerian velocity field $v \in L^1([0,T], C^{p+1}_0(\mathbb{R}^d,\mathbb{R}^d))$ is a Diff^p-motion. This observation suggests that we can model Diff^p-motions through certain Eulerian velocity fields. Since we do not always relate the interval $[0,T]$ to a physical time interval, we term Eulerian velocity fields as deformation vector fields in this dissertation. We are going to investigate under what conditions the flow of a deformation vector field is a Diff^p-motion when the deformation vector field is not determined in advance but coupled with the motion.

2.4 Reproducing Kernel Hilbert Spaces

We will use reproducing kernel Hilbert spaces (RKHSs) continuously embedded in $C^p_0(\mathbb{R}^d,\mathbb{R}^d)$ as our technical and computational tool. We briefly introduce abstract RKHSs and their

properties in Section 2.4.1, then we present tangible RKHSs continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ in Section 2.4.2.

2.4.1 Abstract reproducing kernel Hilbert spaces

A reproducing kernel Hilbert space, in short, is a function space that is a Hilbert space and has a reproducing kernel. The precise definition is the following [5].

Definition 2.4.1. Let X be an abstract set. Denote by F the function space of real-valued functions defined on X. A vector subspace $V \subset F$ is called a *(real-valued) reproducing kernel Hilbert space* if V is a Hilbert space equipped with an inner product $\langle \cdot, \cdot \rangle_V$, and there is a function $k_V : X \times X \to \mathbb{R}$ such that

$$k_V(\cdot, x) \in V \quad \text{for all } x \in X$$

and

$$f(x) = \langle f, k_V(\cdot, x) \rangle_V \quad \text{for all } f \in V \text{ and } x \in X. \tag{2.2}$$

Equation (2.2) is called the reproducing property: the function k_V reproduces all functions in V in the sense that the function value of $f \in V$ at x is reproduced by taking the inner product of f and $k_V(\cdot, x)$[a]. In other words, function values are legit in an RKHS, hence L^2 spaces are not RKHSs. The function k_V is referred to as *the* reproducing kernel of V, which is justified by the following proposition [5].

Proposition 2.4.2.

(i) If a reproducing kernel k_V of V exists, then it is unique.

(ii) Reproducing kernels are symmetric and positive semidefinite[b]*, that is, $k_V(x, y) = k_V(y, x)$ for all $x, y \in X$, and for any finite number of distinct points $\{x_1, \ldots, x_n\} \subset X$ the matrix $\left[\; k_V(x_i, x_j) \; \right]_{n \times n}$ is positive semidefinite.*

(iii) The subspace $\operatorname{span}\{k_V(\cdot, x) : x \in X\}$ *is dense in V.*

(iv) A Hilbert space V is an RKHS on X if and only if all the evaluation functionals are continuous, that is, $\delta_x \in V^*$ for all $x \in X$, where $(\delta_x \mid f) := f(x)$.

[a] Some authors refer to $k_V(x, y) = \langle k_V(\cdot, x), k_V(\cdot, y) \rangle_V$ as the reproducing property, which is equivalent to (2.2), in the sense that $k_V(x, y)$ reproduces the inner product of $k_V(\cdot, x)$ and $k_V(\cdot, y)$. We follow the terminology in [5].

[b] Other terms used in the literature include positive definite functions (no typo on definite) and kernel functions. The reader needs to be careful about this confusing terminology.

Proof. All the proofs rely on the reproducing property (2.2).

(i) Suppose that k'_V is another reproducing kernel, then for all $x \in X$ we have

$$\begin{aligned}\|k_V(\cdot,x) - k'_V(\cdot,x)\|_V^2 &= \langle k_V(\cdot,x), k_V(\cdot,x)\rangle_V - \langle k_V(\cdot,x), k'_V(\cdot,x)\rangle_V \\ &\quad - \langle k'_V(\cdot,x), k_V(\cdot,x)\rangle_V + \langle k'_V(\cdot,x), k'_V(\cdot,x)\rangle_V \\ &= k_V(x,x) - k_V(x,x) - k'_V(x,x) + k'_V(x,x) = 0,\end{aligned}$$

which implies $k_V \equiv k'_V$.

(ii) The symmetry of k_V follows from the symmetry of an inner product:

$$k_V(x,y) = \langle k_V(\cdot,y), k_V(\cdot,x)\rangle_V = \langle k_V(\cdot,x), k_V(\cdot,y)\rangle_V = k_V(y,x).$$

For positive semidefiniteness, let $\alpha_i \in \mathbb{R}$ and $x_i \in X$, $i = 1,\dots,n$. We have

$$\sum_{i,j=1}^{n} \alpha_i\,\alpha_j\,k_V(x_i,x_j) = \Big\langle \sum_{i=1}^{n} \alpha_i\,k_V(\cdot,x_i), \sum_{j=1}^{n} \alpha_j\,k_V(\cdot,x_j)\Big\rangle_V = \Big\| \sum_{i=1}^{n} \alpha_i\,k_V(\cdot,x_i)\Big\|_V^2 \geq 0.$$

(iii) The claim follows from $\{k_V(\cdot,x) : x \in X\}^\perp = \{0\}$. Indeed, if $f \in \{k_V(\cdot,x) : x \in X\}^\perp$, then for all $x \in X$

$$f(x) = \langle f, k_V(\cdot,x)\rangle_V = 0,$$

that is, $f \equiv 0$.

(iv) ($\Rightarrow$) The Cauchy–Schwarz inequality implies

$$|(\delta_x \mid f)| = |f(x)| = |\langle f, k_V(\cdot,x)\rangle_V| \leq \|f\|_V\,\|k_V(\cdot,x)\|_V,$$

which shows that $\delta_x \in V^*$.

($\Leftarrow$) Since $\delta_x \in V^*$, by the Riesz representation theorem, there exists a unique $k_x \in V$ such that $(\delta_x \mid f) = \langle f, k_x\rangle_V$. We then define $k_V(\cdot,x) := k_x(\cdot)$. Since k_V satisfies the requirements of a reproducing kernel, the Hilbert space V is an RKHS.

□

Propositions 2.4.2(i) and 2.4.2(ii) show that we can map an RKHS to a symmetric and positive-semidefnite function. The converse is given by Moore's theorem [5, 78]. Therefore, RKHSs are characterized by symmetric and positive-semidefinite functions.

Theorem 2.4.3 (Moore)**.** *Let $k : X \times X \to \mathbb{R}$ be a symmetric and positive-semidefinite function. There exists a unique RKHS admitting k as the reproducing kernel.*

We refer the reader to [78, Proposition 2.3 and Theorem 2.14] for a complete proof. Actually, Proposition 2.4.2(iii) hints the idea of the proof of existence, which we now present. We first consider the vector space

$$W := \operatorname{span}\{k(\cdot, x) : x \in X\}$$

and a symmetric bilinear form B on W defined by

$$B\Big(\sum_{i=1}^{m} \alpha_i\, k(\cdot, x_i), \sum_{j=1}^{n} \beta_j\, k(\cdot, y_j)\Big) := \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\, \beta_j\, k(x_i, y_j).$$

We then check that B is well defined (a function $f \in W$ may be written as different linear combinations), and a more tricky part, that B is an inner product. Since k is positive semidefinite, we clearly have $B(f, f) \geq 0$, which is what we need to prove the Cauchy–Schwarz *in*equality. If $f = \sum_{i=1}^{n} \alpha_i\, k(\cdot, x_i)$ and $B(f, f) = 0$, it follows that for all $x \in X$

$$\begin{aligned} |f(x)| &= \Big|\sum_{i=1}^{n} \alpha_i\, k(x, x_i)\Big| \\ &= \big|B(f, k(\cdot, x))\big| \qquad \text{(by the definition of the bilinear form)} \\ &\leq \sqrt{B(f, f)}\,\sqrt{B(k(\cdot, x), k(\cdot, x))} \\ &= 0, \end{aligned}$$

which implies $f \equiv 0$. Let V be the completed Hilbert space in which W is dense. The final step is to identify elements in the abstract V with real-valued functions defined on X.

Proposition 2.4.2(iv) allows us to extend the definition of RKHSs from real-valued functions to Hilbert-space-valued functions. Since we will use Hilbert-space-valued RKHSs, $\mathbb{R}^d$-valued RKHSs to be more precise, at the slightest level, we refer the reader to [78, Chapter 6] for more details. Here we scratch the surface.

Definition 2.4.4. Let X be an abstract set and H be a Hilbert space. Denote by F the function space of H-valued functions defined on X. A vector subspace $V \subset F$ is called a *(H-valued) reproducing kernel Hilbert space* if V is a Hilbert space and all the evaluation operators are continuous, that is, $\delta_x \in \mathscr{L}(V, H)$ for all $x \in X$, where $\delta_x f := f(x)$.

The reproducing kernel $k_V : X \times X \to \mathscr{L}(H, H)$ is defined by $k_V(x, y) := \delta_x \delta_y^*$, where $\delta_y^* \in \mathscr{L}(H, V)$ is the adjoint operator of δ_y. To motivate this definition, given $h \in H$, we denote $k_V(\cdot, x)\, h := \delta_x^* h \in V$, then we have

$$\langle f(x), h\rangle_H = \langle \delta_x f, h\rangle_H = \langle f, \delta_x^* h\rangle_V = \langle f, k_V(\cdot, x)\, h\rangle_V,$$

which is the reproducing property of Hilbert-space-valued RKHSs. With some modifications of the statements and proofs, Proposition 2.4.2 and Moore's theorem (Theorem 2.4.3) can be carried over to Hilbert-space-valued RKHSs.

Proposition 2.4.2(ii) shows that every reproducing kernel is positive semidefinite. We will see that it is more convenient to work with positive definite reproducing kernels, which are characterized by the following proposition [78, Theorem 3.6].

Proposition 2.4.5. *Let V be a real-valued RKHS on X with the reproducing kernel k_V. Then the following are equivalent.*

(i) The reproducing kernel k_V is positive definite, that is, for any finite number of distinct points $\{x_1, \ldots, x_n\} \subset X$ the matrix $\left[\ k_V(x_i, x_j)\ \right]_{n\times n}$ is positive definite.

(ii) For any finite number of distinct points $\{x_1, \ldots, x_n\} \subset X$, there exist $f_1, \ldots, f_n \in V$ such that

$$f_i(x_j) = \begin{cases} 1, & \text{if } i = j\,; \\ 0, & \text{if } i \neq j\,. \end{cases}$$

Proof. ($\Rightarrow$) Let e_i be the $n \times 1$ vector whose ith element is one and all other elements are zero. Since the matrix $\left[\ k_V(x_i, x_j)\ \right]_{n\times n}$ is symmetric and positive definite, hence invertible, we can find $\alpha_{i,1}, \ldots, \alpha_{i,n} \in \mathbb{R}$ satisfying the linear system

$$\begin{bmatrix} & & \\ & k_V(x_i, x_j) & \\ & & \end{bmatrix} \begin{bmatrix} \alpha_{i,1} \\ \vdots \\ \alpha_{i,n} \end{bmatrix} = e_i.$$

Define $f_i := \alpha_{i,1}\, k_V(\cdot, x_1) + \cdots + \alpha_{i,n}\, k_V(\cdot, x_n) \in V$, then we have

$$f_i(x_j) = \alpha_{i,1}\, k_V(x_j, x_1) + \cdots + \alpha_{i,n}\, k_V(x_j, x_n) = \begin{cases} 1, & \text{if } i = j\,; \\ 0, & \text{if } i \neq j\,. \end{cases}$$

($\Leftarrow$) Observe that $\sum_{i,j=1}^n \alpha_i\, \alpha_j\, k_V(x_i, x_j) = \| \sum_{i=1}^n \alpha_i\, k_V(\cdot, x_i)\|_V^2 = 0$ if and only if $\sum_{i=1}^n \alpha_i\, k_V(\cdot, x_i) \equiv 0$. With the f_i's corresponding to x_i's, we have

$$\alpha_j = \sum_{i=1}^n \alpha_i\, f_j(x_i) = \sum_{i=1}^n \alpha_i\, \langle f_j, k_V(\cdot, x_i)\rangle_V = \Big\langle f_j, \sum_{i=1}^n \alpha_i\, k_V(\cdot, x_i)\Big\rangle_V = 0,$$

hence the matrix $\left[\ k_V(x_i, x_j)\ \right]_{n\times n}$ is positive definite. □

We close this section with a soft interpolation problem, which will be used in the next section. Let V be a real-valued RKHS on X. Given distinct points $\{x_i\}_{i=1}^n \subset X$ and values

$\{y_i\}_{i=1}^n \subset \mathbb{R}$, we fix $\gamma > 0$ and consider the soft interpolation problem

$$\min_{f \in V} \ \frac{\gamma}{2} \|f\|_V^2 + \frac{1}{2} \sum_{i=1}^n |f(x_i) - y_i|^2.$$

In words, a minimizer is the best $f \in V$, in the sense of the objective function, such that $f(x_i) \approx y_i$, $i = 1, \ldots, n$. It can be shown (see [104, Lemma 8.1 and the discussion after Theorem 8.2]) that the unique minimizer is of the form

$$f = \sum_{i=1}^n \alpha_i \, k_V(\cdot, x_i)$$

for some $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, n$, although the representation may not be unique. We denote $\alpha := \left[\ \alpha_i \ \right]_{n \times 1}$, $K := \left[\ k_V(x_i, x_j) \ \right]_{n \times n}$, and $y := \left[\ y_i \ \right]_{n \times 1}$, then the problem becomes

$$\min_{\alpha \in \mathbb{R}^n} \ \frac{\gamma}{2} \, \alpha^\top K \, \alpha + \frac{1}{2} \left| K\alpha - y \right|^2.$$

Since K is symmetric (Proposition 2.4.2(ii)), a minimizer of this equivalent problem satisfies

$$\gamma \, K\alpha + K \, (K\alpha - y) = 0, \tag{2.3}$$

which implies that the values of the unique minimizer at x_i's are given by

$$\left[\ f(x_i) \ \right]_{n \times 1} = K\alpha = (\gamma I_n + K)^{-1} K y,$$

where I_n is the n-by-n identity matrix. Note that $\gamma I_n + K$ is invertible since K is symmetric and positive semidefinite (Proposition 2.4.2(ii)). If the reproducing kernel k_V is positive definite, which implies that K is invertible, then from (2.3) we further have

$$\alpha = (\gamma I_n + K)^{-1} y.$$

2.4.2 Reproducing kernel Hilbert spaces embedded in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$

We first construct reproducing kernel Hilbert spaces continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R})$, then we proceed to $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$. Every Hilbert space H continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R})$ is an RKHS since

$$|(\delta_x \mid f)| = |f(x)| \leq \|f\|_{p,\infty} \leq C \, \|f\|_H.$$

The claim now follows from Proposition 2.4.2(iv). Hence we turn to build a Hilbert space continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R})$. Moreover, we prefer a classical Hilbert space that we are familiar with rather than an arbitrary Hilbert space.

The Sobolev embedding theorem [2, Theorem 4.12 (part I, case A)] states that for every integer $m > p + \frac{d}{2}$, the Sobolev space $H^m(\mathbb{R}^d)$ is continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R})$. The conventional norm on $H^m(\mathbb{R}^d)$ is given by

$$\|f\|_{H^m} := \left(\sum_{|\alpha| \le m} \|\partial^\alpha f\|_{L^2}^2 \right)^{1/2},$$

where α is a multi-index. Plancherel's theorem shows that

$$\|f\|_{H^m}^2 = \int_{\mathbb{R}^d} \left(\sum_{|\alpha| \le m} |\xi|^{2\alpha} \right) |\hat{f}|^2 \, d\xi,$$

where $\hat{f} \in L^2(\mathbb{R}^d)$ is the Fourier transform of $f \in L^2(\mathbb{R}^d)$. Given $\sigma > 0$, since there are constants C, $C' > 0$ such that

$$C \, (1 + \sigma^2 |\xi|^2)^m \le \sum_{|\alpha| \le m} |\xi|^{2\alpha} \le C' \, (1 + \sigma^2 |\xi|^2)^m,$$

it follows that the Sobolev space $H^m(\mathbb{R}^d)$ is also characterized by

$$H^m(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) : (1 + \sigma^2 |\xi|^2)^{m/2} \, \hat{f} \in L^2(\mathbb{R}^d) \right\}$$

with the equivalent norm

$$|||f|||_{H^m} := \left\| \, (1 + \sigma^2 |\xi|^2)^{m/2} \, \hat{f} \, \right\|_{L^2}.$$

For the purpose of obtaining an explicit expression of the reproducing kernel of $H^m(\mathbb{R}^d)$, we equip $H^m(\mathbb{R}^d)$ with the equivalent norm $|||\cdot|||_{H^m}$ and denote its inner product by

$$\langle f, g \rangle_{H^m} := \int_{\mathbb{R}^d} (1 + \sigma^2 |\xi|^2)^m \, \hat{f}(\xi) \, \overline{\hat{g}(\xi)} \, d\xi.$$

We will discuss the role of σ in the following paragraphs.

Assuming $m > p + \frac{d}{2}$, we now compute the reproducing kernel k_{H^m} of the RKHS $H^m(\mathbb{R}^d)$. The reproducing property gives us

$$f(x) = \langle f, k_{H^m}(\cdot, x) \rangle_{H^m} = \int_{\mathbb{R}^d} (1 + \sigma^2 |\xi|^2)^m \, \hat{f}(\xi) \, \overline{\hat{k}_{H^m}(\xi, x)} \, d\xi. \tag{2.4}$$

On the other hand, we have the Fourier inversion formula

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) \, e^{i \, 2\pi \, \xi^\top x} \, d\xi. \tag{2.5}$$

Since equations (2.4) and (2.5) hold for all $f \in H^m(\mathbb{R}^d)$, we obtain

$$\hat{k}_{H^m}(\xi, x) = \frac{e^{-i\,2\pi\,\xi^\top x}}{(1+\sigma^2|\xi|^2)^m},$$

which, recall that $k_{H^m}(\cdot, x) \in H^m(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$, implies

$$k_{H^m}(y, x) = \int_{\mathbb{R}^d} \frac{e^{i\,2\pi\,\xi^\top (y-x)}}{(1+\sigma^2|\xi|^2)^m}\, d\xi.$$

Using the modified Bessel function of the second kind K_ν, the above integral can be written as (see [104, page 233]):

$$k_{H^m}(x, y) = C_{m,d,\sigma} \left(\frac{|x-y|}{\sigma}\right)^{m-d/2} K_{m-d/2}\left(\frac{|x-y|}{\sigma}\right), \tag{2.6}$$

where $C_{m,d,\sigma}$ does not depend on x and y. After normalizing by a constant such that $k_{H^m}(x, x) = 1$, if $d = 2$, using the same notation for the normalized k_{H^m}, we then have

$$k_{H^m}(x, y) = \frac{1}{2^{m-2}\,(m-2)!} \left(\frac{|x-y|}{\sigma}\right)^{m-1} K_{m-1}\left(\frac{|x-y|}{\sigma}\right) \qquad \text{for } m \geq 2.$$

If $d = 3$, we further have explicit formulas

$$k_{H^2}(x, y) = \exp\left(-\frac{|x-y|}{\sigma}\right);$$

$$k_{H^3}(x, y) = \exp\left(-\frac{|x-y|}{\sigma}\right)\left(1 + \frac{|x-y|}{\sigma}\right);$$

$$k_{H^4}(x, y) = \exp\left(-\frac{|x-y|}{\sigma}\right)\left(1 + \frac{|x-y|}{\sigma} + \frac{1}{3}\left(\frac{|x-y|}{\sigma}\right)^2\right);$$

$$k_{H^5}(x, y) = \exp\left(-\frac{|x-y|}{\sigma}\right)\left(1 + \frac{|x-y|}{\sigma} + \frac{2}{5}\left(\frac{|x-y|}{\sigma}\right)^2 + \frac{1}{15}\left(\frac{|x-y|}{\sigma}\right)^3\right);$$

$$k_{H^6}(x, y) = \exp\left(-\frac{|x-y|}{\sigma}\right)\left(1 + \frac{|x-y|}{\sigma} + \frac{3}{7}\left(\frac{|x-y|}{\sigma}\right)^2 + \frac{2}{21}\left(\frac{|x-y|}{\sigma}\right)^3 + \frac{1}{105}\left(\frac{|x-y|}{\sigma}\right)^4\right).$$

Up to a constant, the kernel k_{H^m} is also known as the Matérn kernel. For $d = 2$, there exist numerical methods to approximate $K_\nu(x)$, $\nu \in \mathbb{N}$. Unfortunately, $\lim_{x\to 0^+} K_\nu(x) = \infty$, even though $\lim_{x\to 0^+} x^\nu K_\nu(x) < \infty$. We provide a procedure in Appendix A.1 to directly approximate $x^\nu K_\nu(x)$, $\nu \in \mathbb{N}$, by modifying an existing method for $K_\nu(x)$ [29, 87] and bypassing the singularity of K_ν at 0. In addition, we notice from the exponent in (2.6) that

$$m - \frac{d}{2} = \left(m - \frac{1}{2}\right) - \frac{d-1}{2},$$

which combining with (2.6) says that the kernel k_{H^m} in $\mathbb{R}^d$ restricted to $\mathbb{R}^{d-1}$ is the kernel

$k_{H^{m-1/2}}$ in $\mathbb{R}^{d-1}$. This is not surprising in view of the trace theorem. Thus we can actually compute k_{H^m} for all $m = 2 + \frac{n}{2}$, where $n \in \mathbb{N} \cup \{0\}$. Figure 2.4 plots the reproducing kernel k_{H^m} with $m = 2, 3, \ldots, 6$ and $\sigma = 1$.

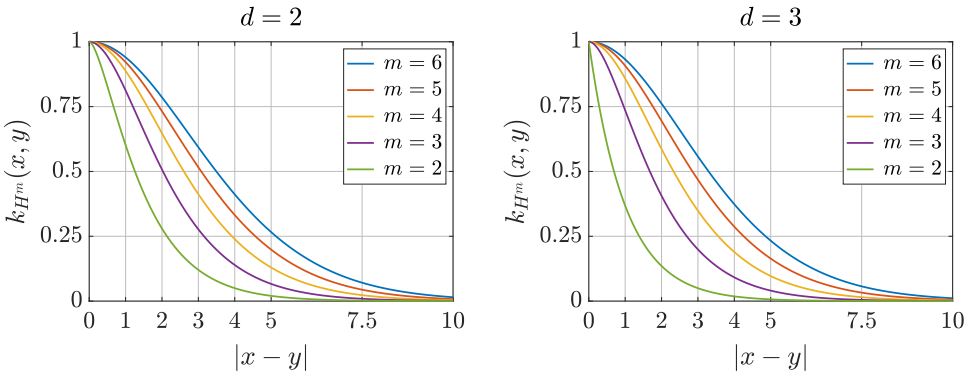


Figure 2.4: Plots of the reproducing kernel k_{H^m} of $H^m(\mathbb{R}^d)$ with $\sigma = 1$.

The effect of σ can be observed through the soft interpolation problem. In the soft interpolation problem, recall that we minimize

$$\frac{\gamma}{2} |||f|||_{H^m}^2 + \frac{1}{2} \sum_{i=1}^{n} |f(x_i) - y_i|^2 = \frac{\gamma}{2} \left\| (1 + \sigma^2 |\xi|^2)^{m/2} \hat{f} \right\|_{L^2}^2 + \frac{1}{2} \sum_{i=1}^{n} |f(x_i) - y_i|^2.$$

Thus, qualitatively, σ serves as the weight on the penalties of derivatives. If σ is small, we tolerate functions with large derivatives; if σ is large, we are inclined to smoother functions. Figure 2.5 illustrates this viewpoint on σ quantitatively. A heuristic rule of thumb is to choose σ such that the union of balls centered at x_i's with radius σ is connected for each cluster of x_i's, although the radius as a function of σ may be adjusted according to Figure 2.4.

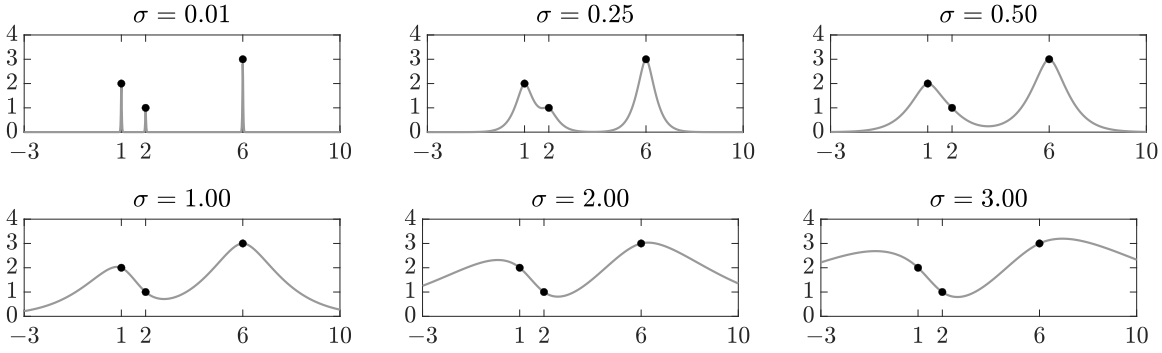


Figure 2.5: Soft interpolation of 3 data points by a function in $H^2(\mathbb{R}) \hookrightarrow C_0^1(\mathbb{R}, \mathbb{R})$ with $\gamma = 10^{-6}$. The heuristic suggests using $0.5 < \sigma < 2$ if we view the x_i's as 2 clusters ($\{1, 2\}$ and $\{6\}$) and using $\sigma > 2$ if all x_i's are lumped in 1 cluster.

We have derived the computable reproducing kernel of the RKHS $H^m(\mathbb{R}^d)$, which is continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R})$. Based on $H^m(\mathbb{R}^d)$, we can construct a Hilbert space continuously embedded in $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ by treating each coordinate as an individual $H^m(\mathbb{R}^d)$. Since the construction is not specific to $H^m(\mathbb{R}^d)$, we let V be any real-valued RKHS on a set X. We now consider the Hilbert space

$$V^d := \big\{(f_1, \ldots, f_d) : f_i \in V,\ i = 1, \ldots, d\big\}$$

equipped with the inner product

$$\langle f, g\rangle_{V^d} := \sum_{i=1}^{d} \langle f_i, g_i\rangle_V,$$

where $f = (f_1, \ldots, f_d)$ and $g = (g_1, \ldots, g_d)$. We show that V^d is an $\mathbb{R}^d$-valued RKHS on X and compute its reproducing kernel. Note that

$$|\delta_x f|^2 = \sum_{i=1}^{d} \big(f_i(x)\big)^2 = \sum_{i=1}^{d} \langle f_i, k_V(\cdot, x)\rangle_V^2 \le \sum_{i=1}^{d} \Big(\|f_i\|_V^2\, \|k_V(\cdot, x)\|_V^2\Big) = k_V(x, x)\, \|f\|_{V^d}^2,$$

which shows that $\delta_x \in \mathscr{L}(V^d, \mathbb{R}^d)$ for all $x \in X$ and hence V^d is an $\mathbb{R}^d$-valued RKHS by Definition 2.4.4. Recall that $k_{V^d}(x, y) = \delta_x \delta_y^*$ (page 14), so next we compute δ_y^*. Let $f = (f_1, \ldots, f_d) \in V^d$ and $h = (h_1, \ldots, h_d) \in \mathbb{R}^d$, by the definition of an adjoint operator, we have

$$\langle f, \delta_y^* h\rangle_{V^d} = \langle \delta_y f, h\rangle_{\mathbb{R}^d} = \sum_{i=1}^{d} f_i(y)\, h_i = \sum_{i=1}^{d} \langle f_i, k_V(\cdot, y)\rangle_V\, h_i = \langle f, k_V(\cdot, y)\, h\rangle_{V^d}.$$

It follows that $k_{V^d}(x, y)\, h = \delta_x \delta_y^*\, h = k_V(x, y)\, h$, or

$$k_{V^d}(x, y) = k_V(x, y)\, id \ \in \mathscr{L}(\mathbb{R}^d, \mathbb{R}^d) \cong \mathbb{R}^{d\times d}, \tag{2.7}$$

where $id : \mathbb{R}^d \to \mathbb{R}^d$ is the identity function.

We remark some properties of the real-valued RKHS $H^m(\mathbb{R}^d)$ and the $\mathbb{R}^d$-valued RKHS $(H^m(\mathbb{R}^d))^d$. Since there exists a $C^\infty(\mathbb{R}^d)$ function which is compactly supported in a neighborhood around $x \in \mathbb{R}^d$, it follows from Proposition 2.4.5 that the reproducing kernel of $H^m(\mathbb{R}^d)$ is positive definite. Another property we now examine is that a change of rectangular coordinates is an isometry in $(H^m(\mathbb{R}^d))^d$. Let $y := Rx + a$ be a change of rectangular coordinates, where $R \in \mathbb{R}^{d\times d}$ is a rotation and $a \in \mathbb{R}^d$ is a translation. Given $v : \mathbb{R}^d \to \mathbb{R}^d$, we define

$$w(y) := R\, v\big(R^\top (y - a)\big),$$

which is v after the change of rectangular coordinates. We say that a change of rectangular coordinates is an isometry in $(H^m(\mathbb{R}^d))^d$ in the sense that $\|v\|_{(H^m)^d} = \|w\|_{(H^m)^d}$. We will explain why this is a desired property in the next section. Here we check that this property holds. We write v and w in coordinates as $v(x) = (v_1(x), \ldots, v_d(x))$ and $w(y) = (w_1(y), \ldots, w_d(y))$ and define an intermediate vector field $v'(y) := \big(v_1\big(R^\top(y - a)\big), \ldots, v_d\big(R^\top(y - a)\big)\big)$, which will help the following derivation. Since

$$\hat{v}'_i(\xi) = \int_{\mathbb{R}^d} v'_i(y)\, e^{-i\,2\pi\,\xi^\top y}\, dy = \int_{\mathbb{R}^d} v_i(x)\, e^{-i\,2\pi\,\xi^\top (Rx + a)}\, dx = e^{-i\,2\pi\,\xi^\top a}\ \hat{v}_i(R^\top \xi),$$

we have $|\hat{v}'_i(\xi)| = |\hat{v}_i(R^\top \xi)|$. We denote $\hat{v} := (\hat{v}_1, \ldots, \hat{v}_d)$, same for w and v', and note that $w = R\,v'$ implies $\hat{w} = R\,\hat{v}'$. It follows that

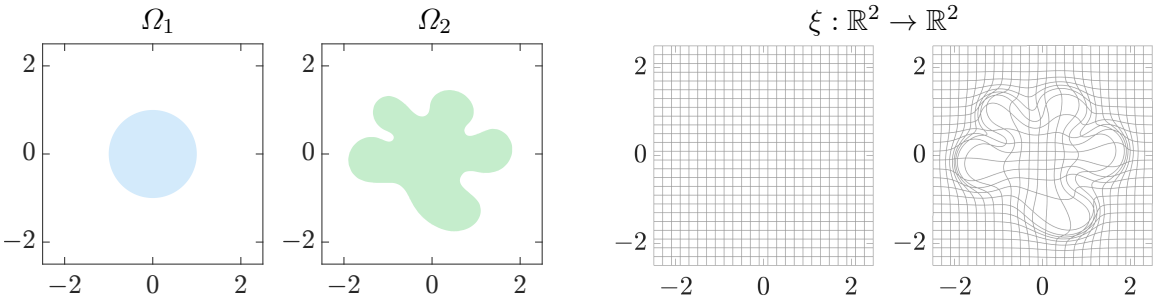
$$\begin{aligned}
\|w\|^2_{(H^m)^d} &= \sum_{i=1}^{d} |||w_i|||^2_{H^m} = \int_{\mathbb{R}^d} (1 + \sigma^2|\xi|^2)^m\, |\hat{w}(\xi)|^2\, d\xi \\
&= \int_{\mathbb{R}^d} (1 + \sigma^2|\xi|^2)^m\, |R\,\hat{v}'(\xi)|^2\, d\xi = \int_{\mathbb{R}^d} (1 + \sigma^2|\xi|^2)^m\, |\hat{v}'(\xi)|^2\, d\xi \\
&= \int_{\mathbb{R}^d} (1 + \sigma^2|\xi|^2)^m\, |\hat{v}(R^\top \xi)|^2\, d\xi = \int_{\mathbb{R}^d} (1 + \sigma^2|R\,\xi|^2)^m\, |\hat{v}(\xi)|^2\, d\xi \\
&= \int_{\mathbb{R}^d} (1 + \sigma^2|\xi|^2)^m\, |\hat{v}(\xi)|^2\, d\xi = \|v\|^2_{(H^m)^d},
\end{aligned}$$

which establishes the isometry.

2.5 Large Deformation Diffeomorphic Metric Mapping

Large deformation diffeomorphic metric mapping (LDDMM) [11] is a versatile framework for diffeomorphic registration of shapes. Given two shapes Ω_1 and Ω_2, LDDMM provides a Diff^p-deformation ξ such that $\xi(\Omega_1) \approx \Omega_2$. Figure 2.6 shows an example of LDDMM: Figure 2.6(a) shows input shapes Ω_1 and Ω_2, while Figure 2.6(b) shows the Diff^p-deformation ξ output by LDDMM. Since ξ is a diffeomorphism, we notice in Figure 2.6(b) that parallel lines are mapped to smooth nonintersecting curves. We next introduce how LDDMM generates a Diff^p-deformation for diffeomorphic shape registration.

The idea of LDDMM is to generate ξ not in one step, but through a Diff^p-motion φ on a time interval $[0, T]$ and let $\xi := \varphi(T)$, the Diff^p-deformation at the end of the Diff^p-motion. We recall from Section 2.3 that the flow of a deformation vector field $v \in L^1([0,T], C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d))$ is a Diff^p-motion. In addition, let V be a Hilbert space continuously embedded in $C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$, for example, we can let $V = (H^m(\mathbb{R}^d))^d$ with $m > p + 1 + \frac{d}{2}$



(a) Input of LDDMM.

(b) Output of LDDMM.

Figure 2.6: Input and output of LDDMM.

(see Section 2.4.2). LDDMM considers the minimization problem

$$\min_{v' \in L^2([0,T],V)} \left(\frac{1}{2} \int_0^T \|v'(t)\|_V^2 \, dt + \rho(\varphi(T, \Omega_1), \Omega_2) \right) \tag{2.8}$$

subject to

$$\varphi(t,x) = x + \int_0^t v'(s, \varphi(s,x)) \, ds \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d,$$

where $\rho(\cdot,\cdot)$ is a function measuring the discrepancy between two shapes. With additional regularity assumptions on ρ, applying the direct method of calculus of variations shows the existence of a minimizer for the LDDMM minimization problem (2.8) due to the Hilbert space structure of V. We can now summarize the procedure of LDDMM for diffeomorphic shape registration. Given two shapes Ω_1 and Ω_2, we denote a minimizer for (2.8) by v and denote its corresponding Diff^p-motion by φ_v. A Diff^p-deformation which registers Ω_1 to Ω_2 is then given by $\xi := \varphi_v(T)$. For the two shapes in Figure 2.6(a), we set $T = 1$ and present in Figure 2.7 a critical point v of the minimization problem (2.8) and its corresponding Diff^p-motion φ_v. The deformed grids should be interpreted as functions from the regular grid to the deformed grids. The Diff^p-deformation at $t = 1.00 = T$ is the output we have seen in Figure 2.6(b). Figure 2.8 verifies that the registered shape $\xi(\Omega_1)$ is close to Ω_2 in the sense that they have similar boundaries.

When we register a shape Ω_1 to another shape Ω_2 using LDDMM, there is a Diff^p-motion of Ω_1 behind the scenes so that Ω_1 is the beginning of the motion while Ω_2 is the target that the motion tries to reach. Henceforth, we shall call Ω_1 an initial shape and Ω_2 a target shape and denote them by Ω_0 and Ω_{targ} respectively. Recall that in Section 2.4.2 we observe a property of $(H^m(\mathbb{R}^d))^d$ that a change of rectangular coordinates is an isometry in $(H^m(\mathbb{R}^d))^d$. Now we can see why this is a desired property. This property implies that the LDDMM minimization problem (2.8) is independent of the choice of rectangular coordinate

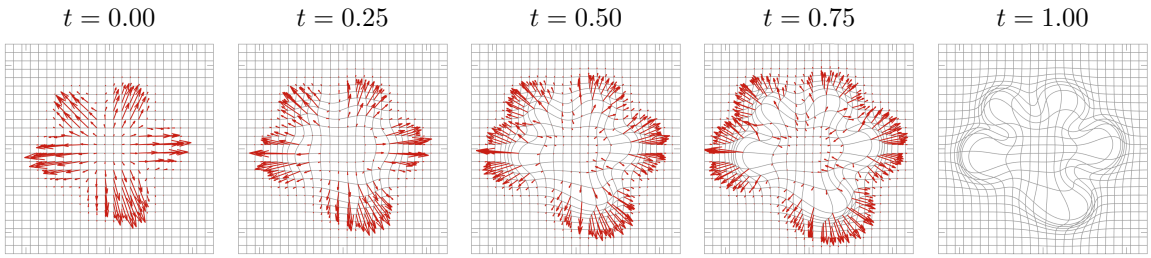
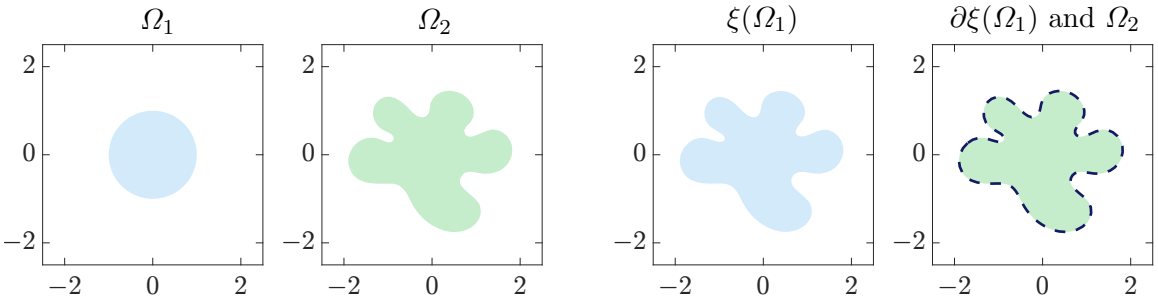


Figure 2.7: A critical point of the LDDMM minimization problem. Shown in the figures is a critical deformation vector field v and the corresponding Diff^p-motion φ_v.



(a) Input of LDDMM.

(b) Registration result.

Figure 2.8: LDDMM diffeomorphic registration.

system. In this sense, we recover shapes as compact sets in the Euclidean affine space. Besides, it would be awkward to obtain a different registration result simply because we did not place shapes in some standard orientation at the origin.

The applications of LDDMM go beyond shapes we consider here. Notice that the three factors in the LDDMM framework are a class of objects, an action of the diffeomorphism group on the class of objects, and a function measuring the discrepancy between two objects. Given two objects $\mathcal{O}_0$ and $\mathcal{O}_{\text{targ}}$, the LDDMM minimization problem then formally becomes

$$\min_{v' \in L^2([0,T],V)} \left(\frac{1}{2} \int_0^T \|v'(t)\|_V^2 \, dt + \rho(\varphi(T) \cdot \mathcal{O}_0, \mathcal{O}_{\text{targ}}) \right)$$

subject to

$$\varphi(t,x) = x + \int_0^t v'(s, \varphi(s,x)) \, ds \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d.$$

The class of objects in the previous discussion is the class of shapes $\mathscr{S}$ with the action $\xi \cdot \Omega := \xi(\Omega)$. A discrepancy function can be, for example, the volume of the symmetric difference of two sets. We can also consider a class of images, or real-valued functions on $\mathbb{R}^d$, with the action $\xi \cdot I := I \circ \xi^{-1}$. For a class of vector fields, possible actions include $\xi \cdot v := (D\xi \; v) \circ \xi^{-1}$ and $\xi \cdot v := (D\xi^{-\top} \, v) \circ \xi^{-1}$. We refer the reader to [104, Chapter 9] for

a wide range of objects, actions, and discrepancy functions.

2.6 Problem Description

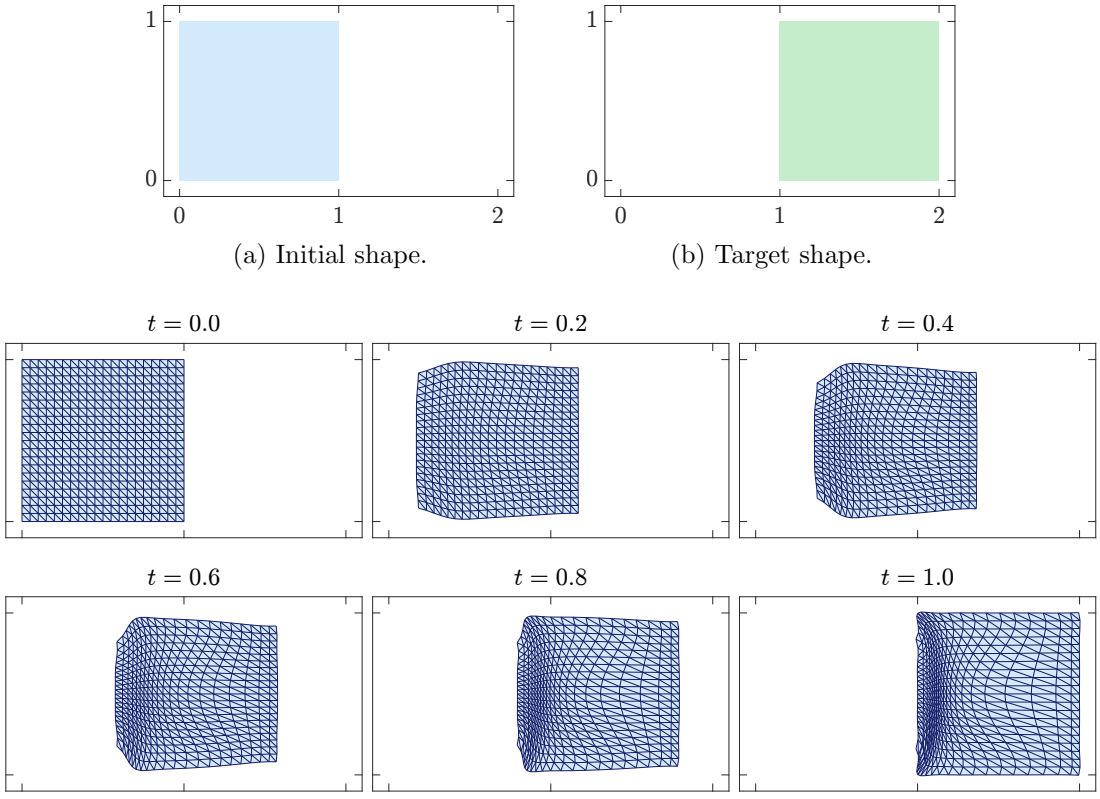
While LDDMM registers two shapes via a Diff^p-motion, the Diff^p-motion may not have a physical interpretation. For example, when we register shapes extracted from brain images of different individuals using LDDMM, there is no Diff^p-motion linking the two individual brains in reality. However, if we know that the target shape indeed comes from the initial shape through a Diff^p-motion, we can expect that incorporating prior knowledge into deformation vector fields would enhance registration results. Figure 2.9 demonstrates this idea. Figures 2.9(a) and 2.9(b) show our initial and target shapes, which are two adjacent squares. Although the most straightforward Diff^p-motion is a simple translation, there are in fact infinitely many Diff^p-motions from the initial shape to the target shape; Figure 2.9(c) shows one possible Diff^p-motion. The distorted mesh may conflict with our understanding that the internal structure should remain almost unchanged. If this is the case, it would be better to model deformation vector fields as the form of velocity fields of rigid motions. More importantly, our motivation is to reveal an approximately true motion simply from an initial shape and an asymmetrical target shape. This possibility resides in a more faithful description of deformation vector fields.

We propose to investigate when deformation vector fields can be modeled explicitly through a parameter. Let V be a Hilbert space continuously embedded in $C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$. In addition, let $\varphi \in C([0,T], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$ and $\theta : [0,T] \to Y$, where Y is a Banach space. We assume that the deformation vector field at the current time takes the form

$$v(t) := \mathcal{M}(\varphi_{|[0,t]}, \theta(t)), \tag{2.9}$$

where $\mathcal{M} : \bigcup_{t \in [0,T]} C([0,t], \mathrm{Diff}_{id}^p(\mathbb{R}^d)) \times Y \to V$ is an abstract model which depends on a motion on some time interval $[0,t] \subset [0,T]$ and a Y-valued parameter. When a deformation vector field is of the form (2.9), we say that it is an interpretable deformation vector field, i.e., $v(t)$ can be interpreted through the model $\mathcal{M}$ and the parameter $\theta(t)$. In (2.9), the model assumption and the dependency on a parameter are sensible, while the dependency on an arbitrary motion should be understood in the context of its (coupled) flow. Note that the "flow" of (2.9) is a solution to the initial value problem

$$\varphi(t,x) = x + \int_0^t \mathcal{M}(\varphi_{|[0,s]}, \theta(s))(\varphi(s,x))\, ds \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d,$$



(a) Initial shape. (b) Target shape.

(c) A feasible registration process.

Figure 2.9: Diffeomorphic registration with unmodeled deformation vector fields.

which reveals the dependency of the deformation vector field at the current time on the history of its motion up to the current time. We will examine sufficient conditions on $\mathcal{M}$ for the existence and uniqueness of solutions to this initial value problem and then consider the minimization problem

$$\min_{\theta \in \Theta} \left(\int_0^T \Lambda(\varphi(t), \theta(t))\, dt + \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}}) \right) \tag{2.10}$$

subject to

$$\varphi(t, x) = x + \int_0^t \mathcal{M}(\varphi|_{[0,s]}, \theta(s))(\varphi(s, x))\, ds \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d.$$

We will also give sufficient conditions such that a minimizer of the above minimization problem exists. This abstract formulation will be realized by two examples in the next chapter, where we present motions corresponding to critical points of the minimization problem (2.10) through a series of numerical experiments.

Chapter 3

Two Examples

The purpose of this chapter is to illustrate our abstract framework of interpretable deformation vector fields. We will focus on concepts and leave many loose ends to be tied up in the following chapters. In other words, we provisionally assume everything works. In Section 3.1, we consider the shape registration problem with shapes composed of multiple connected components. Moreover, we assume that the motion of each connected component is rigid while the motion of the whole space is Diff^p. In Section 3.2, we consider a model of quasi-elastic shapes atrophied by a chemical whose spread in the shape is governed by a reaction-diffusion equation. We then abstract these two examples under the framework of interpretable deformation vector fields in Section 3.3, which serves as an appetizer for the next chapter.

3.1 Piecewise-rigid Motion

Our first example assumes that shapes go through piecewise-rigid motions. Let $\Omega = \bigcup_{i=1}^{N} \Omega_i$ be our shape (see (2.1)). We say that a Diff^p-motion is piecewise rigid if it is a rigid motion of each shape component Ω_i, $i = 1, \ldots, N$. Figure 3.1 illustrates one piecewise-rigid motion with three shape components. Since a piecewise-rigid motion is also a Diff^p-motion, all the shape components will not have contacts at any time during the motion. We will formulate this problem in Section 3.1.1 and demonstrate a series of numerical experiments in Section 3.1.2.

Piecewise-rigid motion is a natural model when registering physical shapes enclosing bones. A wide range of literature in medical image registration has addressed this direction. There are at least three approaches for registering images containing bones. The first approach registers each bone, or rigid component, separately and then "fill the blanks"

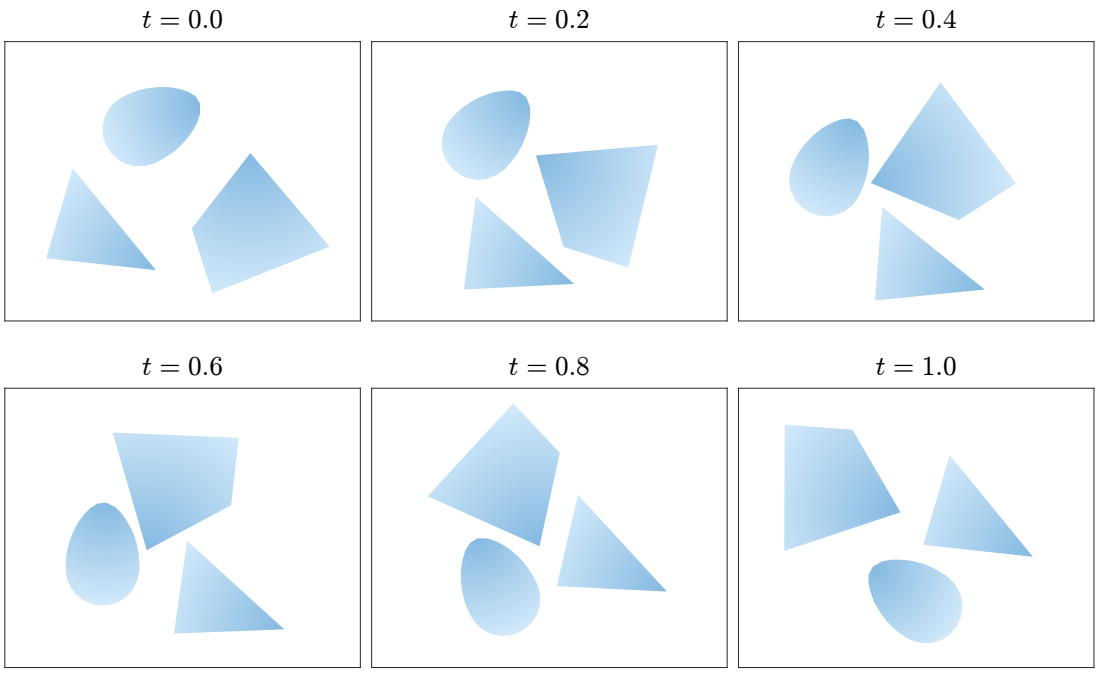


Figure 3.1: Illustration of piecewise-rigid motions. The color shading is to facilitate the orientation tracing.

by a non-rigid deformation [63, 54, 69, 24]. Some problems may occur when blending rigid deformations with a non-rigid deformation. In [63, 69, 24], a non-rigid deformation is generated from a weighted average of rigid deformations, where the weights are functions of the inverse distance to rigid components. This weighted average of rigid deformations implies that the material is more rigid when it is closer to rigid components. However, this continuity in stiffness may not be true in some applications. In [54], on the other hand, a non-rigid deformation is given by a thin-plate spline based on landmarks extracted from the boundaries of registered rigid components. Although the boundaries of rigid components are rigidly registered using a thin-plate spline, the regions within the boundaries are not rigid anymore. One needs to evaluate if this behavior meets the assumption of applications. The second approach penalizes non-rigid deformations via the violation of rigidity on rigid components [64, 86, 92], that is, via the amount of deviation of the right Cauchy–Green deformation tensor from the identity on rigid components. The third approach utilizes the articulated structure of bones [77, 8, 37, 58, 102]. This approach requires an articulated atlas, which may not be conveniently accessible. Our method differs from all the three approaches mentioned above. Following the idea of LDDMM, we aim to find a piecewise-rigid motion φ such that $\varphi(T, \Omega_0) \approx \Omega_{\text{targ}}$. Our method guarantees that the piecewise-rigid deformation $\varphi(T)$ is diffeomorphic, which is a desired property not always addressed in the

literature at the continuous level. We will demonstrate our method using shapes, though this method can also be applied to images with a suitable discrepancy function. We also refer the reader to [52] for more numerical results.

3.1.1 Problem formulation

To formulate the problem, we first cite the following characterizations of rigid motions [49, pages 49 and 69].

Proposition 3.1.1 (Characterizations of rigid motions)**.** *Let φ be a* Diff^p*-motion and v be the corresponding deformation vector field. Then the following are equivalent.*

(i) The motion φ is rigid, i.e., $|\varphi(t,x) - \varphi(t,x')| = |x - x'|$ for all $t \in [0,T]$ and all $x, x' \in \mathbb{R}^d$.

(ii) The motion φ admits the representation

$$\varphi(t,x') = \varphi(t,x) + R(t)\,(x' - x)$$

for all $t \in [0,T]$ and all $x, x' \in \mathbb{R}^d$, where $R(t) \in \mathbb{R}^{d\times d}$ is a rotation, i.e., $R(t)$ is orthogonal and $\det R(t) = 1$.

(iii) The deformation vector field v admits the representation

$$v(t,x') = v(t,x) + W(t)\,(x' - x) \tag{3.1}$$

for almost every $t \in [0,T]$ and all $x, x' \in \mathbb{R}^d$, where $W(t) \in \mathbb{R}^{d\times d}$ is skew-symmetric.

(iv) The derivative $Dv(t,x)$ is skew-symmetric for almost every $t \in [0,T]$ and all $x \in \mathbb{R}^d$.

Proposition 3.1.1(i) is the definition of rigid motions. We formulate the problem using Proposition 3.1.1(iii) in this section and provide one alternative using Proposition 3.1.1(iv) in Section 5.1.2 from a theoretical aspect and another alternative using Proposition 3.1.1(ii) in Section 6.1 from a numerical aspect. We remark that Proposition 3.1.1(iv) does not require $Dv(t,x)$ to be skew-symmetric *and constant in space*, which makes it an easy equivalent when formulating rigid motions. Rigid motions of $\Omega \subset \mathbb{R}^d$ can also be characterized by suitable restrictions in Proposition 3.1.1.

We proceed to formulate the problem using Proposition 3.1.1(iii). We first derive an expression of deformation vector fields of rigid motions of $\mathbb{R}^d$, then we return to our problem of piecewise-rigid motions. We let $x = 0$ in (3.1) and define $u(t) := v(t,0)$, which leads to

$$v(t,x') = u(t) + W(t)\,x'. \tag{3.2}$$

It is clear that (3.2) implies (3.1), and hence they are equivalent. Moreover, since $W(t)$ is skew-symmetric, there exists $\omega(t) \in \mathbb{R}$ if $d = 2$ or $\omega(t) \in \mathbb{R}^3$ if $d = 3$ such that

$$v(t, x') = \begin{cases} u(t) + \omega(t) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x', & \text{if } d = 2\,; \\ u(t) + \omega(t) \times x', & \text{if } d = 3\,, \end{cases} \tag{3.3}$$

where $\times$ stands for the cross product. For all $\theta := (u, \omega)$, where $u \in \mathbb{R}^d$ and $\omega \in \mathbb{R}^{\frac{d(d-1)}{2}}$, the expression (3.3) then inspires the definition of a linear operator $\mathcal{V} : \mathbb{R}^{\frac{d(d+1)}{2}} \to C(\mathbb{R}^d, \mathbb{R}^d)$ given by

$$(\mathcal{V}\,\theta)(x) := \begin{cases} u + \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x, & \text{if } d = 2\,; \\ u + \omega \times x, & \text{if } d = 3\,, \end{cases}$$

which characterizes deformation vector fields of rigid motions. Now we are ready to formulate our problem of piecewise-rigid motions. Given an initial shape $\Omega_0 = \bigcup_{i=1}^N \Omega_i$ and a target shape $\Omega_{\text{targ}} = \bigcup_{i=1}^N \Omega_i'$, we let $\theta_i(t) := (u_i(t), \omega_i(t)) \in \mathbb{R}^{\frac{d(d+1)}{2}}$ be the parameter of a rigid deformation vector field for the i-th shape component at time t and define $\theta(t) := (\theta_1(t), \ldots, \theta_N(t)) \in \mathbb{R}^{\frac{Nd(d+1)}{2}}$. In addition, we let $k := \frac{Nd(d+1)}{2}$ be the total number of parameters in the following for conciseness. The problem of piecewise-rigid motions is stated as

$$\min_{\theta \in L^2([0,T], \mathbb{R}^k)} \left(\frac{1}{2} \int_0^T \left(\|v(t)\|_V^2 + |\theta(t)|^2 \right) dt + \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}}) \right)$$

subject to

$$\begin{cases} \varphi(t, x) = x + \displaystyle\int_0^t v(s, \varphi(s, x))\, ds \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d \\ v(t) = \displaystyle\operatorname*{arg\,min}_{v' \in V} \left(\frac{\gamma}{2} \|v'\|_V^2 + \frac{1}{2} \sum_{i=1}^N \int_{\varphi(t, \Omega_i)} |v' - \mathcal{V}\,\theta_i(t)|^2\, dx \right) \end{cases}, \tag{3.4}$$

where $\gamma > 0$ is a specified small value for technical reasons (see Lemma 4.1.13). The term $\frac{1}{2} \int_0^T \|v(t)\|_V^2\, dt$ in the objective function quantifies the roughness of deformation vector fields, while the term $\frac{1}{2} \int_0^T |\theta(t)|^2\, dt$ resembles the kinetic energy of shapes. In words, we try to find a piecewise-rigid motion which is as smooth as possible and costs as small kinetic energy as possible such that $\varphi(T, \Omega_0) \approx \Omega_{\text{targ}}$. For any diagonal matrix D with positive diagonal entries, a more general form $\frac{1}{2} \int_0^T \theta(t)^\top D\, \theta(t)\, dt$ can be used for the kinetic energy. We choose D to be the identity matrix for simplicity. We will have more discussions on

this problem formulation, including technical and numerical difficulties, in later sections (Sections 5.1 and 6.1). Here we strengthen that $v(t) \in V \hookrightarrow C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ is a smooth vector field with the property $v(t,x) \approx (\mathcal{V}\,\theta_i(t))(x)$ for $x \in \varphi(t, \Omega_i)$, i.e., the motion of each shape component is almost rigid.

3.1.2 Numerical results

We present numerical results of a slightly different problem which will be stated in Section 6.1. This slightly different problem can be thought of as the limiting case of (3.4) when $\gamma \to 0$, i.e., the motion of each shape component is exactly rigid. All the 2D experiments used $V = (H^5(\mathbb{R}^d))^d \hookrightarrow C_0^3(\mathbb{R}^d, \mathbb{R}^d)$, and the 3D experiment used $V = (H^6(\mathbb{R}^d))^d \hookrightarrow C_0^4(\mathbb{R}^d, \mathbb{R}^d)$. The final time was set to $T = 1$. The kernel width σ for V, which will be denoted by σ_V, and the time step size Δt for the forward Euler method will be indicated in each experiment. We used the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm to find a critical point of the minimization problem with the starting point $\theta \equiv 0$ corresponding to $\varphi_\theta(t) = id$ for all t unless stated otherwise. We chose the simple forward Euler method considering the steeply increased complexity in implementation. Nevertheless, whether the optimization can benefit from higher order methods is an interesting topic, which will not be addressed in this dissertation.

We first examine the registration problem without any correspondence between the initial and the target shape. This can be achieved by the minimization problem (3.4) if we adopt a discrepancy function ρ that does not require any correspondence. In particular, we consider varifold pseudo-metrics [27]. Let

$$\nu(\partial\Omega, \partial\Omega') := \int_{\partial\Omega}\int_{\partial\Omega'} \frac{1}{\left(1 + |x - x'|^2/\sigma_\rho^2\right)^2} \Big(n(x)^\top n'(x)\Big)^2 d\sigma'\, d\sigma,$$

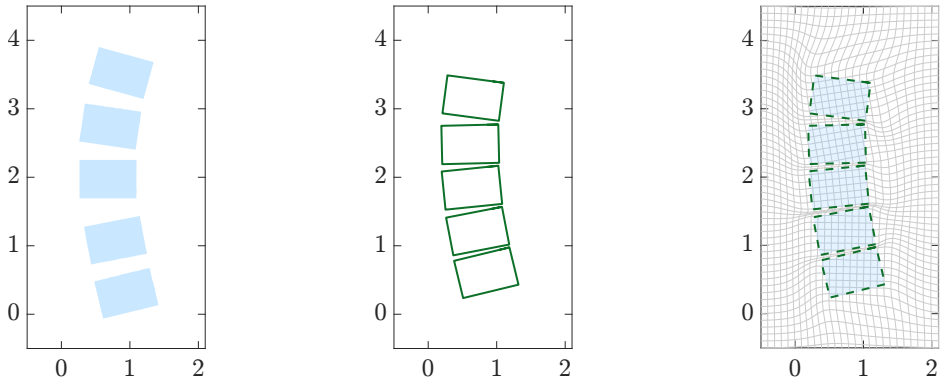
where $\sigma_\rho > 0$ is a parameter for precision, n and n' are unit tangent vector fields for 2D and unit normal vector fields for 3D on $\partial\Omega$ and $\partial\Omega'$, and σ and σ' are volume measures of $\partial\Omega$ and $\partial\Omega'$ respectively. We then define

$$\rho(\Omega, \Omega') := w_\rho\Big(\nu(\partial\Omega, \partial\Omega) - 2\nu(\partial\Omega, \partial\Omega') + \nu(\partial\Omega', \partial\Omega')\Big), \tag{3.5}$$

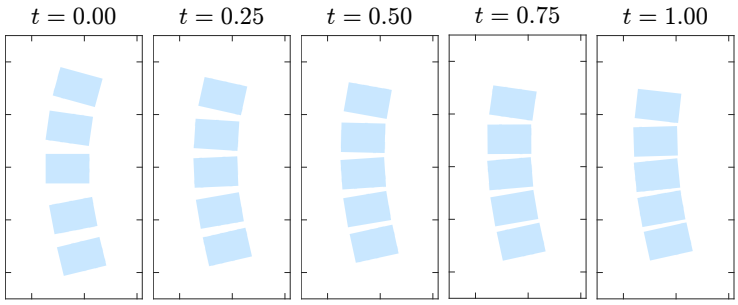
where $w_\rho > 0$ is a weighting parameter. Notice that the discrepancy function ρ defined above only measures the difference between the boundaries $\partial\Omega$ and $\partial\Omega'$, which is necessary and sufficient for the shapes we consider since $\partial\Omega = \partial\Omega'$ if and only if $\Omega = \Omega'$. Moreover, this discrepancy function ρ is "continuous," which will be made precise in Definition 4.1.4 and Remark 4.1.9. The reader may loosely interpret the continuity of ρ as $\rho(\Omega, \Omega')$ only

changes a little when Ω "changes a little," where the change is quantified by a deformation applying to Ω. We also remark that $\nu(\partial\Omega, \partial\Omega')$, and hence $\rho(\partial\Omega, \partial\Omega')$, is well defined for our shapes whose boundaries are rectifiable (see (2.1)).

With all the ingredients in place, we now supply inputs Ω_0 and Ω_{targ} to the minimization problem (3.4) using the discrepancy function (3.5) and investigate the output θ, a critical point. We present an example of a conceptual spine in Figure 3.2. Figures 3.2(a) and 3.2(b) are our Ω_0 and $\partial\Omega_{\text{targ}}$; Figure 3.2(c) shows the deformation $\varphi_\theta(T)$ corresponding to a critical point θ; Figure 3.2(d) shows $\varphi_\theta(t, \Omega_0)$ at various time t. Since the motion is piecewise rigid, we expect that grid lines within shape components remain undeformed. Figure 3.2(c) might not show this property clearly; we encourage the reader to check this property in the following examples. We emphasize that the only input data are Ω_0 and Ω_{targ}, that is, we do not assume the knowledge of any correspondence between the initial and the target shape. In this example, the positions of the initial and the target shape are quite close, and we may say that the registration result meets our expectation. Next we increase the difference between the initial and the target shape as shown in Figure 3.3. We can see from Figure 3.3(c) that $\varphi(T, \Omega_0)$ and Ω_{targ} are registered. However, Figure 3.3(d) reveals that the corresponding piecewise-rigid motion may not be the most intuitive one. Even if we change rectangles to other asymmetric shapes, as shown in Figure 3.4, which provide the information of orientation, we still have a difficult minimization problem since we do not have any correspondence to directly hint a minimization procedure that rotating the top quadrangle 180 degrees will lead to a better minimizer. Figure 3.5 shows another example indicating that we may need more information than the boundary of a target shape when the difference between the initial and the target shape is large. Suppose that we have a rough prior of expected motions, which may be provided either by users or previous data, then one possible approach is to utilize the expected motion as the starting point for a minimization procedure. We will see the effect of starting points in the next paragraph.

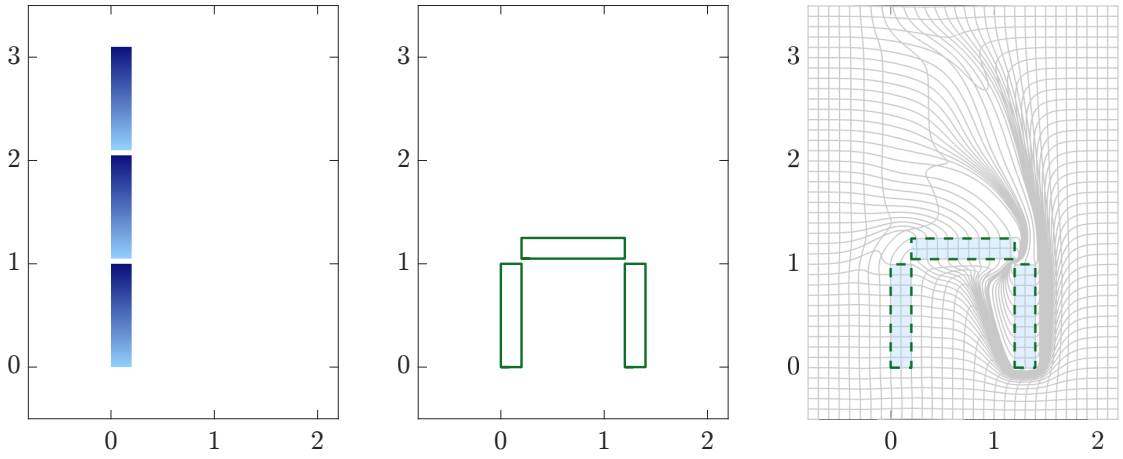


(a) Initial shape. (b) Boundary of the target shape. (c) Registration result.



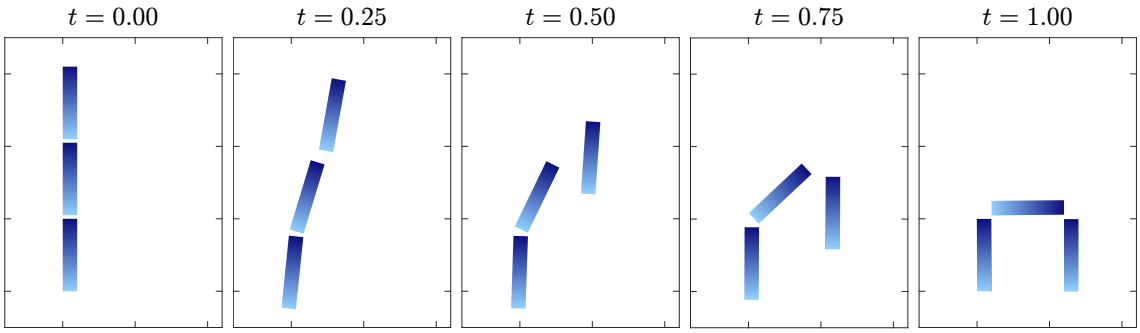
(d) Registration process.

Figure 3.2: Registration using varifold pseudo-metrics without correspondence. In (c), the blue regions represent the deformed shape at the final time $t = 1$, while the dashed lines indicate the boundary of the target shape. The norm of the gradient is $4.05\,\mathrm{e}{-05}$. ($\sigma_V = 0.1$, $\Delta t = 0.005$, $\sigma_\rho = 0.15$, $w_\rho = 500$)



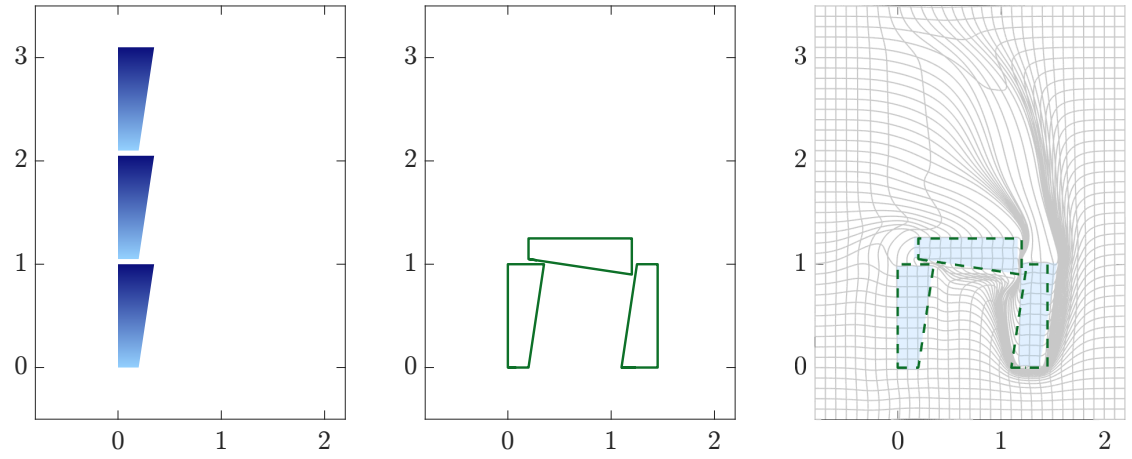
(a) Initial shape. (b) Boundary of the target shape. (c) Registration result.

Figure 3.3: Registration using varifold pseudo-metrics without correspondence. In (a), the height of gaps between rectangles is 0.05. The color shading is to facilitate orientation tracing. The norm of the gradient is $1.24\,\mathrm{e}-05$. ($\sigma_V = 0.06$, $\Delta t = 0.005$, $\sigma_\rho = 0.2$, $w_\rho = 500$)

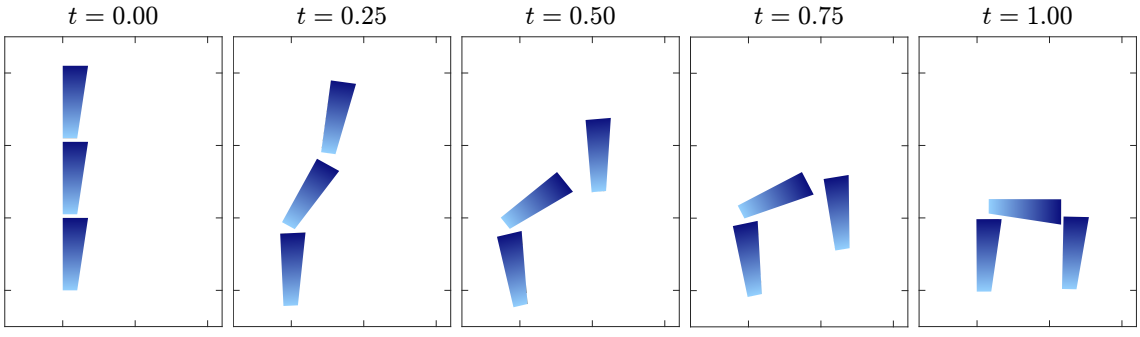


(d) Registration process.

Figure 3.3: (Continued.)

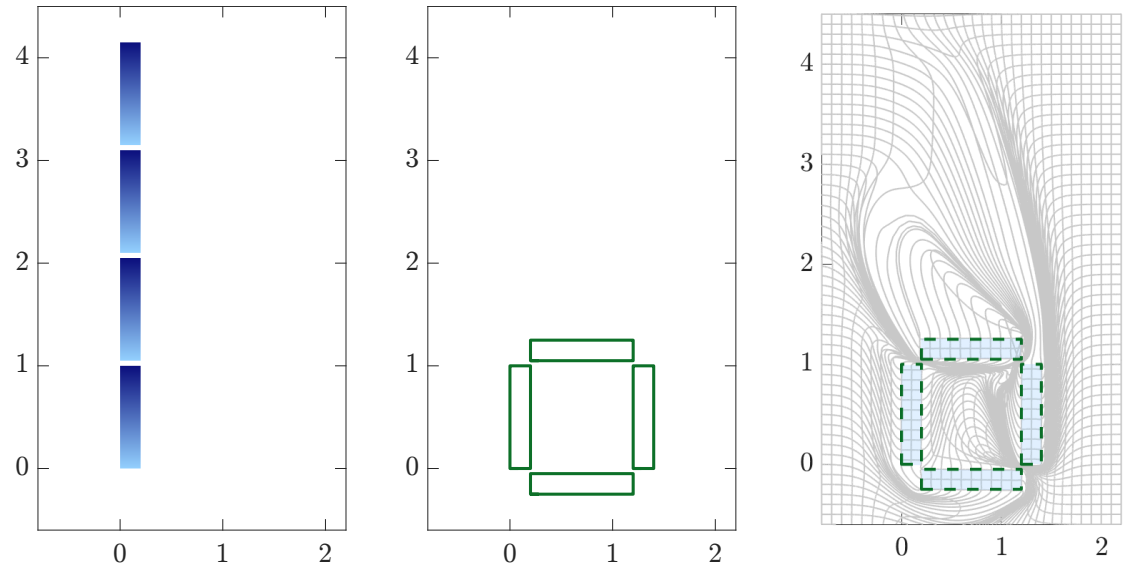


(a) Initial shape. (b) Boundary of the target shape. (c) Registration result.

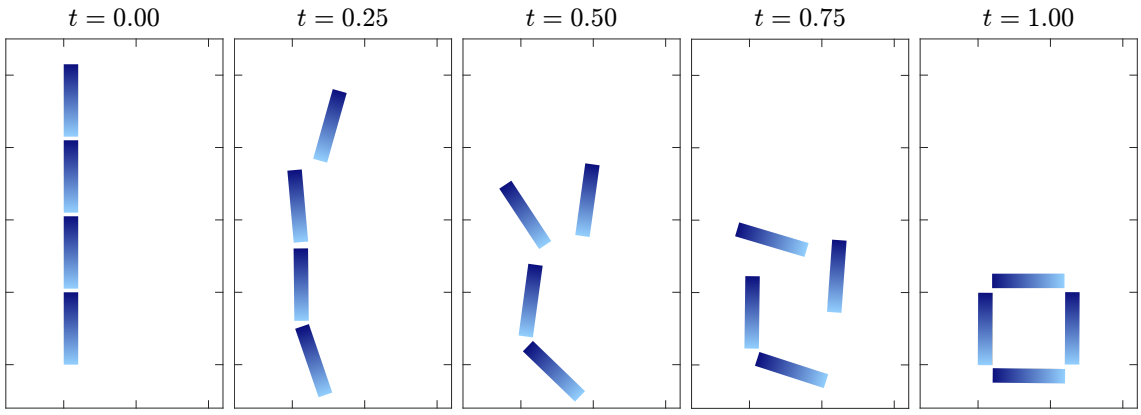


(d) Registration process.

Figure 3.4: Registration using varifold pseudo-metrics without correspondence. The color shading is to facilitate orientation tracing. The norm of the gradient is $3.29\,\mathrm{e}{-05}$. ($\sigma_V = 0.06$, $\Delta t = 0.005$, $\sigma_\rho = 0.4$, $w_\rho = 500$)



(a) Initial shape. (b) Boundary of the target shape. (c) Registration result.



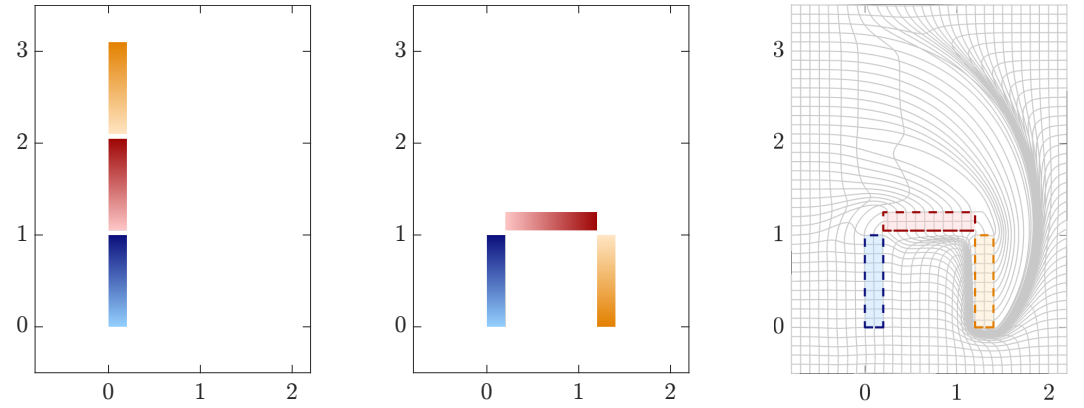
(d) Registration process.

Figure 3.5: Registration using varifold pseudo-metrics without correspondence. The color shading is to facilitate orientation tracing. The norm of the gradient is $1.23\,\mathrm{e}{-04}$. ($\sigma_V = 0.06$, $\Delta t = 0.005$, $\sigma_\rho = 0.2$, $w_\rho = 500$)

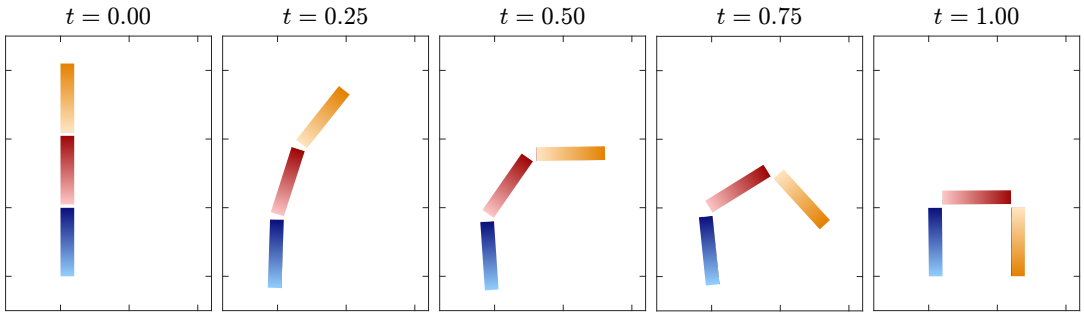
Next we consider the situation when we are given a correspondence between the initial shape $\Omega_0 = \bigcup_{i=1}^N \Omega_i$ and the target shape $\Omega_{\text{targ}} = \bigcup_{i=1}^N \Omega_i'$, i.e., we know rotations $R_i \in \mathbb{R}^{d\times d}$ and translations $a_i \in \mathbb{R}^d$ such that $\{R_i\, x + a_i : x \in \Omega_i\} = \Omega_i'$, for $i = 1, \ldots, N$. Provided we have a sufficient number of corresponding points, for example, the positions of two diagonal vertices of a rectangular shape component in the initial and the target shape, such rotations and translations may be obtained from the orthogonal Procrustes analysis. The reader may wonder what remains to be registered if we have a correspondence. We recall that our goal is a deformation on $\mathbb{R}^d$, so a correspondence on Ω_0 is not sufficient for this purpose. In addition, knowing a correspondence is equivalent to the setting in the literature when the registration of individual rigid components is complete. The correspondence enables us to adopt the weighted ℓ^2-distance as the discrepancy function, that is,

$$\rho(\Omega, \Omega') := \frac{w_\rho}{2} \sum_{i=1}^{M} |q_i - q_i'|^2,$$

where $(q_i)_{i=1}^M$ and $(q_i')_{i=1}^M$ are corresponding discretized nodes of Ω and Ω' respectively. Figures 3.6 and 3.7 show the comparisons with Figures 3.3 and 3.5 when we have a correspondence. We observe that in Figures 3.6(a) and 3.6(b) although the positions of the blue rectangles are the same, there is no guarantee that the blue rectangle would stay static during the motion as shown in Figure 3.6(d). If we know that some components will remain at their initial positions, we should include this information by letting $\theta_i(t) = 0$ for all t. Another interesting point in Figure 3.6(d) is that we did not constrain the "joints" of rectangles to keep them close to each other; this is a consequence of the smoothness of deformation vector fields. However, the joints could still be broken as can be seen in Figure 3.7(d). Also in Figure 3.7(d), we notice that the shape at time $t = 1$ does not match the target precisely. This could be explained by Figure 3.7(c) which shows that those non-intersecting curves prevent a better match. We now examine the effect of starting points for minimization. The only difference between the two examples in Figures 3.7 and 3.8 is the starting point for minimization: Figure 3.7 starts at $\theta(t) = 0$ for all t which corresponds to $\varphi_\theta(t) = id$ for all t, while $\varphi_\theta(t, \Omega_0)$ corresponding to the starting point θ for Figure 3.8 is shown in Figure 3.8(d). We can see from Figure 3.8(e) that the motion of the found critical point resembles the motion of the starting point. The results in Figures 3.7 and 3.8 exhibit two numerical critical points, of which Figure 3.8 has a smaller objective function value. Based on our observations in 2D, we finally present a 3D example shown in Figure 3.9.

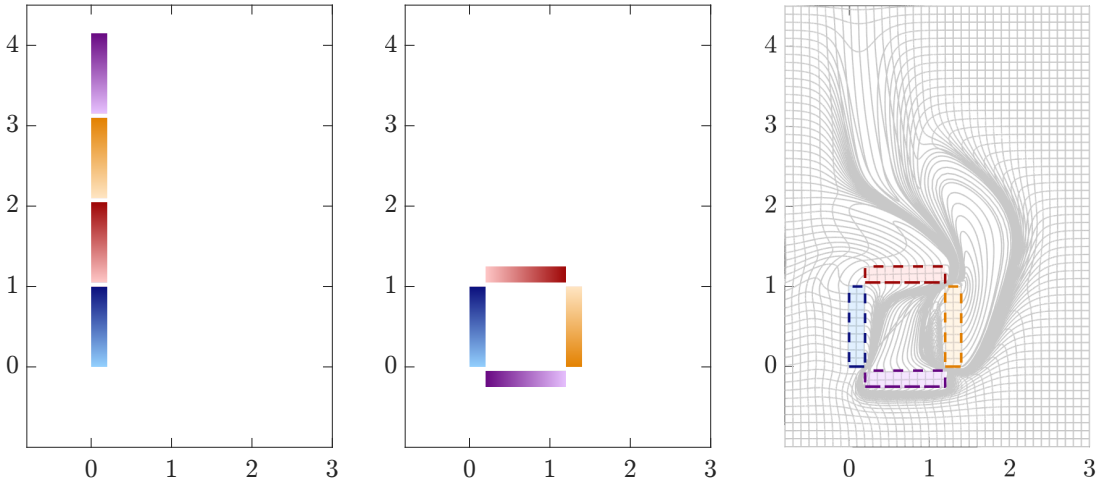


(a) Initial shape. (b) Target shape. (c) Registration result.



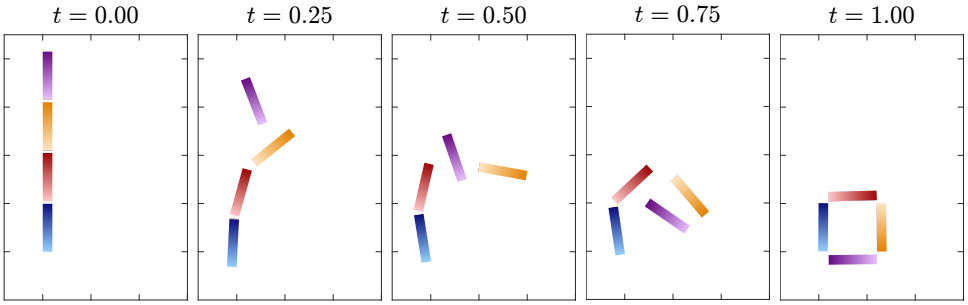
(d) Registration process.

Figure 3.6: Registration using the ℓ^2-distance given a correspondence. In (a) and (b), the color shows the point-to-point correspondence between the initial and the target shape. The norm of the gradient is $3.08\,\mathrm{e}{-05}$. ($\sigma_V = 0.06$, $\Delta t = 0.005$, $w_\rho = 100$)



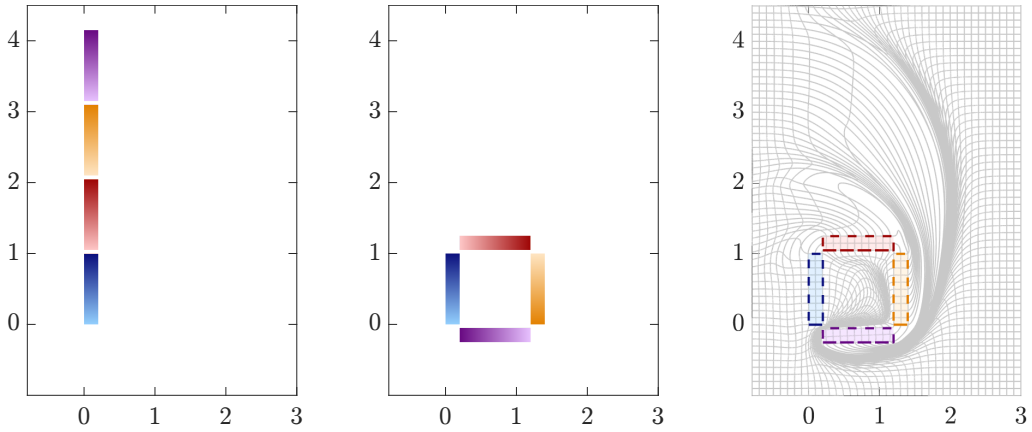
(a) Initial shape. (b) Target shape. (c) Registration result.

Figure 3.7: Registration using the ℓ^2-distance given a correspondence shown by the color. The norm of the gradient is $3.00\,\mathrm{e}{-05}$. ($\sigma_V = 0.06$, $\Delta t = 0.0005$, $w_\rho = 100$)

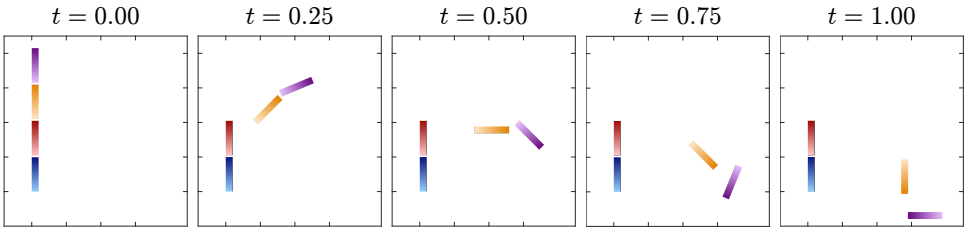


(d) Registration process.

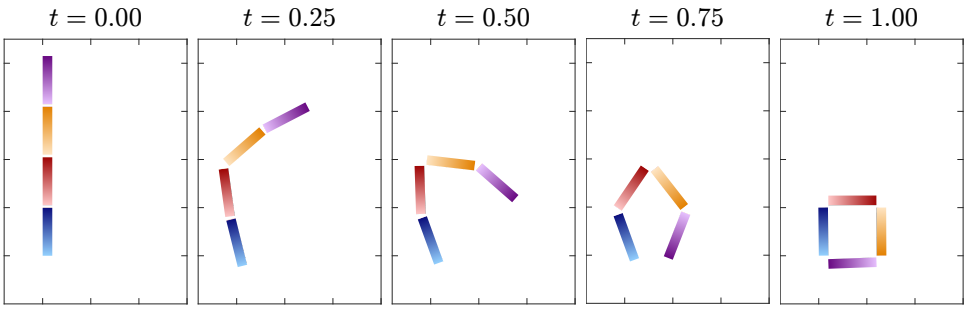
Figure 3.7: (Continued.)



(a) Initial shape. (b) Target shape. (c) Registration result.

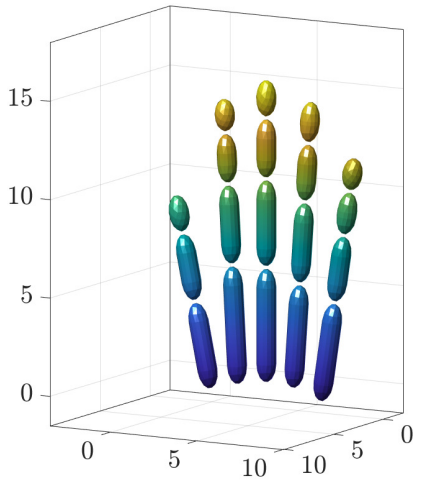


(d) Starting position for optimization.

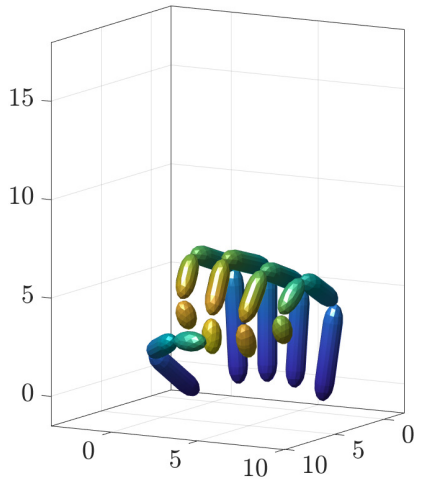


(e) Registration process.

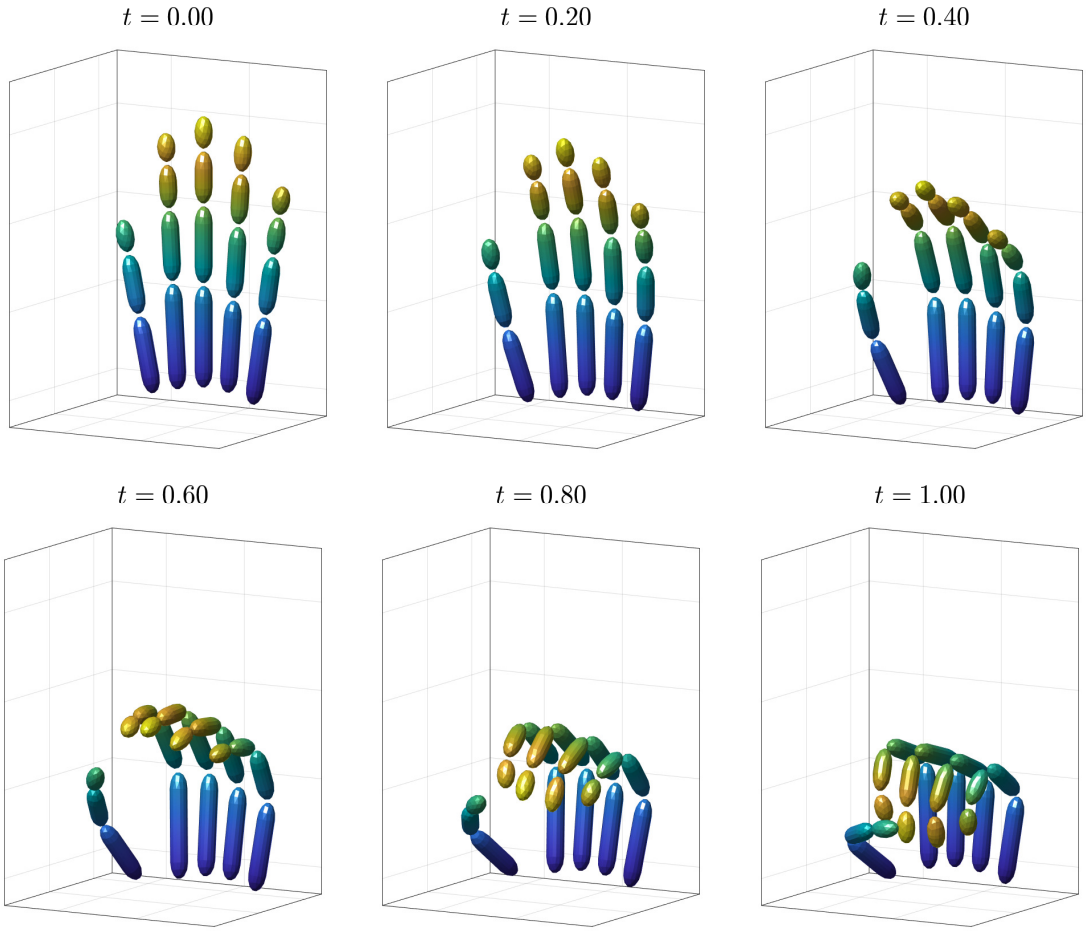
Figure 3.8: Registration using the ℓ^2-distance given a correspondence shown by the color. The norm of the gradient is $2.72\,\mathrm{e}{-05}$. ($\sigma_V = 0.06$, $\Delta t = 0.0005$, $w_\rho = 100$)



(a) Initial shape.



(b) Target shape.



(c) Registration process.

Figure 3.9: Registration using the ℓ^2-distance given a correspondence shown by the color. The norm of the gradient is $1.38\,\mathrm{e}{-03}$. ($\sigma_V = 0.35$, $\Delta t = 0.005$, $w_\rho = 10$)

3.2 Biological Atrophy Modeling

Our second example is motivated by cerebral atrophy due to a brain disease. The goal is to locate the origin of brain disease progression from shapes of cerebral cortex before and after atrophy. Since we only assume initial and final shapes of cortex, it is necessary to incorporate the information of disease progression and model the process of brain atrophy. To model the process of brain atrophy caused by a disease, we model the cerebral cortex as a layered elastic material in Section 3.2.1, then in Section 3.2.2 we model the disease progression as a chemical propagation governed by a reaction-diffusion equation on a moving shape, and finally we model how the chemical atrophies the brain cortex in Section 3.2.3. Numerical experiments will be presented in Section 3.2.4.

The three ingredients in our atrophy modeling are the material of the cerebral cortex, the progression of a brain disease, and the cortex deformation caused by a disease. In the literature, it was common to model the cerebral cortex as a homogeneous isotropic elastic material [83, 93, 21, 57, 96, 12]. The laminar organization of cortex, however, seemed to be less discussed; in the modeling of brain folding, the structure of cortical layers was considered in [67], and different growth rates parallel and perpendicular to cortical layers were adopted in [12]. As for the disease progression, many mathematical models were proposed to describe the disease progression on a *fixed* domain based on the disease mechanism and validation. For example, Achdou et al. [1] considered a Smoluchowski equation for the diffusion and aggregation of a toxic chemical; Kulason et al. [60] modeled the disease activity by a reaction-diffusion equation; Bertsch et al. [14] proposed a model including different mechanisms of two toxic chemicals and their interaction. We refer the reader to [22] for a more thorough review. Finally, the shape evolution due to a time-dependent distribution of an internal chemical has been examined in at least two directions: one direction only focused on the evolution of the shape boundary [31, 40, 90], while the other direction aimed at the evolution of the whole shape [23, 100, 33, 20]. In particular, we point out that in Bressan and Lewicka [20] tissue growth was modeled by an evolving quasi-elastic body. The body is assumed to be composed of chemical-producing cells which are passively transported by the motion. The density of chemical-producing cells then affects the distribution of chemical, which further determines local volume changes of the body and hence influences the deformation vector field according to the elasticity. The model of Bressan and Lewicka [20] has similar structure to ours but the main difference is that they assumed that the time scale of the chemical diffusion is small, while we assume that the evolution of chemical is a long time behavior. To the best of our knowledge, our work is the first attempt to include

all the three ingredients for the modeling of cerebral atrophy. Although some aspects of our model are oversimplified, we anticipate that our work will encourage the development of more sophisticated models in this direction.

3.2.1 Layered elastic shapes

Since we are modeling the cerebral cortex (see Figure 3.10), we want to include the following three properties in our model of layered elastic shapes. First, we need to have a mathematical description of layered shapes. Second, due to the structure of cortical layers and cortical columns, we want to differentiate the elasticity along layers from the elasticity transversal to layers. This property will make it possible to say how easy it is to stretch the material along layers, or how hard it is to reduce the thickness of the material. Third, we want the elasticity to be "isotropic along layers" because of the homogeneity of cortical layers. We will give precise description of this property after we present the mathematical formulation.

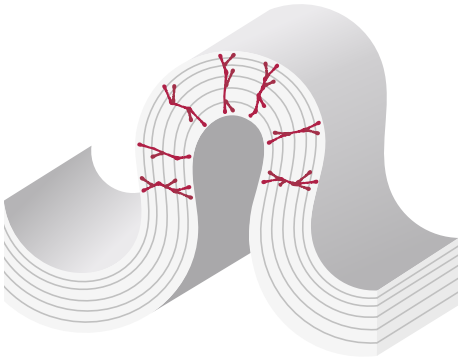
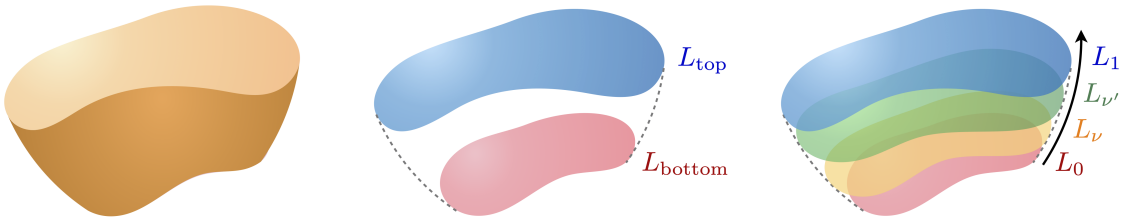


Figure 3.10: Illustration of cerebral cortical layers and cortical columns.

We now introduce our model of layered shapes. Suppose that a shape Ω has two surfaces (or two curves in 2D) $L_{\text{bottom}}, L_{\text{top}} \subset \partial\Omega$ as bottom and top layers (see Figure 3.11(b)). Moreover, suppose that we are given a diffeomorphism $\Phi : [0,1] \times L_{\text{bottom}} \to \Omega$ such that $\Phi(0, L_{\text{bottom}}) = L_{\text{bottom}}$ and $\Phi(1, L_{\text{bottom}}) = L_{\text{top}}$. Note that $\Phi(\nu, L_{\text{bottom}}) =: L_\nu$ is a surface for each $\nu \in [0,1]$ (see Figure 3.11(c)). We refer to Φ as a layered structure of Ω. We say that Ω is a layered shape if Ω has a layered structure. According to this definition of layers, a layered shape is composed of uncountably many layers without thickness, i.e., $\Omega = \bigcup_{\nu \in [0,1]} L_\nu$, which might be contrary to the reader's notion of finitely many layers with thickness. A layered structure Φ then induces a transversal unit vector field $S := \frac{\partial_\nu \Phi}{|\partial_\nu \Phi|} \circ \Phi^{-1}$ on the shape (see Figure 3.12). This transversal vector field, together with tangent planes

of layers, will be useful to describe layered elasticity in the next paragraph.



(a) A shape with a layered structure shown in (b) and (c).

(b) Bottom and top layers of the shape.

(c) Bottom, two intermediate, and top layers of the shape.

Figure 3.11: Illustration of layered shapes.

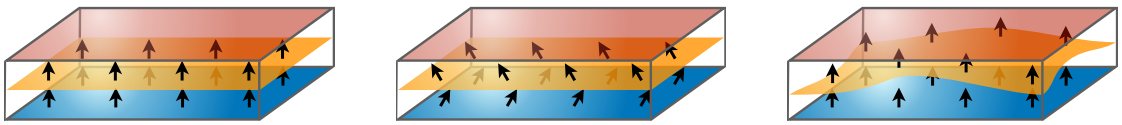


Figure 3.12: Different layered structures of the same rectangular cuboid. Shown in the figures are top layer, one middle layer, bottom layer, and the transversal vector field.

Before we proceed to layered elasticity, we briefly recall classical linear elasticity [49, Chapter X]. We restrict our discussion in the following to $\mathbb{R}^3$ for clarity; the case in $\mathbb{R}^2$ is similar and simpler. Denote by $\Sigma_2(\mathrm{Sym}_3(\mathbb{R}), \mathrm{Sym}_3(\mathbb{R}))$ the space of symmetric bilinear forms on the space of 3-by-3 symmetric matrices. We say that $\mathcal{E} : \Omega \to \Sigma_2(\mathrm{Sym}_3(\mathbb{R}), \mathrm{Sym}_3(\mathbb{R}))$ is an elasticity tensor (field) on Ω if $\mathcal{E}$ is positive definite almost everywhere, i.e., for almost every $x \in \Omega$, we have $\mathcal{E}(x)(A, A) \geq 0$ and $\mathcal{E}(x)(A, A) = 0$ if and only if $A = 0$. For example, the elasticity tensor of a homogeneous isotropic linear elastic material is given by

$$\mathcal{E}_{\text{iso}}(x)(A, B) := \lambda \operatorname{tr}(A) \operatorname{tr}(B) + 2\mu \operatorname{tr}(A^\top B), \tag{3.6}$$

where λ and μ are the Lamé parameters such that $\mu > 0$ and $2\mu + 3\lambda > 0$, which ensure the positive definiteness [49, page 196]. For an elastic shape Ω described by an elasticity tensor $\mathcal{E}$, suppose that the residual stress in Ω vanishes, then a small deformation, i.e., $\xi = id + u$ with $\|Du\|_\infty \ll 1$, will induce an elastic energy

$$\frac{1}{2} \int_\Omega \mathcal{E}(\varepsilon_u, \varepsilon_u)\, dx := \frac{1}{2} \int_\Omega \mathcal{E}(x)(\varepsilon_u(x), \varepsilon_u(x))\, dx,$$

where $\varepsilon_u := \frac{1}{2}\left(Du + Du^\top\right)$ is the infinitesimal strain tensor. Since $\|Du\|_\infty \ll 1$, we observe

that $\varepsilon_u = \frac{1}{2}\big(Du + Du^\top\big) \approx \frac{1}{2}\big(D\xi^\top D\xi - I_3\big)$. Denoting the coordinate unit vectors by e_1, e_2, e_3, we can then write

$$\varepsilon_u = I_3^\top \Big(\frac{1}{2}\big(Du + Du^\top\big) \Big) I_3 \approx \begin{bmatrix} e_1, & e_2, & e_3 \end{bmatrix}^\top \frac{1}{2}\Big(D\xi^\top D\xi - I_3\Big) \begin{bmatrix} e_1, & e_2, & e_3 \end{bmatrix}, \quad (3.7)$$

which says that the diagonal entries of ε_u describe the change of length of coordinate unit vectors, while the off-diagonal entries of ε_u describe the change of angle between coordinate unit vectors. For our layered elastic shapes, however, we want to describe changes with respect to tangent planes and transversal directions rather than coordinate directions. To this end, given a shape Ω with a layered structure Φ, we let T_1 and T_2 be orthogonal unit vector fields on Ω such that $T_1|_{L_\nu}$ and $T_2|_{L_\nu}$ are tangent to layers L_ν for all $\nu \in [0,1]$. Since $\Phi : [0,1] \times L_{\text{bottom}} \to \Omega$ is a diffeomorphism, it follows that the tangential unit vector fields T_1, T_2, and the transversal unit vector field $S = \frac{\partial_\nu \Phi}{|\partial_\nu \Phi|} \circ \Phi^{-1}$ are linearly independent vector fields on Ω. We form a frame (field) $F := \begin{bmatrix} T_1, & T_2, & S \end{bmatrix}$ and let $\zeta_u := F^\top \varepsilon_u F$. Comparing ζ_u with (3.7), we deduce that the entries of ζ_u describe the changes with respect to the directions T_1, T_2, and S. For a layered elastic shape undergoing a small deformation $\xi = id + u$, we define its layered elasticity tensor by

$$\mathcal{E}_\Phi(\varepsilon_u, \varepsilon_u) := \lambda_{\tan}\big(\zeta_{11} + \zeta_{22}\big)^2 + \mu_{\tan}\big(\zeta_{11}^2 + \zeta_{22}^2 + 2\,\zeta_{12}^2\big) + \mu_{\text{tsv}}\,\zeta_{33}^2 + 2\mu_{\text{ang}}\big(\zeta_{13}^2 + \zeta_{23}^2\big),$$

where $\lambda_{\tan}$, $\mu_{\tan}$, μ_{tsv}, μ_{ang} are constants, and ζ_{ij} is the ij-th element of $\zeta_u = F^\top \varepsilon_u F$. A sufficient condition for $\mathcal{E}_\Phi$ to be positive definite is that $\lambda_{\tan} \geq 0$ and $\mu_{\tan}, \mu_{\text{tsv}}, \mu_{\text{ang}} > 0$. We illustrate the behavior of a layered elastic shape in Figure 3.13. Recall that one property of layered elastic shapes we want to have is "isotropic along layers." Specifically, "isotropic along layers" means that the layered elasticity tensor $\mathcal{E}_\Phi$ is independent of the choice of orthogonal tangential unit vector fields T_1 and T_2. This property is indeed true. Let N be a unit vector field on Ω that is normal to layers. We show in Proposition A.2.2 that the layered elasticity tensor can also be written as

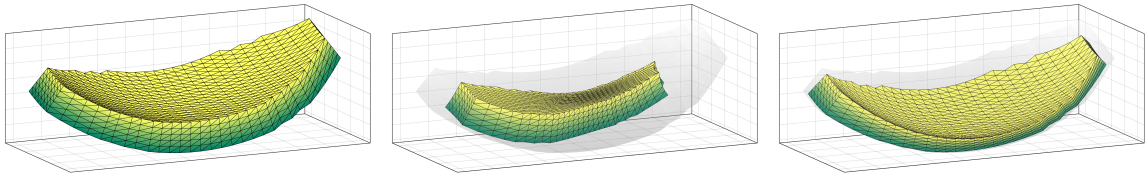
$$\begin{aligned} \mathcal{E}_\Phi(\varepsilon_u, \varepsilon_u) = {} & \lambda_{\tan}\Big(\operatorname{tr}(\varepsilon_u) - N^\top \varepsilon_u N\Big)^2 + \mu_{\tan}\Big(\operatorname{tr}(\varepsilon_u^2) - 2\,|\varepsilon_u N|^2 + (N^\top \varepsilon_u N)^2\Big) \\ & + \mu_{\text{tsv}}\,(S^\top \varepsilon_u S)^2 + 2\mu_{\text{ang}}\Big(|\varepsilon_u S|^2 - (N^\top \varepsilon_u S)^2\Big). \end{aligned} \quad (3.8)$$

In other words, the layered elasticity tensor only depends on tangent planes of layers and the transversal vector field, i.e., the layered structure, which also explains our notation $\mathcal{E}_\Phi$. We further remark that, for a layered shape Ω with a layered structure Φ, a deformed shape $\xi(\Omega)$ by a Diff^p-deformation ξ is also a layered shape whose layered structure becomes $\xi * \Phi : [0,1] \times \xi(L_{\text{bottom}}) \to \xi(\Omega)$ defined by $(\nu, x) \mapsto \xi\big(\Phi(\nu, \xi^{-1}(x))\big)$. In addition, if we

assume that $\xi(\Omega)$ possesses the elasticity tensor $\mathcal{E}_{\xi * \Phi}$ and that the residual stress in $\xi(\Omega)$ vanishes, then a small deformation $id + w$ away from $\xi(\Omega)$ will generate an elastic energy

$$\frac{1}{2} \int_{\xi(\Omega)} \mathcal{E}_{\xi * \Phi}(\varepsilon_w, \varepsilon_w)\, dx. \tag{3.9}$$

We will need this formulation and the change of references and zero-stress states when we model the atrophy of layered elastic shapes in Section 3.2.3.



(a) A simulated layered shape. (b) Tangential deformation. (c) Transversal deformation.

Figure 3.13: Responses to the same shrinking force under different layered elasticity parameters. In (b), $\mu_{\text{tan}} = 0.02\,\mu_{\text{tsv}}$. In (c), $\mu_{\text{tsv}} = 0.02\,\mu_{\text{tan}}$.

3.2.2 Chemical propagation on a moving shape

We now model the dynamics of chemical propagation via a reaction-diffusion equation on a moving shape. Given a shape Ω and a Diff^p-motion $\varphi \in C([0,t], \text{Diff}^p_{id}(\mathbb{R}^d))$, assume that the concerned attribute of the chemical can be represented by a real-valued function $\tau_E : \bigcup_{s\in[0,t]} \{(s,x) : x \in \varphi(s,\Omega)\} \to \mathbb{R}$, where the subscript E stands for the Eulerian description. We derive a partial differential equation for τ_E in this section assuming τ_E is sufficiently regular. Its weak formulation, which is more technically involved, will be presented in Section 5.2.2. Let $U_{\varphi(s)} : \varphi(s,\Omega) \to \mathbb{R}^{d\times d}$ be a matrix (field) such that $U_{\varphi(s)}(x)$ is symmetric and positive definite for all $x \in \varphi(s,\Omega)$. For every part $\Pi \subset \Omega$, we assume that the diffusion flux at time $s \in [0,t]$ is given by

$$\int_{\partial\varphi(s,\Pi)} \big(-U_{\varphi(s)} \nabla \tau_E(s)\big)^\top n(s)\, d\sigma,$$

where $n(s) = n(\varphi(s), \Pi)$ is the unit outward normal on $\partial\varphi(s,\Pi)$. Since the matrix $U_{\varphi(s)}$ describes the diffusion on $\varphi(s,\Omega)$, we will refer to it as the Eulerian diffusion matrix. Suppose that the change of τ_E over the region $\varphi(s,\Pi)$ is caused only by the diffusion flux and the reaction within the region, then we can write

$$\frac{d}{ds} \int_{\varphi(s,\Pi)} \tau_E(s)\, dx = -\bigg(\int_{\partial\varphi(s,\Pi)} \big(-U_{\varphi(s)} \nabla \tau_E(s)\big)^\top n(s)\, d\sigma \bigg) + \int_{\varphi(s,\Pi)} R\big(\tau_E(s)\big)\, dx,$$

where the given function $R : \mathbb{R} \to \mathbb{R}$ describes the reaction. We make a change of variables and obtain

$$\begin{aligned}
&\frac{d}{ds}\int_{\Pi} \tau_E(s)\circ\varphi(s)\,\det D\varphi(s)\,dx \\
&= \int_{\partial\Pi} \Big(\big(U_{\varphi(s)}\nabla\tau_E(s)\big)\circ\varphi(s)\Big)^{\top} \det D\varphi(s)\,\big(D\varphi(s)^{-\top}n(0)\big)\,d\sigma \\
&\quad + \int_{\Pi} R\big(\tau_E(s)\circ\varphi(s)\big)\,\det D\varphi(s)\,dx \\
&= \int_{\Pi} \operatorname{div}\bigg(\big(D\varphi(s)\big)^{-1}\Big(\big(U_{\varphi(s)}\nabla\tau_E(s)\big)\circ\varphi(s)\Big)\det D\varphi(s)\bigg)\,dx \\
&\quad + \int_{\Pi} R\big(\tau_E(s)\circ\varphi(s)\big)\,\det D\varphi(s)\,dx.
\end{aligned}$$

Since $\Pi \subset \Omega$ is arbitrary, the above equation leads to

$$\begin{aligned}
&\frac{\partial}{\partial s}\Big(\tau_E(s)\circ\varphi(s)\,\det D\varphi(s)\Big) \qquad (3.10)\\
&= \operatorname{div}\bigg(\big(D\varphi(s)\big)^{-1}\Big(\big(U_{\varphi(s)}\nabla\tau_E(s)\big)\circ\varphi(s)\Big)\det D\varphi(s)\bigg) + R\big(\tau_E(s)\circ\varphi(s)\big)\det D\varphi(s).
\end{aligned}$$

Let $\tau(s) := \tau_E(s)\circ\varphi(s)$ be the Lagrangian description of the chemical. Note that

$$\nabla\tau(s) = D\varphi(s)^{\top}\big(\nabla\tau_E(s)\circ\varphi(s)\big),$$

which implies

$$\nabla\tau_E(s)\circ\varphi(s) = D\varphi(s)^{-\top}\nabla\tau(s). \qquad (3.11)$$

Plugging (3.11) into (3.10) yields

$$\frac{\partial}{\partial s}\Big(\tau(s)\,\det D\varphi(s)\Big) = \operatorname{div}\Big(W_{\varphi(s)}\,\nabla\tau(s)\,\det D\varphi(s)\Big) + R\big(\tau(s)\big)\det D\varphi(s), \qquad (3.12)$$

where

$$W_{\varphi(s)} = \big(D\varphi(s)\big)^{-1}\big(U_{\varphi(s)}\circ\varphi(s)\big)\,D\varphi(s)^{-\top}. \qquad (3.13)$$

In addition to the partial differential equation (3.12), we assume for simplicity that the chemical is fully contained in the moving shape, which is enforced by the Neumann boundary condition, or the natural boundary condition in the language of weak formulation, given by

$$\big(U_{\varphi(s)}\nabla\tau_E(s)\big)^{\top}n(s) = 0 \quad \text{ on } \partial\varphi(s,\Omega)$$

and equivalently

$$\big(W_{\varphi(s)}\,\nabla\tau(s)\big)^{\top}n(0) = 0 \quad \text{ on } \partial\Omega.$$

We adopt a simple bump function as our initial condition, where the bump function is a compactly supported C^1 radial function parametrized by its center $c \in \mathbb{R}^d$, radius $r > 0$, and height $h > 0$. Precisely, we let $Q : \mathbb{R}^d \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to L^2(\Omega)$ be defined by

$$Q(\theta)(x) = Q(c, r, h)(x) := h \left(\frac{|x - c|^2}{r^2} - 1 \right)^2 \mathbb{1}_{B(c,r)}(x), \tag{3.14}$$

whose graph is shown in Figure 3.14 when $d = 2$. To keep our discussion in a general setting, we denote the number of parameters by k, i.e., $k = d + 2$ for this particular choice, and state the initial condition as $\tau(0) = Q(\theta)$, where $Q : \Theta \subset \mathbb{R}^k \to L^2(\Omega)$ is any specified mapping.

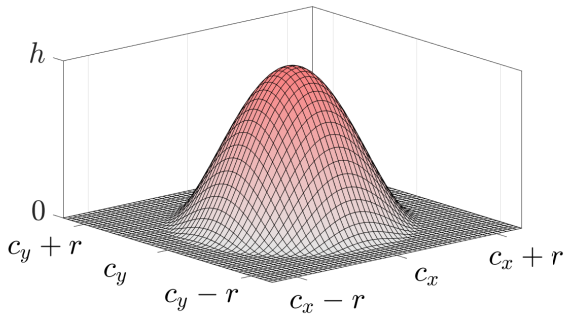


Figure 3.14: Parametrized initial condition of chemical.

We summarize the strong formulation of chemical propagation on a moving shape as follows:

$$\begin{cases} \partial_s \big(\tau \det D\varphi\big)(s) = \operatorname{div}\Big(W_{\varphi(s)} \nabla\tau(s) \det D\varphi(s)\Big) + R(\tau(s)) \det D\varphi(s) & \text{in } (0,t] \times \Omega^o \\ \big(W_{\varphi(s)} \nabla\tau(s)\big)^\top n(0) = 0 & \text{on } [0,t] \times \partial\Omega \\ \tau(0) = Q(\theta) & \text{on } \Omega^o \end{cases},$$

where Ω^o denotes the interior of a compact set Ω. With additional regularity assumptions, we will prove in Section 5.2.3 that the weak formulation of this problem has a unique solution for every $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ and $\theta \in \Theta$.

3.2.3 Atrophy model

In this section, we model the evolution of layered elastic shapes atrophied by an internal chemical propagation. We have described layered elastic shapes in Section 3.2.1 and chemical propagation on a moving shape in Section 3.2.2. However, one thing is missing: the cause of a moving shape. We fill this missing link by modeling how the chemical affects

a layered elastic shape and close the loop of chemical propagation, atrophy, and layered elastic shapes. We illustrate the atrophy model we want to achieve in Figure 3.15. One way to model the atrophy effect is via a force density induced by the chemical, while another way is via the action of chemical on a displacement, which is the approach we take. We remark that the first approach models the effect by a vector field, while the second approach models the effect by a functional.

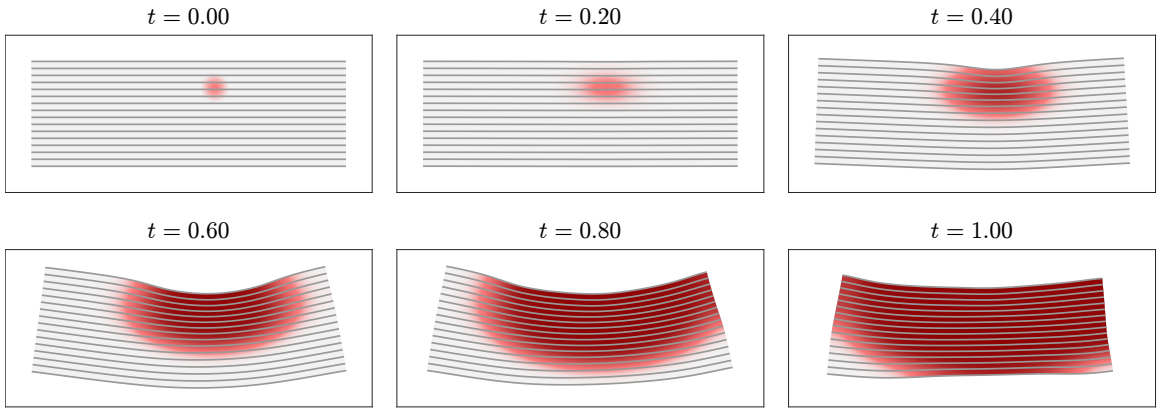


Figure 3.15: Illustration of the evolution model of a layered elastic shape atrophied by an internal chemical propagation.

We now consider a layered elastic shape Ω with a layered structure Φ and an elasticity tensor $\mathcal{E}_\Phi$ and quantify the effect, or the work done, by a chemical distribution τ within Ω when a virtual displacement u' occurs during a time interval Δt. Assuming Δt and Du' are small, we model the virtual work done upon the shape by

$$\left(\int_\Omega \alpha(\tau)\,(-\operatorname{div} u')\,dx\right)\Delta t, \tag{3.15}$$

where we have supposed that we are given a function $\alpha : \mathbb{R} \to \mathbb{R}_{\geq 0}$ describing the strength of atrophy with respect to τ. Thus contributing factors to a greater positive work include a severer atrophy $\alpha(\tau)$, a greater loss of volume $-\operatorname{div} u'$, and a longer time duration Δt. Since we assume Δt is small, in (3.15) we only consider the initial shape Ω and the initial chemical distribution τ and ignore the shape deformation and chemical propagation in this small time duration. In addition, since Du' is small, we use $-\operatorname{div} u'$ to approximate the volume loss. We provide another two interpretations of this model of work. Note that we can write (3.15) as

$$\left(\int_\Omega \alpha(\tau)\,(-\operatorname{div} u')\,dx\right)\Delta t = \left(\int_\Omega \left\langle -\alpha(\tau)\,I_d,\, \varepsilon_{u'}\right\rangle_F dx\right)\Delta t,$$

where $\varepsilon_{u'} = \frac{1}{2}\left(Du' + Du'^{\top}\right)$ is the infinitesimal strain tensor and $\langle A, B\rangle_F := \sum_{i,j} a_{ij}\, b_{ij}$

is the Frobenius inner product. This form says that the atrophy behaves like a pressure acting on the infinitesimal strain tensor and further suggests a possible generalization by considering a more general atrophy tensor than $-\alpha(\tau)\, I_d$. We can also write (3.15) by the divergence theorem as

$$\left(\int_\Omega \alpha(\tau)\,(-\operatorname{div} u')\,dx\right)\Delta t = \left(\int_\Omega \nabla\big(\alpha(\tau)\big)^\top u'\,dx + \int_{\partial\Omega}\big(-\alpha(\tau)\,n\big)^\top u'\,d\sigma\right)\Delta t, \qquad (3.16)$$

where n is the outward unit normal on $\partial\Omega$. Thus, in terms of force density, the work model (3.15) is actually equivalent to a force system which is composed of a body force $\nabla\big(\alpha(\tau)\big)\,\Delta t$ on Ω and a surface force $-\alpha(\tau)\,n\;\Delta t$ on $\partial\Omega$.

Next we discuss the response of a layered elastic shape to the work done modeled in the previous paragraph. We assume that the stress is zero in Ω. According to the principle of minimum potential energy, suppose that the equilibrium is reached after a time duration Δt, then the true displacement u from the zero-stress state Ω to the equilibrium satisfies

$$u = \underset{u' \in (H^1(\Omega))^d}{\arg\min}\left(\frac{1}{2}\int_\Omega \mathcal{E}_\Phi(\varepsilon_{u'},\varepsilon_{u'})\,dx - \left(\int_\Omega \alpha(\tau)\,(-\operatorname{div} u')\,dx\right)\Delta t\right). \qquad (3.17)$$

If $\widetilde{u}$ is another minimizer, we have the relation $\widetilde{u} = u + w$, where $Dw^\top = -Dw$ [49, page 208]. We can see this problem in a more classical way. From (3.16), the right hand side of (3.17) can be written as

$$u = \underset{u' \in (H^1(\Omega))^d}{\arg\min}\left(\frac{1}{2}\int_\Omega \mathcal{E}_\Phi(\varepsilon_{u'},\varepsilon_{u'})\,dx - \left(\int_\Omega \nabla\big(\alpha(\tau)\big)^\top u'\,dx + \int_{\partial\Omega}\big(-\alpha(\tau)\,n\big)^\top u'\,d\sigma\right)\Delta t\right),$$

which is a pure traction boundary value problem in linear elasticity. This problem has a solution if and only if the sum of forces and the sum of torques are zero, that is,

$$\int_\Omega \nabla\big(\alpha(\tau)\big)\,dx + \int_{\partial\Omega}\big(-\alpha(\tau)\,n\big) = 0$$

and

$$\begin{cases} \displaystyle\int_\Omega \nabla\big(\alpha(\tau)\big)^\top \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x\;dx + \int_{\partial\Omega}\big(-\alpha(\tau)\,n\big)^\top \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x\;dx = 0, & \text{if } d = 2\,; \\ \displaystyle\int_\Omega \nabla\big(\alpha(\tau)\big) \times x\;dx + \int_{\partial\Omega}\big(-\alpha(\tau)\,n\big) \times x\;dx = 0, & \text{if } d = 3\,. \end{cases}$$

These two requirements are indeed satisfied, which can be seen coordinate-wise from (3.16) by taking, for example, $u'(x) = (1,0,0)$ and $u'(x) = u'(x_1,x_2,x_3) = (0,x_3,-x_2)$ for the first coordinate when $d = 3$. However, a solution of pure traction problems is not unique and is up to a displacement w such that $Dw^\top = -Dw$. In order to have a unique solution

and, more importantly, to eventually prove the existence of a Diff^p-motion, we adopt an approach similar to the one in [16] by introducing a regularity term and consider

$$u = \underset{u' \in (H^m(\mathbb{R}^d))^d}{\arg\min} \left(\frac{\gamma}{2} \|u'\|^2_{(H^m(\mathbb{R}^d))^d} + \frac{1}{2} \int_\Omega \mathcal{E}_\Phi(\varepsilon_{u'}, \varepsilon_{u'})\, dx - \left(\int_\Omega \alpha(\tau)\, (-\operatorname{div} u')\, dx \right) \Delta t \right). \tag{3.18}$$

The objective function on the right hand side is strictly convex in u' and has a unique minimizer.

A decisive factor in (3.18) is the time duration Δt to reach an equilibrium. Since the time scale to reach an elastic equilibrium is much shorter than the time scale of progressive atrophy over several years, we consider a quasi-steady-state assumption [89]: we assume that an equilibrium is reached in any infinitesimal time interval. This assumption leads to

$$du = \underset{u' \in (H^m(\mathbb{R}^d))^d}{\arg\min} \left(\frac{\gamma}{2} \|u'\|^2_{(H^m(\mathbb{R}^d))^d} + \frac{1}{2} \int_\Omega \mathcal{E}_\Phi(\varepsilon_{u'}, \varepsilon_{u'})\, dx - \left(\int_\Omega \alpha(\tau)\, (-\operatorname{div} u')\, dx \right) dt \right).$$

We also observe that the minimizer du is linear in dt, that is, we can write $du = v\, dt$ and v satisfies

$$v = \underset{u' \in (H^m(\mathbb{R}^d))^d}{\arg\min} \left(\frac{\gamma}{2} \|u'\|^2_{(H^m(\mathbb{R}^d))^d} + \frac{1}{2} \int_\Omega \mathcal{E}_\Phi(\varepsilon_{u'}, \varepsilon_{u'})\, dx - \int_\Omega \alpha(\tau)\, (-\operatorname{div} u')\, dx \right), \tag{3.19}$$

which gives the velocity $v = \frac{du}{dt}$. In deriving (3.19), we remind the reader that we assume the stress is zero in Ω.

We are in a position to put all the elements in the context of a moving shape. For a layered elastic shape Ω_0 at time 0 with a layered structure Φ, an elasticity tensor $\mathcal{E}_\Phi$, and a chemical distribution $\tau(0) = \tau_E(0)$, we assume that the stress in Ω_0 is zero, then from (3.19) the deformation at any infinitesimal time dt is given by $\xi := id + v\, dt$. At time dt, our shape becomes $\xi(\Omega_0)$ with a layered structure $\xi * \Phi$ (see page 42). We also assume that the elasticity tensor is changed to $\mathcal{E}_{\xi*\Phi}$ with the same elasticity parameters λ_{tan}, μ_{tan}, μ_{tsv}, and μ_{ang} as $\mathcal{E}_\Phi$. In addition, after the shape attains the equilibrium at time dt, we assume that the stress is then fully relaxed and $\xi(\Omega_0)$ becomes a new zero-stress state. Since $\xi(\Omega_0)$ has zero stress, its velocity is again determined by (3.19) but with the deformed shape $\xi(\Omega_0)$, the updated elasticity tensor $\mathcal{E}_{\xi*\Phi}$, and the propagated chemical distribution $\tau_E(dt)$ in Eulerian description. In other words, the velocity at time t should satisfy

$$\begin{aligned} v(t) = \underset{u' \in (H^m(\mathbb{R}^d))^d}{\arg\min} \Bigg(& \frac{\gamma}{2} \|u'\|^2_{(H^m(\mathbb{R}^d))^d} \\ & + \frac{1}{2} \int_{\varphi(t,\,\Omega_0)} \mathcal{E}_{\varphi(t)*\Phi}(\varepsilon_{u'}, \varepsilon_{u'})\, dx - \int_{\varphi(t,\,\Omega_0)} \alpha\big(\tau_E(t)\big)\, (-\operatorname{div} u')\, dx \Bigg). \end{aligned} \tag{3.20}$$

We emphasize that the reference shape and the zero-stress state at time t has become $\varphi(t, \Omega_0)$ rather than the initial shape Ω_0. We summarize the assumptions we have made in obtaining (3.20): 1. We model the virtual work by (3.15); 2. We introduce a regularity term to ensure a unique minimizer in (3.18) and a Diff^p-motion; 3. We assume that the time to reach an elastic equilibrium is arbitrarily small; 4. We assume that the stress in Ω_0 is zero and that the equilibrium becomes a new zero-stress state; 5. We assume that the layered elasticity tensor $\mathcal{E}_\Phi$ after a deformation ξ becomes $\mathcal{E}_{\xi * \Phi}$. Although the existence of residual stress in biological tissues is widely known [72, 101, 76, 44, 36], we adopt the assumption 4 to demonstrate a simplified model as a proof of concept.

We can now give a full statement of the atrophy problem. Suppose that we are given a layered elastic shape Ω_0 with a layered structure Φ and an elasticity tensor $\mathcal{E}_\Phi$, along with a target shape Ω_{targ}. Our problem of atrophy is then stated as

$$\min_{\theta \in \Theta} \ \rho(\varphi(T, \Omega_0), \Omega_{\mathrm{targ}}) \tag{3.21}$$

subject to

$$\begin{cases} \varphi(t,x) = x + \displaystyle\int_0^t v(s, \varphi(s,x))\, ds \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d \\ v(t) = \displaystyle\operatorname*{arg\,min}_{v' \in V} \Bigg(\frac{\gamma}{2} \|v'\|_V^2 \\ \qquad\qquad + \frac{1}{2} \displaystyle\int_{\varphi(t,\,\Omega_0)} \mathcal{E}_{\varphi(t) * \Phi}(\varepsilon_{v'}, \varepsilon_{v'})\, dx - \int_{\varphi(t,\,\Omega_0)} \alpha\big(\tau(t) \circ \varphi(t)^{-1}\big)\, (-\operatorname{div} v')\, dx \Bigg) \\ \begin{cases} \partial_t \big(\tau \det D\varphi\big)(t) = \operatorname{div}\Big(W_{\varphi(t)}\, \nabla\tau(t)\, \det D\varphi(t)\Big) + R(\tau(t))\, \det D\varphi(t) & \text{in } (0,T] \times \Omega^o \\ \big(W_{\varphi(t)}\, \nabla\tau(t)\big)^\top n(0) = 0 & \text{on } [0,T] \times \partial\Omega \\ \tau(0) = Q(\theta) & \text{on } \Omega^o \end{cases} \end{cases},$$

where

$$W_{\varphi(t)} = \big(D\varphi(t)\big)^{-1} \big(U_{\varphi(t)} \circ \varphi(t)\big)\ D\varphi(t)^{-\top}$$

and $U_{\varphi(t)}(x)$ is symmetric and positive definite for all $x \in \varphi(t, \Omega_0)$. We consider a specific form of the Eulerian diffusion matrix:

$$U_{\varphi(t)} := r_{\tan} \big(I_d - N_{\varphi(t)}\, N_{\varphi(t)}^\top\big) + r_{\mathrm{tsv}}\, S_{\varphi(t)}\, S_{\varphi(t)}^\top, \tag{3.22}$$

where $r_{\tan}, r_{\mathrm{tsv}} > 0$ are fixed constants and $N_{\varphi(t)}, S_{\varphi(t)}$ are unit normal and transversal vector fields on $\varphi(t, \Omega_0)$ respectively. Note that if $S_{\varphi(t)} = N_{\varphi(t)}$, then this form of $U_{\varphi(t)}$ specifies the diffusion speed along the tangential and the normal directions. It is possible to adopt a more sophisticated form of $U_{\varphi(t)}$ that specifies diffusion speeds along the tangential

and the transversal directions. We use the simple form (3.22) as an approximation.

3.2.4 Numerical results

We examine numerical results of a 2D simulation using synthetic data and a 3D simulation using real data. Both experiments use $V = (H^5(\mathbb{R}^d))^d \hookrightarrow C_0^3(\mathbb{R}^d, \mathbb{R}^d)$. The time step size Δt for the forward Euler method is 0.01. The reaction and atrophy functions we used in the experiments are shown in Figure 3.16 and defined in Section A.3. Both reaction and atrophy functions are C^2 and compactly supported on $[\tau_{\min}, \tau_{\max}] = [0.01, 1]$. Other parameters will be mentioned in each experiment.

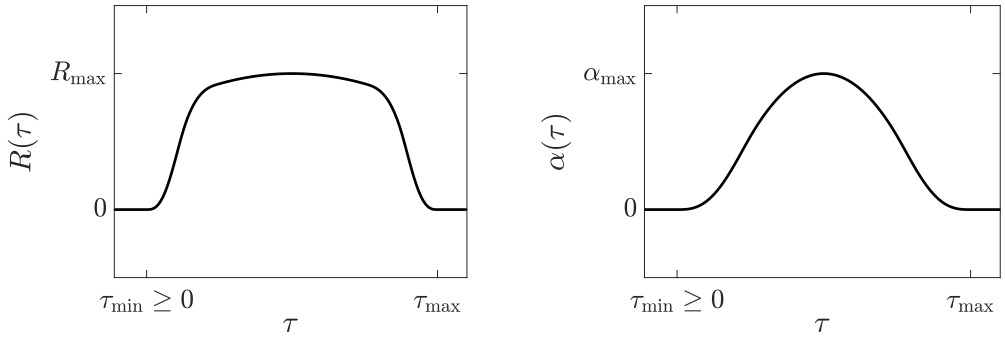


Figure 3.16: Reaction and atrophy functions.

In our 2D simulation, the initial layered shape is shown in Figure 3.17. We first look at different chemical propagations and shape evolutions when we vary the parameters c and h of the initial chemical distribution (see Figure 3.14)

$$Q(\theta)(x) = Q(c, r, h)(x) := h \left(\frac{|x - c|^2}{r^2} - 1 \right)^2 \mathbb{1}_{B(c,r)}(x).$$

Then we let $\Omega_{\text{targ}} := \varphi_{\theta_{\text{true}}}(T_{\text{true}}, \Omega_0)$ be the target shape with a known ground truth $(\theta_{\text{true}}, T_{\text{true}})$ and plot the objective function $J(\theta; T) = J(c_x, c_y, r, h; T) := \rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}})$ using a varifold pseudo-metric (3.5). The parameters for this simulation are as follows: the kernel width is $\sigma_V = 0.2$, and the regularization weight is $\gamma = 0.01$; the elasticity parameters are $\mu_{\text{tan}} = 3$, $\mu_{\text{tsv}} = 3$, and $\mu_{\text{ang}} = 15$; the diffusion speeds are $r_{\text{tan}} = 0.625$ and $r_{\text{tsv}} = 0.125$; the maximum of reaction and diffusion functions equals to $R_{\max} = \alpha_{\max} = 4$; the parameters of the varifold pseudo-metric are $w_\rho = 1$ and $\sigma_\rho = 0.1$.

Figure 3.18 presents the chemical propagation and shape evolution when $\theta = (c_x, c_y, r, h) = (0.5, 0, 0.3, 0.2)$. After $t = 1$, the shape will remain static since $\alpha(\tau) = 0$ for $\tau \geq \tau_{\max}$, that is, there will be no work done on the shape, hence no deformation. This behavior

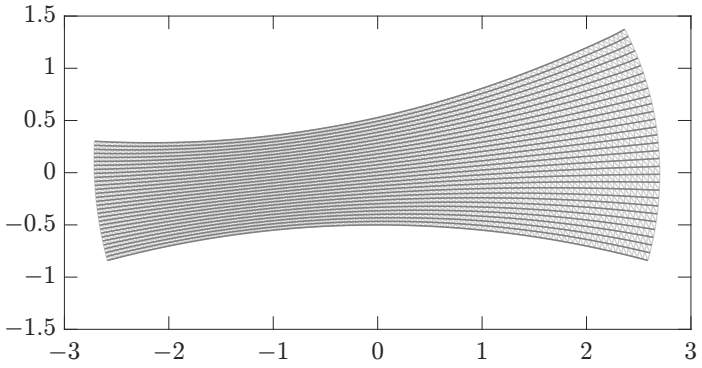


Figure 3.17: Mesh and layers of a simulated shape.

models that an extremely ill shape has no room to be further atrophied. We next see the results when we vary h in Figure 3.19. Figure 3.19(a) repeats the results in Figure 3.18 when $h = 0.2$ for comparison. When we decrease h to 0.05 in Figure 3.19(b), the small amount of initial chemical simply diffuses without causing a visible deformation. When we increase h to 0.35 in Figure 3.19(c), we have a severer deformation as expected. The effect of varying r is similar to the result of varying h. We then start the chemical propagation at different c in Figure 3.20. It is interesting to notice, especially in Figure 3.20(b), that the shape was neither "moved" nor "rotated," which could be the consequence of the zero sum of forces and torques in our model of work. Recall that we did not impose displacement conditions on the boundary. We can thus imagine that a downward force would send the shape to negative infinity, which did not happen in Figure 3.20(b).

Next we investigate the possibility to retrieve the ground truth parameter. Figure 3.21 shows the input data for the optimization problem (3.21). Figure 3.21(a) is the same as Figure 3.17. We let $\theta_{\text{true}} = (0.5, 0, 0.3, 0.2)$ and $T_{\text{true}} = 0.5$, whose shape evolution was presented in Figure 3.18. Figure 3.21(b) shows the extracted top and bottom layers from $\varphi_{\theta_{\text{true}}}(T_{\text{true}}, \Omega_0)$ shown in Figure 3.18. We did not include the side boundary in the data because the side boundary is usually noisy or even arbitrary in real data due to the shape acquisition process; we would like to see if the information from top and bottom layers is sufficient for this kind of thin shapes we are interested in. Within the parameter θ_{true}, the center $c_{\text{true}} = (0.5, 0)$ is what we care about most. Figures 3.22 to 3.24 plot the objective function $J(c_x, c_y, r, h; T)$ with respect to (c_x, c_y) under various r, h, and T. The first row of Figures 3.22 and 3.23 shows a roughly constant value $\rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}) \approx \rho(\Omega_0, \Omega_{\text{targ}})$ due to a small amount of initial chemical and diffusion, similar to the case in Figure 3.19(b). Except the first row of Figures 3.22 and 3.23, we observe a unique global

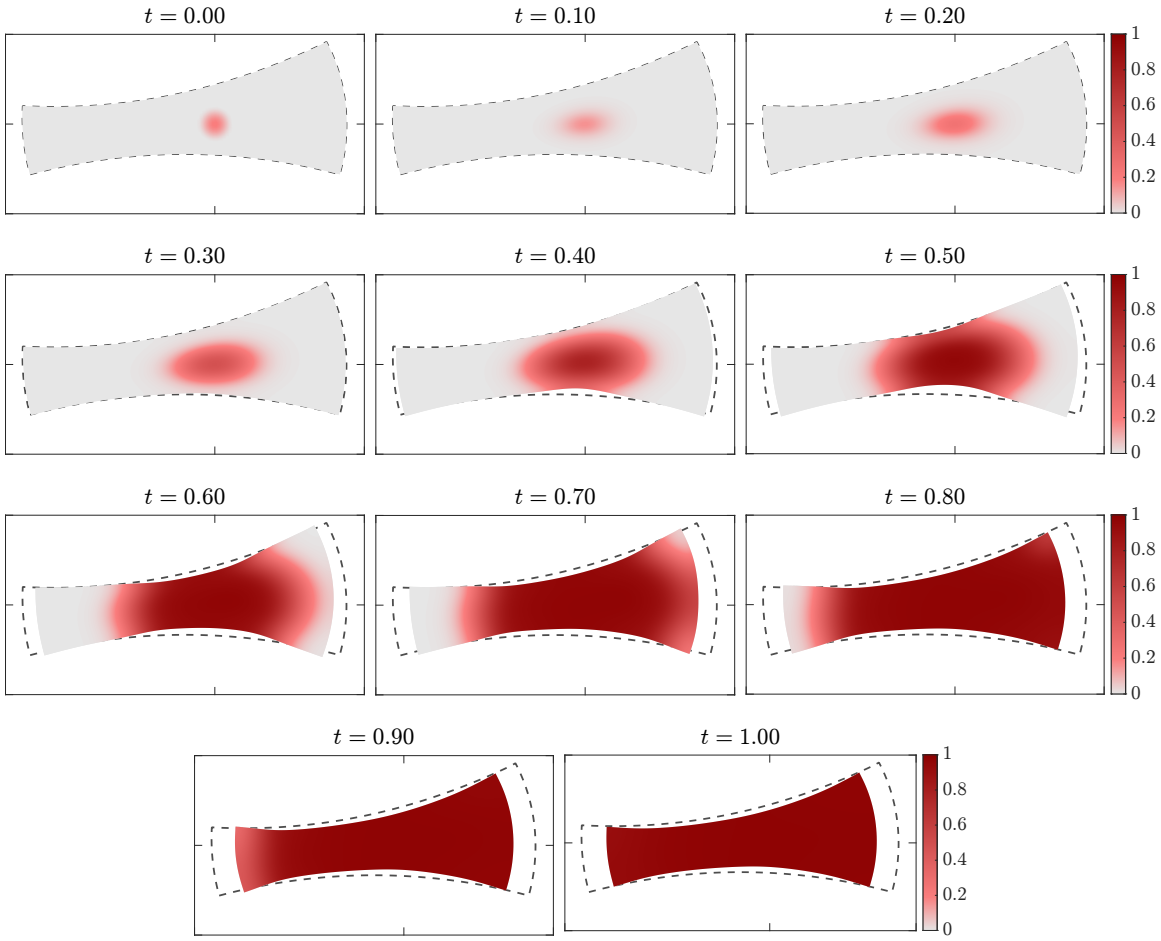
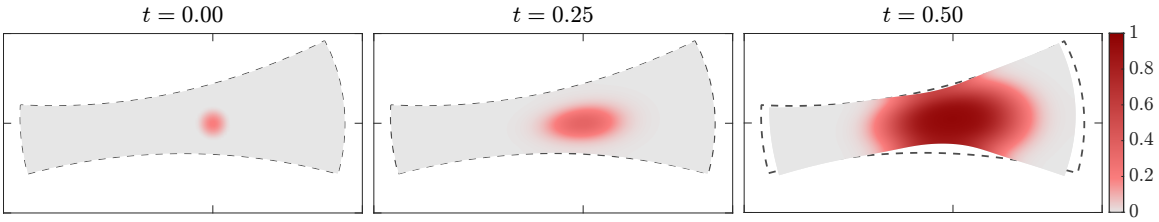
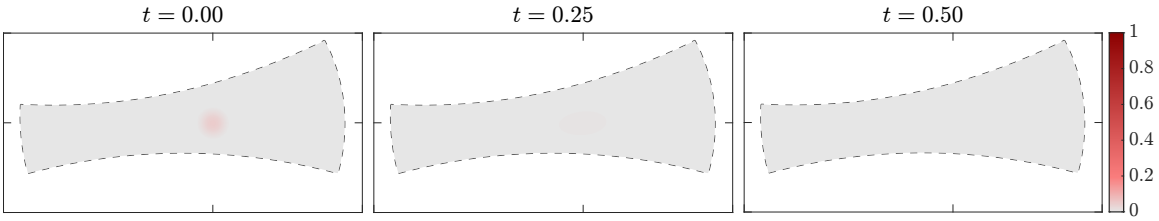


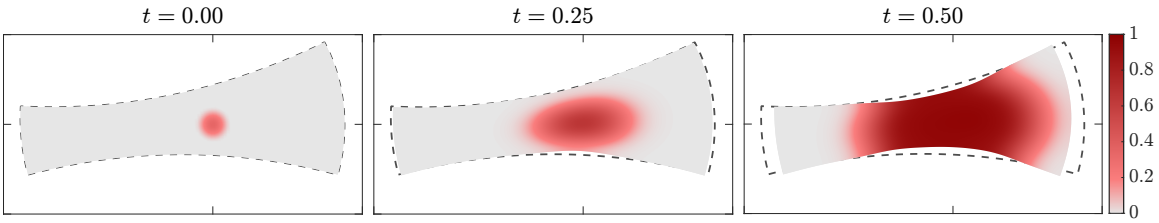
Figure 3.18: Chemical propagation and shape deformations. The color represents the Eulerian chemical distribution $\tau_E(t)$. The maximum of $\tau_E(0)$ is 0.2, and the support of reaction and atrophy functions is $[\tau_{\min}, \tau_{\max}] = [0.01, 1]$. The tangential diffusion speed is five times than the transversal diffusion speed.



(a) $h = 0.2$

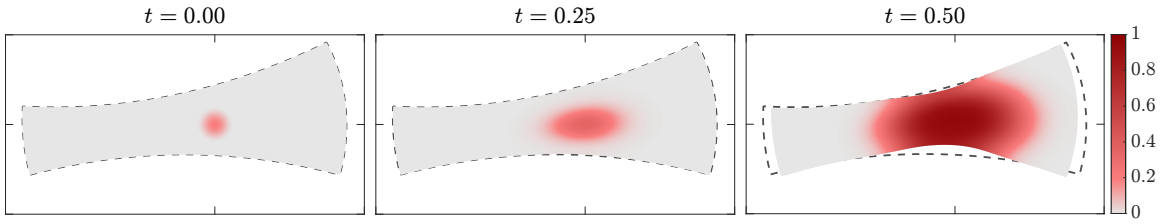


(b) $h = 0.05$

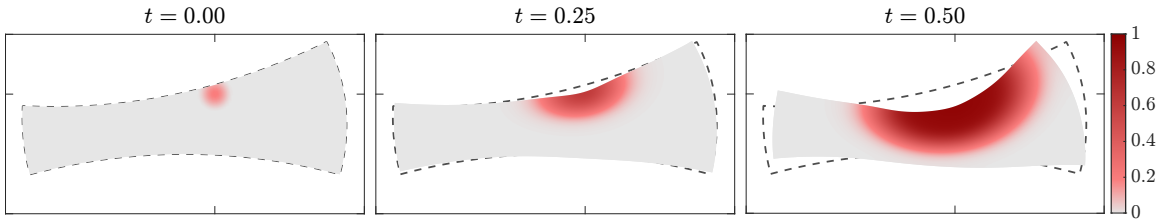


(c) $h = 0.35$

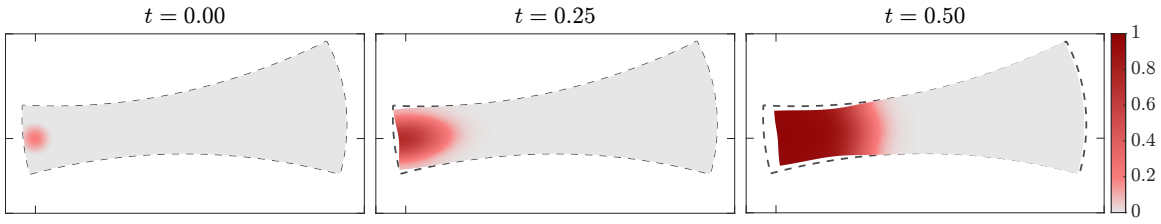
Figure 3.19: Comparison when we vary the height h of the initial chemical distribution. The color represents the Eulerian chemical distribution $\tau_E(t)$. The support of reaction and atrophy functions is $[\tau_{\min}, \tau_{\max}] = [0.01, 1]$. The tangential diffusion speed is five times than the transversal diffusion speed.



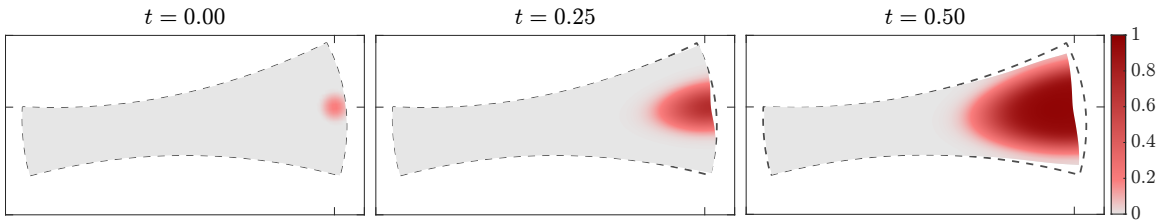
(a) $c = (0.5, 0)$



(b) $c = (0.5, 0.5)$



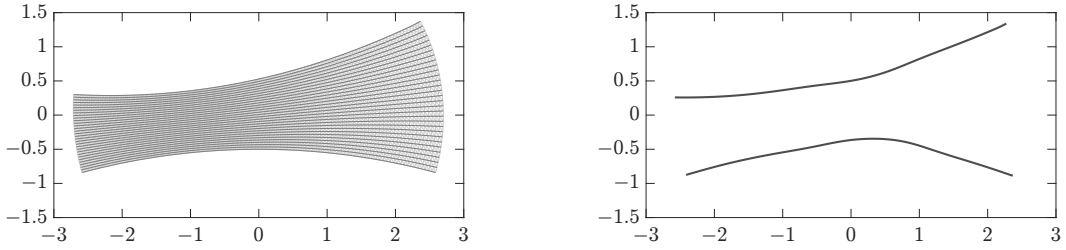
(c) $c = (-2.5, -0.25)$



(d) $c = (2.5, 0.3)$

Figure 3.20: Comparison when we vary the center c of the initial chemical distribution. The color represents the Eulerian chemical distribution $\tau_E(t)$. The maximum of $\tau_E(0)$ is 0.2, and the support of reaction and atrophy functions is $[\tau_{\min}, \tau_{\max}] = [0.01, 1]$. The tangential diffusion speed is five times than the transversal diffusion speed.

minimizer around c_{true}. Thus we can expect $\arg\min_c J(c, r, h; T) \approx c_{\text{true}}$ when r, h, and T are close to true parameters. However, the optimization problem would be challenging if we optimize (c, r, h, T) altogether as we can see from the close objective function values of those global minimizers.



(a) Initial shape.

(b) Top and bottom layers of the target shape.

Figure 3.21: Data for the optimization problem.

In our 3D simulation, we used 3D data derived from the BIOCARD dataset [70], which is a longitudinal study of Alzheimer's disease. More precisely, the initial shape shown in Figure 3.25 was obtained by computing a shape average [65] of scans of the entorhinal cortex of subjects who were cognitively normal when the MRI scans were acquired. The layered structure of the initial shape was inferred using the algorithm in [82, 105]. The parameters of this simulation are as follows: the kernel width is $\sigma_V = 0.5$, and the regularization weight is $\gamma = 0.01$; the elasticity parameters are $\lambda_{\text{tan}} = 5$, $\mu_{\text{tan}} = 5$, $\mu_{\text{tsv}} = 5$, and $\mu_{\text{ang}} = 25$; the diffusion speeds are $r_{\text{tan}} = 1$ and $r_{\text{tsv}} = 0.2$; the maximum of reaction and diffusion functions equals to $R_{\max} = \alpha_{\max} = 15$; the parameters of the varifold pseudo-metric are $w_\rho = 1$ and $\sigma_\rho = 0.1$. Following the same process as in the 2D experiment, Figure 3.26 demonstrates the shape evolution with $\theta = (c_x, c_y, c_z, r, h) = (0,\ 4,\ 0.5,\ 0.6,\ 0.2)$. We then extracted top and bottom layers from the shape at $t = 0.4$ in Figure 3.26; the extracted layers for the optimization problem is shown in Figure 3.27(b). Figure 3.28 plots the objective function $J(c_x, c_y, c_z, r_{\text{true}}, h_{\text{true}}; T_{\text{true}})$ evaluated at nodes. The plots of the objective function suggest that we could have a unique global minimizer around c_{true}, although the optimization would be difficult since the coarse and nonuniform mesh shown in Figure 3.25 implies a rough or even possibly discontinuous objective function. We also remark that whether propagation or pure diffusion occurs depends on the fineness of mesh. For example, if the mesh is so coarse that the support of the initial chemical distribution is totally contained within one discretized tetrahedron, then we may conclude wrongly that propagation did not happen. It is important that the mesh size is fine enough with respect to both the expected radius

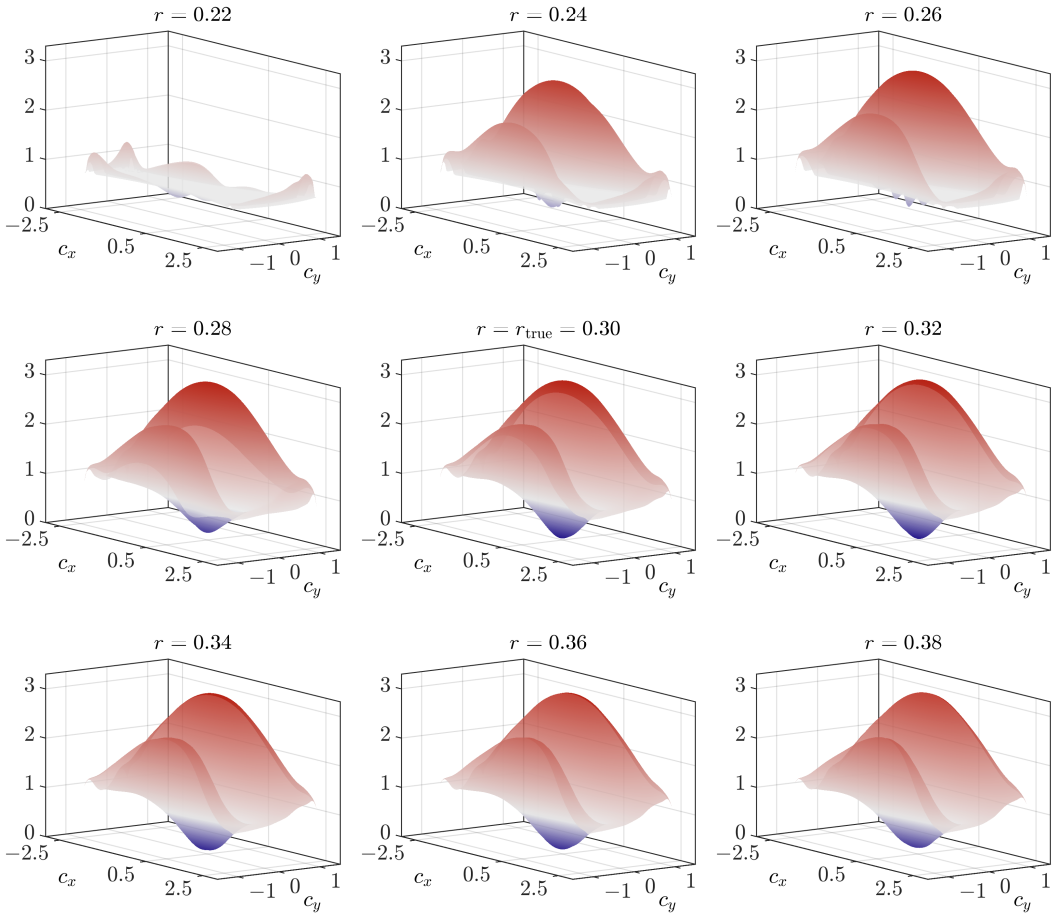


Figure 3.22: Objective function $J(c_x, c_y, r, h; T)$ when $h = h_{\text{true}}$ and $T = T_{\text{true}}$. The true center is $c_{\text{true}} = (0.5, 0)$.

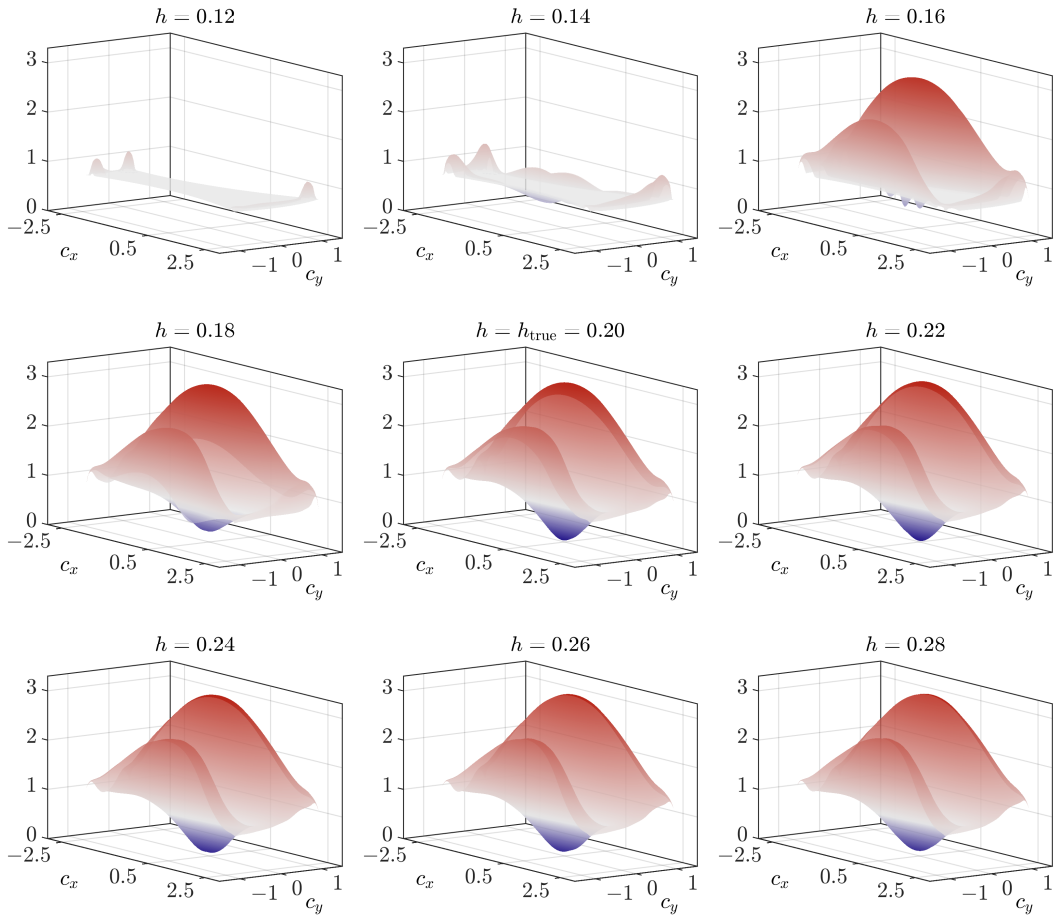


Figure 3.23: Objective function $J(c_x, c_y, r, h; T)$ when $r = r_{\text{true}}$ and $T = T_{\text{true}}$. The true center is $c_{\text{true}} = (0.5, 0)$.

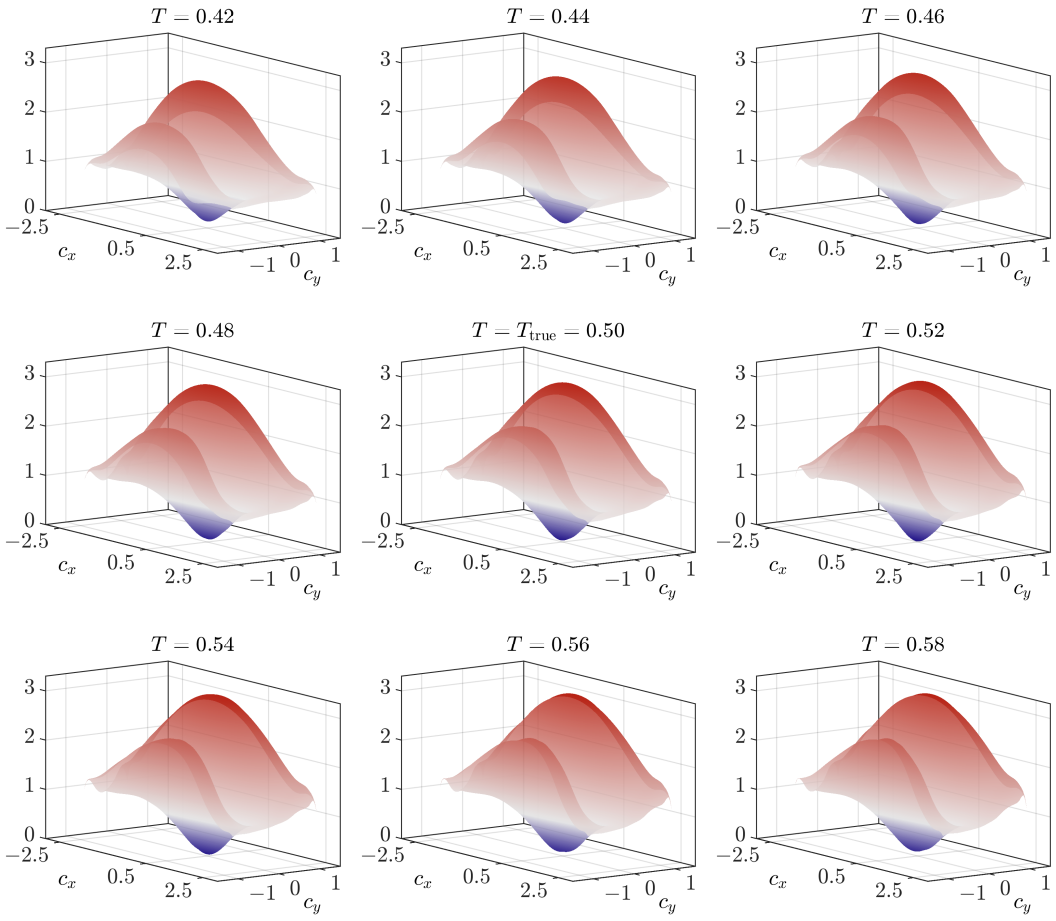


Figure 3.24: Objective function $J(c_x, c_y, r, h; T)$ when $r = r_{\text{true}}$ and $h = h_{\text{true}}$. The true center is $c_{\text{true}} = (0.5, 0)$.

of the support of the initial chemical distribution and the propagation speed.

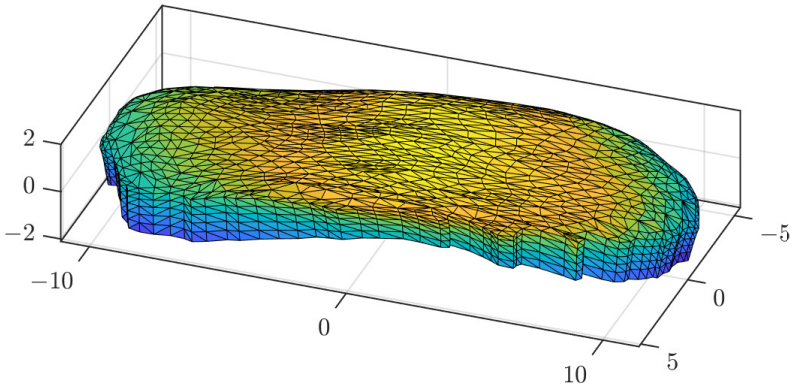


Figure 3.25: Entorinal cortex averaged over multiple subjects from the BIOCARD dataset.

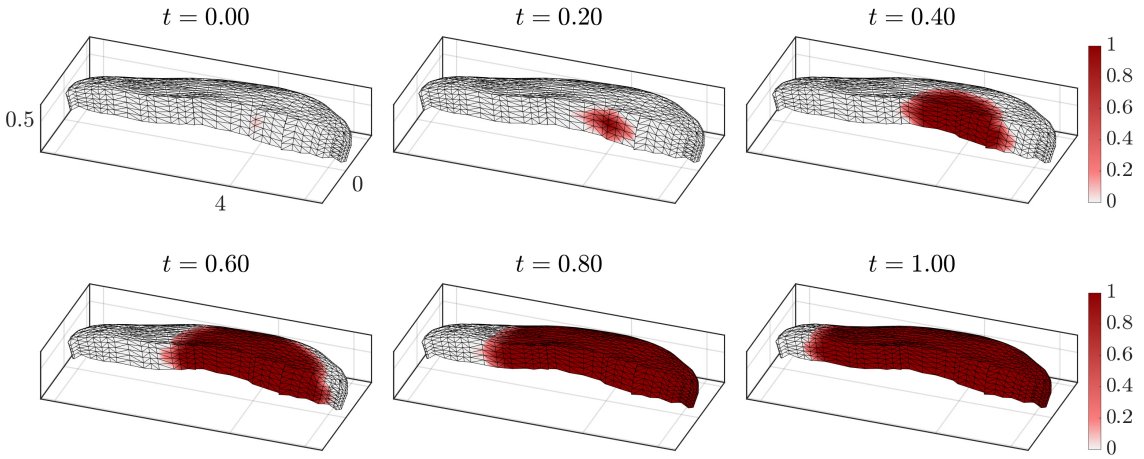
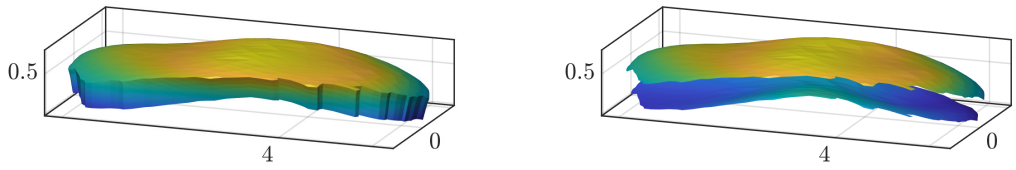


Figure 3.26: Simulated chemical propagation and deformations of entorhinal cortex. The color represents the Eulerian chemical distribution $\tau_E(t)$. The maximum of $\tau_E(0)$ is 0.2, and the support of reaction and atrophy functions is $[\tau_{\min}, \tau_{\max}] = [0.01, 1]$. The tangential diffusion speed is five times than the transversal diffusion speed.

3.3 Abstraction of the Two Examples

In this section, we recast our model of deformation vector fields of the two examples into the abstract form $\mathcal{M} : \bigcup_{t \in [0,T]} C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d)) \times Y \to V$. In the next chapter, we will develop theorems via this abstract form. The existence of minimizers for the minimization problems in our two examples will then follow as special cases of the theorems.



(a) Initial shape. (b) Top and bottom layers of the target shape.

Figure 3.27: Data for the optimization problem.

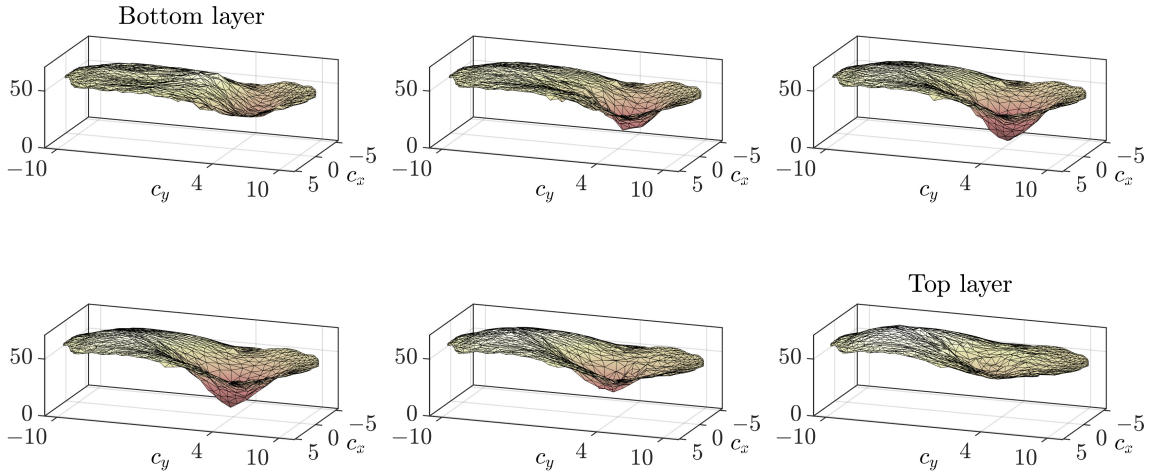


Figure 3.28: Objective function values evaluated at nodes. The ground truth $c_{\text{true}} = (0, 4, 0.5)$ is between the third the the fourth layer.

3.3.1 Piecewise-rigid motion

We recall from (3.4) that the model of deformation vector fields in this example is given by

$$\mathcal{M}_R(\varphi_{|[0,t]}, \theta(t)) := \operatorname*{arg\,min}_{v \in V} \left(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} \sum_{i=1}^{N} \int_{\varphi(t,\, \Omega_i)} |v - \mathcal{V}\, \theta_i(t)|^2 \, dx \right). \tag{3.23}$$

Note that the model (3.23) can also be written as

$$\mathcal{M}_R(\varphi_{|[0,t]}, \theta(t)) = \operatorname*{arg\,min}_{v \in V} \left(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} \int_{\varphi(t,\, \Omega_0)} |v|^2 \, dx - \sum_{i=1}^{N} \int_{\varphi(t,\, \Omega_i)} (\mathcal{V}\, \theta_i(t))^\top v \, dx \right).$$

For technical reasons (see page 114), let $\Omega^* \subset \mathbb{R}^d$ be a fixed ball centered at the origin with a very large radius, and let $\chi : \mathbb{R}^d \to [0, 1]$ be a C^∞ cutoff function of compact support such that $\chi_{|\Omega^*} \equiv 1$. With the integrand of the last term multiplied by χ, using the same

notation, we instead consider

$$\mathcal{M}_R(\varphi|_{[0,t]},\theta(t)) = \operatorname*{arg\,min}_{v\in V}\Big(\frac{\gamma}{2}\|v\|_V^2 + \frac{1}{2}\int_{\varphi(t,\Omega_0)}|v|^2\,dx - \sum_{i=1}^{N}\int_{\varphi(t,\Omega_i)}\chi\ (\mathcal{V}\,\theta_i(t))^\top v\,dx\Big)$$
$$=: \operatorname*{arg\,min}_{v\in V}\Big(\frac{\gamma}{2}\|v\|_V^2 + \frac{1}{2}\,(\mathcal{A}_\varphi^t\,v \mid v) - (\beta^t_{\varphi,\theta(t)} \mid v)\Big), \tag{3.24}$$

where $\mathcal{A}_\varphi^t \in \mathscr{L}(V,V^*)$ and $\beta^t_{\varphi,\theta(t)} \in V^*$. The peculiar notation $\mathcal{A}_\varphi^t$ and $\beta^t_{\varphi,\theta(t)}$ will be explained in Section 4.1. We notice that $(\mathcal{A}_\varphi^t\,v \mid v) \geq 0$ for all $v \in V$. Let $k = \frac{Nd(d+1)}{2}$ be the total number of parameters as before. The problem of this example can then be stated as

$$\min_{\theta\in L^2([0,T],\mathbb{R}^k)}\left(\frac{1}{2}\int_0^T\Big(\|v(t)\|_V^2 + |\theta(t)|^2\Big)\,dt + \rho(\varphi(T,\Omega_0),\Omega_{\text{targ}})\right) \tag{3.25}$$

subject to

$$\varphi(t,x) = x + \int_0^t \mathcal{M}_R(\varphi|_{[0,s]},\theta(s))(\varphi(s,x))\,ds \quad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^d.$$

3.3.2 Biological atrophy modeling

We recall from (3.21) that our problem is stated as

$$\min_{\theta\in\Theta}\ \rho(\varphi(T,\Omega_0),\Omega_{\text{targ}})$$

subject to

$$\begin{cases}
\varphi(t,x) = x + \displaystyle\int_0^t v(s,\varphi(s,x))\,ds \quad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^d \\
v(t) = \displaystyle\operatorname*{arg\,min}_{v'\in V}\Big(\frac{\gamma}{2}\|v'\|_V^2 \\
\qquad\qquad + \displaystyle\frac{1}{2}\int_{\varphi(t,\Omega_0)}\mathcal{E}_{\varphi(t)*\Phi}(\varepsilon_{v'},\varepsilon_{v'})\,dx - \int_{\varphi(t,\Omega_0)}\alpha\big(\tau(t)\circ\varphi(t)^{-1}\big)\,(-\operatorname{div}v')\,dx\Big) \\
\begin{cases}
\partial_t\big(\tau\,\det D\varphi\big)(t) = \operatorname{div}\Big(W_{\varphi(t)}\,\nabla\tau(t)\,\det D\varphi(t)\Big) + R(\tau(t))\,\det D\varphi(t) & \text{in } (0,T]\times\Omega^o \\
\big(W_{\varphi(t)}\,\nabla\tau(t)\big)^\top n(0) = 0 & \text{on } [0,T]\times\partial\Omega \\
\tau(0) = Q(\theta) & \text{on } \Omega^o
\end{cases}
\end{cases},$$

where

$$W_{\varphi(t)} = \big(D\varphi(t)\big)^{-1}\big(U_{\varphi(t)}\circ\varphi(t)\big)\ D\varphi(t)^{-\top}$$

and

$$U_{\varphi(t)} := r_{\tan}\big(I_d - N_{\varphi(t)}\,N_{\varphi(t)}^\top\big) + r_{\text{tsv}}\,S_{\varphi(t)}\,S_{\varphi(t)}^\top.$$

For the same technical reasons (see page 114) as the problem of piecewise-rigid motions, we let $\Omega^* \subset \mathbb{R}^d$ be a fixed ball centered at the origin with a very large radius, and let $\chi : \mathbb{R}^d \to [0,1]$ be a C^∞ cutoff function of compact support such that $\chi_{|\Omega^*} \equiv 1$. We then define our model of deformation vector fields as

$$\mathcal{M}_A(\varphi_{|[0,t]}, \theta) := \operatorname*{arg\,min}_{v \in V} \Big(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} (\mathcal{A}_\varphi^t v \mid v) - (\beta_{\varphi,\theta}^t \mid v) \Big), \tag{3.26}$$

where

$$(\mathcal{A}_\varphi^t v \mid v) := \int_{\varphi(t,\,\Omega_0)} \mathcal{E}_{\varphi(t)\,*\,\Phi}(\varepsilon_v, \varepsilon_v)\, dx$$

and

$$(\beta_{\varphi,\theta}^t \mid v) := \int_{\varphi(t,\,\Omega_0)} \chi \;\; \alpha\big(\tau(t) \circ \varphi(t)^{-1}\big)\, (-\operatorname{div} v)\, dx. \tag{3.27}$$

Since an elasticity tensor is positive definite almost everywhere, we have $(\mathcal{A}_\varphi^t v \mid v) \geq 0$ for all $v \in V$. We recall our framework of interpretable deformation vector fields

$$v(t) = \mathcal{M}(\varphi_{|[0,t]}, \theta(t)). \tag{2.9}$$

In our atrophy model $\mathcal{M}_A$ (3.26), we identify $\theta \in \mathbb{R}^k$ with a constant function, thus this framework (2.9) still applies. In summary, let $\Theta \subset \mathbb{R}^k$ be a compact set. Our atrophy problem is written as

$$\min_{\theta \in \Theta} \;\; \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}})$$

subject to

$$\varphi(t,x) = x + \int_0^t \mathcal{M}_A(\varphi_{|[0,s]}, \theta)(\varphi(s,x))\, ds \quad \text{ for all } (t,x) \in [0,T] \times \mathbb{R}^d.$$

Chapter 4

Core Theorems

We develop sufficient conditions for the existence of minimizers of the problem

$$\min_{\theta \in \Theta} \left(\int_0^T \Lambda(\varphi(t), \theta(t))\, dt + \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}}) \right)$$

subject to

$$\varphi(t, x) = x + \int_0^t \mathcal{M}(\varphi_{|[0,\, s]}, \theta(s))(\varphi(s, x))\, ds \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d.$$

We collect statements and remarks of our theorems in Section 4.1 to conduct our train of thought. Then we prepare a necessary technicality, Faà di Bruno's formula, in Section 4.2. It would be a good time to recall properties of the Bochner integral in Section 2.2 before reading Section 4.3. After those preparations, we will be ready to prove our theorems in Section 4.3. We assume that V is a separable Hilbert space continuously embedded in $C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ throughout this chapter (see Section 2.4.2 for the existence of V). In addition, it would be helpful to keep in mind our notation of hierarchical time intervals: we will use $t \in [0, T]$, $s \in [0, t]$, and $s' \in [0, s]$.

4.1 Statements

Theorem 4.1.1. *Let $\mathcal{M} : \bigcup_{t \in [0,\, T]} C([0, t], \mathrm{Diff}^p_{id}(\mathbb{R}^d)) \times Y \to V$ be a model of deformation vector fields. Given $\theta : [0, T] \to Y$ defined almost everywhere, denote the deformation vector field associated to $\varphi \in C([0, t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $t \le T$, by*

$$v_\varphi(s) := \mathcal{M}(\varphi_{|[0,\, s]}, \theta(s)).$$

We consider the initial value problem

$$\varphi(t,x) = x + \int_0^t v_\varphi(s,\varphi(s,x))\,ds \quad \text{for all } x \in \mathbb{R}^d. \tag{4.1}$$

Suppose that $v_\varphi : [0,t] \to V$ *is strongly measurable for all* $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$. *Moreover, suppose that* $\varphi \mapsto v_\varphi$ *is locally Lipschitz in the sense that given a fixed* $\overline{\varphi} \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, *there exist* $r(\overline{\varphi}) > 0$ *and* $f_{\overline{\varphi}} \in L^1([0,T])$ *such that*

$$\|v_\varphi(s) - v_\psi(s)\|_V \le f_{\overline{\varphi}}(s) \sup_{s' \in [0,\,s]} \|\varphi(s') - \psi(s')\|_{p,\infty}$$

for all $\varphi, \psi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ *with* $\sup_{s \in [0,t]} \|\varphi(s) - \overline{\varphi}(s)\|_{p,\infty} \le r$ *and* $\sup_{s \in [0,t]} \|\psi(s) - \overline{\varphi}(s)\|_{p,\infty} \le r$ *and for almost every* $s \in [0,t]$.

(i) If there exists $g \in L^1([0,T])$ *such that the deformation vector field satisfies*

$$\|v_\varphi(s)\|_V \le g(s)\left(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_{p,\infty}\right) \tag{4.2}$$

for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ *and for almost every* $s \in [0,t]$, *then the initial value problem* (4.1) *has a unique maximal solution either in* $C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ *for some* $T' \le T$ *or in* $C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$.

(ii) If there exists $g \in L^1([0,T])$ *such that the deformation vector field satisfies*

$$\|v_\varphi(s)\|_V \le g(s)\left(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_{\infty}\right) \tag{4.3}$$

for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ *and for almost every* $s \in [0,t]$, *then the initial value problem* (4.1) *has a unique solution in* $C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$.

(See the proof on page 75.)

Remark 4.1.2. Although there is no explicit restriction on $\theta : [0,T] \to Y$, an implicit restriction comes from the integrability of $f_{\overline{\varphi}}$ and g, which may depend on θ. For example, if $g(s) = \|\theta(s)\|_Y$, then θ has to be (Bochner) integrable. We also remark that v_φ only needs to be defined almost everywhere, so a θ defined almost everywhere is valid.

Remark 4.1.3. It is clear that $f_{\overline{\varphi}}$ and g are positive almost everywhere.

We require a regularity assumption, appearing in the next proposition, on the discrepancy function ρ. To this end, we define a seminorm on a compact set $\Omega \subset \mathbb{R}^d$ by

$$\|v\|^\Omega_{p,\infty} := \sum_{j=0}^{p} \max_{x \in \Omega} |D^j v(x)|.$$

We remind the reader that a shape $\Omega \in \mathscr{S}$ is a compact subset of $\mathbb{R}^d$ (see (2.1)). We also recall that $\xi(\mathscr{S}) \subset \mathscr{S}$ for all $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$.

Definition 4.1.4. We say that a discrepancy function $\rho : \mathscr{S} \times \mathscr{S} \to \mathbb{R}_{\geq 0}$ is continuous with respect to $\|\cdot\|_{p,\infty}$ if for all $\Omega, \Omega' \in \mathscr{S}$ and all sequences $(\xi_n)_{n=1}^{\infty} \subset \mathrm{Diff}^p_{id}(\mathbb{R}^d)$ such that $\|\xi_n - \xi\|^{\Omega}_{p,\infty} \to 0$ for some $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$, one has

$$\rho(\xi_n(\Omega), \Omega') \to \rho(\xi(\Omega), \Omega').$$

In addition, we use the following terminology to simplify the statement in the next proposition, and more importantly, to emphasize conditions that are solely affected by the choice of Θ and $\mathcal{M}$.

Definition 4.1.5. Let Θ be a weakly closed subset of a reflexive Banach space and $\mathcal{M}$ be a model of deformation vector fields. We say that Θ and $\mathcal{M}$ are compatible if:

- For all $\theta \in \Theta$, the initial value problem (4.1) has a unique solution in $C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$.
- For all $\Omega \in \mathscr{S}$ and for all sequence $(\theta_n)_{n=1}^{\infty} \subset \Theta$ such that $\theta_n \rightharpoonup \theta$, we have

$$\|\varphi_{\theta_n}(t) - \varphi_{\theta}(t)\|^{\Omega}_{p,\infty} \to 0 \quad \text{for all } t \in [0,T],$$

  where φ_{θ_n} and φ_{θ} are the unique solutions of θ_n and θ respectively.

We recall that a function $f : X \to \mathbb{R}$ defined on a normed vector space X is called coercive if $f(x) \to +\infty$ when $\|x\|_X \to \infty$.

Proposition 4.1.6. *Let Θ be a weakly closed subset of a reflexive Banach space and $\mathcal{M}$ be a model of deformation vector fields. Suppose that Θ and $\mathcal{M}$ are compatible. Given $\Omega_0, \Omega_{\mathrm{targ}} \in \mathscr{S}$, we denote the unique solution to the initial value problem* (4.1) *corresponding to $\theta \in \Theta$ by φ_θ and consider the minimization problem*

$$\min_{\theta \in \Theta} \left(\int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt + \rho(\varphi_\theta(T, \Omega_0), \Omega_{\mathrm{targ}}) \right). \tag{4.4}$$

Assume that:

(1) Either the objective function is coercive or Θ is bounded.

(2) The function $\theta \mapsto \int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt$ is weakly sequentially lower semicontinuous.

(3) The discrepancy function ρ is continuous with respect to $\|\cdot\|_{p,\infty}$.

Then the minimization problem (4.4) *has a minimizer.*

(See the proof on page 82.)

Remark 4.1.7. The reader may recognize the program of the direct method of calculus of variations in the above proposition.

Remark 4.1.8. The condition on Θ can be relaxed to a weakly sequentially closed subset of a vector space which is the dual of a Banach space. In practice, Θ is usually a strongly closed ball in a Hilbert space, even the whole Hilbert space, or a closed subset of a finite-dimensional vector space.

Remark 4.1.9. It can be shown that varifold pseudo-metrics, the discrepancy functions we used in Chapter 3, are continuous with respect to $\|\cdot\|_{p,\infty}$, $p \geq 1$ [26, Proposition 6]. Another example of discrepancy function is the volume of the symmetric difference between two sets, i.e., $\rho(\Omega, \Omega') := w_\rho \operatorname{vol}(\Omega \triangle \Omega')$, which is continuous with respect to $\|\cdot\|_{p,\infty}$, $p \geq 0$. However, the volume of a symmetric difference is less attractive from the computational viewpoint: it is not trivial to compute the volume of a symmetric difference accurately, not to mention that the volume of a symmetric difference is not differentiable with respect to discretized nodes, whereas function values and gradients are driving forces in most minimization procedures. The above-mentioned discrepancy functions do not require a point-to-point correspondence between two shapes. If there is a point-to-point correspondence within a subclass of shapes, for example, a class of parametrized shapes $\{\Omega_\alpha = f_\alpha(\Psi) : f_\alpha \text{ is one-to-one and onto}\} \subset \mathscr{S}$ on a fixed $\Psi \subset \mathbb{R}^d$, then, restricted to the subclass, we can consider $\rho(\Omega_\alpha, \Omega_\beta) := w_\rho \int_\Psi |f_\alpha - f_\beta|^2 \, dx$, which is an analogue of the ℓ^2-distance of discretized points. This discrepancy function is continuous with respect to $\|\cdot\|_{p,\infty}$, $p \geq 0$.

Now we focus on sufficient conditions on $\mathcal{M}$ such that Θ and $\mathcal{M}$ are compatible. To state the theorem, we define

$$\overline{\mathfrak{B}}(id, r) := \{\xi \in \operatorname{Diff}^p_{id}(\mathbb{R}^d) : \|\xi - id\|_{p,\infty} \leq r \text{ and } \|\xi^{-1} - id\|_{p,\infty} \leq r\}.$$

Note that $\overline{\mathfrak{B}}(id, r)$ is a subset of $\operatorname{Diff}^p_{id}(\mathbb{R}^d)$, which we equip with the metric $d_{p,\infty}(\xi, \eta) = \|\xi - \eta\|_{p,\infty}$. On the other hand, if we equip $\operatorname{Diff}^p_{id}(\mathbb{R}^d)$ with a different metric $\mathfrak{d}_{p,\infty}(\xi, \eta) := \max\{\|\xi - \eta\|_{p,\infty}, \|\xi^{-1} - \eta^{-1}\|_{p,\infty}\}$, then $\overline{\mathfrak{B}}(id, r)$ can be interpreted as a closed ball of $\operatorname{Diff}^p_{id}(\mathbb{R}^d)$, as our notation suggests. This notation will not lead to confusion because $d_{p,\infty}$ and $\mathfrak{d}_{p,\infty}$ generate the same topology on $\operatorname{Diff}^p_{id}(\mathbb{R}^d)$. Indeed, it is clear that $id : (\operatorname{Diff}^p_{id}(\mathbb{R}^d), \mathfrak{d}_{p,\infty}) \to (\operatorname{Diff}^p_{id}(\mathbb{R}^d), d_{p,\infty})$ is continuous. It can also be shown that $\xi \mapsto \xi^{-1}$ is continuous from $(\operatorname{Diff}^p_{id}(\mathbb{R}^d), d_{p,\infty})$ to $(\operatorname{Diff}^p_{id}(\mathbb{R}^d), d_{p,\infty})$, which implies that $id : (\operatorname{Diff}^p_{id}(\mathbb{R}^d), d_{p,\infty}) \to (\operatorname{Diff}^p_{id}(\mathbb{R}^d), \mathfrak{d}_{p,\infty})$ is continuous. Since $d_{p,\infty}$ and $\mathfrak{d}_{p,\infty}$ generate the

same topology on $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$, the meaning of $C([0,t], \overline{\mathfrak{B}}(id, r))$ in the following theorem is unambiguous.

Theorem 4.1.10. *Let Θ be a weakly closed subset of a reflexive Banach space B and $\mathcal{M}$ be a model of deformation vector fields. For all $\Omega \in \mathscr{S}$, suppose that:*

- *The mapping $s \mapsto \mathcal{M}(\varphi_{|[0,\,s]}, \theta(s))$ is strongly measurable for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ and $\theta \in \Theta$.*
- *The model is bounded:*
  For all $\theta \in \Theta$, there exists $f_\theta \in L^2([0,T])$ such that
  $$\|\mathcal{M}(\varphi_{|[0,\,s]}, \theta(s))\|_V \leq f_\theta(s) \left(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_\infty\right)$$
  for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ and for almost every $s \in [0,t]$. Moreover, for all $m > 0$, there exists a constant $F_m > 0$ such that $\|\theta\|_B \leq m$ implies $\|f_\theta\|_{L^2} \leq F_m$.
- *The model is Lipschitz in φ:*
  For all $r > 0$ and $\theta \in \Theta$, there exists $g_{r,\theta} \in L^2([0,T])$ such that
  $$\|\mathcal{M}(\varphi_{|[0,\,s]}, \theta(s)) - \mathcal{M}(\psi_{|[0,\,s]}, \theta(s))\|_V \leq g_{r,\theta}(s) \sup_{s' \in [0,\,s]} \|\varphi(s') - \psi(s')\|^\Omega_{p,\infty}$$
  for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ and for almost every $s \in [0,t]$. Moreover, for all $m > 0$, there exists a constant $G_{r,m} > 0$ such that $\|\theta\|_B \leq m$ implies $\|g_{r,\theta}\|_{L^2} \leq G_{r,m}$.
- *The model is continuous in θ:*
  If $(\theta_n)_{n=1}^\infty \subset \Theta$ and $\theta_n \rightharpoonup \theta$, then for $0 \leq j \leq p$
  $$\int_0^t D^j \Big(\mathcal{M}(\varphi_{|[0,\,s]}, \theta_n(s))(\varphi(s,x))\Big)\, ds \to \int_0^t D^j \Big(\mathcal{M}(\varphi_{|[0,\,s]}, \theta(s))(\varphi(s,x))\Big)\, ds$$
  for all $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $t \in [0,T]$, and $x \in \Omega$.

Then Θ and $\mathcal{M}$ are compatible.

(See the proof on page 84.)

Remark 4.1.11. The last assumption, the continuity of the model in θ, appears really raw, and it is; this is exactly one step in the proof. It is the weak condition $\theta_n \rightharpoonup \theta$ that prevents us from modifying this form. There are two "simple" situations in which this assumption may hold. The first one is that

$$\theta' \mapsto \int_0^t D^j \Big(\mathcal{M}(\varphi_{|[0,\,s]}, \theta'(s))(\varphi(s,x))\Big)\, ds$$

is bounded and linear, and the second one is that Θ is in a finite-dimensional vector space. Our two examples represent these two cases. This assumption is a challenge that strictly restricts models we can validate.

Denote by $\mathcal{K}_V$ the inverse of the duality map of the Hilbert space V, i.e., $(\mathcal{K}_V^{-1} v \mid v) = \|v\|_V^2$ and $\|\mathcal{K}_V^{-1} v\|_{V^*} = \|v\|_V$. We further consider models of deformation vector fields taking the form

$$\mathcal{M}(\varphi_{|[0,t]}, \theta(t)) = (\gamma \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1} \beta_{\varphi,\theta(t)}^t. \tag{4.5}$$

This form is in the following setting. The bounded linear operator $\mathcal{A}_\varphi^t \in \mathscr{L}(V, V^*)$ depends on φ "up to" time t (the notation $\mathcal{A}_{\varphi_{|[0,t]}}$ is too cumbersome) and is nonnegative in the sense that $(\mathcal{A}_\varphi^t v \mid v) \geq 0$ for all $v \in V$. The linear functional $\beta_{\varphi,\theta(t)}^t \in V^*$ depends on both $\varphi_{|[0,t]}$ and $\theta(t)$. This form of model usually comes from a regularized energy since

$$\begin{aligned}\mathcal{M}(\varphi_{|[0,t]}, \theta(t)) &= (\gamma \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1} \beta_{\varphi,\theta(t)}^t \\ &= \operatorname*{arg\,min}_{v \in V} \left(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} (\mathcal{A}_\varphi^t v \mid v) - (\beta_{\varphi,\theta(t)}^t \mid v) \right).\end{aligned}$$

Thus this form will be referred to as the energy form. Notice that the models of our two examples in the previous chapter are of the energy form (see (3.24) and (3.26)). Theorem 4.1.10 immediately gives sufficient conditions on the operators $\mathcal{A}_\varphi^t$ and $\beta_{\varphi,\theta(t)}^t$.

Corollary 4.1.12. *Let Θ be a weakly closed subset of a reflexive Banach space B. Suppose that the model of deformation vector fields $\mathcal{M}$ is of the energy form* (4.5)*. For all $\Omega \in \mathscr{S}$, we also suppose that:*

- *The mapping $s \mapsto \mathcal{A}_\varphi^s$ is in $C([0,t], \mathscr{L}(V, V^*))$ for all $\varphi \in C([0,t], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$. The mapping $s \mapsto (\beta_{\varphi,\theta(s)}^s \mid v)$ is Lebesgue measurable for all $\varphi \in C([0,t], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$, $\theta \in \Theta$, and $v \in V$.*
- *For all $\theta \in \Theta$, there exists $f_\theta \in L^2([0,T])$ such that*

  $$\|\beta_{\varphi,\theta(s)}^s\|_{V^*} \leq f_\theta(s) \left(1 + \sup_{s' \in [0,\, s]} \|\varphi(s') - id\|_\infty\right)$$

  for all $\varphi \in C([0,t], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$ and for almost every $s \in [0,t]$. Moreover, for all $m > 0$, there exists a constant $F_m > 0$ such that $\|\theta\|_B \leq m$ implies $\|f_\theta\|_{L^2} \leq F_m$.
- *For all $r > 0$ and $\theta \in \Theta$, there exist $\ell_r > 0$ and $g_{r,\theta} \in L^2([0,T])$ such that*

  $$\|\mathcal{A}_\varphi^s - \mathcal{A}_\psi^s\|_{\mathscr{L}(V, V^*)} \leq \ell_r \sup_{s' \in [0,\, s]} \|\varphi(s') - \psi(s')\|_{p,\infty}^\Omega$$

and that

$$\|\beta^s_{\varphi,\theta(s)} - \beta^s_{\psi,\theta(s)}\|_{V^*} \le g_{r,\theta}(s) \sup_{s' \in [0,s]} \|\varphi(s') - \psi(s')\|^{\Omega}_{p,\infty}$$

for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ *and for almost every* $s \in [0,t]$. *Moreover, for all* $m > 0$, *there exists a constant* $G_{r,m} > 0$ *such that* $\|\theta\|_B \le m$ *implies* $\|g_{r,\theta}\|_{L^2} \le G_{r,m}$.

- *If* $(\theta_n)_{n=1}^{\infty} \subset \Theta$ *and* $\theta_n \rightharpoonup \theta$, *then for* $0 \le j \le p$

$$\int_0^t D^j\Big(\mathcal{M}(\varphi|_{[0,s]}, \theta_n(s))(\varphi(s,x))\Big)\, ds \to \int_0^t D^j\Big(\mathcal{M}(\varphi|_{[0,s]}, \theta(s))(\varphi(s,x))\Big)\, ds$$

for all $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $t \in [0,T]$, *and* $x \in \Omega$.

Then Θ *and* $\mathcal{M}$ *are compatible.*

(See the proof on page 89.)

The lemma below will be especially useful when we work on the energy form (4.5). We will also need it later in Chapter 5.

Lemma 4.1.13. *For every nonnegative* $\mathcal{L} \in \mathscr{L}(V, V^*)$, *i.e.,* $(\mathcal{L}v \mid v) \ge 0$ *for all* $v \in V$, *we have* $(\gamma \mathcal{K}_V^{-1} + \mathcal{L})^{-1} \in \mathscr{L}(V^*, V)$ *and*

$$\big\|(\gamma \mathcal{K}_V^{-1} + \mathcal{L})^{-1}\big\|_{\mathscr{L}(V^*,V)} \le \frac{1}{\gamma}.$$

(See the proof on page 88.)

4.2 Faà di Bruno's formula

We now derive a version of Faà di Bruno's formula, an expression of $D^p(v \circ \varphi)$ (see also [104, Section 7.1] and [55]), in order to estimate $\|v \circ \varphi\|_{p,\infty}$ and $\|v \circ \varphi - v \circ \psi\|_{p,\infty}$. Only Lemma 4.2.1 and Corollary 4.2.4 in this section will be used later. The reader may skip this section if he or she is willing to accept Lemma 4.2.1 and Corollary 4.2.4.

We first brute-force expand $D^p(v \circ \varphi)$ for small p. In the following, h_1, h_2, and h_3 are arbitrary vectors in $\mathbb{R}^d$. The first derivative is given by the chain rule:

$$\Big(D(v \circ \varphi)(x)\Big)(h_1) = \Big(Dv\big(\varphi(x)\big)\Big)\big(D\varphi(x)\, h_1\big).$$

We use the product rule for the second derivative

$$\begin{aligned}\Big(D^2(v \circ \varphi)(x)\Big)(h_1, h_2) = {} & \Big(D^2 v\big(\varphi(x)\big)\Big)\big(D\varphi(x)\, h_1, D\varphi(x)\, h_2\big) \\ & + \Big(Dv\big(\varphi(x)\big)\Big)\big(D^2\varphi(x)(h_1, h_2)\big),\end{aligned}$$

and the third derivative

$$\begin{aligned}\Big(D^3(v\circ\varphi)(x)\Big)(h_1,h_2,h_3) = &\Big(D^3v\big(\varphi(x)\big)\Big)\big(D\varphi(x)\,h_1, D\varphi(x)\,h_2, D\varphi(x)\,h_3\big)\\ &+\Big(D^2v\big(\varphi(x)\big)\Big)\big(D^2\varphi(x)(h_1,h_3), D\varphi(x)\,h_2\big)\\ &+\Big(D^2v\big(\varphi(x)\big)\Big)\big(D\varphi(x)\,h_1, D^2\varphi(x)(h_2,h_3)\big)\\ &+\Big(D^2v\big(\varphi(x)\big)\Big)\big(D^2\varphi(x)(h_1,h_2), D\varphi(x)\,h_3\big)\\ &+\Big(Dv\big(\varphi(x)\big)\Big)\big(D^3\varphi(x)(h_1,h_2,h_3)\big).\end{aligned}$$

For higher order derivatives, we shall use the multi-index notation and a little combinatorics. Let $\alpha := (\alpha_1,\dots,\alpha_q)$ be a q-tuple. We define $|\alpha| := |\alpha_1| + \cdots + |\alpha_q|$ and

$$D^{|\alpha|}\varphi(x)\,h_\alpha := D^{|\alpha|}\varphi(x)(h_{\alpha_1},\dots,h_{\alpha_q}).$$

Furthermore, we denote by $\mathcal{P}(p,k)$ the collection of partitions of $\{1,\dots,p\}$ into k nonempty unlabeled groups. Using the multi-index notation, we can write, for example,

$$\begin{aligned}\mathcal{P}(4,2) = \Big\{&\big((1),(2,3,4)\big),\ \big((2),(1,3,4)\big),\ \big((3),(1,2,4)\big),\ \big((4),(1,2,3)\big),\\ &\big((1,2),(3,4)\big),\ \big((1,3),(2,4)\big),\ \big((1,4),(2,3)\big)\Big\},\end{aligned}$$

where each partition is a 2-tuple formed by 2 multi-indices. We form partitions by ordered tuples rather than unordered sets to avoid ambiguity in the following formula. Note that the number of partitions is given by the Stirling number of the second kind, i.e., $|\mathcal{P}(p,k)| = \left\{{p \atop k}\right\}$. These notations enable us to derive higher order derivatives.

Lemma 4.2.1 (Faà di Bruno's formula)**.** *Let $v \in C^p(\mathbb{R}^d,\mathbb{R}^d)$ and $\varphi \in C^p(\mathbb{R}^d,\mathbb{R}^d)$, then*

$$\begin{aligned}&\Big(D^p(v\circ\varphi)(x)\Big)(h_1,\dots,h_p)\\ &=\sum_{k=1}^{p}\sum_{(I_1,\dots,I_k)\,\in\,\mathcal{P}(p,\,k)}\Big(D^kv\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_k|}\varphi(x)\,h_{I_k}\big).\end{aligned}$$

Proof. We have proved the cases when $p = 1,2,3$. We prove the general case $p \in \mathbb{N}$ by induction. Observe that a partition in $\mathcal{P}(p+1,k)$ can be formed in two ways depending on if $p+1$ is a singleton in the partition. If $p+1$ is a singleton, the partition is of the form

$$\big(I_1,\dots,I_{k-1},(p+1)\big),\quad (I_1,\dots,I_{k-1}) \in \mathcal{P}(p,k-1);$$

if $p+1$ is not a singleton, the partition is of the form

$$\big(I_1,\dots,I_{i-1},(I_i,p+1),I_{i+1},\dots,I_k\big),\quad (I_1,\dots,I_k)\in\mathcal{P}(p,k).$$

We let $\mathcal{P}(p,0)=\emptyset$ so that the observation also holds for $k=1$. This observation and the induction hypothesis then yield

$$
\begin{aligned}
&\Big(D^{p+1}(v\circ\varphi)(x)\Big)(h_1,\dots,h_p,h_{p+1})\\
&=\sum_{k=1}^{p}\sum_{(I_1,\dots,I_k)\in\mathcal{P}(p,k)}\Bigg(\Big(D^{k+1}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_k|}\varphi(x)\,h_{I_k},D\varphi(x)\,h_{p+1}\big)\\
&\qquad\qquad+\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|+1}\varphi(x)(h_{I_1},h_{p+1}),\dots,D^{|I_k|}\varphi(x)\,h_{I_k}\big)\\
&\qquad\qquad+\ \cdots\\
&\qquad\qquad+\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_k|+1}\varphi(x)(h_{I_k},h_{p+1})\big)\Bigg)\\
&=\Big(D^{p+1}v\big(\varphi(x)\big)\Big)\big(D\varphi(x)\,h_1,\dots,D\varphi(x)\,h_{p+1}\big)\\
&\quad+\sum_{k=1}^{p}\Bigg(\sum_{(I_1,\dots,I_{k-1})\in\mathcal{P}(p,k-1)}\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_{k-1}|}\varphi(x)\,h_{I_{k-1}},D\varphi(x)\,h_{p+1}\big)\\
&\qquad+\sum_{(I_1,\dots,I_k)\in\mathcal{P}(p,k)}\Bigg(\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|+1}\varphi(x)(h_{I_1},h_{p+1}),\dots,D^{|I_k|}\varphi(x)\,h_{I_k}\big)\\
&\qquad\qquad+\ \cdots\\
&\qquad\qquad+\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_k|+1}\varphi(x)(h_{I_k},h_{p+1})\big)\Bigg)\Bigg)\\
&=\Big(D^{p+1}v\big(\varphi(x)\big)\Big)\big(D\varphi(x)\,h_1,\dots,D\varphi(x)\,h_{p+1}\big)\\
&\quad+\sum_{k=1}^{p}\sum_{(I_1,\dots,I_k)\in\mathcal{P}(p+1,k)}\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_k|}\varphi(x)\,h_{I_k}\big)\\
&=\sum_{k=1}^{p+1}\sum_{(I_1,\dots,I_k)\in\mathcal{P}(p+1,k)}\Big(D^{k}v\big(\varphi(x)\big)\Big)\big(D^{|I_1|}\varphi(x)\,h_{I_1},\dots,D^{|I_k|}\varphi(x)\,h_{I_k}\big),
\end{aligned}
$$

which completes the proof. □

Since $|\varphi(x)|\to\infty$ as $|x|\to\infty$ for $\varphi\in id+C_0^p(\mathbb{R}^d,\mathbb{R}^d)$, the above lemma immediately gives us $v\circ\varphi\in C_0^p(\mathbb{R}^d,\mathbb{R}^d)$ and the following estimate of $\|v\circ\varphi\|_{p,\infty}$.

Corollary 4.2.2. *Let $v\in C_0^p(\mathbb{R}^d,\mathbb{R}^d)$ and $\varphi\in id+C_0^p(\mathbb{R}^d,\mathbb{R}^d)$, then $v\circ\varphi\in C_0^p(\mathbb{R}^d,\mathbb{R}^d)$ and*

$$
\|v\circ\varphi\|_{p,\infty}\le\left(1+\sum_{n=1}^{p}\sum_{k=1}^{n}\left\{{n\atop k}\right\}(1+\|\varphi-id\|_{p,\infty})^k\right)\|v\|_{p,\infty},
$$

where $\left\{ {n \atop k} \right\} = |\mathcal{P}(n,k)|$ *is the Stirling number of the second kind.*

To estimate $\|v \circ \varphi - v \circ \psi\|_{p,\infty}$, we first observe the case when $p = 1$:

$$\begin{aligned}
\|v \circ \varphi - v \circ \psi\|_{1,\infty} &= \|v \circ \varphi - v \circ \psi\|_\infty + \|(Dv \circ \varphi)\, D\varphi - (Dv \circ \psi)\, D\psi\|_\infty \\
&\leq \|v \circ \varphi - v \circ \psi\|_\infty + \|(Dv \circ \varphi)\, D\varphi - (Dv \circ \psi)\, D\varphi\|_\infty \\
&\qquad + \|(Dv \circ \psi)\, D\varphi - (Dv \circ \psi)\, D\psi\|_\infty \\
&\leq \|Dv\|_\infty \|\varphi - \psi\|_\infty + \|D^2 v\|_\infty \|\varphi - \psi\|_\infty \|D\varphi\|_\infty \\
&\qquad + \|Dv\|_\infty \|D\varphi - D\psi\|_\infty \\
&\leq \|v\|_{2,\infty} \|\varphi - \psi\|_{1,\infty} (3 + \|\varphi - id\|_{1,\infty}).
\end{aligned}$$

With the expression of $D^p(v \circ \varphi)(x)$ given by Lemma 4.2.1, we can generalize the above procedure to higher orders.

Corollary 4.2.3. *Let* $v \in C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ *and* $\varphi, \psi \in id + C_0^p(\mathbb{R}^d, \mathbb{R}^d)$*, then*

$$\begin{aligned}
&\|v \circ \varphi - v \circ \psi\|_{p,\infty} \\
&\leq \left(1 + \sum_{n=1}^{p} \sum_{k=1}^{n} \left\{ {n \atop k} \right\} \left((1 + \|\varphi - id\|_{p,\infty})^k \right.\right. \\
&\qquad \left.\left. + \sum_{i=0}^{k-1} (1 + \|\varphi - id\|_{p,\infty})^{k-1-i} (1 + \|\psi - id\|_{p,\infty})^i \right) \right) \|v\|_{p+1,\infty} \|\varphi - \psi\|_{p,\infty},
\end{aligned}$$

where $\left\{ {n \atop k} \right\} = |\mathcal{P}(n,k)|$ *is the Stirling number of the second kind.*

The takeaway of Corollaries 4.2.2 and 4.2.3 is that

$$\|v \circ \varphi\|_{p,\infty} \leq q(\|\varphi - id\|_{p,\infty}) \|v\|_{p,\infty}$$

and

$$\|v \circ \varphi - v \circ \psi\|_{p,\infty} \leq q'(\|\varphi - id\|_{p,\infty}, \|\psi - id\|_{p,\infty}) \|v\|_{p+1,\infty} \|\varphi - \psi\|_{p,\infty},$$

where q and q' are polynomials of degree p. We will only use the following form of Corollaries 4.2.2 and 4.2.3.

Corollary 4.2.4. *Let* $v \in C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ *and* $\varphi, \psi \in id + C_0^p(\mathbb{R}^d, \mathbb{R}^d)$*. If* $\|\varphi - id\|_{p,\infty} \leq r$ *and* $\|\psi - id\|_{p,\infty} \leq r$*, then*

$$\|v \circ \varphi\|_{p,\infty} \leq C_r \|v\|_{p,\infty}$$

and

$$\|v \circ \varphi - v \circ \psi\|_{p,\infty} \leq C_r' \|v\|_{p+1,\infty} \|\varphi - \psi\|_{p,\infty},$$

where C_r and C_r' only depend on r for a fixed p.

4.3 Proofs

We first prove the following technical lemmas related to the strong measurability.

Lemma 4.3.1. *Suppose that $v : [0,T] \to C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ is strongly measurable. If $\varphi \in C([0,T], id + C_0^p(\mathbb{R}^d, \mathbb{R}^d))$, then $t \mapsto v(t) \circ \varphi(t)$ is strongly measurable from $[0,T]$ to $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$.*

Proof. Since v is strongly measurable, there is a sequence of simple functions $v_n : [0,T] \to C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ that converges to v at almost every $t \in [0,T]$. On the other hand, since φ is uniformly continuous on $[0,T]$, we can construct a sequence of simple functions $\varphi_n : [0,T] \to id + C_0^p(\mathbb{R}^d, \mathbb{R}^d)$ that converges to φ uniformly, for example, define the value of a simple function on an interval $[k\delta, (k+1)\,\delta]$ by the value $\varphi(k\delta)$ at the left endpoint. Note that $t \mapsto v_n(t) \circ \varphi_n(t)$ is a simple function from $[0,T]$ to $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$. Now we show that $v_n(t) \circ \varphi_n(t)$ converges to $v(t) \circ \varphi(t)$ for almost every $t \in [0,T]$. The uniform convergence of φ_n to φ implies for every $t \in [0,T]$ and n large enough

$$\|\varphi_n(t) - id\|_{p,\infty} \le \|\varphi(t) - id\|_{p,\infty} + 1 \le \sup_{t \in [0,T]} \|\varphi(t) - id\|_{p,\infty} + 1 =: r.$$

It follows from Corollary 4.2.4 that

$$\begin{aligned} &\|v_n(t) \circ \varphi_n(t) - v(t) \circ \varphi(t)\|_{p,\infty} \\ &\le \|v_n(t) \circ \varphi_n(t) - v(t) \circ \varphi_n(t)\|_{p,\infty} + \|v(t) \circ \varphi_n(t) - v(t) \circ \varphi(t)\|_{p,\infty} \\ &\le C_r \, \|v_n(t) - v(t)\|_{p,\infty} + C_r' \, \|v(t)\|_{p+1,\infty} \, \|\varphi_n(t) - \varphi(t)\|_{p,\infty} \to 0 \end{aligned}$$

for almost every $t \in [0,T]$. □

Lemma 4.3.2. *Let B and B' be Banach spaces, and let $F \in C([0,T], \mathscr{L}(B, B'))$. If $v : [0,T] \to B$ is strongly measurable, then $t \mapsto F(t)\, v(t)$ is strongly measurable from $[0,T]$ to B'.*

The proof is exactly the same as the proof in the previous lemma by replacing $(\varphi_n)_{n=1}^\infty$ with a sequence of simple functions $(F_n)_{n=1}^\infty$ that converges to F uniformly. We leave details to the reader. Finally, Pettis' theorem (Theorem 2.2.3) immediately gives the following lemma.

Lemma 4.3.3. *Let B be a separable and reflexive Banach space. Given $\mu : [0,T] \to B^*$, if $t \mapsto (\mu(t) \mid b)$ is Lebesgue measurable for all $b \in B$, then μ is strongly measurable.*

The estimates in Corollary 4.2.4 will be used frequently in the proofs without further reference to. We also recall that $V \hookrightarrow C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$, thus there exists $c_V > 0$ such that $\|v\|_{p+1,\infty} \leq c_V \|v\|_V$. We remind the reader that the value of generic constants C_a may change from equation to equation.

Some of our proofs are modularized by multiple claims. In this case, the end of the proof of a claim will be marked by $\boxslash$, while the end of the whole proof will be indicated by the symbol $\square$ as usual.

Now we start the proof of the existence and uniqueness of solutions.

Theorem 4.1.1. *Let $\mathcal{M} : \bigcup_{t \in [0,T]} C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d)) \times Y \to V$ be a model of deformation vector fields. Given $\theta : [0,T] \to Y$ defined almost everywhere, denote the deformation vector field associated to $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $t \leq T$, by*

$$v_\varphi(s) := \mathcal{M}(\varphi_{|[0,s]}, \theta(s)).$$

We consider the initial value problem

$$\varphi(t,x) = x + \int_0^t v_\varphi(s, \varphi(s,x))\, ds \quad \text{for all } x \in \mathbb{R}^d. \tag{4.1}$$

Suppose that $v_\varphi : [0,t] \to V$ is strongly measurable for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$. Moreover, suppose that $\varphi \mapsto v_\varphi$ is locally Lipschitz in the sense that given a fixed $\overline{\varphi} \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, there exist $r(\overline{\varphi}) > 0$ and $f_{\overline{\varphi}} \in L^1([0,T])$ such that

$$\|v_\varphi(s) - v_\psi(s)\|_V \leq f_{\overline{\varphi}}(s) \sup_{s' \in [0,s]} \|\varphi(s') - \psi(s')\|_{p,\infty}$$

for all $\varphi, \psi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ with $\sup_{s \in [0,t]} \|\varphi(s) - \overline{\varphi}(s)\|_{p,\infty} \leq r$ and $\sup_{s \in [0,t]} \|\psi(s) - \overline{\varphi}(s)\|_{p,\infty} \leq r$ and for almost every $s \in [0,t]$.

(i) If there exists $g \in L^1([0,T])$ such that the deformation vector field satisfies

$$\|v_\varphi(s)\|_V \leq g(s) \left(1 + \sup_{s' \in [0,s]} \|\varphi(s') - id\|_{p,\infty}\right) \tag{4.2}$$

for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ and for almost every $s \in [0,t]$, then the initial value problem (4.1) has a unique maximal solution either in $C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ for some $T' \leq T$ or in $C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$.

(ii) If there exists $g \in L^1([0,T])$ such that the deformation vector field satisfies

$$\|v_\varphi(s)\|_V \leq g(s) \left(1 + \sup_{s' \in [0,s]} \|\varphi(s') - id\|_{\infty}\right) \tag{4.3}$$

for all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ *and for almost every* $s \in [0,t]$*, then the initial value problem* (4.1) *has a unique solution in* $C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$.

Proof. (i) We are going to prove that there exists a unique $\varphi \in C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfying

$$\varphi(t) = id + \int_0^t v_\varphi(s) \circ \varphi(s)\, ds, \tag{4.6}$$

where the integral is the Bochner integral. We first check that the Bochner integral in (4.6) is well defined under the assumption (4.2). Given any $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, since $v_\varphi : [0,t] \to V$ is strongly measurable, the assumption $V \hookrightarrow C^{p+1}_0(\mathbb{R}^d, \mathbb{R}^d)$ and Lemma 4.3.1 imply that $s \mapsto v_\varphi(s) \circ \varphi(s)$ is strongly measurable from $[0,t]$ to $C^p_0(\mathbb{R}^d, \mathbb{R}^d)$. From Theorem 2.2.7, the Bochner integral is well defined if and only if $s \mapsto \|v_\varphi(s) \circ \varphi(s)\|_{p,\infty}$ is Lebesgue integrable on $[0,t]$. Indeed, we have

$$\begin{aligned}
\int_0^t \|v_\varphi(s) \circ \varphi(s)\|_{p,\infty}\, ds &\le \int_0^t C_{\varphi(s)}\ c_V\, \|v_\varphi(s)\|_V\, ds \\
&\le \int_0^t C_{\varphi(s)}\ c_V\ g(s) \left(1 + \sup_{s' \in [0,s]} \|\varphi(s') - id\|_{p,\infty}\right) ds \qquad \text{(by (4.2))} \\
&\le C_\varphi\ c_V \left(1 + \sup_{s \in [0,t]} \|\varphi(s) - id\|_{p,\infty}\right) \|g\|_{L^1} \ < \ \infty,
\end{aligned}$$

where C_φ is a polynomial in $\sup_{s \in [0,t]} \|\varphi(s) - id\|_{p,\infty} < \infty$.

The following claim shows that a unique solution to the formulation using the Bochner integral is equivalent to a unique solution to the original initial value problem. Hence we can focus on (4.6) afterwards.

Claim 4.3.4. *Under the assumption* (4.2)*, a function* $\varphi \in C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ *is a solution to the initial value problem*

$$\varphi(t) = id + \int_0^t v_\varphi(s) \circ \varphi(s)\, ds, \tag{4.6}$$

where the integral is the Bochner integral, if and only if φ *is a solution to the initial value problem*

$$\varphi(t,x) = x + \int_0^t v_\varphi(s, \varphi(s,x))\, ds \quad \text{for all } (t,x) \in [0,T') \times \mathbb{R}^d, \tag{4.1}$$

where the integral is the Lebesgue integral.

Proof. ($\Rightarrow$) Suppose that $\varphi \in C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfies (4.6). Since for all $x \in \mathbb{R}^d$ the evaluation operator $\delta_x : C^p_0(\mathbb{R}^d, \mathbb{R}^d) \to \mathbb{R}^d$ is bounded and linear, Theorem 2.2.8 implies

that for all $t \in [0, T')$ and $x \in \mathbb{R}^d$

$$\begin{aligned}\varphi(t,x) = \delta_x\big(\varphi(t)\big) &= \delta_x\bigg(id + \int_0^t v_\varphi(s) \circ \varphi(s)\, ds\bigg) \\ &= \delta_x\big(id\big) + \int_0^t \delta_x\Big(v_\varphi(s) \circ \varphi(s)\Big)\, ds \\ &= x + \int_0^t v_\varphi(s, \varphi(s,x))\, ds,\end{aligned}$$

which is (4.1).

($\Leftarrow$) Now suppose that $\varphi \in C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfies (4.1). The assumption (4.2) and the paragraph before the claim together show that the Bochner integral in

$$\widetilde{\varphi}(t) := id + \int_0^t v_\varphi(s) \circ \varphi(s)\, ds$$

is well defined. It follows, as in the forward direction, that for all $x \in \mathbb{R}^d$

$$\widetilde{\varphi}(t,x) = x + \int_0^t v_\varphi(s, \varphi(s,x))\, ds = \varphi(t,x),$$

that is, $\widetilde{\varphi}(t) \equiv \varphi(t)$. Thus we have

$$\varphi(t) = id + \int_0^t v_\varphi(s) \circ \varphi(s)\, ds,$$

which shows that φ is a solution of (4.6). ◪

To prove (4.6), we set the scheme to apply the Banach fixed point theorem. For any $\varphi \in C([0,t_0], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ with $\varphi(0) = id$ and $\psi \in C([t_0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ with $\psi(t_0) = id$, we define the extension $\varphi \oplus \psi \in C([0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ by

$$(\varphi \oplus \psi)(t) := \begin{cases} \varphi(t), & \text{if } t \in [0,t_0]\,; \\ \psi(t) \circ \varphi(t_0), & \text{if } t \in (t_0, t_0+\eta]\,. \end{cases}$$

Suppose that we have obtained a unique $\varphi_{t_0} \in C([0,t_0], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfying (4.6) up to time t_0, which is true right at the beginning with $\varphi_0(0) = id$. If we can show that there exist $\eta(\varphi_{t_0}) > 0$ and a unique $\varphi \in C([t_0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfying

$$\varphi(t) = id + \int_{t_0}^t v_{\varphi_{t_0} \oplus \varphi}(s) \circ \varphi(s)\, ds, \tag{4.7}$$

then since $w \mapsto w \circ \varphi_{t_0}(t_0)$ is in $\mathscr{L}(C^p_0(\mathbb{R}^d, \mathbb{R}^d), C^p_0(\mathbb{R}^d, \mathbb{R}^d))$, the extension $\varphi_{t_0} \oplus \varphi$ will satisfy (4.6) by Theorem 2.2.8, that is, $\varphi_{t_0} \oplus \varphi$ is the unique solution on $[0, t_0+\eta]$. Our proof for part (i) will be complete by extending the unique solution repeatedly to the maximal interval of existence.

We now apply the Banach fixed point theorem to prove the existence of $\eta(\varphi_{t_0}) > 0$ and the existence of a unique $\varphi \in C([t_0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfying (4.7). Denote $\xi_{t_0} := \varphi_{t_0}(t_0)$. Since $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ is open in $id + C^p_0(\mathbb{R}^d, \mathbb{R}^d)$ and $\varphi \mapsto v_\varphi$ is locally Lipschitz, by defining $\overline{\varphi} \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ as

$$\overline{\varphi}(t) := \begin{cases} \varphi_{t_0}(t), & \text{if } t \in [0, t_0]\,; \\ \xi_{t_0}, & \text{if } t \in (t_0, T]\,, \end{cases}$$

we deduce that there exist $r(\overline{\varphi}) > 0$ and $f_{\overline{\varphi}} \in L^1([0,T])$ such that $\overline{B}(id, r) = \{\xi' \in id + C^p_0(\mathbb{R}^d, \mathbb{R}^d) : \|\xi' - id\|_{p,\infty} \le r\}$ is contained in $\mathrm{Diff}^p_{id}(\mathbb{R}^d)$ and such that

$$\begin{aligned} \|v_{\varphi_{t_0} \oplus \varphi}(t) - v_{\varphi_{t_0} \oplus \psi}(t)\|_V &\le f_{\overline{\varphi}}(t) \sup_{s \in [0,t]} \|(\varphi_{t_0} \oplus \varphi)(s) - (\varphi_{t_0} \oplus \psi)(s)\|_{p,\infty} \\ &= f_{\overline{\varphi}}(t) \sup_{s \in [t_0,t]} \|\varphi(s) \circ \xi_{t_0} - \psi(s) \circ \xi_{t_0}\|_{p,\infty} \end{aligned} \tag{4.8}$$

for all $\varphi, \psi \in C([t_0, t_0+\eta], \overline{B}(id, r))$ with $\varphi(t_0) = \psi(t_0) = id$ and for almost every $t \in [0, t_0+\eta]$. Note that the dependency of $r(\overline{\varphi})$ and $f_{\overline{\varphi}}$ on $\overline{\varphi}$ is through φ_{t_0}, so $r(\overline{\varphi})$ and $f_{\overline{\varphi}}$ in fact only depend on φ_{t_0}. Hence we will write $r(\varphi_{t_0})$ and $f_{\varphi_{t_0}}$ instead. With φ_{t_0} and r fixed, we define the iterate mapping $\Gamma : C([t_0, t_0+\eta], \overline{B}(id, r)) \to C([t_0, t_0+\eta], id + C^p_0(\mathbb{R}^d, \mathbb{R}^d))$ by

$$\Gamma(\varphi)(t) := id + \int_{t_0}^t v_{\varphi_{t_0} \oplus \varphi}(s) \circ \varphi(s)\, ds. \tag{4.9}$$

The Bochner integral is well defined as shown in the beginning of the proof. The domain $C([t_0, t_0+\eta], \overline{B}(id, r))$ and the codomain $C([t_0, t_0+\eta], id + C^p_0(\mathbb{R}^d, \mathbb{R}^d))$ of Γ are both equipped with the metric

$$d(\varphi, \psi) := \sup_{t \in [t_0, t_0+\eta]} \|\varphi(t) - \psi(t)\|_{p,\infty},$$

which renders the domain a complete metric space. We are going to choose an $\eta(\varphi_{t_0}) > 0$ such that Γ is a well-defined contraction, then the Banach fixed point theorem will imply the existence of a unique fixed point $\varphi \in C([t_0, t_0+\eta], \overline{B}(id, r)) \subset C([t_0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ that satisfies (4.7).

We first find a sufficient condition on η such that the range of the iterate mapping Γ is still in its domain. Theorem 2.2.9 and the assumption (4.2) yield

$$\begin{aligned} &\|\Gamma(\varphi)(t) - id\|_{p,\infty} \\ &\le \int_{t_0}^{t_0+\eta} \|v_{\varphi_{t_0} \oplus \varphi}(t) \circ \varphi(t)\|_{p,\infty}\, dt \end{aligned}$$

$$
\begin{aligned}
&\le \int_{t_0}^{t_0+\eta} C_r\, c_V\, \|v_{\varphi_{t_0}\oplus\varphi}(t)\|_V\, dt \\
&\le \int_{t_0}^{t_0+\eta} C_r\, c_V\ g(t) \Big(1 + \sup_{s\,\in\,[0,\,t_0]} \|\varphi_{t_0}(s) - id\|_{p,\infty} + \sup_{s\,\in\,[t_0,\,t]} \|\varphi(s)\circ\xi_{t_0} - id\|_{p,\infty}\Big)\, dt \\
&\le \int_{t_0}^{t_0+\eta} C_r\, c_V\ g(t) \Big(1 + \sup_{s\,\in\,[0,\,t_0]} \|\varphi_{t_0}(s) - id\|_{p,\infty} \\
&\qquad\qquad + \sup_{s\,\in\,[t_0,\,t]} \|(\varphi(s) - id)\circ\xi_{t_0}\|_{p,\infty} + \|\xi_{t_0} - id\|_{p,\infty}\Big)\, dt \\
&\le C_r\, c_V \left(1 + \sup_{s\,\in\,[0,\,t_0]} \|\varphi_{t_0}(s) - id\|_{p,\infty} + C_{\xi_{t_0}}\, r + \|\xi_{t_0} - id\|_{p,\infty}\right) \int_{t_0}^{t_0+\eta} g(t)\, dt \\
&\le C_{\varphi_{t_0}} \int_{t_0}^{t_0+\eta} g(t)\, dt,
\end{aligned}
$$

where we have used the dependency of r on φ_{t_0} in the last step. Thus a sufficient condition such that $\Gamma(\varphi) \in C([t_0, t_0+\eta], \overline{B}(id, r))$ is

$$
\int_{t_0}^{t_0+\eta} g(t)\, dt \;\le\; \frac{r(\varphi_{t_0})}{C_{\varphi_{t_0}}}. \tag{4.10}
$$

Our next step is to find a sufficient condition on η such that the iterate mapping Γ is a contraction. For all $\varphi, \psi \in C([t_0, t_0+\eta], \overline{B}(id, r))$, we have

$$
\begin{aligned}
d(\Gamma(\varphi), \Gamma(\psi)) &= \sup_{t\,\in\,[t_0,\,t_0+\eta]} \|\Gamma(\varphi)(t) - \Gamma(\psi)(t)\|_{p,\infty} \\
&\le \int_{t_0}^{t_0+\eta} \|v_{\varphi_{t_0}\oplus\varphi}(t)\circ\varphi(t) - v_{\varphi_{t_0}\oplus\psi}(t)\circ\psi(t)\|_{p,\infty}\, dt \\
&\le \int_{t_0}^{t_0+\eta} \Big(\|v_{\varphi_{t_0}\oplus\varphi}(t)\circ\varphi(t) - v_{\varphi_{t_0}\oplus\psi}(t)\circ\varphi(t)\|_{p,\infty} \\
&\qquad + \|v_{\varphi_{t_0}\oplus\psi}(t)\circ\varphi(t) - v_{\varphi_{t_0}\oplus\psi}(t)\circ\psi(t)\|_{p,\infty}\Big)\, dt \\
&\le \int_{t_0}^{t_0+\eta} \Big(C_r\, \|v_{\varphi_{t_0}\oplus\varphi}(t) - v_{\varphi_{t_0}\oplus\psi}(t)\|_{p,\infty} \\
&\qquad + C'_r\, \|v_{\varphi_{t_0}\oplus\psi}(t)\|_{p+1,\infty}\, \|\varphi(t) - \psi(t)\|_{p,\infty}\Big)\, dt.
\end{aligned}
$$

We use (4.8) for the first term and (4.2) for the second term and continue the above inequality:

$$
\begin{aligned}
&d(\Gamma(\varphi), \Gamma(\psi)) \\
&\le C_r \int_{t_0}^{t_0+\eta} \Big(c_V\ f_{\varphi_{t_0}}(t) \sup_{s\,\in\,[t_0,\,t]} \|\varphi(s)\circ\xi_{t_0} - \psi(s)\circ\xi_{t_0}\|_{p,\infty} \\
&\qquad + c_V\ g(t) \Big(1 + \sup_{s\,\in\,[0,\,t_0]} \|\varphi_{t_0}(s) - id\|_{p,\infty}
\end{aligned}
$$

$$+ \sup_{s \in [t_0,\, t]} \|\varphi(s) \circ \xi_{t_0} - id\|_{p,\infty}\Big)\ \|\varphi(t) - \psi(t)\|_{p,\infty}\Big)\, dt$$

$$\leq \left(C'_{\varphi_{t_0}} \int_{t_0}^{t_0+\eta} \big(f_{\varphi_{t_0}}(t) + g(t)\big)\, dt \right) d(\varphi, \psi).$$

It follows that a sufficient condition ensuring Γ is a contraction is

$$C'_{\varphi_{t_0}} \int_{t_0}^{t_0+\eta} \big(f_{\varphi_{t_0}}(t) + g(t)\big)\, dt \ < \ 1. \tag{4.11}$$

We conclude from the sufficient conditions (4.10) and (4.11) that for any $\eta > 0$ satisfies

$$\int_{t_0}^{t_0+\eta} \big(f_{\varphi_{t_0}}(t) + g(t)\big)\, dt \ < \ \min\left\{ \frac{r(\varphi_{t_0})}{C_{\varphi_{t_0}}},\ \frac{1}{C'_{\varphi_{t_0}}} \right\},$$

there exists a unique $\varphi \in C([t_0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ which solves

$$\varphi(t) = id + \int_{t_0}^{t} v_{\varphi_{t_0} \oplus \varphi}(s) \circ \varphi(s)\, ds. \tag{4.7}$$

Our proof of part (i) is complete.

(ii) Since the condition (4.3) of part (ii) implies the condition (4.2) of part (i), there exists a solution either on $[0, T')$ for some $T' \leq T$ or on $[0, T]$. Suppose on the contrary that we only have a solution $\varphi \in C([0, T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$. We are going to reach a contradiction by showing that $\lim_{t \uparrow T'} \varphi(t) \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$, thus $[0, T')$ is not the maximal interval of existence. To this end, we first show that the deformation vector field of the solution φ is integrable given the condition (4.3). One technical detail is that $t \mapsto \|v_\varphi(t)\|_{p+1,\infty}$ is Lebesgue measurable due to the assumption of the strong measurability of v_φ and Corollary 2.2.4.

Claim 4.3.5. *The deformation vector field $v_\varphi : [0, T') \to C^{p+1}_0(\mathbb{R}^d, \mathbb{R}^d)$ of the solution φ is integrable, i.e., $\int_0^{T'} \|v_\varphi(t)\|_{p+1,\infty}\, dt < \infty$.*

Proof. The solution φ satisfies

$$\varphi(t, x) = x + \int_0^t v_\varphi(s, \varphi(s, x))\, ds \quad \text{ for all } (t, x) \in [0, T') \times \mathbb{R}^d. \tag{4.1}$$

For every $t' \in [0, t]$, the equation (4.1) and the condition (4.3) then give

$$\|\varphi(t') - id\|_\infty \leq \int_0^t \|v_\varphi(s)\|_\infty\, ds \leq \int_0^t c_V\ g(s) \left(1 + \sup_{s' \in [0,\, s]} \|\varphi(s') - id\|_\infty\right) ds,$$

which, by taking the supremum over $[0, t]$ and Grönwall's lemma, implies

$$\sup_{t' \in [0,\, t]} \|\varphi(t') - id\|_\infty \leq c_V\, \|g\|_{L^1}\, \exp\Big(c_V \|g\|_{L^1}\Big) \quad \text{ for all } t \in [0, T'). \tag{4.12}$$

Combining the condition (4.3) and the inequality (4.12), the deformation vector field of the solution is bounded by

$$\begin{aligned}\int_0^{T'} \|v_\varphi(s)\|_{p+1,\infty}\, ds &\le \int_0^{T'} c_V\ g(s)\left(1 + c_V\,\|g\|_{L^1}\,\exp\Big(c_V\|g\|_{L^1}\Big)\right) ds \\ &\le C\,\|g\|_{L^1} < \infty. \end{aligned}$$

◪

Now we fix v_φ and consider the initial value problem

$$\psi(t) = id + \int_0^t v_\varphi(s) \circ \psi(s)\, ds. \tag{4.13}$$

Since $\int_0^{T'} \|v_\varphi(t)\|_{p+1,\infty}\, dt < \infty$ by Claim 4.3.5, it follows from Theorem A.2.3 that (4.13) has a unique solution $\psi \in C([0,T'], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$. We observe that $\varphi \in C([0,T'), \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ is also a solution of (4.13). By the uniqueness and continuity, we know that $\lim_{t\uparrow T'} \varphi(t) = \lim_{t\uparrow T'} \psi(t) = \psi(T') \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$. In other words, the interval $[0,T')$ is not the maximal interval of existence, a contradiction. □

Under the conditions of Theorem 4.1.1(ii), we can extend Claim 4.3.5 and further show that the unique solution and its inverse is uniformly bounded in time and that $D^p\varphi(t)$ is Lipschitz continuous. These properties will be used in the proof of Theorem 4.1.10.

Proposition 4.3.6. *Suppose that all assumptions of Theorem 4.1.1(ii) hold. Given $\theta : [0,T] \to Y$, there exist $r(\theta)$ and $\ell(\theta)$ such that the unique solution $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ satisfies*

$$\|\varphi(t) - id\|_{p,\infty} \le r \quad \text{and} \quad |D^p\varphi(t,x) - D^p\varphi(t,y)| \le \ell\ |x-y|$$

for all $t \in [0,T]$ and all $x, y \in \mathbb{R}^d$. Moreover, we also have

$$\|\varphi(t)^{-1} - id\|_{p,\infty} \le r \quad \text{for all } t \in [0,T].$$

Proof. We have proved in Claim 4.3.5 that

$$\max\left\{\|\varphi(t) - id\|_\infty, \int_0^T \|v_\varphi(s)\|_{p+1,\infty}\, ds\right\} \le C.$$

Now we bound $\|D\varphi(t) - I_d\|_\infty$ and $\|D^n\varphi(t)\|_\infty$, $2 \le n \le p$, successively. Since φ is p-times differentiable in space, the solution φ satisfies

$$D\varphi(t,x) = I_d + \int_0^t D\big(v_\varphi(s,\varphi(s,x))\big)\, ds \quad \text{for all } (t,x) \in [0,T]\times\mathbb{R}^d, \tag{4.14}$$

where I_d is the d-by-d identity matrix, and for $2 \le n \le p$

$$|D^n\varphi(t,x)| \le \int_0^t \left|D^n\big(v_\varphi(s,\varphi(s,x))\big)\right| ds \quad \text{for all } (t,x) \in [0,T]\times\mathbb{R}^d. \tag{4.15}$$

From (4.14) we have

$$|D\varphi(t,x) - I_d| \le \int_0^t \|Dv_\varphi(s)\|_\infty \Big(\|D\varphi(s) - I_d\|_\infty + 1\Big)\, ds.$$

Grönwall's lemma and $\int_0^T \|v_\varphi(s)\|_{p+1,\infty}\, ds \le C$ then gives $\|D\varphi(t) - I_d\|_\infty \le C'$. It follows that $\|D\varphi(t)\|_\infty \le C' + 1$. For $\|D^n\varphi(t)\|_\infty$, $2 \le n \le p$, we proceed by induction. Suppose that

$$\max\Big\{\|D\varphi(t)\|_\infty, \ldots, \|D^{n-1}\varphi(t)\|_\infty\Big\} \le C'',$$

then the inequality (4.15), Faà di Bruno's formula (Lemma 4.2.1), the bound for the integral of vector field, and the induction hypothesis yield

$$|D^n\varphi(t,x)| \le \widetilde{C} + \int_0^t \|Dv_\varphi(s)\|_\infty \, \|D^n\varphi(s)\|_\infty \, ds.$$

Hence we obtain $\|D^n\varphi(t)\|_\infty \le \widetilde{C}'$ again by Grönwall's lemma. This completes the proof of the first part that there exists $r(\theta) > 0$ such that $\|\varphi(t) - id\|_{p,\infty} \le r$.

For the second part on the Lipschitz continuity, since

$$\max\Big\{\|D\varphi(t)\|_\infty, \ldots, \|D^p\varphi(t)\|_\infty, \int_0^T \|v_\varphi(s)\|_{p+1,\infty}\, ds\Big\} \le C,$$

Faà di Bruno's formula (Lemma 4.2.1) then implies

$$\begin{aligned}
|D^p\varphi(t,x) - D^p\varphi(t,y)| &\le \int_0^t \left|D^p\big(v_\varphi(s,\varphi(s,x)) - D^p\big(v_\varphi(s,\varphi(s,y))\big)\right| ds \\
&\le C'\,|x-y| + \int_0^t \|Dv_\varphi(s)\|_\infty \, |D^p\varphi(t,x) - D^p\varphi(t,y)|\, ds.
\end{aligned}$$

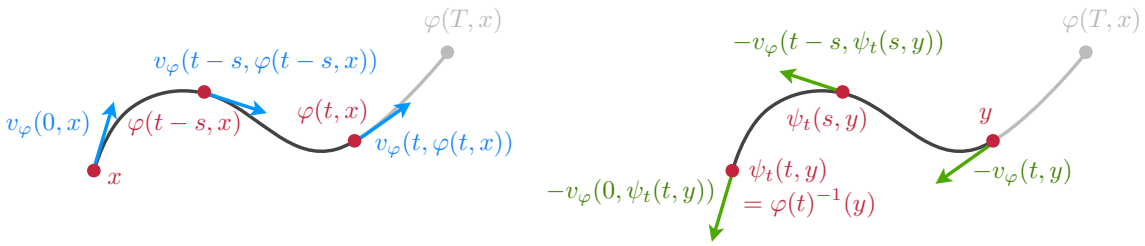
Thus, by Grönwall's lemma, we have $|D^p\varphi(t,x) - D^p\varphi(t,y)| \le \ell\,|x-y|$ for some $\ell(\theta) > 0$.

To bound the inverse, we fix a time $t \in [0,T]$ and consider the initial value problem

$$\psi(s,y) = y + \int_0^s \Big(-v_\varphi\big(t-s', \psi(s',y)\big)\Big)\, ds'$$

on the time interval $[0,t]$. Since $\int_0^t \|-v_\varphi(t-s')\|_{p+1,\infty}\, ds' = \int_0^t \|v_\varphi(s)\|_{p+1,\infty}\, ds < \infty$, Theorem A.2.3 shows that there exists a unique solution $\psi_t \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$. Following the same analysis as in Claim 4.3.5 and above, we then deduce that $\|\psi_t(s) - id\|_{p,\infty} \le r$ for all $s \in [0,t]$. Observe from Figure 4.1 that $\psi_t(t) = \varphi(t)^{-1}$. By varying $t \in [0,T]$, we acquire a family of solutions $\psi_t \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ to the corresponding initial value problems

and conclude that $\|\varphi(t)^{-1} - id\|_{p,\infty} = \|\psi_t(t) - id\|_{p,\infty} \leq r$ for all $t \in [0,T]$. □



(a) Trajectory of v_φ from time 0 to t.

(b) Trajectory of $-v_\varphi$ from time t to 0.

Figure 4.1: Illustration of $\psi_t(t) = \varphi(t)^{-1}$.

For completeness, we present the standard direct method of calculus of variations to prove the existence of minimizers.

Proposition 4.1.6. *Let Θ be a weakly closed subset of a reflexive Banach space and $\mathcal{M}$ be a model of deformation vector fields. Suppose that Θ and $\mathcal{M}$ are compatible. Given $\Omega_0, \Omega_{\text{targ}} \in \mathscr{S}$, we denote the unique solution to the initial value problem (4.1) corresponding to $\theta \in \Theta$ by φ_θ and consider the minimization problem*

$$\min_{\theta \in \Theta} \left(\int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt + \rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}) \right). \tag{4.4}$$

Assume that:

(1) Either the objective function is coercive or Θ is bounded.

(2) The function $\theta \mapsto \int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt$ is weakly sequentially lower semicontinuous.

(3) The discrepancy function ρ is continuous with respect to $\|\cdot\|_{p,\infty}$.

Then the minimization problem (4.4) has a minimizer.

Proof. Denote the objective function by

$$J(\theta) := \int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt + \rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}).$$

Let $(\theta_n)_{n=1}^\infty \subset \Theta$ be a minimizing sequence, i.e., $J(\theta_n) \to \inf_{\theta' \in \Theta} J(\theta')$. The assumption (1) implies that the minimizing sequence $(\theta_n)_{n=1}^\infty$ is bounded. Since Θ is a weakly closed subset of a reflexive Banach space, we can extract a subsequence, still denoted by $(\theta_n)_{n=1}^\infty$, such that $\theta_n \rightharpoonup \theta \in \Theta$. The assumption (2) leads to

$$\int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt \leq \liminf_{n \to \infty} \int_0^T \Lambda(\varphi_{\theta_n}(t), \theta_n(t))\, dt,$$

while the assumption (3) and $\|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^{\Omega} \to 0$ from the compatibility assumption together give

$$\rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}) = \lim_{n\to\infty} \rho(\varphi_{\theta_n}(T, \Omega_0), \Omega_{\text{targ}}).$$

It follows that

$$\begin{aligned} J(\theta) &= \int_0^T \Lambda(\varphi_\theta(t), \theta(t))\, dt + \rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}) \\ &\leq \liminf_{n\to\infty} \int_0^T \Lambda(\varphi_{\theta_n}(t), \theta_n(t))\, dt + \lim_{n\to\infty} \rho(\varphi_{\theta_n}(T, \Omega_0), \Omega_{\text{targ}}) \\ &= \lim_{n\to\infty} J(\theta_n) = \inf_{\theta' \in \Theta} J(\theta'), \end{aligned}$$

which shows that $J(\theta) = \inf_{\theta' \in \Theta} J(\theta')$, that is, $\theta \in \Theta$ is a minimizer. □

Now we prove the sufficient conditions on $\mathcal{M}$ such that $\mathcal{M}$ and Θ are compatible. We recall

$$\overline{\mathfrak{B}}(id, r) = \{\xi \in \text{Diff}_{id}^p(\mathbb{R}^d) : \|\xi - id\|_{p,\infty} \leq r \text{ and } \|\xi^{-1} - id\|_{p,\infty} \leq r\}$$

and the definition of the seminorm

$$\|v\|_{p,\infty}^{\Omega} = \sum_{j=0}^{p} \max_{x \in \Omega} |D^j v(x)|.$$

Theorem 4.1.10. *Let Θ be a weakly closed subset of a reflexive Banach space B and $\mathcal{M}$ be a model of deformation vector fields. For all $\Omega \in \mathscr{S}$, suppose that:*

- *The mapping $s \mapsto \mathcal{M}(\varphi_{|[0,\,s]}, \theta(s))$ is strongly measurable for all $\varphi \in C([0,t], \text{Diff}_{id}^p(\mathbb{R}^d))$ and $\theta \in \Theta$.*
- *The model is bounded:*
  For all $\theta \in \Theta$, there exists $f_\theta \in L^2([0,T])$ such that

  $$\|\mathcal{M}(\varphi_{|[0,\,s]}, \theta(s))\|_V \leq f_\theta(s) \left(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_\infty\right)$$

  for all $\varphi \in C([0,t], \text{Diff}_{id}^p(\mathbb{R}^d))$ and for almost every $s \in [0,t]$. Moreover, for all $m > 0$, there exists a constant $F_m > 0$ such that $\|\theta\|_B \leq m$ implies $\|f_\theta\|_{L^2} \leq F_m$.
- *The model is Lipschitz in φ:*
  For all $r > 0$ and $\theta \in \Theta$, there exists $g_{r,\theta} \in L^2([0,T])$ such that

  $$\|\mathcal{M}(\varphi_{|[0,\,s]}, \theta(s)) - \mathcal{M}(\psi_{|[0,\,s]}, \theta(s))\|_V \leq g_{r,\theta}(s) \sup_{s' \in [0,\,s]} \|\varphi(s') - \psi(s')\|_{p,\infty}^{\Omega}$$

  for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ and for almost every $s \in [0,t]$. Moreover, for all $m > 0$,

there exists a constant $G_{r,m} > 0$ such that $\|\theta\|_B \le m$ implies $\|g_{r,\theta}\|_{L^2} \le G_{r,m}$.

- *The model is continuous in θ:*

  If $(\theta_n)_{n=1}^{\infty} \subset \Theta$ and $\theta_n \rightharpoonup \theta$, then for $0 \le j \le p$

  $$\int_0^t D^j\Big(\mathcal{M}(\varphi|_{[0,\,s]}, \theta_n(s))(\varphi(s,x))\Big)\,ds \to \int_0^t D^j\Big(\mathcal{M}(\varphi|_{[0,\,s]}, \theta(s))(\varphi(s,x))\Big)\,ds$$

  for all $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $t \in [0,T]$, and $x \in \Omega$.

Then Θ and $\mathcal{M}$ are compatible.

Proof. It is clear that the assumptions are sufficient to invoke Theorem 4.1.1(ii) for each $\theta \in \Theta$, thus the initial value problem (4.1) has a unique solution in $C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ for all $\theta \in \Theta$. Let $(\theta_n)_{n=1}^{\infty} \subset \Theta$ and $\theta_n \rightharpoonup \theta$. Now we prove that

$$\|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|^{\Omega}_{p,\infty} \to 0 \quad \text{ for all } t \in [0,T].$$

Since $\theta_n \rightharpoonup \theta$, there exists $m > 0$ such that $\|\theta_n\|_B \le m$ for all n and $\|\theta\|_B \le m$. Proposition 4.3.6 then implies that there exists $r(m) > 0$ such that $\varphi_{\theta_n}, \varphi_\theta \in C([0,T], \overline{\mathfrak{B}}(id, r))$ for all n and that there exists $\ell(\theta) > 0$ such that

$$|D^p\varphi_\theta(t,x) - D^p\varphi_\theta(t,y)| \le \ell(\theta)\,|x-y| \tag{4.16}$$

for all $t \in [0,T]$ and all $x, y \in \mathbb{R}^d$. Let f_θ and $g_{r,\theta}$ be L^2 functions that satisfy the assumptions. For a fixed $t \in [0,T]$, note that

$$\begin{aligned}
&\varphi_{\theta_n}(t,x) - \varphi_\theta(t,x)\\
&= \int_0^t \Big(\mathcal{M}(\varphi_{\theta_n}|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x)) - \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta(s))(\varphi_\theta(s,x))\Big)\,ds\\
&= \int_0^t \Big(\mathcal{M}(\varphi_{\theta_n}|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x)) - \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x))\Big)\,ds\\
&\quad + \int_0^t \Big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x)) - \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_\theta(s,x))\Big)\,ds\\
&\quad + \int_0^t \Big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_\theta(s,x)) - \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta(s))(\varphi_\theta(s,x))\Big)\,ds\\
&=: I_{1,n}(t,x) + I_{2,n}(t,x) + I_{3,n}(t,x).
\end{aligned}$$

Taking $\|\cdot\|^{\Omega}_{p,\infty}$ on both sides, we are going to show that

$$\|I_{1,n}(t)\|^{\Omega}_{p,\infty} + \|I_{2,n}(t)\|^{\Omega}_{p,\infty} \le C_m \left(\int_0^t \Big(\sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|^{\Omega}_{p,\infty}\Big)^2 ds\right)^{\frac{1}{2}}, \tag{4.17}$$

and that

$$\lim_{n\to\infty} \|I_{3,n}(t)\|_{p,\infty}^{\Omega} = 0 \quad \text{for all } t \in [0,T]. \tag{4.18}$$

Identities (4.17) and (4.18) will then lead to $\|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^{\Omega} \to 0$ for all $t \in [0,T]$.

We estimate $\|I_{1,n}(t)\|_{p,\infty}^{\Omega}$ as follows.

$$\begin{aligned}
&\|I_{1,n}(t)\|_{p,\infty}^{\Omega} \\
&= \left\| \int_0^t \Big(\mathcal{M}(\varphi_{\theta_n}|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x)) - \mathcal{M}(\varphi_{\theta}|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x)) \Big)\, ds \right\|_{p,\infty}^{\Omega} \\
&\le \int_0^t \left\| \mathcal{M}(\varphi_{\theta_n}|_{[0,\,s]}, \theta_n(s)) \circ \varphi_{\theta_n}(s) - \mathcal{M}(\varphi_{\theta}|_{[0,\,s]}, \theta_n(s)) \circ \varphi_{\theta_n}(s) \right\|_{p,\infty} ds \\
&\le \int_0^t C_r\ c_V\ g_{r,\theta_n}(s) \sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|_{p,\infty}^{\Omega}\ ds \\
&\le C_r\ c_V\ \|g_{r,\theta_n}\|_{L^2} \left(\int_0^t \Big(\sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|_{p,\infty}^{\Omega} \Big)^2 ds \right)^{\frac{1}{2}} \\
&\le C_m \left(\int_0^t \Big(\sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|_{p,\infty}^{\Omega} \Big)^2 ds \right)^{\frac{1}{2}},
\end{aligned}$$

where we have used the assumption that $\|\theta_n\|_B \le m$ implies $\|g_{r,\theta_n}\|_{L^2} \le G_{r,m}$ and the dependency of r on m in the last step.

For $\|I_{2,n}(t)\|_{p,\infty}^{\Omega}$, a straightforward adaptation of Corollay 4.2.4 with the seminorm $\|\cdot\|_{p,\infty}^{\Omega}$ gives

$$\|v \circ \varphi - v \circ \psi\|_{p,\infty}^{\Omega} \le C_r\, \|v\|_{p+1,\infty}\, \|\varphi - \psi\|_{p,\infty}^{\Omega}.$$

It follows that

$$\begin{aligned}
&\|I_{2,n}(t)\|_{p,\infty}^{\Omega} \\
&= \left\| \int_0^t \Big(\mathcal{M}(\varphi_{\theta}|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta_n}(s,x)) - \mathcal{M}(\varphi_{\theta}|_{[0,\,s]}, \theta_n(s))(\varphi_{\theta}(s,x)) \Big)\, ds \right\|_{p,\infty}^{\Omega} \\
&\le \int_0^t \left\| \mathcal{M}(\varphi_{\theta}|_{[0,\,s]}, \theta_n(s)) \circ \varphi_{\theta_n}(s) - \mathcal{M}(\varphi_{\theta}|_{[0,\,s]}, \theta_n(s)) \circ \varphi_{\theta}(s) \right\|_{p,\infty}^{\Omega} ds \\
&\le \int_0^t C_r\ c_V\ f_{\theta_n}(s) \Big(1 + \sup_{s' \in [0,\,s]} \|\varphi_\theta(s') - id\|_\infty \Big) \sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|_{p,\infty}^{\Omega}\ ds \\
&\le C_r\ c_V\ (1+r)\ \|f_{\theta_n}\|_{L^2} \left(\int_0^t \Big(\sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|_{p,\infty}^{\Omega} \Big)^2 ds \right)^{\frac{1}{2}} \\
&\le C_m \left(\int_0^t \Big(\sup_{s' \in [0,\,s]} \|\varphi_{\theta_n}(s') - \varphi_\theta(s')\|_{p,\infty}^{\Omega} \Big)^2 ds \right)^{\frac{1}{2}}.
\end{aligned}$$

We proceed to show that $\lim_{n\to\infty} \|I_{3,n}(t)\|_{p,\infty}^{\Omega} = 0$ for all $t \in [0,T]$, that is,

$$\sum_{j=0}^{p} \max_{x\,\in\,\Omega} \left| \int_0^t D^j \Big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_\theta(s,x)) - \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta(s))(\varphi_\theta(s,x)) \Big)\, ds \right| \to 0.$$

Denote the first term by

$$u_n^{(j)}(x) := \int_0^t D^j \Big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_\theta(s,x)) \Big)\, ds$$

and the second term by

$$u^{(j)}(x) := \int_0^t D^j \Big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta(s))(\varphi_\theta(s,x)) \Big)\, ds,$$

then it is equivalent to show that $u_n^{(j)}(x) \to u^{(j)}(x)$ uniformly on the compact set Ω for each $0 \le j \le p$. The pointwise convergence is given by the assumption. We aim to prove the uniform boundedness and equicontinuity of those sequences, so as to invoke the Arzelà–Ascoli theorem. The sequences $(u_n^{(j)})_{n=1}^{\infty}$, $0 \le j \le p$, are uniformly bounded since

$$\sum_{j=0}^{p} |u_n^{(j)}(x)| \le \int_0^t C_r\, c_V \big\| \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s)) \big\|_V\, ds \le C_r\, c_V\, (1+r)\, (\sqrt{T} F_m).$$

Recall from (4.16) that $\ell(\theta)$ is a Lipschitz constant for $D^p\varphi_\theta$. It follows that

$$\widetilde{\ell}(\theta) := \max\{\|D\varphi_\theta\|_\infty, \dots, \|D^p\varphi_\theta\|_\infty, \ell(\theta)\}$$

is a Lipschitz constant for all $\varphi_\theta, D\varphi_\theta, \dots, D^p\varphi_\theta$. The equicontinuity of the sequences $(u_n^{(j)})_{n=1}^{\infty}$, $0 \le j \le p$, now follows from

$$\begin{aligned}
|u_n^{(j)}(x) - u_n^{(j)}(y)| \le \int_0^t \Bigg(& \Big| \Big(D^j \big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s)) \big)(\varphi_\theta(s,x)) \Big) \Big(D^j \varphi_\theta(s,x) \Big) \\
& \quad - \Big(D^j \big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s)) \big)(\varphi_\theta(s,y)) \Big) \Big(D^j \varphi_\theta(s,x) \Big) \Big| \\
& \quad + \Big| \Big(D^j \big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s)) \big)(\varphi_\theta(s,y)) \Big) \Big(D^j \varphi_\theta(s,x) \Big) \\
& \quad - \Big(D^j \big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s)) \big)(\varphi_\theta(s,y)) \Big) \Big(D^j \varphi_\theta(s,y) \Big) \Big| \Bigg)\, ds \\
\le \int_0^t & \Big(c_V\ f_{\theta_n}(s)\, (1+r)\ \widetilde{\ell}\, |x-y|\ \ (1+r) \\
& \quad + c_V\ f_{\theta_n}(s)\, (1+r)\ \widetilde{\ell}\, |x-y| \Big)\, ds \\
\le\ & C_m\, \widetilde{\ell}\, |x-y|.
\end{aligned}$$

From the Arzelà–Ascoli theorem, we know that every subsequence of $u_n^{(j)}(x)$ has a further

subsequence that converges uniformly to $u^{(j)}(x)$ on the compact set Ω for each $0 \le j \le p$, which shows that $\|u_n^{(j)} - u^{(j)}\|_\infty^\Omega \to 0$ for each $0 \le j \le p$. In other words, we have proved $\lim_{n\to\infty} \|I_{3,n}(t)\|_{p,\infty}^\Omega = 0$ for all $t \in [0,T]$.

In summary, the proved identity (4.17) leads to

$$\begin{aligned} \|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^\Omega &\le \|I_{1,n}(t)\|_{p,\infty}^\Omega + \|I_{2,n}(t)\|_{p,\infty}^\Omega + \|I_{3,n}(t)\|_{p,\infty}^\Omega \\ &\le \|I_{3,n}(t)\|_{p,\infty}^\Omega + C_m \left(\int_0^t \left(\|\varphi_{\theta_n}(s) - \varphi_\theta(s)\|_{p,\infty}^\Omega \right)^2 ds \right)^{\frac{1}{2}}. \end{aligned} \tag{4.19}$$

Squaring (4.19) and using the inequality $(a+b)^2 \le 2\,(a^2+b^2)$, we get

$$\left(\|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^\Omega \right)^2 \le 2 \left(\|I_{3,n}(t)\|_{p,\infty}^\Omega \right)^2 + \int_0^t 2\, C_m^2 \left(\|\varphi_{\theta_n}(s) - \varphi_\theta(s)\|_{p,\infty}^\Omega \right)^2 ds.$$

By Grönwall's lemma, we finally obtain

$$\begin{aligned} &\left(\|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^\Omega \right)^2 \\ &\le 2 \left(\|I_{3,n}(t)\|_{p,\infty}^\Omega \right)^2 + \int_0^t 4 \left(\|I_{3,n}(s)\|_{p,\infty}^\Omega \right)^2 C_m^2 \, \exp\!\big(2\, C_m^2 (t-s)\big) \, ds. \end{aligned} \tag{4.20}$$

Note that

$$\begin{aligned} &\|I_{3,n}(t)\|_{p,\infty}^\Omega \\ &= \left\| \int_0^t \Big(\mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_\theta(s,x)) - \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta(s))(\varphi_\theta(s,x)) \Big)\, ds \right\|_{p,\infty}^\Omega \\ &\le \int_0^T \left(\left\| \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta_n(s))(\varphi_\theta(s,x)) \right\|_{p,\infty} + \left\| \mathcal{M}(\varphi_\theta|_{[0,\,s]}, \theta(s))(\varphi_\theta(s,x)) \right\|_{p,\infty} \right) ds \\ &\le \int_0^T C_r\, c_V \left(f_{\theta_n}(s) + f_\theta(s) \right) (1+r)\, ds \;\le\; C_m, \end{aligned}$$

so $\|I_{3,n}(t)\|_{p,\infty}^\Omega$ is uniformly bounded in n and t. The proved identity (4.18) and the dominated convergence theorem then show that the right-hand side of (4.20) goes to 0 as $n \to \infty$ and thus

$$\|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^\Omega \to 0 \quad \text{ for all } t \in [0,T],$$

which completes the proof. □

We proceed to prove that Θ and $\mathcal{M}$ are compatible when $\mathcal{M}$ is of the energy form (4.5). First we prove a key lemma.

Lemma 4.1.13. *For every nonnegative $\mathcal{L} \in \mathscr{L}(V, V^*)$, i.e., $(\mathcal{L}v \mid v) \ge 0$ for all $v \in V$, we*

have $(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L})^{-1}\in\mathscr{L}(V^*,V)$ *and*

$$\left\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L})^{-1}\right\|_{\mathscr{L}(V^*,V)}\leq\frac{1}{\gamma}.$$

Proof. The linear operator $\gamma\,\mathcal{K}_V^{-1}+\mathcal{L}$ is invertible since the solution of $(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L})\,v=\mu$ for all $\mu\in V^*$ is characterized by the unique minimizer of a strictly convex function, that is,

$$v=\operatorname*{arg\,min}_{v'\in V}\Big(\frac{\gamma}{2}\,\|v'\|_V^2+\frac{1}{2}\,(\mathcal{L}v'\mid v')-(\mu\mid v')\Big).$$

Next we show that

$$\left\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L})^{-1}\,\mu\right\|_V\leq\frac{1}{\gamma}\,\|\mu\|_{V^*}\qquad\text{for all }\mu\in V^*,$$

which is equivalent to

$$\|v\|_V\leq\frac{1}{\gamma}\left\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L})\,v\right\|_{V^*}\qquad\text{for all }v\in V.$$

Denote the identity mapping on V by $id_V:V\to V$. We have

$$\begin{aligned}\left(\frac{1}{\gamma}\left\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L})\,v\right\|_{V^*}\right)^2&=\left(\frac{1}{\gamma}\left\|\mathcal{K}_V\left(\gamma\,\mathcal{K}_V^{-1}+\mathcal{L}\right)v\right\|_V\right)^2\\&=\frac{1}{\gamma^2}\,\|(\gamma\,id_V+\mathcal{K}_V\mathcal{L})\,v\|_V^2\\&=\|v\|_V^2+\frac{1}{\gamma^2}\,\|\mathcal{K}_V\mathcal{L}v\|_V^2+\frac{2}{\gamma}\,\langle v,\mathcal{K}_V\mathcal{L}v\rangle_V\\&=\|v\|_V^2+\frac{1}{\gamma^2}\,\|\mathcal{K}_V\mathcal{L}v\|_V^2+\frac{2}{\gamma}\,(\mathcal{L}v\mid v)\;\geq\;\|v\|_V^2,\end{aligned}$$

where the last inequality follows from $(\mathcal{L}v\mid v)\geq 0$. □

Corollary 4.1.12. *Let Θ be a weakly closed subset of a reflexive Banach space B. Suppose that the model of deformation vector fields $\mathcal{M}$ is of the energy form (4.5). For all $\Omega\in\mathscr{S}$, we also suppose that:*

- *The mapping $s\mapsto\mathcal{A}_\varphi^s$ is in $C([0,t],\mathscr{L}(V,V^*))$ for all $\varphi\in C([0,t],\mathrm{Diff}_{id}^p(\mathbb{R}^d))$. The mapping $s\mapsto(\beta^s_{\varphi,\theta(s)}\mid v)$ is Lebesgue measurable for all $\varphi\in C([0,t],\mathrm{Diff}_{id}^p(\mathbb{R}^d))$, $\theta\in\Theta$, and $v\in V$.*
- *For all $\theta\in\Theta$, there exists $f_\theta\in L^2([0,T])$ such that*

$$\|\beta^s_{\varphi,\theta(s)}\|_{V^*}\leq f_\theta(s)\left(1+\sup_{s'\in[0,\,s]}\|\varphi(s')-id\|_\infty\right)$$

for all $\varphi\in C([0,t],\mathrm{Diff}_{id}^p(\mathbb{R}^d))$ and for almost every $s\in[0,t]$. Moreover, for all $m>0$,

there exists a constant $F_m > 0$ *such that* $\|\theta\|_B \leq m$ *implies* $\|f_\theta\|_{L^2} \leq F_m$.

- *For all* $r > 0$ *and* $\theta \in \Theta$, *there exist* $\ell_r > 0$ *and* $g_{r,\theta} \in L^2([0,T])$ *such that*

$$\|\mathcal{A}^s_\varphi - \mathcal{A}^s_\psi\|_{\mathscr{L}(V,V^*)} \leq \ell_r \sup_{s' \in [0,\,s]} \|\varphi(s') - \psi(s')\|^\Omega_{p,\infty}$$

and that

$$\|\beta^s_{\varphi,\,\theta(s)} - \beta^s_{\psi,\,\theta(s)}\|_{V^*} \leq g_{r,\theta}(s) \sup_{s' \in [0,\,s]} \|\varphi(s') - \psi(s')\|^\Omega_{p,\infty}$$

for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ *and for almost every* $s \in [0,t]$. *Moreover, for all* $m > 0$, *there exists a constant* $G_{r,m} > 0$ *such that* $\|\theta\|_B \leq m$ *implies* $\|g_{r,\theta}\|_{L^2} \leq G_{r,m}$.

- *If* $(\theta_n)_{n=1}^\infty \subset \Theta$ *and* $\theta_n \rightharpoonup \theta$, *then for* $0 \leq j \leq p$

$$\int_0^t D^j\Big(\mathcal{M}(\varphi_{|[0,\,s]}, \theta_n(s))(\varphi(s,x))\Big)\,ds \to \int_0^t D^j\Big(\mathcal{M}(\varphi_{|[0,\,s]}, \theta(s))(\varphi(s,x))\Big)\,ds$$

for all $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $t \in [0,T]$, *and* $x \in \Omega$.

Then Θ *and* $\mathcal{M}$ *are compatible.*

Proof. We check that the assumptions of Theorem 4.1.10 are satisfied.

First we verify that $s \mapsto \mathcal{M}(\varphi_{|[0,\,s]}, \theta(s)) = (\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\,\beta^s_{\varphi,\,\theta(s)}$ is strongly measurable. Since $s \mapsto (\beta^s_{\varphi,\,\theta(s)} \mid v)$ is Lebesgue measurable by assumption, the fact that V is a separable Hilbert space and Lemma 4.3.3 imply that $s \mapsto \beta^s_{\varphi,\,\theta(s)}$ is strongly measurable. Moreover, the assumption that $s \mapsto \mathcal{A}^s_\varphi$ is in $C([0,t], \mathscr{L}(V,V^*))$ and $\big\|(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\big\|_{\mathscr{L}(V^*,V)} \leq \frac{1}{\gamma}$ from Lemma 4.1.13 give us that $s \mapsto (\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}$ is in $C([0,t], \mathscr{L}(V^*,V))$. It follows from Lemma 4.3.2 that $s \mapsto (\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\,\beta^s_{\varphi,\,\theta(s)}$ is strongly measurable.

Next, the model $(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\,\beta^s_{\varphi,\,\theta(s)}$ is bounded because

$$\big\|(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\big\|_{\mathscr{L}(V^*,V)} \leq \frac{1}{\gamma}$$

from Lemma 4.1.13 and the assumption

$$\|\beta^s_{\varphi,\,\theta(s)}\|_{V^*} \leq f_\theta(s)\,\Big(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_\infty\Big).$$

Finally, the model $(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\,\beta^s_{\varphi,\,\theta(s)}$ is Lipschitz in φ since (i) the operator $(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}$ is Lipschitz in φ and we have $\|\beta^s_{\varphi,\,\theta(s)}\|_{V^*} \leq f_\theta(s)\,(1+r)$; (ii) the operator $\beta^s_{\varphi,\,\theta(s)}$ is Lipschitz in φ and $(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}$ is uniformly bounded by $\frac{1}{\gamma}$; (iii) a multiplying function $\widetilde{g}_{r,\theta}$ such that

$$\|(\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\varphi)^{-1}\,\beta^s_{\varphi,\,\theta(s)} - (\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}^s_\psi)^{-1}\,\beta^s_{\psi,\,\theta(s)}\| \leq \widetilde{g}_{r,\theta}(s) \sup_{s' \in [0,\,s]} \|\varphi(s') - \psi(s')\|^\Omega_{p,\infty}$$

is given by

$$\widetilde{g}_{r,\theta}(s) := \frac{1}{\gamma^2}\,\ell_r\,f_\theta(s)\,(1+r) + \frac{1}{\gamma}\,g_{r,\theta}(s).$$

Note that $\widetilde{g}_{r,\theta} \in L^2([0,T])$ and

$$\|\widetilde{g}_{r,\theta}\|_{L^2} \leq \frac{1}{\gamma^2}\,\ell_r\,F_m\,(1+r) + \frac{1}{\gamma}\,G_{r,m}$$

for all $\|\theta\|_B \leq m$. □

Chapter 5

Applications of Core Theorems

In this chapter, we apply our core theorems from Chapter 4 to the problem of piecewise-rigid motions in Section 5.1 and to the problem of atrophy modeling in Section 5.2. We will show that both problems have a minimizer. As in Chapter 4, we assume that V is a separable Hilbert space continuously embedded in $C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d)$ throughout this chapter. We also remind the reader that we write $t \in [0, T]$, $s \in [0, t]$, and $s' \in [0, s]$ for hierarchical time intervals.

5.1 Application to Piecewise-rigid Motion

We first prove the existence of minimizers in Section 5.1.1 under our formulation in Section 3.3.1. Since the formulation in Section 3.3.1 only guarantees an almost piecewise-rigid motion, we provide another formulation, purely of theoretical interest, in Section 5.1.2 for an exact piecewise-rigid motion. We will discuss computational aspects later in Section 6.1.

5.1.1 Formulation using the energy form

We recall the problem from Section 3.3.1:

$$\min_{\theta \in L^2([0,T],\mathbb{R}^k)} \left(\frac{1}{2} \int_0^T \left(\|v(t)\|_V^2 + |\theta(t)|^2 \right) dt + \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}}) \right) \tag{3.25}$$

subject to

$$\varphi(t, x) = x + \int_0^t v(s, \varphi(s, x))\, ds = x + \int_0^t \mathcal{M}_R(\varphi_{|[0,\,s]}, \theta(s))(\varphi(s, x))\, ds.$$

The model of deformation vector fields is given by

$$\begin{aligned}\mathcal{M}_R(\varphi|_{[0,t]}, \theta(t)) &= \operatorname*{arg\,min}_{v \in V} \left(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} \int_{\varphi(t,\,\Omega_0)} |v|^2 \, dx - \sum_{i=1}^{N} \int_{\varphi(t,\,\Omega_i)} \chi \ (\mathcal{V}\,\theta_i(t))^\top v \, dx \right) \\ &=: \operatorname*{arg\,min}_{v \in V} \left(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} \left(\mathcal{A}_\varphi^t \, v \mid v\right) - \left(\beta_{\varphi,\,\theta(t)}^t \mid v\right) \right) \\ &= (\gamma \, \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1} \, \beta_{\varphi,\,\theta(t)}^t,\end{aligned}$$

and $\mathcal{V} : \mathbb{R}^{\frac{d(d+1)}{2}} \to C(\mathbb{R}^d, \mathbb{R}^d)$ is defined by

$$(\mathcal{V}\,\theta')(x) = \begin{cases} u + \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x, & \text{if } d = 2\,; \\ \\ u + \omega \times x, & \text{if } d = 3\,, \end{cases}$$

where $\theta' = (u, \omega)$. We observe that the operators $\mathcal{A}_\varphi^t$ and $\beta_{\varphi,\,\theta(t)}^t$ only depend on $\varphi(t)$, which is certainly covered by the assumption of the dependency on $\varphi|_{[0,\,t]}$. Since $\beta_{\varphi,\,\theta(t)}^t$ is linear in $\theta(t)$ in this case, we change the notation and write $\beta_\varphi^t \, \theta(t) := \beta_{\varphi,\,\theta(t)}^t$. We are ready to show that this problem has a minimizer.

Theorem 5.1.1. *If $p \geq 1$ and the discrepancy function ρ is continuous with respect to $\|\cdot\|_{p,\infty}$, then the minimization problem* (3.25) *has a minimizer.*

Proof. We aim to invoke Proposition 4.1.6. It is clear that the function $\theta \mapsto \frac{1}{2} \int_0^T |\theta(t)|^2 \, dt = \frac{1}{2} \|\theta\|^2_{L^2([0,\,T],\,\mathbb{R}^k)}$ is coercive. Moreover, it is strongly continuous and convex, which implies that it is weakly lower semicontinuous, and hence weakly sequentially lower semicontinuous. It remains to show that $\theta \mapsto \frac{1}{2} \int_0^T \|v(t)\|_V^2 \, dt$ is weakly sequentially lower semicontinuous and that $L^2([0,T], \mathbb{R}^k)$ and $\mathcal{M}_R$ are compatible. We check that the conditions of Corollay 4.1.12 are achieved to show the compatibility, then we prove that $\theta \mapsto \frac{1}{2} \int_0^T \|v(t)\|_V^2 \, dt$ is weakly sequentially lower semicontinuous by showing that $\theta_n \rightharpoonup \theta$ in $L^2([0,\,T],\,\mathbb{R}^k)$ implies $v_n \rightharpoonup v$ in $L^2([0,T], V)$, where v_n and v correspond to θ_n and θ respectively.

We first examine the mappings $s \mapsto \mathcal{A}_\varphi^s$ and $s \mapsto (\beta_\varphi^s \, \theta(s) \mid v)$. Since $p \geq 1$, we can make a change of variables and write

$$\left(\mathcal{A}_\varphi^s \, u \mid v\right) = \int_{\varphi(s,\,\Omega_0)} u^\top v \, dx = \int_{\Omega_0} (u^\top v) \circ \varphi(s) \, \det D\varphi(s) \, dx$$

and

$$\left(\beta_\varphi^s \, \theta(s) \mid v\right) = \sum_{i=1}^{N} \int_{\varphi(s,\,\Omega_i)} \chi \ (\mathcal{V}\,\theta_i(s))^\top v \, dx = \sum_{i=1}^{N} \int_{\Omega_i} \Big(\chi \ (\mathcal{V}\,\theta_i(s))^\top v \Big) \circ \varphi(s) \, \det D\varphi(s) \, dx.$$

If follows from $p \geq 1$, $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, and the measurability of θ that the mapping $s \mapsto \mathcal{A}^s_\varphi$ is in $C([0,t], \mathscr{L}(V,V^*))$ and that the mapping $s \mapsto (\beta^s_\varphi \theta(s) \mid v)$ is Lebesgue measurable.

Now we estimate $\|\beta^s_\varphi \theta(s)\|_{V^*}$. Note that

$$\|\mathcal{V}\,\theta_i(s)\|_\infty^{\varphi(s,\,\Omega_i)} \leq C\,(1 + \|\varphi(s) - id\|_\infty^{\Omega_i})\,|\theta_i(s)| \leq C\,(1 + \|\varphi(s) - id\|_\infty)\,|\theta(s)|.$$

It follows that

$$\begin{aligned} |(\beta^s_\varphi \theta(s) \mid v)| &= \left| \sum_{i=1}^N \int_{\varphi(s,\,\Omega_i)} \chi\ (\mathcal{V}\,\theta_i(s))^\top v\,dx \right| \\ &\leq C\,(1 + \|\varphi(s) - id\|_\infty)\,|\theta(s)|\ \|v\|_\infty\ \|\chi\|_{L^1}. \end{aligned} \tag{5.1}$$

Thus we let $f_\theta(s) := C\,c_V\,|\theta(s)|\ \|\chi\|_{L^1}$ with $F_m := C\,c_V\,\|\chi\|_{L^1}\,m$.

Next we check the Lipschitz conditions. For $\mathcal{A}^s_\varphi$, we obtain after a change of variables that

$$\left|(\mathcal{A}^s_\varphi\,u \mid v) - (\mathcal{A}^s_\psi\,u \mid v)\right| \leq \int_{\Omega_0} \Big|\,(u^\top v) \circ \varphi(s)\,\det D\varphi(s) - (u^\top v) \circ \psi(s)\,\det D\psi(s)\,\Big|\,dx.$$

Note that $A \mapsto \det A$ is a polynomial of degree d in elements of $A \in \mathbb{R}^{d\times d}$. By the mean value theorem, there exists a constant $C_d > 0$ such that

$$|\det A - \det B| \leq C_d\,(|A| + |B|)^{d-1}\,|A - B|$$

for all $A, B \in \mathbb{R}^{d\times d}$. Using standard arguments, we deduce that for all $r > 0$ there exists $\ell_r > 0$ such that

$$\|\mathcal{A}^s_\varphi - \mathcal{A}^s_\psi\|_{\mathscr{L}(V,V^*)} \leq \ell_r\,\|\varphi(s) - \psi(s)\|_{1,\infty}^{\Omega_0} \leq \ell_r \sup_{s'\,\in\,[0,\,s]} \|\varphi(s') - \psi(s')\|_{p,\infty}^{\Omega_0}$$

for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$. As for $\beta^s_\varphi \theta(s)$, similarly we have

$$\begin{aligned} &\left|(\beta^s_\varphi \theta(s) \mid v) - (\beta^s_\psi \theta(s) \mid v)\right| \\ &\leq \sum_{i=1}^N \int_{\Omega_i} \Big|\Big(\chi\ (\mathcal{V}\,\theta_i(s))^\top v\Big) \circ \varphi(s)\,\det D\varphi(s) - \Big(\chi\ (\mathcal{V}\,\theta_i(s))^\top v\Big) \circ \psi(s)\,\det D\psi(s)\Big|\,dx \\ &\leq C_r\,\|\varphi(s) - \psi(s)\|_{1,\infty}^{\Omega_0}\ |\theta(s)|\ \|v\|_V \\ &\leq C_r \sup_{s'\,\in\,[0,s]} \|\varphi(s') - \psi(s')\|_{p,\infty}^{\Omega_0}\ |\theta(s)|\ \|v\|_V \end{aligned} \tag{5.2}$$

for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$. We let $g_{r,\theta}(s) := C_r\,|\theta(s)|$ with $G_{r,m} := C_r\,m$.

We show that $\mathcal{M}_R$ is continuous in θ. We fix $0 \leq j \leq p$ and define a linear operator

$\mathcal{F}_j : L^2([0,T],\mathbb{R}^k) \to \mathbb{R}^{d^{j+1}}$ by

$$\begin{aligned}\mathcal{F}_j\,\theta' &= \int_0^t D^j\Big(\mathcal{M}_R(\varphi_{|[0,\,s]},\theta'(s))(\varphi(s,x))\Big)\,ds\\ &= \int_0^t D^j\big(\big((\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_\varphi^s)^{-1}\,\beta_\varphi^s\,\theta'(s)\big)(\varphi(s,x))\big)\,ds.\end{aligned}$$

In addition, the uniform bound $\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_\varphi^s)^{-1}\|_{\mathscr{L}(V^*,V)} \le \frac{1}{\gamma}$ and the estimate of $\|\beta_\varphi^s\,\theta'(s)\|_{V^*}$ from (5.1) give us

$$\begin{aligned}|\mathcal{F}_j\,\theta'| &\le \int_0^t C_{\varphi(s)}\,\big\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_\varphi^s)^{-1}\,\beta_\varphi^s\,\theta'(s)\big\|_{p,\infty}\,ds\\ &\le C_\varphi\int_0^t c_V\,\frac{1}{\gamma}\;C\,(1+\|\varphi(s)-id\|_\infty)\,|\theta'(s)|\;\;\|\chi\|_{L^1}\,ds \le C_\varphi\,\|\theta'\|_{L^2([0,\,T],\,\mathbb{R}^k)},\end{aligned}$$

where C_φ is a polynomial in $\sup\limits_{s\,\in\,[0,\,T]}\|\varphi(s)-id\|_{p,\infty}$. Since the linear operator $\mathcal{F}_j$ is bounded and its codomain is of finite dimension, it follows at once that $\theta_n \rightharpoonup \theta$ implies $\mathcal{F}_j\,\theta_n \to \mathcal{F}_j\,\theta$, that is,

$$\int_0^t D^j\Big(\mathcal{M}_R(\varphi_{|[0,\,s]},\theta_n(s))(\varphi(s,x))\Big)\,ds \to \int_0^t D^j\Big(\mathcal{M}_R(\varphi_{|[0,\,s]},\theta(s))(\varphi(s,x))\Big)\,ds.$$

We have shown that the conditions of Corollay 4.1.12 are all satisfied, hence $L^2([0,T],\mathbb{R}^k)$ and $\mathcal{M}_R$ are compatible.

Our last step is to show that $\theta_n \rightharpoonup \theta$ in $L^2([0,T],\mathbb{R}^k)$ implies $v_n \rightharpoonup v$ in $L^2([0,T],V)$, where

$$v_n(t) := (\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_{\varphi_{\theta_n}}^t)^{-1}\,\beta_{\varphi_{\theta_n}}^t\,\theta_n(t) \quad\text{ and }\quad v(t) := (\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_{\varphi_\theta}^t)^{-1}\,\beta_{\varphi_\theta}^t\,\theta(t),$$

and φ_{θ_n} and φ_θ are the unique solutions of θ_n and θ respectively, whose existence is due to the compatibility we just proved. To this end, we let $\mu \in \big(L^2([0,T],V)\big)^* \cong L^2([0,T],V) \cong L^2([0,T],V^*)$ and estimate $|(\mu \mid v_n)-(\mu \mid v)|$. We have

$$\begin{aligned}&|(\mu \mid v_n)-(\mu \mid v)|\\ &\le \int_0^T \|\mu(t)\|_{V^*}\;\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_{\varphi_{\theta_n}}^t)^{-1}-(\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_{\varphi_\theta}^t)^{-1}\|_{\mathscr{L}(V^*,V)}\;\|\beta_{\varphi_{\theta_n}}^t\,\theta_n(t)\|_{V^*}\;dt\\ &\quad+\int_0^T \|\mu(t)\|_{V^*}\;\|(\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_{\varphi_\theta}^t)^{-1}\|_{\mathscr{L}(V^*,V)}\;\|\beta_{\varphi_{\theta_n}}^t-\beta_{\varphi_\theta}^t\|_{\mathscr{L}(\mathbb{R}^k,\,V^*)}\;|\theta_n(t)|\;dt\\ &\quad+\left|\int_0^T\Big(\mu(t)\,\Big|\,(\gamma\,\mathcal{K}_V^{-1}+\mathcal{A}_{\varphi_\theta}^t)^{-1}\beta_{\varphi_\theta}^t\,(\theta_n(t)-\theta(t))\Big)dt\right|.\end{aligned}$$

Since $\theta_n \rightharpoonup \theta$, there exist m and $r(m)$ such that $\|\theta_n\|_{L^2([0,\,T],\,\mathbb{R}^k)} \le m$ and $\|\varphi_{\theta_n}(t)-id\|_{p,\infty} \le r$ (see Proposition 4.3.6). The conditions of Corollay 4.1.12 we have checked and the in-

equality (5.2) then lead to

$$
\begin{aligned}
&|(\mu \mid v_n) - (\mu \mid v)| \\
&\leq \int_0^T \|\mu(t)\|_{V^*} \, \frac{1}{\gamma^2} \, \ell_r \sup_{s \in [0,t]} \|\varphi_{\theta_n}(s) - \varphi_\theta(s)\|_{p,\infty}^{\Omega_0} \; f_{\theta_n}(t) \, (1+r) \; dt \\
&\quad + \int_0^T \|\mu(t)\|_{V^*} \, \frac{1}{\gamma} \; C_r \sup_{s \in [0,t]} \|\varphi_{\theta_n}(s) - \varphi_\theta(s)\|_{p,\infty}^{\Omega_0} \; |\theta_n(t)| \; dt \\
&\quad + \left| \int_0^T \Big(\mu(t) \,\Big|\, (\gamma \, \mathcal{K}_V^{-1} + \mathcal{A}_{\varphi_\theta}^t)^{-1} \beta_{\varphi_\theta}^t \, (\theta_n(t) - \theta(t)) \Big) \, dt \right| \\
&\leq C_r \, \|\mu\|_{L^2([0,T],V)} \, (Fm + m) \sup_{t \in [0,T]} \|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^{\Omega_0} \\
&\quad + \left| \int_0^T \Big(\mu(t) \,\Big|\, (\gamma \, \mathcal{K}_V^{-1} + \mathcal{A}_{\varphi_\theta}^t)^{-1} \beta_{\varphi_\theta}^t \, (\theta_n(t) - \theta(t)) \Big) \, dt \right| .
\end{aligned}
$$

The compatibility gives $\sup_{t \in [0,T]} \|\varphi_{\theta_n}(t) - \varphi_\theta(t)\|_{p,\infty}^{\Omega_0} \to 0$. On the other hand, from (5.1) we know that

$$
\theta' \mapsto \int_0^T \Big(\mu(t) \,\Big|\, (\gamma \, \mathcal{K}_V^{-1} + \mathcal{A}_{\varphi_\theta}^t)^{-1} \beta_{\varphi_\theta}^t \, \theta'(t) \Big) \, dt
$$

is a bounded linear functional. Hence the second term in the above inequality also goes to zero. We conclude that $(\mu \mid v_n) \to (\mu \mid v)$, which completes the proof. □

Let $h : L^2([0,T], \mathbb{R}^k) \to \mathbb{R}$ be an arbitrary weakly sequentially lower semicontinuous function. If we change the objective function in the minimization problem of piecewise-rigid motion from (3.25) to

$$
\min_{\theta \in L^2([0,T],\, \mathbb{R}^k)} \left(h(\theta) + \frac{1}{2} \int_0^T |\theta(t)|^2 \, dt + \rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}) \right),
$$

the above proof still works. In particular, if $h \equiv 0$, the corresponding problem has a minimizer. More generally, for any weakly sequentially lower semicontinuous function $\widetilde{h} : L^2([0,T], \mathbb{R}^k) \to \mathbb{R}$ such that $\int_0^T |\theta(t)|^2 \, dt \leq C \, \widetilde{h}(\theta)$, the minimization problem

$$
\min_{\theta \in L^2([0,T],\, \mathbb{R}^k)} \Big(\widetilde{h}(\theta) + \rho(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}) \Big),
$$

has a minimizer. This formulation enables the flexibility of the choice of h or $\widetilde{h}$, at the expense of deviating from piecewise-rigid motions by the regularization $\frac{\gamma}{2} \|v\|_V^2$ in

$$
\mathcal{M}_R(\varphi|_{[0,t]}, \theta(t)) = \operatorname*{arg\,min}_{v \in V} \left(\frac{\gamma}{2} \|v\|_V^2 + \frac{1}{2} \int_{\varphi(t, \Omega_0)} |v|^2 \, dx - \sum_{i=1}^N \int_{\varphi(t, \Omega_i)} \chi \; (\mathcal{V} \, \theta_i(t))^\top v \, dx \right). \tag{3.24}
$$

If one is open to always include the term $\frac{1}{2}\int_0^T \|v(t)\|_V^2\,dt$ in the objective function, exact piecewise-rigid motions can indeed be achieved theoretically, as shown in the next section.

5.1.2 Formulation using constraints

To attain exact piecewise-rigid motions, we now use another characterization of rigid motions, Proposition 3.1.1(iv). For every $\xi \in \mathrm{Diff}^p_{id}(\mathbb{R}^d)$, we define a linear operator $\mathcal{L}_\xi : V \to L^1(\mathbb{R}^d, \mathbb{R}^{d\times d})$ by

$$(\mathcal{L}_\xi\, w)(x) := \mathbb{1}_{\xi(\Omega_0)}(x)\left(Dw(x)^\top + Dw(x)\right)$$

and reformulate the problem as

$$\min_{v\,\in\, L^2([0,T],V)} \left(\frac{1}{2}\int_0^T \|v(t)\|_V^2\,dt + \rho(\varphi(T,\Omega_0),\Omega_{\mathrm{targ}})\right) \tag{5.3}$$

subject to

$$\begin{cases} \varphi(t,x) = x + \displaystyle\int_0^t v(s,\varphi(s,x))\,ds & \text{for all } (t,x)\in[0,T]\times\mathbb{R}^d \\ \mathcal{L}_{\varphi(t)}\,v(t) = 0 & \text{for almost every } t\in[0,T] \end{cases}.$$

The constraint $\mathcal{L}_{\varphi(t)}\,v(t) = 0$ characterizes rigid motions of each connected components of Ω_0, namely, piecewise-rigid motions. If $p \geq 1$, then $\xi \mapsto \mathcal{L}_\xi$ is continuous. In addition, suppose that the discrepancy function ρ is continuous with respect to $\|\cdot\|_{p,\infty}$, then the minimization problem (5.3) has a minimizer according to Theorem 1 in [4].

Similarly, we can add a weakly sequentially lower semicontinuous function $v \mapsto h(v)$ to the objective function, or replace $\frac{1}{2}\int_0^T \|v(t)\|_V^2\,dt$ by a weakly sequentially lower semicontinuous function $v \mapsto \widetilde{h}(v)$ such that $\int_0^T \|v(t)\|_V^2\,dt \leq C\,\widetilde{h}(v)$, and still have the existence of minimizers. However, the formulation in this section requires constrained optimization, which is more difficult to solve numerically. For both formulations in Sections 5.1.1 and 5.1.2, there is another undesired numerical issue, which we will visit in Section 6.1.

5.2 Application to the Atrophy Model

Let $\Theta \subset \mathbb{R}^k$ be a compact set. We recall from Section 3.3.2 that the problem is

$$\min_{\theta\,\in\,\Theta}\ \rho(\varphi(T,\Omega_0),\Omega_{\mathrm{targ}})$$

subject to

$$\varphi(t,x) = x + \int_0^t v(s,\varphi(s,x))\,ds = x + \int_0^t \mathcal{M}_A(\varphi_{|[0,\,s]},\theta)(\varphi(s,x))\,ds.$$

The model of deformation vector fields is given by

$$\begin{aligned}
&\mathcal{M}_A(\varphi_{|[0,\,t]},\theta)\\
&= \operatorname*{arg\,min}_{v\,\in\,V}\Big(\frac{\gamma}{2}\,\|v\|_V^2 + \frac{1}{2}\int_{\varphi(t,\,\Omega_0)}\mathcal{E}_{\varphi(t)}(\varepsilon_v,\varepsilon_v)\,dx - \int_{\varphi(t,\,\Omega_0)}\chi\ \alpha\big(\tau(t)\circ\varphi(t)^{-1}\big)\ (-\operatorname{div} v)\ dx\Big)\\
&=: \operatorname*{arg\,min}_{v\,\in\,V}\Big(\frac{\gamma}{2}\,\|v\|_V^2 + \frac{1}{2}\,(\mathcal{A}_\varphi^t\,v \mid v) - (\beta_{\varphi,\,\theta}^t \mid v)\Big)\\
&= (\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1}\,\beta_{\varphi,\,\theta}^t,
\end{aligned}$$

with $\tau = \mathcal{T}(\varphi_{|[0,\,t]},\theta)$ the unique solution to the system

$$\begin{cases}
\partial_s\big(\tau\ \det D\varphi\big)(s) = \operatorname{div}\Big(W_{\varphi(s)}\,\nabla\tau(s)\ \det D\varphi(s)\Big) + R(\tau(s))\ \det D\varphi(s) & \text{in } (0,t]\times\Omega_0^o\\
\big(W_{\varphi(s)}\,\nabla\tau(s)\ \det D\varphi(s)\big)^\top n_0 = 0 & \text{on } [0,t]\times\partial\Omega_0\\
\tau(0) = Q(\theta) & \text{on } \Omega_0^o
\end{cases},$$

where $W_{\varphi(s)} = \big(D\varphi(s)\big)^{-1}\big(U_{\varphi(s)}\circ\varphi(s)\big)\ D\varphi(s)^{-\top}$ and $U_{\varphi(s)} : \varphi(s,\Omega) \to \mathbb{R}^{d\times d}$ is the Eulerian diffusion matrix (field). In this case, the operator $\mathcal{A}_\varphi^t$ still only depends on $\varphi(t)$, but the operator $\beta_{\varphi,\,\theta}^t$ now depends on $\varphi_{|[0,\,t]}$ because of $\tau(t)$.

Before we prove the existence of minimizers, we need to show that $\tau = \mathcal{T}(\varphi,\theta)$, a solution to the PDE model, exists and is unique for all $\varphi \in C([0,t],\operatorname{Diff}_{id}^p(\mathbb{R}^d))$ and all $\theta\in\Theta$. In particular, we need to specify what we mean by a solution to the PDE model and the codomain of $\mathcal{T}$. We collect necessary tools from [62, Chapter 1 and Chapter 3] in Section 5.2.1 and state the weak formulation of the PDE problem in Section 5.2.2. We prove the existence and uniqueness of weak solutions to the PDE problem in Section 5.2.3, and thus the solution mapping $\mathcal{T}$ is well defined. The existence of minimizers will then be proved in Section 5.2.4.

5.2.1 Abstract parabolic initial value problems

We first generalize the space of real-valued distributions $\mathscr{D}^*\big((0,t)\big)$ to the space of Hilbert-space-valued distributions $\mathscr{D}^*\big((0,t),H\big)$. We define as in [62, Chapter 1, Section 1.3] that

$$\mathscr{D}^*\big((0,t),H\big) := \mathscr{L}\big(\mathscr{D}\big((0,t)\big),H\big).$$

For a locally integrable function $u \in L^1_{\mathrm{loc}}([0,t],H)$, i.e., u is strongly measurable with $\int_a^b \|u(s)\|_H \, ds < \infty$ for all $[a,b] \subset (0,t)$, we can define the corresponding

$$\widetilde{u}(\varphi) := \int_0^t u(s)\,\varphi(s)\,ds \;\in\; H \quad \text{for all } \varphi \in \mathscr{D}\big((0,t)\big)$$

and show that $\widetilde{u} \in \mathscr{D}^*\big((0,t),H\big)$ by replacing the Lebesgue integral with the Bochner integral in the usual argument of real-valued distributions. We thus identify u with $\widetilde{u}$ and obtain $L^1_{\mathrm{loc}}([0,t],H) \subset \mathscr{D}^*\big((0,t),H\big)$. We can also generalize distributional derivatives to Hilbert-space-valued distributions. The derivative $\frac{du}{ds}$ of $u \in \mathscr{D}^*\big((0,t),H\big)$ in the sense of distribution is defined by

$$\frac{du}{ds}(\varphi) := -u(\dot{\varphi}) \;\in\; H \quad \text{for all } \varphi \in \mathscr{D}\big((0,t)\big),$$

which can be shown is still in $\mathscr{D}^*\big((0,t),H\big)$. Since $L^1_{\mathrm{loc}}([0,t],H) \subset \mathscr{D}^*\big((0,t),H\big)$, we can take distributional derivatives for every $u \in L^1_{\mathrm{loc}}([0,t],H)$.

Given a bounded and open set $\Omega \subset \mathbb{R}^d$, we now consider spaces $L^2\big([0,t],H^1(\Omega)\big)$, $L^2\big([0,t],L^2(\Omega)\big)$, and $L^2\big([0,t],H^1(\Omega)^*\big)$; we do not identify $H^1(\Omega)$ with $H^1(\Omega)^*$. We will write $L^2\big([0,t],H^1\big)$, $L^2\big([0,t],L^2\big)$, and $L^2\big([0,t],(H^1)^*\big)$ if there is no confusion of the set Ω. Identifying $L^2\big([0,t],L^2\big)$ with its dual leads to the Hilbert triple

$$L^2\big([0,t],H^1\big) \subset L^2\big([0,t],L^2\big) \subset L^2\big([0,t],(H^1)^*\big),$$

each space being dense in the following one. Note that for every $u \in L^2\big([0,t],H^1\big)$, we have

$$u \in L^2\big([0,t],H^1\big) \subset L^2\big([0,t],(H^1)^*\big) \subset \mathscr{D}^*\big((0,t),(H^1)^*\big),$$

thus we can take its time derivative in the sense of distribution, which will be denoted by $\partial_s u \in \mathscr{D}^*\big((0,t),(H^1)^*\big)$ to differentiate with the strong and pointwise derivative $\dot{u}$. If $\partial_s u$ is more regular, we then have the following "intermediate regular" result [62, Chapter 1, Proposition 2.1 and Theorem 3.1].

Theorem 5.2.1. *If $u \in L^2\big([0,t],H^1\big)$ and $\partial_s u \in L^2\big([0,t],(H^1)^*\big)$, then $u \in C\big([0,t],L^2\big)$.*

With the generalized notion of time derivatives, we introduce the following abstract parabolic initial value problem. Let $\mathcal{L} \in \mathscr{L}\big(L^2\big([0,t],H^1\big),\, L^2\big([0,t],(H^1)^*\big)\big)$ be coercive, i.e., there exists $a > 0$ such that $(\mathcal{L}u \mid u) \geq a\,\|u\|^2_{L^2([0,t],\,H^1)}$ for all $u \in L^2\big([0,t],H^1\big)$. Given $f \in L^2\big([0,t],(H^1)^*\big)$ and $u_0 \in L^2(\Omega)$, we want to solve the parabolic initial value problem

$$\begin{cases} \partial_s u + \mathcal{L}u = f \\ u(0) = u_0 \end{cases}, \tag{5.4}$$

where the first equation is an operational equation in $L^2\big([0,t],(H^1)^*\big)$. The following theorem states that the initial value problem has a unique solution in $L^2\big([0,t],H^1\big)$ [62, Chapter 3, Theorem 1.1 with $\Lambda u = \partial_s u$ and $M = \mathcal{L}$, Section 4.3, and Remark 4.3].

Theorem 5.2.2. *Let $\mathcal{L} \in \mathscr{L}\big(L^2\big([0,t],H^1\big),\, L^2\big([0,t],(H^1)^*\big)\big)$ be coercive. For all $f \in L^2\big([0,t],(H^1)^*\big)$ and $u_0 \in L^2(\Omega)$, the initial value problem* (5.4) *has a unique solution in $L^2\big([0,t],H^1\big)$.*

Remark 5.2.3. If $u \in L^2\big([0,t],H^1\big)$ is a solution to (5.4), then we have $\partial_s u = f - \mathcal{L}u \in L^2\big([0,t],(H^1)^*\big)$. Since $u \in L^2\big([0,t],H^1\big)$ and $\partial_s u \in L^2\big([0,t],(H^1)^*\big)$, Theorem 5.2.1 shows that $u \in C\big([0,t],L^2\big)$. Thus $u(0) = u_0 \in L^2(\Omega)$ makes sense. In fact, it is possible to make sense of the initial condition without invoking Theorem 5.2.1. The fact that a solution $u \in L^2\big([0,t],H^1\big) \subset L^2\big([0,t],(H^1)^*\big)$ and $\partial_s u \in L^2\big([0,t],(H^1)^*\big)$ implies that u has a continuous representative in $C\big([0,t],(H^1)^*\big)$. Hence, as long as $u_0 \in H^1(\Omega)^*$, the initial condition is well defined. We mention Theorem 5.2.1 to emphasize that a solution is actually in $L^2\big([0,t],H^1\big) \cap C\big([0,t],L^2\big)$, which will be needed later.

We close this section with a useful lemma.

Lemma 5.2.4. *Suppose that $u \in L^2\big([0,t],H^1\big)$ and $\partial_s u \in L^2\big([0,t],(H^1)^*\big)$. Then $\partial_s \|u(\cdot)\|_{L^2}^2$ in the sense of distribution is in $L^1([0,t])$ and equals to $s \mapsto 2\,\big((\partial_s u)(s) \mid u(s)\big)_{(H^1)^*,\,H^1}$ for almost every $s \in [0,t]$. It follows that $s \mapsto \|u(s)\|_{L^2}^2$ is in $W^{1,1}([0,t])$ and*

$$\begin{aligned}\|u(s)\|_{L^2}^2 &= \|u(0)\|_{L^2}^2 + \int_0^s \big(\partial_s \|u(\cdot)\|_{L^2}^2\big)(s')\,ds' \\ &= \|u(0)\|_{L^2}^2 + \int_0^s 2\,\big((\partial_s u)(s') \mid u(s')\big)_{(H^1)^*,\,H^1}\,ds'.\end{aligned}$$

Proof. It can be shown [62, Chapter 1, Theorem 2.1] that there exists a sequence $(u_n)_{n=1}^\infty \subset C^\infty([0,t],H^1)$ such that

$$\|u_n - u\|_{L^2([0,t],\,H^1)} \to 0 \quad \text{and} \quad \|\partial_s u_n - \partial_s u\|_{L^2([0,t],\,(H^1)^*)} \to 0.$$

For all $\psi \in \mathscr{D}\big((0,t)\big)$, since $\|u_n(\cdot)\|_{L^2}^2$ is differentiable, we obtain

$$\begin{aligned}\int_0^t \|u(s)\|_{L^2}^2\ \partial_s\psi(s)\,ds &= \lim_{n\to\infty} \int_0^t \|u_n(s)\|_{L^2}^2\ \partial_s\psi(s)\,ds \\ &= -\lim_{n\to\infty} \int_0^t 2\,\big\langle (\partial_s u_n)(s), u_n(s)\big\rangle_{L^2}\,\psi(s)\ ds \\ &= -\lim_{n\to\infty} \int_0^t 2\,\big((\partial_s u_n)(s) \mid u_n(s)\big)_{(H^1)^*,\,H^1}\,\psi(s)\ ds\end{aligned}$$

$$= - \int_0^t 2 \left((\partial_s u)(s) \mid u(s) \right)_{(H^1)^*, H^1} \psi(s) \, ds,$$

which shows that $\big(\partial_s \|u(\cdot)\|_{L^2}^2\big)(s) = 2 \left((\partial_s u)(s) \mid u(s) \right)_{(H^1)^*, H^1}$ for almost every $s \in [0,t]$. □

Remark 5.2.5. For a function $u \in L^2\big([0,t], H^1\big)$ with $\partial_s u \in L^2\big([0,t], (H^1)^*\big)$, Theorem 5.2.1 says that $s \mapsto \|u(s)\|_{L^2}^2$ is continuous. Lemma 5.2.4 says that $s \mapsto \|u(s)\|_{L^2}^2$ is in fact absolutely continuous.

5.2.2 Weak solutions to the PDE model

Assuming τ is sufficiently regular, we have derived in Section 3.2.2 the reaction-diffusion equation on a moving shape $s \mapsto \varphi(s, \Omega)$ driven by the motion $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ as follows:

$$\begin{cases} \partial_s \big(\tau \, \det D\varphi\big)(s) = \mathrm{div}\Big(W_{\varphi(s)} \, \nabla\tau(s) \, \det D\varphi(s)\Big) + R(\tau(s)) \, \det D\varphi(s) & \text{in } (0,t] \times \Omega^o \\ \big(W_{\varphi(s)} \, \nabla\tau(s)\big)^\top n(0) = 0 & \text{on } [0,t] \times \partial\Omega \\ \tau(0) = Q(\theta) & \text{on } \Omega^o \end{cases},$$

where $W_{\varphi(s)} = \big(D\varphi(s)\big)^{-1} \big(U_{\varphi(s)} \circ \varphi(s)\big) \, D\varphi(s)^{-\top}$ with the Eulerian diffusion matrix (field) $U_{\varphi(s)} : \varphi(s, \Omega) \to \mathbb{R}^{d \times d}$, while $R : \mathbb{R} \to \mathbb{R}$ is the reaction function, and $Q : \Theta \subset \mathbb{R}^k \to L^2(\Omega)$ is the parametrized initial condition. Suppose that the reaction function R is bounded. For all τ, $\tau' \in L^2\big([0,t], H^1\big)$, we define $\mathcal{L} : L^2\big([0,t], H^1\big) \to L^2\big([0,t], (H^1)^*\big)$ and $g(\tau) \in L^2\big([0,t], L^2\big) \subset L^2\big([0,t], (H^1)^*\big)$ by

$$\big(\!\big(\mathcal{L}\tau \mid \tau'\big)\!\big) := \int_0^t \big\langle W_{\varphi(s)} \, \nabla\tau(s) \, \det D\varphi(s), \, \nabla\tau'(s) \big\rangle_{L^2} \, ds$$

and

$$g(\tau)(s) := R(\tau(s)) \, \det D\varphi(s) \quad \text{for almost every } s \in [0,t],$$

where $\big(\!\big(\, \cdot \mid \cdot \,\big)\!\big)$ denotes the pairing of $L^2\big([0,t], (H^1)^*\big)$ and $L^2\big([0,t], H^1\big)$. Multiplying the PDE by an arbitrary $\psi \in L^2\big([0,t], H^1\big)$ and integrating by parts, we obtain

$$\begin{aligned} &\big(\!\big(\, \partial_s \big(\tau \, \det D\varphi\big) \mid \psi \, \big)\!\big) \\ &= \int_0^t \big(\partial_s \big(\tau \, \det D\varphi\big)(s) \mid \psi(s)\big)_{(H^1)^*, H^1} \, ds \\ &= - \int_0^t \big\langle W_{\varphi(s)} \, \nabla\tau(s) \, \det D\varphi(s), \, \nabla\psi(s) \big\rangle_{L^2} \, ds + \int_0^t \big(R(\tau(s)) \, \det D\varphi(s) \mid \psi(s)\big)_{(H^1)^*, H^1} \, ds \\ &= -\big(\!\big(\mathcal{L}\tau \mid \psi\big)\!\big) + \big(\!\big(g(\tau) \mid \psi\big)\!\big). \end{aligned}$$

Thus the weak formulation of the problem is given by

$$\begin{cases} \partial_s \big(\tau \, \det D\varphi\big) + \mathcal{L}\tau = g(\tau) \\ \tau(0) = Q(\theta) \end{cases}, \tag{5.5}$$

where the time derivative is in the sense of distribution, and the first equation is an operational equation in $L^2\big([0,t],(H^1)^*\big)$. For technical reasons (see Lemma 5.2.6), we make a change of function $u(s) := e^{-\lambda s}\,\tau(s)\,\det D\varphi(s)$ and, assuming $\varphi \in C([0,t],\mathrm{Diff}^2_{id}(\mathbb{R}^d))$, obtain an equivalent problem

$$\begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(u) \\ u(0) = Q(\theta) \end{cases}, \tag{5.6}$$

where $\mathcal{L}_\lambda : L^2\big([0,t],H^1\big) \to L^2\big([0,t],(H^1)^*\big)$ and $g_\lambda(u) \in L^2\big([0,t],L^2\big) \subset L^2\big([0,t],(H^1)^*\big)$ are defined by

$$\begin{aligned} \big(\!\big(\mathcal{L}_\lambda u \mid u'\big)\!\big) := \int_0^t \Big(& \lambda \left\langle u(s),\, u'(s)\right\rangle_{L^2} + \left\langle W_{\varphi(s)}\,\nabla u(s),\, \nabla u'(s)\right\rangle_{L^2} \\ & - \left\langle u(s)\, W_{\varphi(s)}\, \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)},\, \nabla u'(s)\right\rangle_{L^2} \Big)\, ds \end{aligned} \tag{5.7}$$

and

$$g_\lambda(u)(s) := e^{-\lambda s}\, R\!\left(\frac{e^{\lambda s}\,u(s)}{\det D\varphi(s)}\right) \det D\varphi(s) \quad \text{for almost every } s \in [0,t]. \tag{5.8}$$

In the next section, we will give sufficient conditions on $U_{\varphi(s)}$ (thus $W_{\varphi(s)}$), R, and Q so that the initial value problem (5.6) has a unique solution $u \in L^2\big([0,t],H^1\big)$ for some $\lambda > 0$, and hence the equivalent problem (5.5) has a unique solution $\tau \in L^2\big([0,t],H^1\big)$.

5.2.3 Existence and uniqueness of weak solutions to the PDE model

In this section, we fix a Diff^p-motion $\varphi \in C([0,t],\mathrm{Diff}^p_{id}(\mathbb{R}^d))$ and prove that for some $\lambda > 0$ the initial value problem

$$\begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(u) \\ u(0) = Q(\theta) \end{cases} \tag{5.6}$$

has a unique solution $u \in L^2\big([0,t],H^1\big)$, where $\mathcal{L}_\lambda$ and $g_\lambda(u)$ are defined in (5.7) and (5.8). We proceed in two steps. In the first step, we apply Theorem 5.2.2 to show that there exists $\lambda > 0$ such that the initial value problem

$$\begin{cases} \partial_s u + \mathcal{L}_\lambda u = f \\ u(0) = Q(\theta) \end{cases} \tag{5.9}$$

has a unique solution for *all* $f \in L^2\big([0,t],(H^1)^*\big)$ and $\theta \in \Theta$. We then fix such λ and enter the second step. In the second step, we consider the initial value problem

$$\begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(w) \\ u(0) = Q(\theta) \end{cases}. \tag{5.10}$$

From the first step, for each given w there exists a unique solution u_w to (5.10). We then show that $w \mapsto u_w$ is a contraction and apply the Banach fixed point theorem. Since φ is assumed to be fixed in this section, we view C_φ as a constant C in this section.

We begin our first step. We are going to choose $\lambda > 0$ so that $\mathcal{L}_\lambda$ is bounded and coercive in order to apply Theorem 5.2.2. The almost pointwise coercivity of $\mathcal{L}_\lambda$ is shown in the following lemma, which is the key in most of our later proofs of the PDE model.

Lemma 5.2.6. *Let $p \geq 2$. Suppose that the Eulerian diffusion matrix is symmetric and positive definite and satisfies*

$$\max\Big\{ \sup_{s\in[0,t]} \|U_{\varphi(s)}\|_\infty, \sup_{s\in[0,t]} \|U^{-1}_{\varphi(s)}\|_\infty \Big\} < \infty,$$

then there exist $\lambda(\varphi) > 0$ and $a(\varphi) > 0$ such that for all $u \in L^2\big([0,t],H^1\big)$ and almost every $s \in [0,t]$

$$\big((\mathcal{L}_\lambda u)(s) \mid u(s) \big)_{(H^1)^*,\, H^1} \geq a\, \|u(s)\|^2_{H^1},$$

where $\mathcal{L}_\lambda : L^2\big([0,t],H^1\big) \to L^2\big([0,t],(H^1)^*\big)$ is given by

$$\begin{aligned} ((\mathcal{L}_\lambda u \mid u')) = \int_0^t \Big(& \lambda \left\langle u(s),\, u'(s) \right\rangle_{L^2} + \left\langle W_{\varphi(s)}\, \nabla u(s),\, \nabla u'(s) \right\rangle_{L^2} \\ & - \left\langle u(s)\, W_{\varphi(s)}\, \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)},\, \nabla u'(s) \right\rangle_{L^2} \Big)\, ds. \end{aligned} \tag{5.7}$$

and $W_{\varphi(s)} = \big(D\varphi(s)\big)^{-1} \big(U_{\varphi(s)} \circ \varphi(s)\big)\, D\varphi(s)^{-\top}$.

Proof. The fact that $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, $p \geq 2$, and the assumption on $U_{\varphi(s)}$ give us

$$\begin{aligned} \max\Big\{ & \sup_{s\in[0,t]} \|\varphi(s) - id\|_{2,\infty}, \sup_{s\in[0,t]} \|\varphi(s)^{-1} - id\|_{1,\infty}, \\ & \sup_{s\in[0,t]} \|W_{\varphi(s)}\|_\infty, \sup_{s\in[0,t]} \|W^{-1}_{\varphi(s)}\|_\infty, \Big\} < \infty. \end{aligned}$$

Moreover, we know that there are constants C, $C' > 0$ such that for almost every $s \in [0,t]$

$$\left\langle W_{\varphi(s)}\, \nabla u(s), \nabla u(s) \right\rangle_{L^2} \geq C\, \|\nabla u(s)\|^2_{L^2}$$

and

$$\left\langle u(s)\, W_{\varphi(s)}\, \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)},\, \nabla u(s) \right\rangle_{L^2} \leq C' \,\|u(s)\|_{L^2}\, \|\nabla u(s)\|_{L^2}.$$

Using Lemma A.2.1, it follows that for all $\varepsilon > 0$

$$\begin{aligned}
\big(\,(\mathcal{L}_\lambda u)(s) \mid u(s)\,\big)_{(H^1)^*,\,H^1} &\geq \lambda\,\|u(s)\|_{L^2}^2 + C\,\|\nabla u(s)\|_{L^2}^2 - C'\,\|u(s)\|_{L^2}\,\|\nabla u(s)\|_{L^2} \\
&\geq \lambda\,\|u(s)\|_{L^2}^2 + C\,\|\nabla u(s)\|_{L^2}^2 - C'\left(\frac{1}{2\varepsilon}\,\|u(s)\|_{L^2}^2 + \frac{\varepsilon}{2}\,\|\nabla u(s)\|_{L^2}^2\right) \\
&= \left(\lambda - \frac{C'}{2\varepsilon}\right)\|u(s)\|_{L^2}^2 + \left(C - \frac{C'\varepsilon}{2}\right)\|\nabla u(s)\|_{L^2}^2.
\end{aligned}$$

Since $C > 0$, we can choose a small $\varepsilon > 0$ and a large $\lambda > 0$ such that

$$C - \frac{C'\varepsilon}{2} > 0 \quad \text{and} \quad \lambda - \frac{C'}{2\varepsilon} > 0.$$

Let $a := \min\big\{C - \frac{C'\varepsilon}{2},\, \lambda - \frac{C'}{2\varepsilon}\big\} > 0$, then we continue the above inequality and obtain

$$\big(\,(\mathcal{L}_\lambda u)(s) \mid u(s)\,\big)_{(H^1)^*,\,H^1} \geq a\,\|u(s)\|_{H^1}^2,$$

which completes the proof. □

With the key Lemma 5.2.6, we complete our first step at once by applying Theorem 5.2.2 as the following theorem shows.

Theorem 5.2.7. *Let $p \geq 2$. Suppose that the parametrized initial function Q has values in $L^2(\Omega)$. If the Eulerian diffusion matrix is symmetric and positive definite and satisfies*

$$\max\Big\{\sup_{s\,\in\,[0,\,t]} \|U_{\varphi(s)}\|_\infty,\, \sup_{s\,\in\,[0,\,t]} \|U_{\varphi(s)}^{-1}\|_\infty\Big\} < \infty,$$

then there exists $\lambda(\varphi) > 0$ such that for all $f \in L^2\big([0,t],(H^1)^*\big)$ and $\theta \in \Theta$ the initial value problem

$$\begin{cases} \partial_s\, u + \mathcal{L}_\lambda\, u = f \\ u(0) = Q(\theta) \end{cases} \tag{5.9}$$

has a unique solution in $L^2\big([0,t],H^1\big)$, where $\mathcal{L}_\lambda : L^2\big([0,t],H^1\big) \to L^2\big([0,t],(H^1)^*\big)$ is given by

$$\begin{aligned}
((\mathcal{L}_\lambda\, u \mid u')) = \int_0^t \Big(&\lambda\,\big\langle u(s),\, u'(s)\big\rangle_{L^2} + \big\langle W_{\varphi(s)}\,\nabla u(s),\, \nabla u'(s)\big\rangle_{L^2} \\
&- \left\langle u(s)\, W_{\varphi(s)}\, \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)},\, \nabla u'(s)\right\rangle_{L^2}\Big)\, ds.
\end{aligned} \tag{5.7}$$

and $W_{\varphi(s)} = \big(D\varphi(s)\big)^{-1}\,\big(U_{\varphi(s)} \circ \varphi(s)\big)\; D\varphi(s)^{-\top}$.

Proof. As in the proof of Lemma 5.2.6, we have

$$\max\Big\{ \sup_{s\in[0,t]} \|\varphi(s) - id\|_{2,\infty},\ \sup_{s\in[0,t]} \|\varphi(s)^{-1} - id\|_{1,\infty},\ \sup_{s\in[0,t]} \|W_{\varphi(s)}\|_\infty \Big\} < \infty,$$

which immediately implies $\mathcal{L}_\lambda \in \mathscr{L}\big(L^2\big([0,t],H^1\big),\, L^2\big([0,t],(H^1)^*\big)\big)$ for all $\lambda \in \mathbb{R}$ since

$$\begin{aligned}
&\big|\, (\!(\mathcal{L}_\lambda u \mid u')\!)\,\big| \\
&\le C \int_0^t \Big(\|u(s)\|_{L^2}\, \|u'(s)\|_{L^2} + \|\nabla u(s)\|_{L^2}\, \|\nabla u'(s)\|_{L^2} + \|u(s)\|_{L^2}\, \|\nabla u'(s)\|_{L^2} \Big)\, ds \\
&\le C \int_0^t 3\, \|u(s)\|_{H^1}\, \|u'(s)\|_{H^1}\, ds \\
&\le 3\, C\ \|u\|_{L^2([0,t],\,H^1)}\, \|u'\|_{L^2([0,t],\,H^1)}.
\end{aligned}$$

With $\lambda(\varphi) > 0$ and $a(\varphi) > 0$ chosen by Lemma 5.2.6, it follows that

$$(\!(\mathcal{L}_\lambda u \mid u)\!) \ge \int_0^t a\, \|u(s)\|_{H^1}^2\, ds = a\, \|u\|_{L^2([0,t],\,H^1)}^2,$$

which shows that $\mathcal{L}_\lambda$ is coercive. We conclude that the initial value problem (5.9) has a unique solution in $L^2\big([0,t],H^1\big)$ by invoking Theorem 5.2.2. □

We prepare to enter the second step. With a fixed λ chosen by Lemma 5.2.6, we consider the initial value problem on a subinterval $[a,b] \subset [0,t]$ given $f \in L^2\big([a,b],(H^1)^*\big)$ and $u_a \in L^2(\Omega)$ as follows:

$$\text{on } [a,b]: \ \begin{cases} \partial_s u + \mathcal{L}_\lambda u = f \\ u(a) = u_a \end{cases}. \tag{5.11}$$

In the proof of Lemma 5.2.6, we observe that λ is chosen according to the uniform bound of $\varphi(s)$, $\varphi(s)^{-1}$, $U_{\varphi(s)}$, and $U_{\varphi(s)}^{-1}$ on the entire interval $[0,t]$, so we can still go through the same proof and obtain the same λ when we replace the interval $[0,t]$ by any subinterval in the proof. We summarize this observation into the following corollary.

Corollary 5.2.8. *Let λ be chosen as in Lemma 5.2.6. For all $f \in L^2\big([a,b],(H^1)^*\big)$ and $u_a \in L^2(\Omega)$, the initial value problem* (5.11) *has a unique solution in $L^2\big([a,b],H^1\big)$.*

We present our second step in the following theorem, whose proof is inspired by [41, Section 9.2.1, Theorem 2].

Theorem 5.2.9. *Suppose that the conditions in Theorem 5.2.7 are satisfied, and let λ be chosen accordingly. If the reaction function R is bounded and Lipschitz, then for all $\theta \in \Theta$*

the initial value problem

$$\begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(u) \\ u(0) = Q(\theta) \end{cases} \tag{5.6}$$

has a unique solution $u \in L^2\big([0,t], H^1\big)$*, where* $g_\lambda(u) \in L^2\big([0,t], L^2\big) \subset L^2\big([0,t], (H^1)^*\big)$ *is given by*

$$g_\lambda(u)(s) = e^{-\lambda s}\, R\bigg(\frac{e^{\lambda s}\, u(s)}{\det D\varphi(s)}\bigg)\, \det D\varphi(s) \qquad \textit{for almost every } s \in [0,t]. \tag{5.8}$$

Proof. We first fix an $s_0 \in [0,t]$, and let $\eta > 0$ be arbitrary. Define

$$X := L^2\big([s_0, s_0+\eta], H^1\big) \cap C\big([s_0, s_0+\eta], L^2\big).$$

Given $w \in X$ and $u_{s_0} \in L^2(\Omega)$, we consider the initial value problem

$$\text{on } [s_0, s_0+\eta]: \begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(w) \\ u(s_0) = u_{s_0} \end{cases}. \tag{5.12}$$

Corollary 5.2.8 and Remark 5.2.3 yield a unique solution $u_w \in X$ to the problem (5.12). We can thus define a mapping $\mathcal{F} : X \to X$ by $\mathcal{F}(w) = u_w$. We equip the set X with the complete metric

$$d(u,u') := \sup_{s \in [s_0,\, s_0+\eta]} \|u(s) - u'(s)\|_{L^2}.$$

We are going to show that $\mathcal{F}$ is a contraction when η is small and apply the Banach fixed point theorem.

Let w, $w' \in X$. We want to show that there exists $c < 1$ such that $d(\mathcal{F}(w), \mathcal{F}(w')) \le c\, d(w,w')$. Denote by $u := \mathcal{F}(w)$ and $u' := \mathcal{F}(w')$ the solutions to (5.12) given w and w' respectively, that is, u and u' satisfy

$$\begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(w) \\ u(s_0) = u_{s_0} \end{cases} \quad \text{and} \quad \begin{cases} \partial_s u' + \mathcal{L}_\lambda u' = g_\lambda(w') \\ u'(s_0) = u_{s_0} \end{cases}.$$

Subtracting the two operational equations in $L^2\big([s_0, s_0+\eta], (H^1)^*\big)$ and evaluating at $u-u'$, we obtain that for almost every $s \in [s_0, s_0+\eta]$

$$\begin{aligned} &\Big(\big(\partial_s(u-u')\big)(s) \,\Big|\, (u-u')(s)\Big)_{(H^1)^*,\, H^1} + \Big(\big(\mathcal{L}_\lambda(u-u')\big)(s) \,\Big|\, (u-u')(s)\Big)_{(H^1)^*,\, H^1} \\ &= \Big(\big(g_\lambda(w) - g_\lambda(w')\big)(s) \,\Big|\, (u-u')(s)\Big)_{(H^1)^*,\, H^1}. \end{aligned} \tag{5.13}$$

We recall the almost pointwise coercivity of $\mathcal{L}_\lambda$ from Lemma 5.2.6:

$$\Big(\big(\mathcal{L}_\lambda(u-u')\big)(s) \,\Big|\, (u-u')(s)\Big)_{(H^1)^*,\, H^1} \ge a\, \|(u-u')(s)\|_{H^1}^2. \tag{5.14}$$

With Lemma 5.2.4, the inequality (5.14), the equation (5.13), and Lemma A.2.1 we have

$$\begin{aligned}
&\frac{1}{2}\big(\partial_s \|(u-u')(\cdot)\|_{L^2}^2\big)(s) + a\,\|(u-u')(s)\|_{H^1}^2 \\
&\le \Big(\big(\partial_s(u-u')\big)(s)\,\Big|\,(u-u')(s)\Big)_{(H^1)^*,\,H^1} + \Big(\big(\mathcal{L}_\lambda(u-u')\big)(s)\,\Big|\,(u-u')(s)\Big)_{(H^1)^*,\,H^1} \\
&\le \big\|\big(g_\lambda(w)-g_\lambda(w')\big)(s)\big\|_{(H^1)^*}\,\big\|(u-u')(s)\big\|_{H^1} \\
&\le \frac{1}{2\,(2a)}\,\big\|\big(g_\lambda(w)-g_\lambda(w')\big)(s)\big\|_{(H^1)^*}^2 + \frac{2a}{2}\,\big\|(u-u')(s)\big\|_{H^1}^2. \qquad (5.15)
\end{aligned}$$

Note that the Lipschitz assumption on the reaction function R implies

$$\big\|\big(g_\lambda(w)-g_\lambda(w')\big)(s)\big\|_{(H^1)^*}^2 \;\le\; C\,\big\|\big(g_\lambda(w)-g_\lambda(w')\big)(s)\big\|_{L^2}^2 \;\le\; C\,\|(w-w')(s)\|_{L^2}^2. \quad (5.16)$$

Combining estimates (5.15) and (5.16) gives us

$$\big(\partial_s \|(u-u')(\cdot)\|_{L^2}^2\big)(s) \le C\,\|(w-w')(s)\|_{L^2}^2.$$

Consequently,

$$\begin{aligned}
\|(u-u')(s)\|_{L^2}^2 &= \|(u-u')(s_0)\|_{L^2}^2 + \int_{s_0}^{s} (\partial_s \|(u-u')(\cdot)\|_{L^2}^2)(s')\,ds' \\
&\le 0 \;+ \int_{s_0}^{s} C\,\|(w-w')(s')\|_{L^2}^2\,ds' \\
&\le C\;\eta\;d^2(w,w'),
\end{aligned}$$

which, by taking the supremum over $[s_0, s_0+\eta]$, further gives

$$d(\mathcal{F}(w),\mathcal{F}(w')) = d(u,u') \le \sqrt{C\,\eta}\;d(w,w').$$

Therefore, $\mathcal{F}$ is a contraction when $\sqrt{C\,\eta} < 1$. With such chosen $\eta > 0$, the problem (5.12) has a unique solution in $L^2\big([s_0, s_0+\eta], H^1\big) \cap C([s_0, s_0+\eta], L^2)$.

Since η is independent of $s \in [0,t]$, there exists a unique solution on any subinterval with length less than or equal to η. In particular, we consider N subintervals $I_1 := [0,\eta]$, $I_2 := [\eta/2, 3\eta/2]$, $I_3 := [\eta, 2\eta]$, …, $I_N := [\frac{N-1}{2}\,\eta, t]$, whose union equals to $[0,t]$ and $|I_i| \le \eta$. Denote the unique solution on $[0,\eta]$ with the initial condition $Q(\theta) \in L^2(\Omega)$ by u_1, denote the unique solution on $[\eta/2, 3\eta/2]$ with the initial condition $u_1(\eta/2) \in L^2(\Omega)$ by u_2, and so on. After constructing solutions $u_1, \ldots, u_N$ on subintervals, we let

$$u(s) := \begin{cases} u_1(s), & \text{if } t \in I_1\,; \\ \;\vdots & \\ u_N(s), & \text{if } t \in I_N\,, \end{cases}$$

which, by uniqueness of solutions, is well defined on overlapping intervals. We show that u is a solution to the problem

$$\text{on } [0,t]: \begin{cases} \partial_s u + \mathcal{L}_\lambda u = g_\lambda(u) \\ u(0) = Q(\theta) \end{cases}. \tag{5.6}$$

By construction, we have $u \in L^2\big([0,t], H^1\big) \cap C\big([0,t], L^2\big)$ and $u(0) = Q(\theta)$, so it remains to verify that u satisfies the operational equation in $L^2\big([0,t], (H^1)^*\big)$. Let $\{\rho_i\}_{i=1}^N$ be a C^∞ partition of unity subordinate to the closed intervals $\{I_i\}_{i=1}^N$ with suitable adjustments at endpoints of intervals. For any $\psi \in L^2\big([0,t], H^1\big)$, since $\rho_i\,\psi \in L^2\big(I_i, H^1\big)$, we obtain from the construction of u that

$$\begin{aligned}(\partial_s u + \mathcal{L}_\lambda u \mid \psi) &= \Big(\partial_s u + \mathcal{L}_\lambda u \;\Big|\; \sum_{i=1}^N \rho_i\,\psi\Big) = \sum_{i=1}^N (\partial_s u_i + \mathcal{L}_\lambda u_i \mid \rho_i\,\psi) \\ &= \sum_{i=1}^N (g_\lambda(u_i) \mid \rho_i\,\psi) = \Big(g_\lambda(u) \;\Big|\; \sum_{i=1}^N \rho_i\,\psi\Big) = (g_\lambda(u) \mid \psi),\end{aligned}$$

and we conclude that u is a solution to (5.6). From the uniqueness of solutions on subintervals, it is clear that the solution u is unique, and the proof is complete. □

We summarize the results of this section into the following theorem.

Theorem 5.2.10. *Let $p \geq 2$. Suppose that:*

- *The parametrized initial function Q has values in $L^2(\Omega)$.*
- *For all $\varphi \in C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, the Eulerian diffusion matrix is symmetric and positive definite and satisfies*

$$\max\Big\{ \sup_{s\,\in\,[0,t]} \|U_{\varphi(s)}\|_\infty,\ \sup_{s\,\in\,[0,t]} \|U^{-1}_{\varphi(s)}\|_\infty \Big\} < \infty.$$

- *The reaction function R is bounded and Lipschitz.*

Then the solution mapping

$$\mathcal{T}: \bigcup_{t\,\in\,[0,T]} C([0,t], \mathrm{Diff}^p_{id}(\mathbb{R}^d)) \times \Theta \to \bigcup_{t\,\in\,[0,T]} \Big(L^2\big([0,t], H^1\big) \cap C\big([0,t], L^2\big)\Big)$$

is well defined.

5.2.4 Existence of minimizers

We aim to apply Corollay 4.1.12 as in the problem of piecewise-rigid motion, although verifying the conditions related to $\beta^t_{\varphi,\,\theta}$ becomes much more involved. We first tackle difficulties

arising from the PDE model in Theorem 5.2.11, then we prove the compatibility of Θ and $\mathcal{M}_A$ in Proposition 5.2.14 by checking that the conditions in Corollay 4.1.12 are satisfied. The existence of minimizers will then follow. According to Theorem 5.2.10, we denote the unique solution to the PDE model by $\tau_{\varphi,\theta} := \mathcal{T}(\varphi,\theta)$

Theorem 5.2.11. *Let $p \geq 2$. Suppose that the parametrized initial function $Q : \Theta \to L^2(\Omega)$ and the reaction function R are both bounded and Lipschitz. Given $r > 0$, suppose that there are constants $b(r), \ell(r) > 0$ such that for all $\xi, \zeta \in \overline{\mathfrak{B}}(id, r)$ the Eulerian diffusion matrix is symmetric and positive definite and satisfies*

$$\max\left\{\|U_\xi\|_\infty,\, \|U_\xi^{-1}\|_\infty\right\} \leq b \quad \text{and} \quad \|U_\xi \circ \xi - U_\zeta \circ \zeta\|_\infty \leq \ell\, \|\xi - \zeta\|_{p,\infty}^{\Omega}.$$

Then there exists a constant $c(r) > 0$ such that

$$\|\tau_{\varphi,\theta}(t) - \tau_{\psi,\omega}(t)\|_{L^2} \leq c\left(\sup_{s\in[0,t]} \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} + |\theta - \omega|\right)$$

for all $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ and all $\theta, \omega \in \Theta$.

Proof. We prepare some preliminary estimates in the following two claims. From the assumption of the Eulerian diffusion matrix, Lemma 5.2.6, and Theorem 5.2.9, there exist $\lambda(r) > 0$ and $a(r) > 0$ such that for all $\varphi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ and all $u \in L^2\big([0,t], H^1\big)$

$$\big(\,(\mathcal{L}_{\varphi,\lambda} u)(s) \mid u(s)\,\big)_{(H^1)^*,\, H^1} \geq a\, \|u(s)\|_{H^1}^2 \quad \text{for almost every } s \in [0,t],$$

and such that for all $\varphi \in C([0,t], \overline{\mathfrak{B}}(id, r))$ and $\theta \in \Theta$ there exists a unique solution $u_{\varphi,\theta}$ to the initial value problem

$$\begin{cases} \partial_s u + \mathcal{L}_{\varphi,\lambda} u = g_{\varphi,\lambda}(u) \\ u(0) = Q(\theta) \end{cases}. \tag{5.6}$$

We recall

$$\begin{aligned} (\!(\mathcal{L}_{\varphi,\lambda} u \mid u')\!) = \int_0^t \Big(& \lambda \left\langle u(s),\, u'(s)\right\rangle_{L^2} + \left\langle W_{\varphi(s)}\, \nabla u(s),\, \nabla u'(s)\right\rangle_{L^2} \\ & - \left\langle u(s)\, W_{\varphi(s)}\, \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)},\, \nabla u'(s)\right\rangle_{L^2}\Big)\, ds \end{aligned} \tag{5.7}$$

and $W_{\varphi(s)} = \big(D\varphi(s)\big)^{-1}\, \big(U_{\varphi(s)} \circ \varphi(s)\big)\, D\varphi(s)^{-\top}$, where we have changed the notation from $\mathcal{L}_\lambda$ to $\mathcal{L}_{\varphi,\lambda}$ to accommodate our current discussion on various $\varphi \in C([0,t], \overline{\mathfrak{B}}(id, r))$.

Claim 5.2.12. *For all $\varphi \in C([0,t], \overline{\mathfrak{B}}(id, r))$, all $u \in L^2\big([0,t], H^1\big)$, and almost every*

$s \in [0,t]$, we have

$$\|(\mathcal{L}_{\varphi,\lambda}u)(s) - (\mathcal{L}_{\psi,\lambda}u)(s)\|_{(H^1)^*} \le C_r \, \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \, \|u(s)\|_{H^1}.$$

Proof. For an arbitrary $v \in H^1(\Omega)$, direct computation gives

$$\begin{aligned}
&\left| \Big((\mathcal{L}_{\varphi,\lambda}u)(s) - (\mathcal{L}_{\psi,\lambda}u)(s) \;\Big|\; v \Big)_{(H^1)^*,\,H^1} \right| \\
&\le \left| \big\langle (W_{\varphi(s)} - W_{\psi(s)})\, \nabla u(s),\, \nabla v \big\rangle_{L^2(\Omega)} \right| \\
&\quad + \left| \left\langle u(s) \left(W_{\varphi(s)} \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)} - W_{\psi(s)} \frac{\nabla(\det D\psi(s))}{\det D\psi(s)} \right), \nabla v \right\rangle_{L^2(\Omega)} \right|.
\end{aligned}$$

Since $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$, $p \ge 2$, and $\sup_{s\,\in\,[0,\,t]} \|U_{\varphi(s)}\|_\infty \le b$, for all $s \in [0,t]$ we have

$$\max\left\{ \|W_{\varphi(s)}\|_\infty,\, \|W_{\psi(s)}\|_\infty,\, \left\| \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)} \right\|_\infty,\, \left\| \frac{\nabla(\det D\psi(s))}{\det D\psi(s)} \right\|_\infty \right\} \le C_r$$

and

$$\left\| \frac{\nabla(\det D\varphi(s))}{\det D\varphi(s)} - \frac{\nabla(\det D\psi(s))}{\det D\psi(s)} \right\|_\infty^{\Omega} \le C_r \, \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega}.$$

The Lipschitz assumption of the Eulerian diffusion matrix implies

$$\|W_{\varphi(s)} - W_{\psi(s)}\|_\infty^{\Omega} \le C_r \, \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega}.$$

We conclude that

$$\left| \Big((\mathcal{L}_{\varphi,\lambda}u)(s) - (\mathcal{L}_{\psi,\lambda}u)(s) \;\Big|\; v \Big)_{(H^1)^*,\,H^1} \right| \le C_r \, \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \, \|u(s)\|_{H^1} \, \|v\|_{H^1}.$$

◪

We also recall

$$g_{\varphi,\lambda}(u)(s) = e^{-\lambda s} \, R\!\left(\frac{e^{\lambda s}\, u(s)}{\det D\varphi(s)} \right) \det D\varphi(s) \quad \text{for almost every } s \in [0,t], \tag{5.8}$$

where we have changed the notation from g_λ to $g_{\varphi,\lambda}$.

Claim 5.2.13. *For all $\varphi \in C([0,t], \overline{\mathfrak{B}}(id,r))$ and all $\theta \in \Theta$, we have*

$$\|u_{\varphi,\theta}(s)\|_{L^2} \le C_r \;\; \textit{for all } s \in [0,t] \quad \textit{and} \quad \int_0^t \|u_{\varphi,\theta}(s)\|_{H^1}^2 \, ds \le C_r'.$$

Proof. Since $u_{\varphi,\theta}$ satisfies

$$\begin{cases} (\partial_s u_{\varphi,\theta})(s) + (\mathcal{L}_{\varphi,\lambda}\, u_{\varphi,\theta})(s) = g_{\varphi,\lambda}(u_{\varphi,\theta})(s) & \text{for almost every } s \in [0,t] \\ u_{\varphi,\theta}(0) = Q(\theta) \end{cases},$$

we have from Lemma 5.2.4 and the almost pointwise coercivity of $\mathcal{L}_{\varphi,\lambda}$ that

$$\begin{aligned}
&\frac{1}{2}\left(\partial_s \|u_{\varphi,\theta}(\cdot)\|_{L^2}^2\right)(s) \\
&\leq \Big((\partial_s u_{\varphi,\theta})(s) \;\Big|\; u_{\varphi,\theta}(s)\Big)_{(H^1)^*,\,H^1} + \Big((\mathcal{L}_{\varphi,\lambda}\, u_{\varphi,\theta})(s) \;\Big|\; u_{\varphi,\theta}(s)\Big)_{(H^1)^*,\,H^1} \\
&= \Big\langle g_{\varphi,\lambda}(u_{\varphi,\theta})(s),\, u_{\varphi,\theta}(s)\Big\rangle_{L^2(\Omega)} \\
&\leq \frac{C_r}{2}\,\|R\|_\infty^2\,\mathrm{vol}(\Omega) + \frac{1}{2}\,\|u_{\varphi,\theta}(s)\|_{L^2}^2. \qquad (5.17)
\end{aligned}$$

It follows that

$$\begin{aligned}
\|u_{\varphi,\theta}(s)\|_{L^2}^2 &= \|u_{\varphi,\theta}(0)\|_{L^2}^2 + \int_0^s \left(\partial_s \|u_{\varphi,\theta}(\cdot)\|_{L^2}^2\right)(s')\,ds' \\
&\leq \sup_{\theta\in\Theta} \|Q(\theta)\|_{L^2}^2 + C_r\,\|R\|_\infty^2\,\mathrm{vol}(\Omega)\;T + \int_0^s \|u_{\varphi,\theta}(s')\|_{L^2}^2\,ds',
\end{aligned}$$

so Gronwall's lemma gives

$$\|u_{\varphi,\theta}(s)\|_{L^2} \leq \Big(\sup_{\theta\in\Theta} \|Q(\theta)\|_{L^2}^2 + C_r\,\|R\|_\infty^2\,\mathrm{vol}(\Omega)\;T\Big)\,\exp(T) = C_r', \qquad (5.18)$$

which proves the first part of the claim.

The almost pointwise coercivity of $\mathcal{L}_{\varphi,\lambda}$ and the inequality (5.17) also shows that

$$\begin{aligned}
&\frac{1}{2}\left(\partial_s \|u_{\varphi,\theta}(\cdot)\|_{L^2}^2\right)(s) + a\,\|u_{\varphi,\theta}(s)\|_{H^1}^2 \\
&\leq \Big((\partial_s u_{\varphi,\theta})(s) \;\Big|\; u_{\varphi,\theta}(s)\Big)_{(H^1)^*,\,H^1} + \Big((\mathcal{L}_{\varphi,\lambda}\, u_{\varphi,\theta})(s) \;\Big|\; u_{\varphi,\theta}(s)\Big)_{(H^1)^*,\,H^1} \\
&\leq \frac{C_r}{2}\,\|R\|_\infty^2\,\mathrm{vol}(\Omega) + \frac{1}{2}\,\|u_{\varphi,\theta}(s)\|_{L^2}^2,
\end{aligned}$$

which implies

$$a\,\|u_{\varphi,\theta}(s)\|_{H^1}^2 \leq \frac{C_r}{2}\,\|R\|_\infty^2\,\mathrm{vol}(\Omega) + \frac{1}{2}\,\|u_{\varphi,\theta}(s)\|_{L^2}^2 - \frac{1}{2}\left(\partial_s \|u_{\varphi,\theta}(\cdot)\|_{L^2}^2\right)(s).$$

Integrating over $[0,t]$, Lemma 5.2.4 and the estimate (5.18) then yield

$$\begin{aligned}
&a\int_0^t \|u_{\varphi,\theta}(s)\|_{H^1}^2\,ds \\
&\leq \frac{C_r}{2}\,\|R\|_\infty^2\,\mathrm{vol}(\Omega)\;T + \frac{1}{2}\int_0^t \|u_{\varphi,\theta}(s)\|_{L^2}^2\,ds - \frac{1}{2}\left(\|u_{\varphi,\theta}(t)\|_{L^2}^2 - \|u_{\varphi,\theta}(0)\|_{L^2}^2\right) \\
&\leq \frac{C_r}{2}\,\|R\|_\infty^2\,\mathrm{vol}(\Omega)\;T + \left(\frac{T}{2}+1\right)C_r',
\end{aligned}$$

that is, $\int_0^t \|u_{\varphi,\theta}(s)\|_{H^1}^2\,ds \leq C_r''$, which proves the second part of the claim. ◪

We are now ready to estimate $\|\tau_{\varphi,\theta}(t) - \tau_{\psi,\omega}(t)\|_{L^2}$. Recall the change of function $u(t) = e^{-\lambda t}\,\tau(t)\,\det D\varphi(t)$, thus it suffices to show that

$$\|u_{\varphi,\theta}(t) - u_{\psi,\omega}(t)\|_{L^2} \leq C_r \left(\sup_{s\,\in\,[0,\,t]} \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} + |\theta - \omega| \right)$$

since $\varphi, \psi \in C([0,t], \overline{\mathfrak{B}}(id, r))$, $p \geq 2$, and $\|u_{\varphi,\theta}(t)\|_{L^2}, \|u_{\psi,\omega}(t)\|_{L^2} \leq C_r$ by Claim 5.2.13.

The unique solutions $u_{\varphi,\theta}$ and $u_{\psi,\omega}$ satisfy

$$\begin{cases} (\partial_s u_{\varphi,\theta})(s) + (\mathcal{L}_{\varphi,\lambda}\, u_{\varphi,\theta})(s) = g_{\varphi,\lambda}(u_{\varphi,\theta})(s) & \text{for almost every } s \in [0,t] \\ u_{\varphi,\theta}(0) = Q(\theta) \end{cases}$$

and

$$\begin{cases} (\partial_s u_{\psi,\omega})(s) + (\mathcal{L}_{\psi,\lambda}\, u_{\psi,\omega})(s) = g_{\psi,\lambda}(u_{\psi,\omega})(s) & \text{for almost every } s \in [0,t] \\ u_{\psi,\omega}(0) = Q(\omega) \end{cases}.$$

Lemma 5.2.4 and the almost pointwise coercivity of $\mathcal{L}_{\varphi,\lambda}$ again give

$$\begin{aligned} &\frac{1}{2}\left(\partial_s \|(u_{\varphi,\theta} - u_{\psi,\omega})(\cdot)\|_{L^2}^2\right)(s) + a\,\|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{H^1}^2 \\ &\leq \Big(\big(\partial_s (u_{\varphi,\theta} - u_{\psi,\omega})\big)(s) \;\Big|\; (u_{\varphi,\theta} - u_{\psi,\omega})(s) \Big)_{(H^1)^*,\,H^1} \\ &\quad + \Big(\big(\mathcal{L}_{\varphi,\lambda}\,(u_{\varphi,\theta} - u_{\psi,\omega})\big)(s) \;\Big|\; (u_{\varphi,\theta} - u_{\psi,\omega})(s) \Big)_{(H^1)^*,\,H^1} \\ &= -\Big(\big((\mathcal{L}_{\varphi,\lambda} - \mathcal{L}_{\psi,\lambda})\,u_{\psi,\omega}\big)(s) \;\Big|\; (u_{\varphi,\theta} - u_{\psi,\omega})(s) \Big)_{(H^1)^*,\,H^1} \\ &\quad + \Big\langle g_{\varphi,\lambda}(u_{\varphi,\theta})(s) - g_{\psi,\lambda}(u_{\psi,\omega})(s),\, (u_{\varphi,\theta} - u_{\psi,\omega})(s) \Big\rangle_{L^2(\Omega)}. \end{aligned}$$

From

$$g_{\varphi,\lambda}(u)(s) = e^{-\lambda s}\, R\!\left(\frac{e^{\lambda s}\,u(s)}{\det D\varphi(s)}\right) \det D\varphi(s) \quad \text{for almost every } s \in [0,t] \tag{5.8}$$

and the assumption that R is bounded and Lipschitz, we obtain

$$\|g_{\varphi,\lambda}(u_{\varphi,\theta})(s) - g_{\psi,\lambda}(u_{\psi,\omega})(s)\|_{L^2(\Omega)} \leq C_r \left(\|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2} + \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right).$$

Together with Claim 5.2.12, we continue the above inequality and write

$$\begin{aligned} &\frac{1}{2}\left(\partial_s \|(u_{\varphi,\theta} - u_{\psi,\omega})(\cdot)\|_{L^2}^2\right)(s) + a\,\|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{H^1}^2 \\ &\leq C_r \Big(\|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega}\; \|u_{\psi,\omega}(s)\|_{H^1}\; \|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{H^1} \\ &\qquad + \left(\|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2} + \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega}\right)\; \|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2} \Big) \end{aligned}$$

$$\leq C_r \left(\frac{1}{2\varepsilon} \left(\|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right)^2 \|u_{\psi,\omega}(s)\|_{H^1}^2 + \frac{\varepsilon}{2} \|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{H^1}^2 \right.$$
$$\left. + \|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2}^2 + \frac{1}{2} \left(\|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right)^2 + \frac{1}{2} \|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2}^2 \right).$$

By choosing $\varepsilon > 0$ such that $C_r \frac{\varepsilon}{2} < a$, we have

$$\left(\partial_s \|(u_{\varphi,\theta} - u_{\psi,\omega})(\cdot)\|_{L^2}^2\right)(s)$$
$$\leq C_r \left(\left(\|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right)^2 \left(\|u_{\psi,\omega}(s)\|_{H^1}^2 + 1 \right) + \|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2}^2 \right).$$

Thus

$$\|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2}^2$$
$$= \|(u_{\varphi,\theta} - u_{\psi,\omega})(0)\|_{L^2}^2 + \int_0^s \left(\partial_s \|(u_{\varphi,\theta} - u_{\psi,\omega})(\cdot)\|_{L^2}^2\right)(s')\, ds'$$
$$\leq \|Q(\theta) - Q(\omega)\|_{L^2}^2 + C_r \left(\sup_{s \in [0,t]} \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right)^2 \left(\int_0^t \|u_{\psi,\omega}(s)\|_{H^1}^2 \, ds + T \right)$$
$$+ C_r \int_0^s \|(u_{\varphi,\theta} - u_{\psi,\omega})(s')\|_{L^2}^2 \, ds'.$$

Applying Claim 5.2.13 and Gronwall's lemma leads to

$$\|(u_{\varphi,\theta} - u_{\psi,\omega})(s)\|_{L^2}^2 \leq C_r \left(\|Q(\theta) - Q(\omega)\|_{L^2}^2 + \left(\sup_{s \in [0,t]} \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right)^2 \right).$$

With the Lipschitz assumption of Q and the identity $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, we conclude by letting $s = t$ that

$$\|(u_{\varphi,\theta} - u_{\psi,\omega})(t)\|_{L^2} \leq C_r \left(|\theta - \omega| + \sup_{s \in [0,t]} \|\varphi(s) - \psi(s)\|_{p,\infty}^{\Omega} \right).$$

□

We now prove the existence of minimizers by applying Corollay 4.1.12. We keep the elasticity tensor and the Eulerian diffusion matrix in their abstract forms in the following proposition. The specific forms in our minimization problem will follow as a corollary.

Proposition 5.2.14. *Let $p \geq 2$. Suppose that the parametrized initial function $Q : \Theta \to L^2(\Omega)$, the reaction function R , and the atrophy function α are bounded and Lipschitz. Given $\Omega \in \mathscr{S}$, suppose in addition that:*

- *The function*

$$s \mapsto \int_{\varphi(s,\,\Omega)} \mathcal{E}_{\varphi(s)}(\varepsilon_v, \varepsilon_{v'}) \, dx$$

is continuous for all $\varphi \in C([0,t], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$ and all $v, v' \in V$.

- *For each $r > 0$ there is a constant $\ell_E(r) > 0$ such that*

$$\left| \int_{\xi(\Omega)} \mathcal{E}_\xi(\varepsilon_v, \varepsilon_{v'}) \, dx - \int_{\zeta(\Omega)} \mathcal{E}_\zeta(\varepsilon_v, \varepsilon_{v'}) \, dx \right| \leq \ell_E \, \|\xi - \zeta\|_{p,\infty}^{\Omega}$$

  for all $\xi, \zeta \in \overline{\mathfrak{B}}(id, r)$ and all $v, v' \in V$.

- *For each $r > 0$, there are constants $b(r), \ell_U(r) > 0$ such that the Eulerian diffusion matrix is symmetric and positive definite and satisfies*

$$\max\left\{ \|U_\xi\|_\infty, \|U_\xi^{-1}\|_\infty \right\} \leq b \quad \text{and} \quad \|U_\xi \circ \xi - U_\zeta \circ \zeta\|_\infty \leq \ell_U \, \|\xi - \zeta\|_{p,\infty}^{\Omega}$$

  for all $\xi, \zeta \in \overline{\mathfrak{B}}(id, r)$.

Then Θ and $\mathcal{M}_A$ are compatible.

Proof. We recall that

$$(\mathcal{A}_\varphi^s \, v \mid v') = \int_{\varphi(s,\,\Omega)} \mathcal{E}_{\varphi(s)}(\varepsilon_v, \varepsilon_{v'}) \, dx$$

and

$$(\beta_{\varphi,\,\theta}^s \mid v) = \int_{\varphi(s,\,\Omega)} \chi \;\; \alpha\big(\tau_{\varphi,\,\theta}(s) \circ \varphi(s)^{-1}\big) \;\; (-\operatorname{div} v) \; dx.$$

The conditions in Corollay 4.1.12 related to $\mathcal{A}_\varphi^s$ are satisfied by the given assumptions. In addition, the function $s \mapsto (\beta_{\varphi,\,\theta}^s \mid v)$ is clearly Lebesgue measurable. It remains to show that $\beta_{\varphi,\,\theta}^s$ is bounded and Lipschitz in φ and that the model $\mathcal{M}_A$ is continuous in θ.

Note that $\|\beta_{\varphi,\,\theta}^s\|_{V^*}$ is uniformly bounded by $\|\alpha\|_\infty \, \|\chi\|_{L^1}$. To show that $\varphi \mapsto \beta_{\varphi,\,\theta}^s$ is Lipschitz in $C([0,t], \overline{\mathfrak{B}}(id, r))$, we make a change of variables, use the Lipschitz continuity of α, Theorem 5.2.11, and the boundedness of α, then arrive at the desired result

$$\|\beta_{\varphi,\,\theta}^s - \beta_{\psi,\,\theta}^s\|_{V^*} \leq C_r \sup_{s' \,\in\, [0,\,s]} \|\varphi(s') - \psi(s')\|_{p,\infty}^{\Omega}.$$

Finally, assume that $\theta_n \to \theta$ in $\Theta \subset \mathbb{R}^k$, then for $0 \leq j \leq p$

$$\begin{aligned}
&\left| \int_0^t D^j \Big(\mathcal{M}_A(\varphi|_{[0,\,s]}, \theta_n(s))(\varphi(s,x)) \Big) \, ds - \int_0^t D^j \Big(\mathcal{M}_A(\varphi|_{[0,\,s]}, \theta(s))(\varphi(s,x)) \Big) \, ds \right| \\
&\leq \int_0^t C_\varphi \; c_V \left\| (\gamma \, \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^s)^{-1} \, (\beta_{\varphi,\,\theta_n}^s - \beta_{\varphi,\,\theta}^s) \right\|_V ds \\
&\leq C_\varphi \, |\theta_n - \theta| \to 0,
\end{aligned}$$

where the last inequality follows from $\left\| (\gamma \, \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^s)^{-1} \right\|_{\mathscr{L}(V^*,V)} \leq \frac{1}{\gamma}$ and Theorem 5.2.11. Therefore, by Corollay 4.1.12, Θ and $\mathcal{M}_A$ are compatible. □

Since our layered elasticity $\mathcal{E}_{\xi * \Phi}$ (page 43) and the specific form of the Eulerian diffusion

matrix (3.22) satisfy the conditions in Proposition 5.2.14, we have established the existence of minimizers for the problem of atrophy.

Corollary 5.2.15. *Let $p \geq 2$ and let $\Theta \subset \mathbb{R}^k$ be a compact set. Suppose that the parametrized initial function $Q : \Theta \to L^2(\Omega)$, the reaction function R, and the atrophy function α are bounded and Lipschitz. If the discrepancy function ρ is continuous with respect to $\|\cdot\|_{p,\infty}$, then the minimization problem*

$$\min_{\theta \in \Theta} \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}})$$

subject to

$$\varphi(t,x) = x + \int_0^t \mathcal{M}_A(\varphi_{|[0,\,s]}, \theta)(\varphi(s,x))\, ds \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d$$

has a minimizer.

We remark that the conditions of Proposition 5.2.14 are satisfied in particular when the elasticity tensor and the Eulerian diffusion matrix are unaffected by deformations. Examples include the isotropic elasticity tensor $\mathcal{E}_\xi := \mathcal{E}_{\text{iso}}$ (see (3.6)) and the isotropic diffusion matrix $U_\xi := r\, I_d$. More generally, the conditions of Proposition 5.2.14 can be adjusted to encompass history-dependent elasticity tensor $\mathcal{E}_{\varphi_{|[0,t]}}$ and diffusion matrix $U_{\varphi_{|[0,t]}}$, e.g., when elasticity and diffusion are altered by chemical propagation.

We recall from (3.24) and (3.27) that in our two examples we adjust our formulation by a cutoff function χ. The reason is that in order to apply Theorem 4.1.1(ii), the deformation vector fields must satisfy

$$\|v_\varphi(s)\|_V \leq g(s) \Big(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_\infty\Big). \tag{4.3}$$

If we can prove the long-time existence and uniqueness of solutions under the relaxed condition

$$\|v_\varphi(s)\|_V \leq g(s) \Big(1 + \sup_{s' \in [0,\,s]} \|\varphi(s') - id\|_{1,\infty}\Big), \tag{5.19}$$

then we can get rid of this artificial modification. The difficulty of this relaxation in our current proof arises in Claim 4.3.5 when we try to prove $\int_0^{T'} \|v_\varphi(t)\|_{p+1,\infty}\, dt < \infty$. If we only have (5.19), then

$$\varphi(t,x) = x + \int_0^t v_\varphi(s, \varphi(s,x))\, ds \quad \text{for all } x \in \mathbb{R}^d \tag{4.1}$$

will imply, for every $t' \in [0, t]$,

$$\begin{aligned}\|\varphi(t') - id\|_{1,\infty} &\leq \int_0^t \|v_\varphi(s) \circ \varphi(s)\|_{1,\infty}\, ds \\ &\leq \int_0^t C_{\varphi(s)}\ c_V\ g(s) \left(1 + \sup_{s' \in [0,\, s]} \|\varphi(s') - id\|_{1,\infty}\right) ds.\end{aligned}$$

Because of the additional term $C_{\varphi(s)}$, we cannot invoke Grönwall's lemma and follow the same approach as the one in Claim 4.3.5 to first bound $\sup_{s' \in [0,\, s]} \|\varphi(s') - id\|_{1,\infty}$ and then obtain $\int_0^{T'} \|v_\varphi(t)\|_{p+1,\infty}\, dt < \infty$. We will need a different technique to carry out the proof under the relaxed condition (5.19).

Chapter 6

Computation

This chapter is devoted to the objective function evaluation of our two examples. We discuss numerical issues and the objective function evaluation of piecewise-rigid motion in Secion 6.1. In Section 6.2, we present in detail the objective function evaluation of the atrophy model. Finally, we cover our GPU implementation to accelerate computation in Section 6.3.

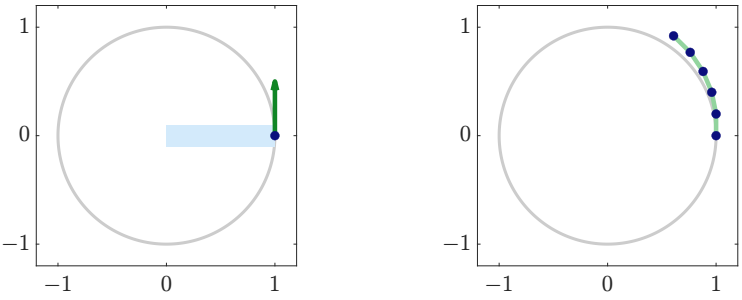
6.1 Numerical Formulation of Piecewise-rigid Motion

We have discussed the formulation of piecewise-rigid motion using the energy form in Section 5.1.1 and using constraints Section 5.1.2. However, both formulations suffer from an unwelcome numerical error illustrated in Figure 6.1. With a fixed time-dependent deformation vector field v corresponding to a piecewise-rigid motion, we recall that the motion is the solution to the initial value problem

$$\varphi(t,x) = x + \int_0^t v(s,\varphi(s,x))\,ds.$$

Hence, although φ is a piecewise-rigid motion theoretically, any numerical error introduced in the time quadrature will damage the rigidity. Imagine that a ball turns into an egg shape after it goes through a "rigid motion." In this section, we consider a numerical formulation which guarantees the rigidity of shape components and empirically leads to a piecewise-rigid motion.

We first discretize our initial shape $\Omega_0 = \bigcup_{i=1}^N \Omega_i$ into m nodes $\bigcup_{i=1}^N \{q_{j,0}^{(i)}\}_{j=1}^{m_i}$, where $m = m_1 + \cdots + m_N$ and $q_{j,0}^{(i)} \in \mathbb{R}^d$. We recall from Section 3.1.1 that for a parameter $\theta = (u,\omega)$ composed of an Eulerian linear velocity of the origin and an angular velocity, the



(a) A rectangle rotating counterclockwise rigidly.

(b) Numerical trajectory of a landmark on the rectangle.

Figure 6.1: Rigidity is deteriorated by numerical error.

rigid velocity field is given by a linear operator

$$(\mathcal{V}\,\theta)(x) = \begin{cases} u + \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x, & \text{if } d = 2\,; \\ u + \omega \times x, & \text{if } d = 3\,. \end{cases}$$

Let $(\theta^{(1)}(t), \ldots, \theta^{(N)}(t))$ be the parameter for each shape component at time t. We have seen that we may lose rigidity numerically if we compute the position of nodes at time t by

$$q_j^{(i)}(t) = q_{j,0}^{(i)} + \int_0^t \big(\mathcal{V}\,\theta^{(i)}(s)\big)\big(q_j^{(i)}(s)\big)\,ds.$$

To obtain a piecewise-rigid motion numerically, now we use our last characterization of rigid motions, Proposition 3.1.1(ii), which says that φ is a rigid motion if and only if

$$\varphi(t, x') = \varphi(t, x) + R(t)\,(x' - x)$$

for all $t \in [0, T]$ and all $x, x' \in \mathbb{R}^d$, where $R(t) \in \mathbb{R}^{d\times d}$ is a rotation. We fix x to be a chosen center of rotation c_0 and define $c(t) := \varphi(t, c_0)$, then the above equation is equivalent to

$$\varphi(t, x') = c(t) + R(t)\,(x' - c_0), \tag{6.1}$$

which motivates the definition of the following Lagrange position mapping

$$\mathcal{P}(c, \alpha, x, c_0) := c + \mathcal{R}(\alpha)\,(x - c_0),$$

where, if $d = 2$, then $\alpha \in \mathbb{R}$ and

$$\mathcal{R}(\alpha) := \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix};$$

if $d = 3$, then $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ and

$$\mathcal{R}(\alpha) := \begin{bmatrix} \cos\alpha_3 & -\sin\alpha_3 & 0 \\ \sin\alpha_3 & \cos\alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha_2 & 0 & -\sin\alpha_2 \\ 0 & 1 & 0 \\ \sin\alpha_2 & 0 & \cos\alpha_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{bmatrix}.$$

With our discretized initial shape $\bigcup_{i=1}^{N} \{q_{j,0}^{(i)}\}_{j=1}^{m_i}$ and the chosen rotation centers $\{c_0^{(i)}\}_{i=1}^{N}$, given $\theta(t) = \big(u^{(1)}(t), \omega^{(1)}(t), \ldots, u^{(N)}(t), \omega^{(N)}(t)\big)$ composed of Lagrange linear velocities of rotation centers and angular velocities for each shape component, we compute the positions of nodes by

$$\begin{cases} c^{(i)}(t) = c_0^{(i)} + \displaystyle\int_0^t u^{(i)}(s)\, ds \\ \alpha^{(i)}(t) = \displaystyle\int_0^t \omega^{(i)}(s)\, ds \\ q_j^{(i)}(t) = \mathcal{P}\big(c^{(i)}(t), \alpha^{(i)}(t), q_{j,0}^{(i)}, c_0^{(i)}\big) \end{cases}.$$

The initial angle is set to be 0 so that the deformation is identity at time 0, that is,

$$\mathcal{P}\big(c^{(i)}(0),\, \alpha^{(i)}(0),\, x,\, c_0^{(i)}\big) = \mathcal{P}\big(c_0^{(i)},\, 0,\, x,\, c_0^{(i)}\big) = x.$$

We notice that in this approach the positions of nodes at time t are computed directly by rotation and translation of the initial nodes through the position mapping $\mathcal{P}$, thus we are assured that the rigidity is preserved.

Next we consider the objective function evaluation. Recall that our minimization problem is

$$\min_{\theta \in L^2([0,T],\mathbb{R}^k)} \left(\frac{1}{2} \int_0^T \Big(\|v(t)\|_V^2 + |\theta(t)|^2 \Big)\, dt + \rho\big(\varphi_\theta(T, \Omega_0), \Omega_{\text{targ}}\big) \right)$$

with $v(t,x) \approx (\mathcal{V}\,\theta_i(t))(x)$ for $x \in \varphi_\theta(t, \Omega_i)$. Here we focus on the computation of $\|v(t)\|_V^2$. Before we discuss this computation, we first show how to compute velocities at nodes using our new formulation. We have from (6.1) that

$$\dot\varphi(t, x') = \dot c(t) + \dot R(t)\,(x' - c_0),$$

which gives rise to the definition of the Lagrange velocity mapping

$$\dot{\mathcal{P}}(u, \omega, \alpha, x, c_0) := u + \dot{\mathcal{R}}(\omega, \alpha)\,(x - c_0),$$

where, if $d = 2$, then $\omega, \alpha \in \mathbb{R}$ and

$$\dot{\mathcal{R}}(\omega, \alpha) := \omega \begin{bmatrix} -\sin\alpha & -\cos\alpha \\ \cos\alpha & -\sin\alpha \end{bmatrix};$$

if $d = 3$, then $\omega = (\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3$, $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$, and

$$\begin{aligned}\dot{\mathcal{R}}(\omega, \alpha) := \omega_1 &\begin{bmatrix} \cos\alpha_3 & -\sin\alpha_3 & 0 \\ \sin\alpha_3 & \cos\alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha_2 & 0 & -\sin\alpha_2 \\ 0 & 1 & 0 \\ \sin\alpha_2 & 0 & \cos\alpha_2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\alpha_1 & -\cos\alpha_1 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \end{bmatrix} \\ + \omega_2 &\begin{bmatrix} \cos\alpha_3 & -\sin\alpha_3 & 0 \\ \sin\alpha_3 & \cos\alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\sin\alpha_2 & 0 & -\cos\alpha_2 \\ 0 & 0 & 0 \\ \cos\alpha_2 & 0 & -\sin\alpha_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{bmatrix} \\ + \omega_3 &\begin{bmatrix} -\sin\alpha_3 & -\cos\alpha_3 & 0 \\ \cos\alpha_3 & -\sin\alpha_3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\alpha_2 & 0 & -\sin\alpha_2 \\ 0 & 1 & 0 \\ \sin\alpha_2 & 0 & \cos\alpha_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{bmatrix}. \end{aligned}$$

We hence compute velocities at nodes by

$$\begin{cases} \alpha^{(i)}(t) = \displaystyle\int_0^t \omega^{(i)}(s)\, ds \\ v_j^{(i)}(t) = \dot{\mathcal{P}}\big(u^{(i)}(t), \omega^{(i)}(t), \alpha^{(i)}(t), q_{j,0}^{(i)}, c_0^{(i)}\big) \end{cases}.$$

With velocities at nodes at hand, we now examine the discretized minimization problem

$$\min_{\theta \in L^2([0,T], \mathbb{R}^k)} \left(\frac{1}{2} \int_0^T \Big(\|v(t)\|_V^2 + |\theta(t)|^2 \Big)\, dt + \rho\Big(\{q_j^{(i)}(T)\}_{i,j}, \{q_{j,\text{targ}}^{(i)}\}_{i,j} \Big) \right).$$

After discretization, note that ρ becomes a function of $\{q_j^{(i)}(T)\}_{i,j}$, instead of a function of $\varphi_\theta(T, \Omega_0)$. Thus only the end points of trajectories of nodes will affect the discretized discrepancy function. It follows that only velocities at nodes, which affect the trajectories of nodes, will affect the discretized discrepancy function. Thus the discretized minimization problem is equivalent to

$$\min_{\theta \in L^2([0,T], \mathbb{R}^k)} \left(\frac{1}{2} \int_0^T \Big(\|v^*(t)\|_V^2 + |\theta(t)|^2 \Big)\, dt + \rho\Big(\{q_j^{(i)}(T)\}_{i,j}, \{q_{j,\text{targ}}^{(i)}\}_{i,j} \Big) \right),$$

where

$$v^*(t) = \operatorname*{arg\,min}_{v' \in V} \Big\{ \|v'\|_V^2 : v'\big(q_j^{(i)}(t)\big) = v_j^{(i)}(t) \text{ for all } i, j \Big\}.$$

Suppose that the reproducing kernel k_V of V is positive definite. Similar to the soft interpolation problem at the end of Section 2.4.1, it can be shown (see [104, Lemma 8.6, Lemma 8.7, Theorem 8.8]) that

$$v^*(t)(\cdot) = \sum_{i=1}^{N} \sum_{j=1}^{m_i} k_V\big(\cdot, q_j^{(i)}(t)\big)\, \beta_j^{(i)}(t),$$

where

$$\begin{bmatrix} \beta_1^{(1)}(t) \\ \beta_2^{(1)}(t) \\ \vdots \\ \beta_{m_N}^{(N)}(t) \end{bmatrix} = \begin{bmatrix} k_V\big(q_1^{(1)}(t), q_1^{(1)}(t)\big) & k_V\big(q_1^{(1)}(t), q_2^{(1)}(t)\big) & \cdots & k_V\big(q_1^{(1)}(t), q_{m_N}^{(N)}(t)\big) \\ k_V\big(q_2^{(1)}(t), q_1^{(1)}(t)\big) & k_V\big(q_2^{(1)}(t), q_2^{(1)}(t)\big) & \cdots & k_V\big(q_2^{(1)}(t), q_{m_N}^{(N)}(t)\big) \\ \vdots & \vdots & & \vdots \\ k_V\big(q_{m_N}^{(N)}(t), q_1^{(1)}(t)\big) & k_V\big(q_{m_N}^{(N)}(t), q_2^{(1)}(t)\big) & \cdots & k_V\big(q_{m_N}^{(N)}(t), q_{m_N}^{(N)}(t)\big) \end{bmatrix}^{-1} \begin{bmatrix} v_1^{(1)}(t) \\ v_2^{(1)}(t) \\ \vdots \\ v_{m_N}^{(N)}(t) \end{bmatrix},$$

which will be denoted by $\boldsymbol{\beta}(t) = K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)$. It follows that

$$\|v^*(t)\|_V^2 = \boldsymbol{\beta}(t)^\top K_V(\boldsymbol{q}(t))\,\boldsymbol{\beta}(t) = \boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t). \tag{6.2}$$

In summary, given a discretized initial shape $\bigcup_{i=1}^N \{q_{j,0}^{(i)}\}_{j=1}^{m_i}$, rotation centers $\{c_0^{(i)}\}_{i=1}^N$, and a discretized target shape $\bigcup_{i=1}^N \{q_{j,\mathrm{targ}}^{(i)}\}_{j=1}^{m_i'}$, the numerical formulation of the problem of piecewise-rigid motion discretized in space is written as

$$\min_{\theta \in L^2([0,T],\mathbb{R}^k)} \left(\frac{1}{2} \int_0^T \Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t) + |\theta(t)|^2 \Big)\, dt + \rho\Big(\{q_j^{(i)}(T)\}_{i,j}, \{q_{j,\mathrm{targ}}^{(i)}\}_{i,j} \Big) \right)$$

subject to

$$\begin{cases} c^{(i)}(t) = c_0^{(i)} + \displaystyle\int_0^t u^{(i)}(s)\, ds \\ \alpha^{(i)}(t) = \displaystyle\int_0^t \omega^{(i)}(s)\, ds \\ q_j^{(i)}(t) = \mathcal{P}\big(c^{(i)}(t), \alpha^{(i)}(t), q_{j,0}^{(i)}, c_0^{(i)}\big) \\ v_j^{(i)}(t) = \dot{\mathcal{P}}\big(u^{(i)}(t), \omega^{(i)}(t), \alpha^{(i)}(t), q_{j,0}^{(i)}, c_0^{(i)}\big) \end{cases},$$

where $\theta(t) = \big(u^{(1)}(t), \omega^{(1)}(t), \ldots, u^{(N)}(t), \omega^{(N)}(t)\big)$. We include the computation of gradients in Section A.4 for the interested reader.

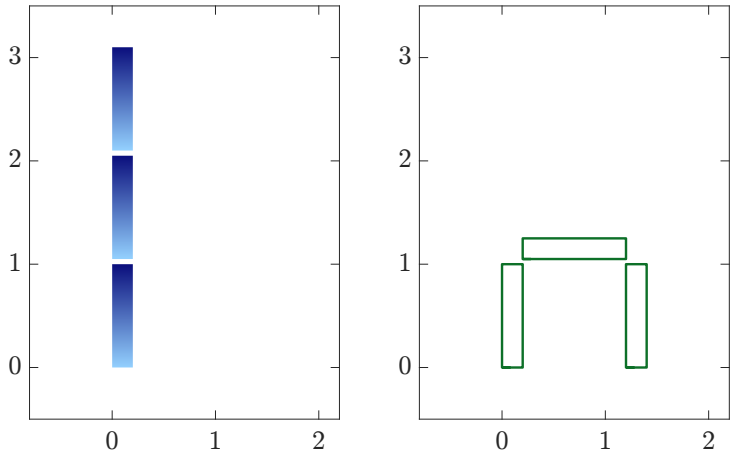
We must be careful that not all θ gives a Diff^p-motion in our discretized minimization problem. For example, there exists θ that makes two nodes move closer and hit each other. It is the term $\int_0^T \boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\, dt$ in the objective function that numerically rules out θ causing collisions of nodes and yields an empirical Diff^p-motion. In other words, we cannot drop the term $\int_0^T \boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\, dt$ in our discretized formulation; this is in sharp contrast with the energy form formulation (see the discussion at the end of Section 5.1.1). There is always a tension between rigidity and flexibility. We also remark on the choice of rotation centers. Theoretically, we can conveniently set all rotation centers to the origin; practically, the choice of rotation centers affect critical points found by a minimization procedure. First, if we set all rotation centers to the origin, we may encounter a badly scaled minimization problem: under the same rotation angle, the angular traveling arc length of a shape component away from the origin is longer than the one that is close to the

origin. That is, the same angular velocity would affect shape components disproportionately. Second, centroids of shape components may not always be a good choice for rotation centers. In Figure 6.2, we revisit one of our experiments with different rotation centers. From Figures 6.2(c) and 6.2(d), we can see that rotation centers affect critical points found. Thus a sensible choice of rotation centers could facilitate the minimization process. Rotation centers of numerical experiments in Section 3.1.2 were all placed at joints.

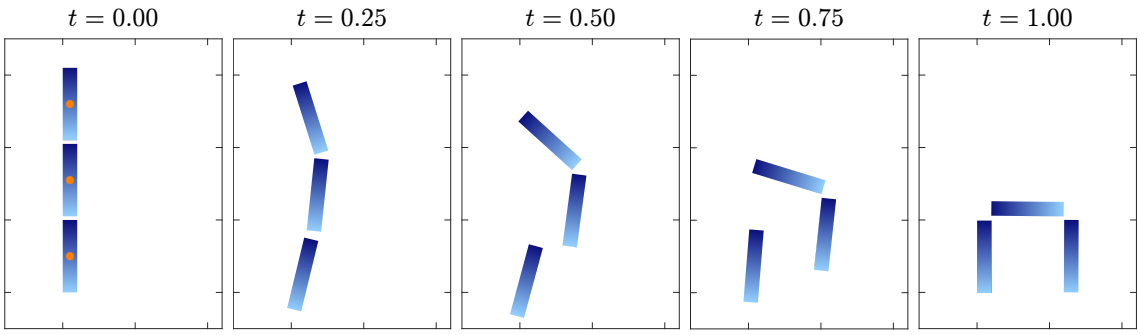
We add one more remark. An inconvenient fact is that kernel matrices are ill conditioned in general. Table 6.1 shows the condition number of $K_{H^m(\mathbb{R}^2)}(\boldsymbol{q})$ when $\boldsymbol{q}$ is composed of 21×21 equally spaced nodes on $[0,1]\times[0,1]$. For example, if we use $K_{H^5(\mathbb{R}^2)}(\boldsymbol{q})$ with $\sigma = 0.1$, which is not an uncommon choice, we can only expect the computation $K_{H^5(\mathbb{R}^2)}(\boldsymbol{q})^{-1}\,\boldsymbol{v}$ to be approximately in single precision since $\mathrm{cond}(K_{H^5(\mathbb{R}^2)}(\boldsymbol{q})) = 4.87\,\mathrm{e}{+}08$. We can avoid matrix inversion in LDDMM by considering an equivalent formulation

$$\begin{aligned}
&\min_{\boldsymbol{v}\,\in\, L^2([0,T],\,\mathbb{R}^{nd})} \left(\frac{1}{2}\int_0^T \boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\;dt + \rho\big(\boldsymbol{q}(T), \boldsymbol{q}_{\mathrm{targ}}\big) \right) \\
= &\min_{\boldsymbol{\alpha}\,\in\, L^2([0,T],\,\mathbb{R}^{nd})} \left(\frac{1}{2}\int_0^T \boldsymbol{\alpha}(t)^\top K_V(\boldsymbol{q}(t))\,\boldsymbol{\alpha}(t)\;dt + \rho\big(\boldsymbol{q}(T), \boldsymbol{q}_{\mathrm{targ}}\big) \right),
\end{aligned}$$

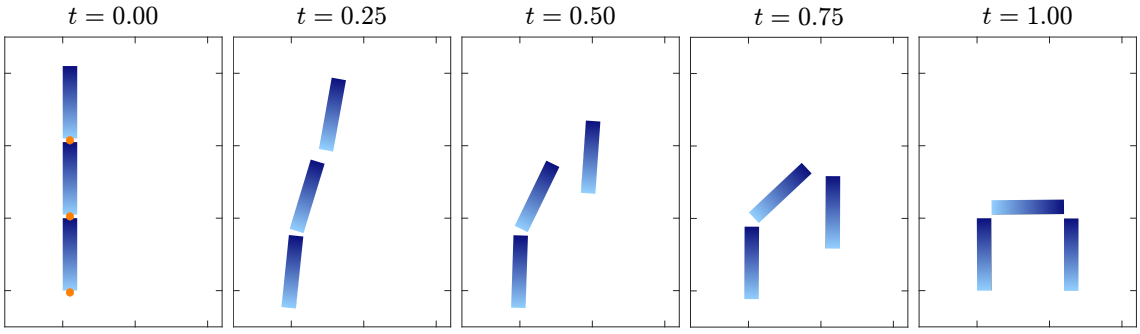
where n is the number of discretized nodes and $\boldsymbol{v}(t) = K_V(\boldsymbol{q}(t))\,\boldsymbol{\alpha}(t)$. At the continuous level, the equivalent formulation means that we shift the attention from $v(t) \in V$ to $\alpha(t) \in V^*$. However, since a model of deformation vector fields usually describes vector fields in V, it is inevitable that we need to solve the linear system $K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)$ if the objective function contains $\|v(t)\|_V^2$ or if we use $\|v(t)\|_V^2$ as a regularization term in the model (see (6.2) and (6.7)). Moreover, ill-conditioned kernel matrices also affect the accuracy of computed gradients if the evaluation of the objective function involves kernel matrix inversion. Although one might consider modifying $K_V(\boldsymbol{q}(t))$, for example, adding a weighted identity matrix and using $K_V(\boldsymbol{q}(t)) + c\,I_{nd}$ instead, the value c will not be small in order to improve the condition number. In other words, the numerical results using modified kernel matrices will deviate from its theoretical statement and lead to wrong conclusions. As the equivalent formulation of LDDMM suggests, it would be interesting to see if there is a way to directly model deformation vector fields in V^*.



(a) Initial shape. (b) Boundary of the target shape.



(c) Registration process when rotation centers are placed at centroids, shown in orange at $t = 0$.



(d) Registration process when rotation centers are placed at joints, shown in orange at $t = 0$.

Figure 6.2: Effect of the choice of rotation centers.

Table 6.1: The condition number of $K_{H^m(\mathbb{R}^2)}(\boldsymbol{q})$ with $\boldsymbol{q}$ composed of 21×21 equally spaced nodes on $[0,1] \times [0,1]$.

	$\mathrm{cond}(K_{H^m(\mathbb{R}^2)}(\boldsymbol{q}))$			
σ	$m=3$	$m=4$	$m=5$	$m=6$
0.05	1.94 e+03	3.93 e+04	7.80 e+05	1.54 e+07
0.10	9.03 e+04	6.67 e+06	4.87 e+08	3.55 e+10
0.15	7.91 e+05	1.24 e+08	1.94 e+10	3.05 e+12
0.20	3.45 e+06	9.20 e+08	2.47 e+11	6.68 e+13

6.2 Objective Function Evaluation of the Atrophy Model

Our problem of the atrophy model is

$$\min_{\theta \in \Theta} \ \rho(\varphi(T, \Omega_0), \Omega_{\mathrm{targ}})$$

subject to

$$\begin{cases} \varphi(t,x) = x + \displaystyle\int_0^t v(s, \varphi(s,x))\, ds \\ v(t) = (\gamma\, \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1}\, \beta_{\varphi,\theta}^t \\ \tau(t) = \mathcal{T}(\varphi_{|[0,t]}, \theta) \end{cases},$$

where

$$(\mathcal{A}_\varphi^t\, u \mid w) = \int_{\varphi(t,\, \Omega_0)} \mathcal{E}_{\varphi(t) * \Phi}(\varepsilon_u, \varepsilon_w)\, dx$$

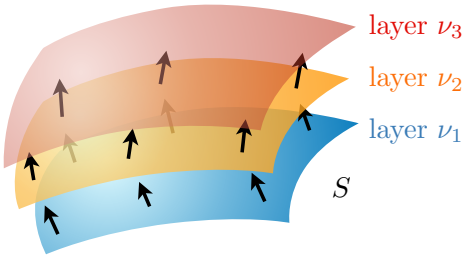
and

$$(\beta_{\varphi,\theta}^t \mid w) = \int_{\varphi(t,\, \Omega_0)} \chi\ \ \alpha\big(\tau(t) \circ \varphi(t)^{-1}\big)\ (-\mathrm{div}\, w)\ dx.$$

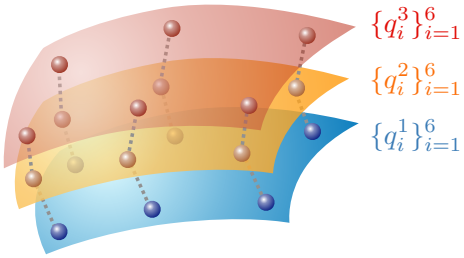
To evaluate the objective function $\rho(\varphi(T, \Omega_0), \Omega_{\mathrm{targ}})$, we elaborate on how to compute $v(t)$. We first describe the discretization of layered shapes in Section 6.2.1. We then present computations of the elastic energy $(\mathcal{A}_\varphi^t\, u \mid w)$ in Section 6.2.2 and the work $(\beta_{\varphi,\theta}^t \mid w)$ in Section 6.2.3; computations of $\mathcal{A}_\varphi^t\, u$ and $\beta_{\varphi,\theta}^t$ are also derived in respective sections. We formulate mass and stiffness matrices in Section 6.2.4 so that we can compute τ by the finite element method. Moreover, we present a specialized preconditioned conjugate gradient method in Section 6.2.5 which reduces the computation in $(\gamma\, \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1}\, \beta_{\varphi,\theta}^t$, the last step to obtain $v(t)$. Finally, we summarize our discretization scheme in time and space in Section 6.2.6. We present computations for $d = 3$ and mention necessary changes for $d = 2$.

6.2.1 Discretization of layered shapes

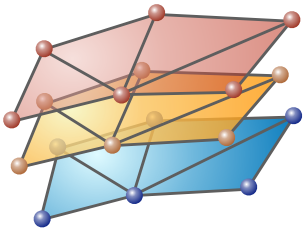
Given a layered shape Ω with a layered structure Φ (see Section 3.2.1), we discretize Ω into a set of nodes $\bigcup_{\ell=1}^{L}\{q_i^\ell\}_{i=1}^{N}$ according to its layered structure as described as follows. Based on the layered structure, we sample and identify discretized layers $\{\Phi(\nu_\ell, L_{\text{bottom}})\}_{\ell=1}^{L}$ (see Figures 6.3(a)). To be consistent with the layered structure, all discretized layers are further discretized into the same number of nodes $\{q_i^\ell\}_{i=1}^{N}$. Moreover, the vectors $q_i^{\ell+1} - q_i^\ell$ are parallel to the transversal vector $\partial_\nu\Phi(\nu_\ell, q_i^1)$ at q_i^ℓ for all i and ℓ (see Figures 6.3(a) and 6.3(b)). Note that nodes $\{q_i^1\}_{i=1}^{N}$ are on the bottom layer, and nodes $\{q_i^L\}_{i=1}^{N}$ are on the top layer. Since Φ is a diffeomorphism, the same triangulation structure can be applied to each layer if distances between layers are small (see Figure 6.3(c)). It follows that $\{q_{i_1}^\ell, q_{i_2}^\ell, q_{i_3}^\ell, q_{i_1}^{\ell+1}, q_{i_2}^{\ell+1}, q_{i_3}^{\ell+1}\}$ forms a triangular prism for any triangular face (i_1, i_2, i_3) of one layer. Those prisms between the first and second layers are further split into tetrahedra (see Figure 6.4(a)) without adding vertices using the procedure introduced in [35], which guarantees consistent triangular faces across adjacent prisms. To ensure the same tetrahedralization structure between consecutive layers, the tetrahedralization between the first and second layers is then replicated to prisms between consecutive upper layers (see Figure 6.3(d)).



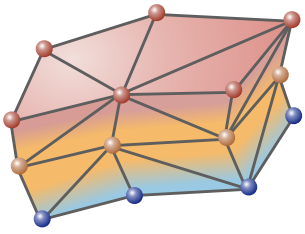
(a) Layers and the transversal vector field given by the layered structure.



(b) Discretized nodes according to the layers and the transversal vector field.



(c) The same triangulation structure applied to each layer.



(d) The same tetrahedralization structure applied to volumes between consecutive layers.

Figure 6.3: Tetrahedralization of layered shapes.

6.2.2 Computation of elastic energy

We recall the layered elastic energy defined by the layered elasticity tensor (3.8):

$$
\begin{aligned}
&(\mathcal{A}_\varphi^t\, u \mid w) \\
&= \int_{\varphi(t,\,\Omega_0)} \mathcal{E}_{\varphi(t)\,*\,\Phi}(\varepsilon_u, \varepsilon_w)\, dx \\
&= \int_{\varphi(t,\,\Omega_0)} \Big(\lambda_{\tan} \Big(\mathrm{tr}(\varepsilon_u) - N_{\varphi(t)}^\top\, \varepsilon_u\, N_{\varphi(t)} \Big) \Big(\mathrm{tr}(\varepsilon_w) - N_{\varphi(t)}^\top\, \varepsilon_w\, N_{\varphi(t)} \Big) \\
&\qquad\qquad + \mu_{\tan} \Big(\mathrm{tr}(\varepsilon_u\, \varepsilon_w) - 2\, N_{\varphi(t)}^\top\, \varepsilon_u\, \varepsilon_w\, N_{\varphi(t)} + (N_{\varphi(t)}^\top\, \varepsilon_u\, N_{\varphi(t)})(N_{\varphi(t)}^\top\, \varepsilon_w\, N_{\varphi(t)}) \Big) \\
&\qquad\qquad + \mu_{\mathrm{tsv}}\, (S_{\varphi(t)}^\top\, \varepsilon_u\, S_{\varphi(t)})(S_{\varphi(t)}^\top\, \varepsilon_w\, S_{\varphi(t)}) \\
&\qquad\qquad + 2\, \mu_{\mathrm{ang}} \Big(S_{\varphi(t)}^\top\, \varepsilon_u\, \varepsilon_w\, S_{\varphi(t)} - (N_{\varphi(t)}^\top\, \varepsilon_u\, S_{\varphi(t)})(N_{\varphi(t)}^\top\, \varepsilon_w\, S_{\varphi(t)}) \Big) \Big)\, dx, \qquad (6.3)
\end{aligned}
$$

where $\varepsilon_u = \frac{1}{2}\left(Du + Du^\top\right)$ and $\varepsilon_w = \frac{1}{2}\left(Dw + Dw^\top\right)$ are infinitesimal strain tensors, $N_{\varphi(t)}$ is a unit vector field normal to layers of $\varphi(t, \Omega_0)$, and $S_{\varphi(t)} = \frac{(D\varphi(t)\, S)}{|(D\varphi(t)\, S)|} \circ \varphi(t)^{-1}$ is the unit transversal vector field according to the layered structure $\varphi(t) * \Phi$. After discretizing the layered shape $\varphi(t, \Omega_0)$ into a union of tetrahedra (see Section 6.2.1), which could be tetrahedra evolved from time 0, we compute the integral (6.3) by summing over those tetrahedra. Thus we now focus the computation on one single tetrahedron. Notice that we need $N_{\varphi(t)}$, $S_{\varphi(t)}$, Du, and Dw to evaluate (6.3). Recall that the tetrahedralization procedure in Section 6.2.1 splits triangular prisms into tetrahedra (see Figure 6.4(a)). Given a tetrahedron, we approximate $N_{\varphi(t)}$ as the average of normals of the two bases of the corresponding prism, and $S_{\varphi(t)}$ is approximated as the average of three sides of the corresponding prism. To be more precise, let the "upward-pointing" unit normals of two bases of the prism be N_1 and N_2 (see Figure 6.4(b)), and let the unit transversals from three sides of the prism be S_1, S_2, and S_3 (see Figure 6.4(c)). The vectors $N_{\varphi(t)}$ and $S_{\varphi(t)}$ of the three tetrahedra split from this prism are approximated by

$$
N_{\varphi(t)} \approx \frac{N_1 + N_2}{|N_1 + N_2|} \quad \text{and} \quad S_{\varphi(t)} \approx \frac{S_1 + S_2 + S_3}{|S_1 + S_2 + S_3|}.
$$

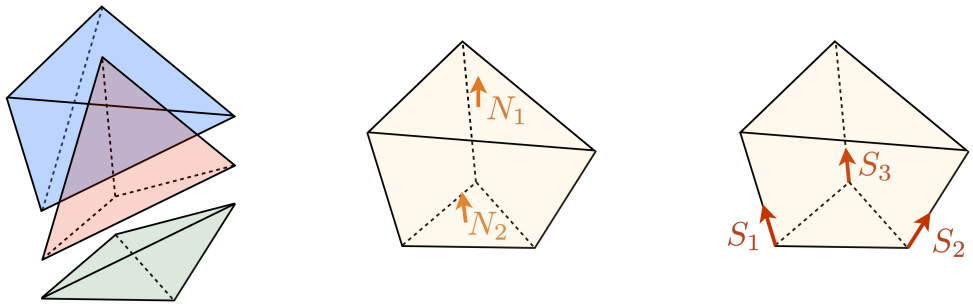
For the approximation of Du, denote the positions of the four nodes at the vertices of a tetrahedron by q_0, q_1, q_2, q_3, and denote $u(q_0)$, $u(q_1)$, $u(q_2)$, $u(q_3)$ by u_0, u_1, u_2, u_3. Within a tetrahedron $\mathcal{T}$, we approximate Du by the derivative of the unique affine transformation that maps q_i to u_i, $i = 0, \ldots, 3$. Precisely, we write the unique affine transformation as

$$
\overline{u}(x) := a + A\, x,
$$

where $a \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d\times d}$ are determined by $\overline{u}(q_i) = u_i$, $i = 0, \dots, 3$. Thus, for all $x \in \mathcal{T}$, we approximate Du by

$$\begin{aligned} Du(x) &\approx D\overline{u}(x) = A \\ &= \begin{bmatrix} u_1 - u_0, & u_2 - u_0, & u_3 - u_0 \end{bmatrix} \begin{bmatrix} q_1 - q_0, & q_2 - q_0, & q_3 - q_0 \end{bmatrix}^{-1} \\ &=: (Du)_{\mathcal{T}}. \end{aligned}$$

The approximation of Dw is computed in exactly the same way. With the approximations of $N_{\varphi(t)}$, $S_{\varphi(t)}$, Du, and Dw, we are now able to compute the discretized elastic energy $(\mathcal{A}^t_\varphi u \mid w)$ using (6.3). We denote the discretized elastic energy by $\boldsymbol{u}^\top A^t_{\boldsymbol{q}} \boldsymbol{w}$, where $\boldsymbol{u}$ and $\boldsymbol{w}$ should be interpreted as a column vector containing all $u(q_i)$'s and $w(q_i)$'s evaluated at nodes. Although a more correct notation for $A^t_{\boldsymbol{q}}$ should be $A_{\boldsymbol{q}(t)}$, we chose this notation due to its resemblance in appearance with the continuous counterpart $\mathcal{A}^t_\varphi$.



(a) Three tetrahedra split from one prism.

(b) Normal vectors of the two bases.

(c) Transversal vectors from the three sides.

Figure 6.4: Illustration of the approximated normal and transversal vectors of tetrahedra split from the same prism.

Next we show how to compute $A^t_{\boldsymbol{q}} \boldsymbol{w} = \partial_{\boldsymbol{u}}\big(\boldsymbol{u}^\top A^t_{\boldsymbol{q}} \boldsymbol{w}\big)$. We denote the discretized ε_u by $\boldsymbol{\varepsilon_u} := \frac{1}{2}\left((Du)_{\mathcal{T}} + (Du)_{\mathcal{T}}^\top\right)$. Define

$$U := \begin{bmatrix} u_1 - u_0, & u_2 - u_0, & u_3 - u_0 \end{bmatrix} \quad \text{and} \quad Q := \begin{bmatrix} q_1 - q_0, & q_2 - q_0, & q_3 - q_0 \end{bmatrix},$$

so $(Du)_{\mathcal{T}} = UQ^{-1}$. Since $\operatorname{tr}(\boldsymbol{\varepsilon_u}) = \sum_{i=1}^3 e_i^\top \boldsymbol{\varepsilon_u}\, e_i$, where e_i is the canonical basis of $\mathbb{R}^3$, we only need to have an expression of $\partial_{u_i}(a^\top \boldsymbol{\varepsilon_u}\, b)$ for arbitrary $a, b \in \mathbb{R}^3$ in order to compute $\partial_{\boldsymbol{u}}(\boldsymbol{u}^\top A^t_{\boldsymbol{q}} \boldsymbol{w})$ (see (6.3)). Note that

$$a^\top \boldsymbol{\varepsilon_u}\, b = a^\top \left(\frac{1}{2}\left(UQ^{-1} + Q^{-\top}U^\top\right)\right) b = \frac{1}{2}\operatorname{tr}\Big(Q^{-1}(ba^\top + ab^\top)\, U\Big),$$

which gives

$$\partial_{u_0}\Big(a^\top \boldsymbol{\varepsilon}_{\boldsymbol{u}}\, b\Big) = \left(-\frac{1}{2}\,\mathbb{1}_3^\top Q^{-1}(ba^\top + ab^\top)\right)^\top$$

and

$$\partial_{u_i}\Big(a^\top \boldsymbol{\varepsilon}_{\boldsymbol{u}}\, b\Big) = \left(\frac{1}{2}\left(Q^{-1}(ba^\top + ab^\top)\right)_{i*}\right)^\top \quad \text{for } i = 1, 2, 3,$$

where $\mathbb{1}_3$ denotes the 3-by-1 all-one vector, and $(A)_{i*}$ denotes the ith row of a matrix A. Note that the index $i = 0, 1, 2, 3$ is local in one tetrahedron and differs from the global index running through discretized nodes.

Let k be the global index running through discretized nodes. When we compute $\partial_{u_k}(\boldsymbol{u}^\top A_{\boldsymbol{q}}^t\, \boldsymbol{w})$ by summing contributions from tetrahedra, we only need to take into account those tetrahedra having q_k as a vertex. Other tetrahedra do not have u_k involved in our computation of $\boldsymbol{u}^\top A_{\boldsymbol{q}}^t\, \boldsymbol{w}$. This information can be precomputed right after we generate the tetrahedralization.

6.2.3 Computation of work

In this section, we show the computation of the work

$$\begin{aligned}(\beta_{\varphi,\,\theta}^t \mid w) &= \int_{\varphi(t,\,\Omega_0)} \chi\ \alpha\big(\tau(t)\circ\varphi(t)^{-1}\big)\ (-\operatorname{div} w)\ dx \\ &= -\int_{\varphi(t,\,\Omega_0)} \chi\ \alpha\big(\tau(t)\circ\varphi(t)^{-1}\big)\ \operatorname{tr}(Dw)\ dx.\end{aligned}$$

Similar to the previous section, we discretize $\varphi(t, \Omega_0)$ into a union of tetrahedra and focus our computation on one single tetrahedron. In a single tetrahedron $\mathcal{T}$, we evaluate $\chi\ \alpha\big(\tau(t)\circ \varphi(t)^{-1}\big)$ at the nodes of $\mathcal{T}$ and denote the average of those values by $\alpha_{\mathcal{T}}$. The derivative Dw is approximated in the same way as in Section 6.2.2, that is, $(Dw)_{\mathcal{T}} = WQ^{-1}$, where

$$W = \begin{bmatrix} w_1 - w_0, & w_2 - w_0, & w_3 - w_0 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} q_1 - q_0, & q_2 - q_0, & q_3 - q_0 \end{bmatrix}.$$

The contribution of a single tetrahedron $\mathcal{T}$ in the full integral is then given by

$$\alpha_{\mathcal{T}} \operatorname{tr}\big((Dw)_{\mathcal{T}}\big)\operatorname{vol}(\mathcal{T}) = \alpha_{\mathcal{T}} \operatorname{tr}\big(WQ^{-1}\big)\,\frac{1}{6}\,|\det Q|\,. \tag{6.4}$$

We denote the discretized version of the work by $\big(b_{\boldsymbol{q},\,\theta}^t\big)^\top \boldsymbol{w}$.

To obtain $b_{\boldsymbol{q},\,\theta}^t = \partial_{\boldsymbol{w}}\big(\big(b_{\boldsymbol{q},\,\theta}^t\big)^\top \boldsymbol{w}\big)$, we have from (6.4) that

$$\partial_{w_0}\big(\alpha_{\mathcal{T}} \operatorname{tr}\big(WQ^{-1}\big)\operatorname{vol}(\mathcal{T})\big) = -\alpha_{\mathcal{T}} \operatorname{vol}(\mathcal{T})\left(\mathbb{1}_3^\top Q^{-1}\right)^\top$$

and

$$\partial_{w_i}\big(\alpha_{\mathcal{T}}\,\mathrm{tr}\big(WQ^{-1}\big)\,\mathrm{vol}(\mathcal{T})\big) = \alpha_{\mathcal{T}}\,\mathrm{vol}(\mathcal{T})\,\big((Q^{-1})_{i*}\big)^{\top} \quad \text{for } i = 1,2,3,$$

where $\mathbb{1}_3$ denotes the 3-by-1 all-one vector, and $(A)_{i*}$ denotes the ith row of a matrix A.

6.2.4 Mass and stiffness matrices

Our goal in this section is to compute $\tau(t) = \mathcal{T}(\varphi_{|[0,\,t]}, \theta)$ using the finite element method. To this end, the effort will be devoted to obtain expressions of mass and stiffness matrices. Given $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, we recall the weak formulation:

$$\begin{cases} \partial_t\big(\tau\ \det D\varphi\big) + \mathcal{L}\tau = g(\tau) \\ \tau(0) = Q(\theta) \end{cases}, \tag{5.5}$$

where

$$((\mathcal{L}\tau \mid \tau')) = \int_0^T \big\langle W_{\varphi(t)}\,\nabla\tau(t)\ \det D\varphi(t),\ \nabla\tau'(t)\big\rangle_{L^2}\,dt$$

and

$$g(\tau)(t) = R(\tau(t))\ \det D\varphi(t) \quad \text{for almost every } t \in [0,T].$$

According to Proposition A.2.5, the weak formulation (5.5) is equivalent to

$$\begin{aligned} &\frac{d}{dt}\int_{\Omega_0} \tau(t)\ \det D\varphi(t)\ \psi\ dx \\ &= -\int_{\Omega_0}\Big(W_{\varphi(t)}\,\nabla\tau(t)\ \det D\varphi(t)\Big)^{\top}\nabla\psi\ dx + \int_{\Omega_0} R(\tau(t))\ \det D\varphi(t)\ \psi\ dx \end{aligned} \tag{6.5}$$

for almost every $t \in [0,T]$ and all $\psi \in H^1(\Omega_0^o)$. After we discretize Ω_0 into elements with nodes $\{q_{i,0}\}_{i=1}^n$, we introduce piecewise linear basis functions $\{\psi_{i,0}\}_{i=1}^n$ and approximate the solution by $\tau(t,x) = \sum_{i=1}^n \tau_i(t)\,\psi_{i,0}(x)$. Plugging this approximation into (6.5) and letting $\psi = \psi_{i,0}$ for $i = 1,\dots,n$, we arrive at

$$\begin{aligned} &\frac{d}{dt}\left(\left[\ \int_{\Omega_0} \psi_{i,0}\,\psi_{j,0}\ \det D\varphi(t)\ dx\ \right]_{n\times n}\left[\ \tau_i(t)\ \right]_{n\times 1}\right) \\ &= -\left[\ \int_{\Omega_0} \nabla\psi_{i,0}^{\top}\ W_{\varphi(t)}\,\nabla\psi_{j,0}\ \det D\varphi(t)\ dx\ \right]_{n\times n}\left[\ \tau_i(t)\ \right]_{n\times 1} \\ &\qquad + \left[\ \int_{\Omega_0} R\Big(\sum_{k=1}^n \tau_k(t)\,\psi_{k,0}\Big)\psi_{i,0}\ \det D\varphi(t)\ dx\ \right]_{n\times 1}, \end{aligned}$$

which we denote by

$$\frac{d}{dt}\big(\Psi(t)\,\boldsymbol{\tau}(t)\big) = -\Gamma(t)\,\boldsymbol{\tau}(t) + \boldsymbol{r}(t). \tag{6.6}$$

We can then use any time stepping scheme to compute $\boldsymbol{\tau}(t)$ once we have node positions $\boldsymbol{q}(t)$ to approximate $\det D\varphi(t)$. We postpone our full algorithm stepping $\boldsymbol{\tau}(t)$ and $\boldsymbol{q}(t)$ together to Section 6.2.6. Below we provide approximations of entries of the mass matrix $\Psi(t)$, the stiffness matrix $\Gamma(t)$, and the vector $\boldsymbol{r}(t)$ assuming we have node positions $\boldsymbol{q}(t)$. The derivations of these approximations are placed in Secion A.5 for the interested reader.

Since $\varphi \in C([0,T], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$, a fixed triangulation or tetrahedralization structure can be applied to nodes $\{q_i(t)\}_{i=1}^n$ for each $t \in [0,T]$. We denote a triangle or tetrahedron of the discretized shape Ω_0 by $\mathcal{T}_0$ and denote its evolution by $t \mapsto \mathcal{T}(t)$. In addition, we write $i \in \mathcal{T}_0$ if the element $\mathcal{T}_0$ has the i-th node as a vertex. We also denote the piecewise linear basis functions with respect to nodes $\{q_i(t)\}_{i=1}^n$ at time t by $\{\psi_i(t)\}_{i=1}^n$, whose gradient is constant on an element $\mathcal{T}(t)$ and is denoted by $\big(\nabla\psi_i(t)\big)_{\mathcal{T}(t)}$. Suppose that we approximate the Eulerian diffusion matrix $U_{\varphi(t)}$ by $U_{\boldsymbol{q}(t)}$, for example, using the approximated $N_{\varphi(t)}$ and $S_{\varphi(t)}$ in Section 6.2.2 for the Eulerian diffusion matrix (3.22) we consider. If $d = 2$, we have

$$\int_{\mathcal{T}_0} \psi_{i,0}\,\psi_{j,0}\,\det D\varphi(t)\;dx \approx \begin{cases} \dfrac{1}{6}\,\mathrm{vol}\big(\mathcal{T}(t)\big), & \text{if } i,j \in \mathcal{T}_0 \text{ and } i = j\,; \\ \dfrac{1}{12}\,\mathrm{vol}\big(\mathcal{T}(t)\big), & \text{if } i,j \in \mathcal{T}_0 \text{ and } i \neq j\,; \\ 0, & \text{otherwise}\,, \end{cases}$$

$$\begin{aligned} &\int_{\mathcal{T}_0} \nabla\psi_{i,0}^\top\, W_{\varphi(t)}\,\nabla\psi_{j,0}\,\det D\varphi(t)\;dx \\ &\approx \begin{cases} \big(\nabla\psi_i(t)\big)_{\mathcal{T}(t)}^\top\, U_{\boldsymbol{q}(t)}\,\big(\nabla\psi_j(t)\big)_{\mathcal{T}(t)}\,\mathrm{vol}\big(\mathcal{T}(t)\big), & \text{if } i,j \in \mathcal{T}_0\,; \\ 0, & \text{otherwise}\,, \end{cases} \end{aligned}$$

and

$$\begin{aligned} &\int_{\mathcal{T}_0} R\Big(\sum_{k=1}^n \tau_k(t)\,\psi_{k,0}\Big)\psi_{i,0}\,\det D\varphi(t)\;dx \\ &\approx \frac{1}{12}\Big(2R(\tau_i(t)) + R(\tau_j(t)) + R(\tau_k(t))\Big)\,\mathrm{vol}\big(\mathcal{T}(t)\big), \end{aligned}$$

where i, j, k are the indices of the three nodes at vertices of $\mathcal{T}_0$. If $d = 3$, we have

$$\int_{\mathcal{T}_0} \psi_{i,0}\,\psi_{j,0}\,\det D\varphi(t)\;dx \approx \begin{cases} \dfrac{1}{10}\,\mathrm{vol}\big(\mathcal{T}(t)\big), & \text{if } i,j \in \mathcal{T}_0 \text{ and } i = j\,; \\ \dfrac{1}{20}\,\mathrm{vol}\big(\mathcal{T}(t)\big), & \text{if } i,j \in \mathcal{T}_0 \text{ and } i \neq j\,; \\ 0, & \text{otherwise}\,, \end{cases}$$

$$\begin{aligned} &\int_{\mathcal{T}_0} \nabla\psi_{i,0}^\top\, W_{\varphi(t)}\,\nabla\psi_{j,0}\,\det D\varphi(t)\;dx \\ &\approx \begin{cases} \big(\nabla\psi_i(t)\big)_{\mathcal{T}(t)}^\top\, U_{\boldsymbol{q}(t)}\,\big(\nabla\psi_j(t)\big)_{\mathcal{T}(t)}\,\mathrm{vol}\big(\mathcal{T}(t)\big), & \text{if } i,j \in \mathcal{T}_0\,; \\ 0, & \text{otherwise}\,, \end{cases} \end{aligned}$$

and

$$\begin{aligned} &\int_{\mathcal{T}_0} R\Big(\sum_{k=1}^{n} \tau_k(t)\,\psi_{k,0}\Big)\psi_{i,0}\,\det D\varphi(t)\;dx \\ &\approx \frac{1}{20}\Big(2R(\tau_i(t)) + R(\tau_j(t)) + R(\tau_k(t)) + R(\tau_\ell(t))\Big)\,\mathrm{vol}\big(\mathcal{T}(t)\big), \end{aligned}$$

where i, j, k, ℓ are the indices of the four nodes at vertices of $\mathcal{T}_0$. We observe that the data we need to approximate the three integrals are $\{q_i(t)\}_i$ and $\{\tau_i(t)\}_i$; we do not need $\{q_{i,0}\}_i$ even though the integrals are taken on a domain at time 0.

6.2.5 Specialized preconditioned conjugate gradient

We present a preconditioned conjugate gradient (PCG) method specialized to symmetric positive-definite linear systems of the form $(B^{-1} + A)\,x = b$ with B^{-1} symmetric positive definite and A symmetric positive semidefinite. Recall that our deformation vector field is given by

$$v(t) = (\gamma\,\mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1}\,\beta_{\varphi,\theta}^t = \operatorname*{arg\,min}_{v' \in V}\left(\frac{\gamma}{2}\,\|v'\|_V^2 + \frac{1}{2}\,(\mathcal{A}_\varphi^t\,v' \mid v') - (\beta_{\varphi,\theta}^t \mid v')\right).$$

It follows that the discretized deformation vector field is

$$\begin{aligned} \boldsymbol{v}(t) &= \operatorname*{arg\,min}_{\boldsymbol{v}' \in \mathbb{R}^{nd}}\left(\frac{\gamma}{2}\,\boldsymbol{v}'^{\top} K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}' + \frac{1}{2}\,\boldsymbol{v}'^{\top} A_{\boldsymbol{q}}^t\,\boldsymbol{v}' - \big(b_{\boldsymbol{q},\theta}^t\big)^\top \boldsymbol{v}'\right) \\ &= \Big(\gamma\,K_V(\boldsymbol{q}(t))^{-1} + A_{\boldsymbol{q}}^t\Big)^{-1}\,b_{\boldsymbol{q},\theta}^t, \end{aligned} \tag{6.7}$$

to which this specialized PCG applies. This specialized PCG reduces computation compared to the standard PCG and hence reduces running time solving linear systems of this form.

We recall the classical PCG solving $Mx = b$ with a preconditioner P in Algorithm 1. We use $\underline{Pr} \leftarrow Pr$ to denote that the result of Pr is assigned to the *variable* $\underline{Pr}$, so underlines in Algorithm 1 mean that the computation has already been performed. If we apply this classical PCG with some preconditioner directly to the problem $(B^{-1} + A)\, x = b$, note that we will be computing $Md = (B^{-1} + A)\, d$ on line 6 in every iteration, which means that we have the burden of solving $By = d$ in every iteration. Although we can decompose B by the Cholesky factorization before entering PCG, performing forward and backward substitutions for each iteration is still expensive. Surprisingly, we can reduce the number of times solving $By = d$ to just once when we use B as the preconditioner, as shown in Algorithm 2. In Algorithm 2, we introduce the additional variable $\widetilde{d}$ on line 2. Comparing line 2 and 4, we have $\widetilde{d} = B^{-1}\, d$ at the beginning of the while loop. Moreover, this relation is kept by line 14 and 15. Since this loop invariant $\widetilde{d} = B^{-1}\, d$ is maintained, we know that the computation on line 7 is correct. It follows that Algorithm 2 is equivalent to Algorithm 1 under this special setting. We observe that the expensive $B^{-1}x$ only appears on line 1 in Algorithm 2 at the expense of the additional storage of $\widetilde{d}$ and additional scalar multiplication and vector addition on line 14.

6.2.6 Equations of motion discretized in time and space

We have all the tools to compute the discretized objective function

$$J(\theta) := \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}})$$

subject to

$$\begin{cases} \varphi(t,x) = x + \displaystyle\int_0^t v(s, \varphi(s,x))\, ds \\ v(t) = (\gamma\, \mathcal{K}_V^{-1} + \mathcal{A}_\varphi^t)^{-1}\, \beta_{\varphi,\theta}^t \\ \tau(t) = \mathcal{T}(\varphi_{|[0,t]}, \theta) \end{cases} .$$

We summarize the objective function evaluation in Algorithm 3. On line 5 in Algorithm 3, we use a semi-implicit scheme to approximate

$$\frac{d}{dt}\big(\Psi(t)\, \boldsymbol{\tau}(t)\big) = -\Gamma(t)\, \boldsymbol{\tau}(t) + \boldsymbol{r}(t) \tag{6.6}$$

as

$$\frac{\Psi_{t_{i+1}} \boldsymbol{\tau}_{t_{i+1}} - \Psi_{t_i} \boldsymbol{\tau}_{t_i}}{\Delta t} = -\Gamma_{t_{i+1}}\, \boldsymbol{\tau}_{t_{i+1}} + \boldsymbol{r}_{t_i},$$

which gives the expression of $\boldsymbol{\tau}_{t_{i+1}}$ on line 5. We remind the reader that we only need $\boldsymbol{q}(t)$ to compute $\Psi(t)$ and $\Gamma(t)$, but we need $\boldsymbol{q}(t)$ and $\boldsymbol{\tau}(t)$ to compute $\boldsymbol{r}(t)$ (see Section 6.2.4).

Algorithm 1: PCG

Input : M, b, P, x

Output: x

$r \leftarrow b - Mx$

$\underline{Pr} \leftarrow Pr$

$d \leftarrow \underline{Pr}$

$\underline{r^\top Pr} \leftarrow r^\top \underline{Pr}$

while not converge **do**

  $\underline{Md} \leftarrow Md$

  $\alpha \leftarrow \underline{r^\top Pr} \;/\; d^\top \underline{Md}$

  $x \leftarrow x + \alpha\, d$

  $r' \leftarrow r - \alpha\, \underline{Md}$

  $\underline{Pr'} \leftarrow Pr'$

  $\underline{r'^\top Pr'} \leftarrow r'^\top \underline{Pr'}$

  $\beta \leftarrow \underline{r'^\top Pr'} \;/\; \underline{r^\top Pr}$

  $d \leftarrow \underline{Pr'} + \beta\, d$

  $r \leftarrow r'$

  $\underline{r^\top Pr} \leftarrow \underline{r'^\top Pr'}$

end while

Algorithm 2: Specialized PCG

Input : B, A, b, x

Output: x

$r \leftarrow b - (B^{-1} + A)\, x$

$\widetilde{d} \leftarrow r \quad // \, B^{-1} d$

$\underline{Br} \leftarrow Br$

$d \leftarrow \underline{Br}$

$\underline{r^\top Br} \leftarrow r^\top \underline{Br}$

while not converge **do**

  $\underline{Md} \leftarrow \widetilde{d} + Ad \quad // \, (B^{-1} + A)\, d$

  $\alpha \leftarrow \underline{r^\top Br} \;/\; d^\top \underline{Md}$

  $x \leftarrow x + \alpha\, d$

  $r' \leftarrow r - \alpha\, \underline{Md}$

  $\underline{Br'} \leftarrow Br'$

  $\underline{r'^\top Br'} \leftarrow r'^\top \underline{Br'}$

  $\beta \leftarrow \underline{r'^\top Br'} \;/\; \underline{r^\top Br}$

  $\widetilde{d} \leftarrow r' + \beta\, \widetilde{d} \quad // \, B^{-1}\, d$

  $d \leftarrow \underline{Br'} + \beta\, d$

  $r \leftarrow r'$

  $\underline{r^\top Br} \leftarrow \underline{r'^\top Br'}$

end while

Algorithm 3: Objective function evaluation of the atrophy model

Input : θ, $\boldsymbol{q}_0$, $\boldsymbol{q}_{\text{targ}}$

Output: $\rho\big(\boldsymbol{q}_T, \boldsymbol{q}_{\text{targ}}\big)$

$\boldsymbol{\tau}_0 \leftarrow Q(\theta)$

for each time t_i **do**

  $\boldsymbol{v}_{t_i} \leftarrow \Big(\gamma\, K_V(\boldsymbol{q}_{t_i})^{-1} + A^{t_i}_{\boldsymbol{q}}\Big)^{-1} b^{t_i}_{\boldsymbol{q},\theta}$

  $\boldsymbol{q}_{t_{i+1}} \leftarrow \boldsymbol{q}_{t_i} + \Delta t\, \boldsymbol{v}_{t_i}$

  $\boldsymbol{\tau}_{t_{i+1}} \leftarrow \big(\Psi_{t_{i+1}} + \Delta t\, \Gamma_{t_{i+1}}\big)^{-1} \big(\Psi_{t_i}\, \boldsymbol{\tau}_{t_i} + \Delta t\, \boldsymbol{r}_{t_i}\big)$

end for

Compute $\rho\big(\boldsymbol{q}_T, \boldsymbol{q}_{\text{targ}}\big)$.

6.3 GPU Implementation

We now describe our GPU implementation in NVIDIA CUDA (Compute Unified Device Architecture). CUDA is a parallel computing platform which extends standard programming languages like C, C++, and Fortran so that programmers can enjoy both the familiar environment of standard languages and the speedup by parallel computing. According to Flynn's taxonomy, the architecture of CUDA is in the class of single instruction, multiple data (SIMD), or more accurately, single instruction, multiple threads (SIMT). We refer the reader to [74, 73] for a more detailed introduction to the CUDA programming model. Although programmers who are familiar with standard programming languages can easily learn how to write CUDA code, in early generations of GPUs, great effort must be made in order to fit the application into the CUDA memory hierarchy; there is a chance that the CUDA code would be slower than its sequential version if every memory transaction is a non-coalesced global load. Fortunately, this burden has been lifted thanks to the improvement of global memory bandwidth and cache performance in recent generations of GPUs [9], which also means that we need to adjust the optimization strategy of CUDA code accordingly. In Section 6.3.1, we confirm the observation in [9]: fetching data through shared memory may negatively influence the performance in recent GPUs. We then compare several approaches of objective function evaluation in Section 6.3.2. Section 6.3.3 summarizes the overall speedup in objective function evaluation from OpenMP to CUDA. In the following discussions, matrices are stored in column-major order, hence we shall avoid matrix transpose in this section for clarity, except for the first paragraph of Section 6.3.1. All computations were performed in double precision.

6.3.1 Computation of kernel matrices

We recall the definition of kernel matrices from Section 6.1. For $\boldsymbol{q}^\top = \begin{bmatrix} q_1^\top & \cdots & q_n^\top \end{bmatrix}_{1\times nd}$, the corresponding kernel matrix $K_V(\boldsymbol{q})$ is an nd-by-nd block matrix whose ij-th block is $k_V(q_i, q_j) \in \mathscr{L}(\mathbb{R}^d, \mathbb{R}^d) \cong \mathbb{R}^{d\times d}$. Since our numerical experiments use $V = (H^m(\mathbb{R}^d))^d$ exclusively, here we consider the special structure of kernel matrices when the $\mathbb{R}^d$-valued RKHS is V^d with a real-valued RKHS V. We remind the reader that $k_{V^d}(x, y) = k_V(x, y)\, id$ (see (2.7)). It follows that

$$K_{V^d}(\boldsymbol{q}) = K_V(\boldsymbol{q}) \otimes I_d,$$

where $\otimes$ is the Kronecker product. Therefore, when the RKHS is V^d, all we need is $K_V(\boldsymbol{q})$; we do not even need to form the big matrix $K_{V^d}(\boldsymbol{q})$. For example, if we want to compute

$\boldsymbol{a}^\top K_{V^d}(\boldsymbol{q})\,\boldsymbol{b}$, where $\boldsymbol{a}^\top = \begin{bmatrix} a_1^\top & \cdots & a_n^\top \end{bmatrix}_{1\times nd}$ and $\boldsymbol{b}^\top = \begin{bmatrix} b_1^\top & \cdots & b_n^\top \end{bmatrix}_{1\times nd}$, we can rearrange the column vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ into data matrices

$$A = \begin{bmatrix} a_1^\top \\ \vdots \\ a_n^\top \end{bmatrix}_{n\times d} \quad \text{and} \quad B = \begin{bmatrix} b_1^\top \\ \vdots \\ b_n^\top \end{bmatrix}_{n\times d}$$

and notice that

$$\boldsymbol{a}^\top K_{V^d}(\boldsymbol{q})\,\boldsymbol{b} = \langle A, K_V(\boldsymbol{q})\,B\rangle_F,$$

where $\langle\cdot,\cdot,\rangle_F$ is the Frobenius inner product. We thus focus on the computation of kernel matrices when V is a real-valued RKHS. In addition, we will suppress the dependency of kernel matrices on $\boldsymbol{q}$ and simply write K_V.

We now discuss the computation of K_V when V is a real-valued RKHS. In our CUDA implementation, we use a usual tiling method [74, Section 3.2.4] to compute the kernel matrix $K_V \in \mathbb{R}^{n\times n}$ corresponding to a matrix $Q \in \mathbb{R}^{n\times d}$ which stores the positions of n nodes $q_1, \ldots, q_n$. To be more specific, we partition the computation of the kernel matrix K_V into 2D thread blocks and use one thread to compute an element of K_V. Assume that a thread block of size m-by-m is responsible to compute the submatrix formed by rows $i = i_0 + 1, \ldots, i_0 + m$ and columns $j = j_0 + 1, \ldots, j_0 + m$. In this situation, a common guideline is to utilize the fast shared memory: we read the positions of nodes q_i, $i = i_0+1, \ldots, i_0+m$ and q_j, $j = j_0+1, \ldots, j_0+m$ from global memory into shared memory, then all m^2 threads perform computation on shared memory instead of fetching positions from global memory directly. Note that this global memory read is coalesced because of the column-major format of $Q \in \mathbb{R}^{n\times d}$. To test if the guideline is still valid on recent GPUs, we compare the implementation using shared memory with the one using purely global memory on Tesla K40m (Kepler architecture), TITAN Xp (Pascal architecture), and TITAN V (Volta architecture). We generated a fixed Q whose elements are in the interval $[0, 1]$ and computed K_V with $V = H^5(\mathbb{R}^d)$ and $\sigma = 0.1$. Only the function computing K_V was timed by `cudaEventRecord`. We used the median from 50 runs to represent the running time since the variation is comparatively small. Each of the 50 runs was called from a shell script afresh to avoid biased running times due to cache hit. The results using various sizes of thread blocks are presented in Tables 6.2 to 6.4. From Table 6.2 (8-by-8 threads) and Table 6.3 (16-by-16 threads), we see that the running times using shared memory divided by the running times using global memory are all above one. The running times using shared memory are roughly the same as or ever 1.3 to 1.4 times more than the running times using

global memory. In other words, we do not have the expected speedup, usually more than 2 fold in early generations of GPUs, when we use shared memory. Although Table 6.4 (32-by-32 threads) shows some decrease of running times when using shared memory, we need to assess this small performance gain against the additional programming and tuning efforts. These simple experiments show that we must re-evaluate the usage of shared memory rather than counting on it as a guaranteed performance boost.

Table 6.2: The ratio of running times (shared memory divided by global memory) using 8-by-8 thread blocks.

Number of Nodes	2D			3D		
	K40m	TITAN Xp	TITAN V	K40m	TITAN Xp	TITAN V
2,000	1.391	1.056	1.000	1.297	1.037	1.011
4,000	1.425	1.055	1.002	1.316	1.040	1.002
6,000	1.432	1.037	1.001	1.318	1.040	1.001
8,000	1.435	1.037	1.001	1.321	1.041	1.001
10,000	1.436	1.037	1.001	1.321	1.041	1.001

Table 6.3: The ratio of running times (shared memory divided by global memory) using 16-by-16 thread blocks.

Number of Nodes	2D			3D		
	K40m	TITAN Xp	TITAN V	K40m	TITAN Xp	TITAN V
2,000	1.035	1.044	1.016	1.061	1.032	1.042
4,000	1.044	1.044	1.016	1.066	1.031	1.021
6,000	1.044	1.044	1.016	1.063	1.031	1.022
8,000	1.047	1.026	1.015	1.068	1.031	1.019
10,000	1.050	1.026	1.015	1.067	1.031	1.021

6.3.2 Computation of objective functions

We first identify potential parallelism in our two objective functions. In our objective function of piecewise-rigid motion

$$J_R(\theta) := \frac{1}{2}\int_0^T \Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t) + |\theta(t)|^2\Big)\,dt + \rho\Big(\{q_j^{(i)}(T)\}_{i,j}, \{q_{j,\text{targ}}^{(i)}\}_{i,j}\Big),$$

we examine the integral term which is independent of the choice of ρ. The computation of $\boldsymbol{v}(t)$ and $\boldsymbol{q}(t)$ can be embarrassingly parallelized, i.e., one parallel thread is responsible for one node and computes its velocity and position. The computation of $K_V(\boldsymbol{q}(t))$ can be

Table 6.4: The ratio of running times (shared memory divided by global memory) using 32-by-32 thread blocks.

Number of Nodes	2D			3D		
	K40m	TITAN Xp	TITAN V	K40m	TITAN Xp	TITAN V
2,000	0.984	1.036	0.979	0.928	1.026	1.000
4,000	0.981	1.023	1.000	0.928	1.027	0.989
6,000	0.998	1.022	1.000	0.924	1.027	0.991
8,000	0.994	1.022	0.997	0.920	1.027	0.989
10,000	0.996	1.022	0.998	0.916	1.027	0.989

parallelized by the tiling method introduced in Section 6.3.1. To obtain $K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)$, the Cholesky decomposition and forward and backward substitutions can be parallelized. Finally, the inner product $\boldsymbol{v}(t)^\top \big(K_V(\boldsymbol{q}(t))^{-1}\boldsymbol{v}(t)\big)$ and the integral $\int_0^T |\theta(t)|^2\,dt \approx \sum_i |\theta(t_i)|^2\,\Delta t$ can be computed by parallel reduction. As for the atrophy model, our (naïve-looking) objective function is

$$J_A(\theta) := \rho\big(\boldsymbol{q}_T, \boldsymbol{q}_{\text{targ}}\big),$$

which involves heavy works

$$\boldsymbol{v}_{t_i} = \Big(\gamma\, K_V(\boldsymbol{q}_{t_i})^{-1} + A^{t_i}_{\boldsymbol{q}}\Big)^{-1} b^{t_i}_{\boldsymbol{q},\,\theta} \quad \text{and} \quad \boldsymbol{\tau}_{t_{i+1}} = \big(\Psi_{t_{i+1}} + \Delta t\, \Gamma_{t_{i+1}}\big)^{-1}\big(\Psi_{t_i}\,\boldsymbol{\tau}_{t_i} + \Delta t\, \boldsymbol{r}_{t_i}\big).$$

Besides exploiting the aforementioned parallelism, we are going to examine the parallelism in the computation of $A^{t_i}_{\boldsymbol{q}}\,\boldsymbol{v}$, $b^{t_i}_{\boldsymbol{q},\,\theta}$, and matrices assembly for $\Psi_{t_{i+1}}$ and $\Gamma_{t_{i+1}}$ in the remaining of this section. To focus on computation, we will suppress the dependency subscripts and superscripts.

Before we proceed to GPU implementation, we illustrate the unexpected similarity between the computation of $A\boldsymbol{v}$, b, and matrices assembly Ψ_{ij} and Γ_{ij}. We recall the (loose) definitions of $\mathcal{A}v$, β, the mass matrix Ψ, and the stiffness matrix Γ. We have

$$\mathcal{A}v = \partial_w\left(\mathcal{A}v \mid w\right) =: \partial_w\left(\int_\Omega f(x, v(x), w(x))\,dx\right),$$

$$\beta = \partial_w\left(\beta \mid w\right) =: \partial_w\left(\int_\Omega g(x, w(x))\,dx\right),$$

$$\Psi_{ij} =: \int_\Omega h_{ij}(x)\,dx, \quad \text{and} \quad \Gamma_{ij} =: \int_\Omega \widetilde{h}_{ij}(x)\,dx.$$

The functions f, g, h_{ij}, $\widetilde{h}_{ij}$ are introduced as place holders to facilitate the following presentation; their exact definitions are not important here. After discretization, the domain of integration is represented by a union of elements, e.g., triangles or tetrahedra. We denote

the elements by $\{e_k\}_{k=1}^K$ and denote the nodes by $\{q_i\}_{i=1}^n$. In addition, we have derived approximations on one single element from Sections 6.2.2 to 6.2.4 which can be written symbolically as

$$\int_\Omega f(x, v(x), w(x))\,dx \approx \sum_{k=1}^K F^{e_k}, \quad \int_\Omega g(x, w(x))\,dx \approx \sum_{k=1}^K G^{e_k},$$

$$\int_\Omega h_{ij}(x)\,dx \approx \sum_{k=1}^K H_{ij}^{e_k}, \quad \text{and} \quad \int_\Omega \widetilde{h}_{ij}(x)\,dx \approx \sum_{k=1}^K \widetilde{H}_{ij}^{e_k}.$$

Let us further denote the collection of elements adjacent to the node i by $\mathcal{N}_i$ and denote the collection of elements adjacent to the edge ij by $\mathcal{N}_{ij}$. We define $\mathcal{N}_{ii} := \mathcal{N}_i$ and $\mathcal{N}_{ij} := \emptyset$ if $i \neq j$ and there is no edge ij. It follows that the discretized $\mathcal{A}v$, β, and the ij-th entry of mass and stiffness matrices (using the same notation) on triangular or tetrahedral elements are given by

$$(A\boldsymbol{v})_i = \sum_{e\,\in\,\mathcal{N}_i} \partial_{w_i} F^e, \quad b_i = \sum_{e\,\in\,\mathcal{N}_i} \partial_{w_i} G^e, \quad \Psi_{ij} = \sum_{e\,\in\,\mathcal{N}_{ij}} H_{ij}^e, \quad \Gamma_{ij} = \sum_{e\,\in\,\mathcal{N}_{ij}} \widetilde{H}_{ij}^e. \tag{6.8}$$

We have assumed that the sum over the empty set equals zero. The similarity between the computation of $A\boldsymbol{v}$, b, Ψ, and Γ is manifested in (6.8). We remark that, for quadrilaterals, prisms, and bricks, the form of the above expressions of Ψ_{ij} and Γ_{ij} is still correct, but we need to modify the definition of $\mathcal{N}_{ij}$ to accommodate non-edge diagonals within elements. We have seen the computation on one element, that is, $\partial_{w_i} F^e$, $\partial_{w_i} G^e$, H_{ij}^e, and $\widetilde{H}_{ij}^e$, from Sections 6.2.2 to 6.2.4. We now focus on how to compute the expressions in (6.8) in parallel.

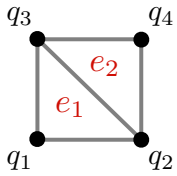
Several approaches for finite element matrices assembly on NVIDIA GPUs have been proposed in the literature. The most straightforward approach is to let one thread compute one nonzero entry of the matrix [17, 25]; we will refer to it as the intuitive method. Although this intuitive approach leads to independent threads, the workload of each thread is unbalanced (the size of $\mathcal{N}_{ij}$ varies), not to mention the thread divergence within a single warp (the instruction depends on indices i and j). The second approach maps one thread to one element: we use one thread to compute all the element data $\{H_{ij}^e : i,\, j \text{ are nodes of } e\}$ related to the element e then accumulate the computed data to the corresponding memory locations. However, race conditions occur during the stage of accumulation. Threads which are responsible for the elements in $\mathcal{N}_{ij}$ compete accumulating to the same memory location of Ψ_{ij}. To remedy race conditions, we may utilize atomic operations, which perform read-modify-write operations without interference from other threads [71, 68]. Note that atomic operations basically serialize the accumulation and may degrade the performance. Another

strategy is coloring the elements so that elements of the same color do not share nodes [42, 71, 25, 106, 68, 59]. With precomputed colors from mesh, each color is then processed in sequence without race conditions, hence there is no need to invoke atomic operations. One problem of the coloring scheme is that elements of the same color usually do not possess data locality, which yields read/write on memory locations that are far apart for threads in one warp. The third way to avoid race conditions is to separate the accumulation from the computation of element data; we will refer to this as the splitting method. We launch one CUDA kernel which maps one thread to one element for the computation of element data, and then launch a second CUDA kernel which maps one thread to one nonzero entry to accumulate computed element data [25, 38]. Since the element data have been computed, the overhead of the second CUDA kernel is smaller than the intuitive method mentioned previously even though they have exactly the same structure. However, there is an additional overhead of writing and reading element data to and from global memory in the splitting method.

Inspired by the intuitive method and the splitting method, we compare four approaches for (6.8) in the objective function evaluation of the atrophy model. We use the intuitive method as the baseline to be compared with the splitting method. To enhance data coalescing, we also adopt auxiliary data proposed in [106, 68]. The idea is to accompany the matrix of node positions (see Figure 6.5(b)) with an additional data matrix where each row stores positions of all nodes of one element (see Figure 6.5(c)) so that in the first CUDA kernel of the splitting method the global memory read from the auxiliary data is coalesced. Moreover, we notice that it is the difference of positions of nodes in one triangular or tetrahedral element that are relevant to the computation not the positions themselves; this observation suggests a different auxiliary data (see Figure 6.5(d)) which reduces the amount of global memory load. However, the memory read of the second CUDA kernel in the splitting method is not coalesced. We summarize the four methods we examined as follows:

- Method 1: The intuitive method without auxiliary data.
- Method 2: The splitting method without auxiliary data.
- Method 3: The splitting method with auxiliary data shown in Figure 6.5(c).
- Method 4: The splitting method with auxiliary data shown in Figure 6.5(d).

We remark that we need to keep the position matrix (Figure 6.5(b)) for Methods 3 and 4 since the computation related to elements is only a part of the objective function evaluation. For example, the position matrix best fits the computation of kernel matrices.



(a) A trivial mesh.

$$\begin{bmatrix} q_1^\top \\ q_2^\top \\ q_3^\top \\ q_4^\top \end{bmatrix}_{n\times d}$$

(b) Common data for Methods 1 to 4.

$$\begin{bmatrix} q_1^\top & q_2^\top & q_3^\top \\ q_2^\top & q_4^\top & q_3^\top \end{bmatrix}_{K\times\left((d+1)\times d\right)}$$

(c) Auxiliary data for Method 3.

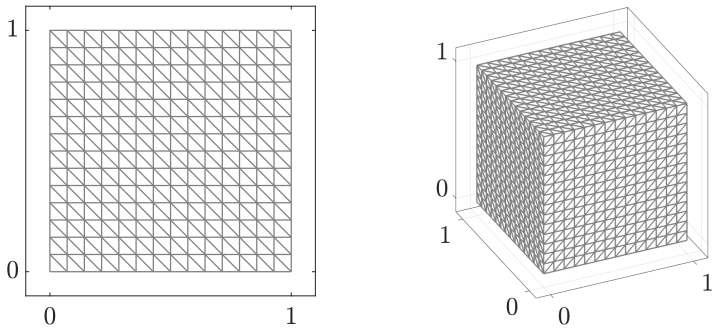
$$\begin{bmatrix} (q_2-q_1)^\top & (q_3-q_1)^\top \\ (q_4-q_2)^\top & (q_3-q_2)^\top \end{bmatrix}_{K\times d^2}$$

(d) Auxiliary data for Method 4.

Figure 6.5: Illustration of data for each method. K is the number of elements.

The 2D and 3D testing data we used as initial shapes are shown in Figure 6.6, and the sizes of testing problems are shown in Table 6.5. Note that the number of nodes of 2D and 3D testing problems are comparable while the number of elements of 3D testing problems are about 2 to 3 times more than the ones of 2D testing problems. To test the effect of auxiliary data in Methods 3 and 4, we consider two arrangements of nodes and elements. Figure 6.7 illustrates the data arrangement for the 2D case; the 3D case is a direct analogy. We observe two localization properties in Figure 6.7(a): for the consecutive elements $\{e_1, e_2, e_3\}$ (the reader may extend the concept to $\{e_1, \ldots, e_{32}\}$), the collective data of node positions for the three elements are $\{q_1, q_2, q_3, q_i, q_{i+1}\}$, which are clustered in two groups in the memory; for the consecutive nodes $\{q_i, q_{i+1}\}$, the collective adjacent elements for the two nodes are $\{e_1, e_2, e_3, e_4, e_k, e_{k+1}, e_{k+2}\}$, which are clustered in two groups in the memory. The first localization property, which we refer to as node localization (with respect to consecutive elements), benefits the first CUDA kernel in the splitting method, whereas the second localization property, which we refer to as element localization (with respect to consecutive nodes), benefits the second CUDA kernel in the splitting method. Note that localization is different from coalescing. As for the other arrangement of nodes and elements, we destroy both localization properties by permuting indices of elements as illustrated in Figure 6.7(b).

We first investigate the effect of the four methods on the running time of Av. Only the function computing Av was timed by `cudaEventRecord`, that is, we excluded the time of generating the auxiliary data in Methods 3 and 4, which is reasonable because

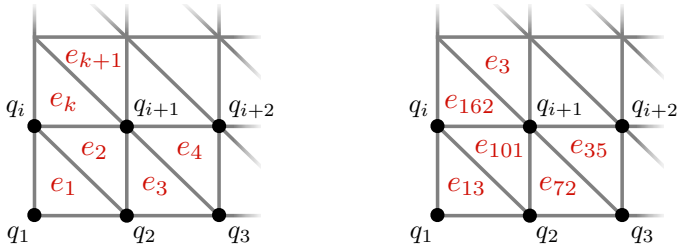


(a) Testing 2D mesh. (b) Testing 3D mesh.

Figure 6.6: Testing data.

Table 6.5: The sizes of testing problems.

2D		3D	
Number of Nodes	Number of Elements	Number of Nodes	Number of Elements
$41^2 = 1{,}681$	3,200	$11^3 = 1{,}331$	6,000
$61^2 = 3{,}721$	7,200	$15^3 = 3{,}375$	16,464
$81^2 = 6{,}561$	12,800	$19^3 = 6{,}859$	34,992



(a) Localized arrangement. (b) Non-localized arrangement (illustration).

Figure 6.7: Nodes and elements arrangement.

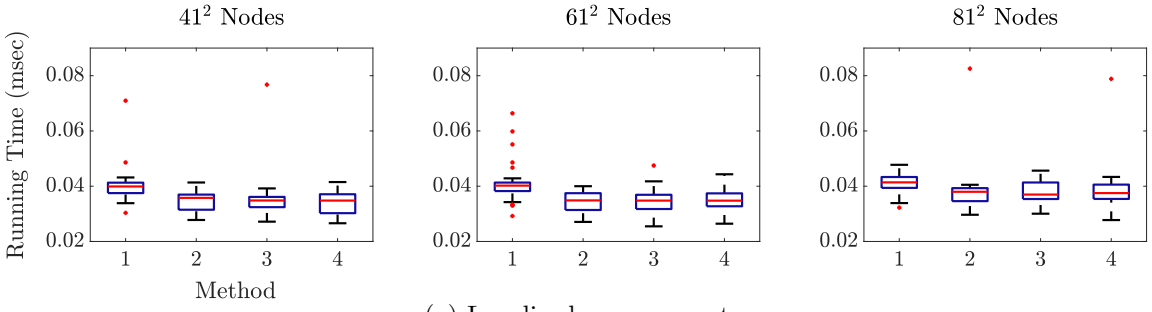
in the context of function evaluation we only update the auxiliary data once when we update positions of nodes, not every time when we compute Av. Every experiment was repeated on TITAN V 50 times, and each of the 50 runs was called from a shell script. The results are shown in Figures 6.8 and 6.9. From Figure 6.8 for 2D problems, we only see that the running times of Method 1 are slightly more than other methods; we have a clearer picture in Figure 6.9 for 3D problems since the computation of element data is more expensive and there are more elements adjacent to one node in 3D than in 2D. It is surprising that in the case of 19^3 nodes in Figure 6.9(a) Method 2 performs better than Method 3 and Method 4. Since we excluded the time of generating the auxiliary data, the better performance of Method 2 suggests that caching performs better than coalescing. The caching performance is reinforced by comparing Method 4 in the case of 19^3 nodes between Figure 6.9(a) and Figure 6.9(b). We summarize the running time of Av from the four methods and experiments of 19^3 nodes as follows.

- Node localization and element localization: 0.060 msec (Fig 6.9(a), Method 2)
- Node coalescing and element localization: 0.065 msec (Fig 6.9(a), Method 4)
- Node coalescing and element non-localization: 0.076 msec (Fig 6.9(b), Method 4)
- Node non-localization and element non-localization: 0.093 msec (Fig 6.9(b), Method 2)

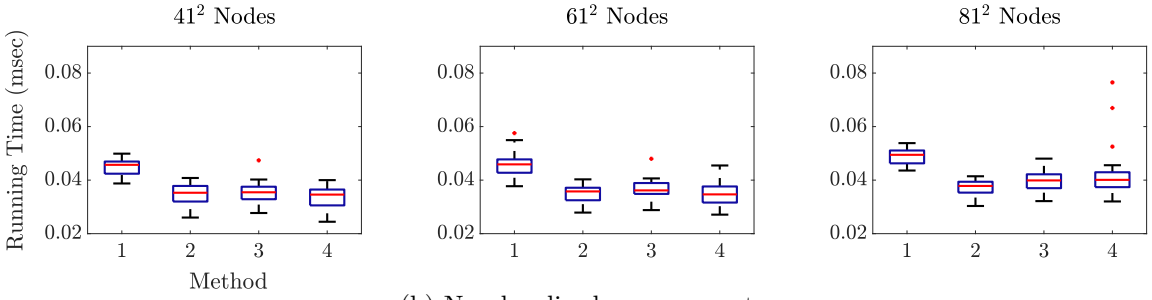
The running times of objective function evaluation are shown in Figures 6.10 and 6.11. We can see a similar pattern as the running time of Av. When we examine the speedup from Method 1 to Method 4 in Figure 6.11(b), however, the speedup is only slightly above 1 as compared to the over 2.5 times speedup in Figure 6.9(b). This is because matrix-matrix multiplications dominate the computation time of objective function evaluation especially when the number of nodes is large. In other words, to gain a further speedup requires a better implementation of matrix-matrix multiplication than the cuBLAS library. Although the speedup of objective function evaluation looks marginal, we remind the reader that this is the speedup from CUDA to CUDA. We report the overall, and more satisfying, speedup from OpenMP to CUDA in the next section.

6.3.3 Speedup of objective function evaluation

We summarize the speedup of objective function evaluation in Tables 6.6 and 6.7, in which we present the median of running times in 50 runs called from a shell script. In both objective functions, we used the testing data in Figure 6.6 and non-localized arrangement Figure 6.7(b) as our initial shapes with 51 discretized time points. The PCG* in Table 6.7

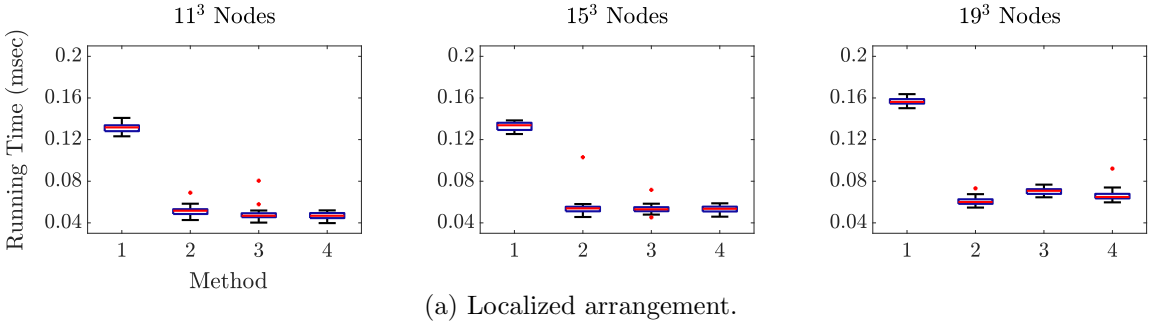


(a) Localized arrangement.

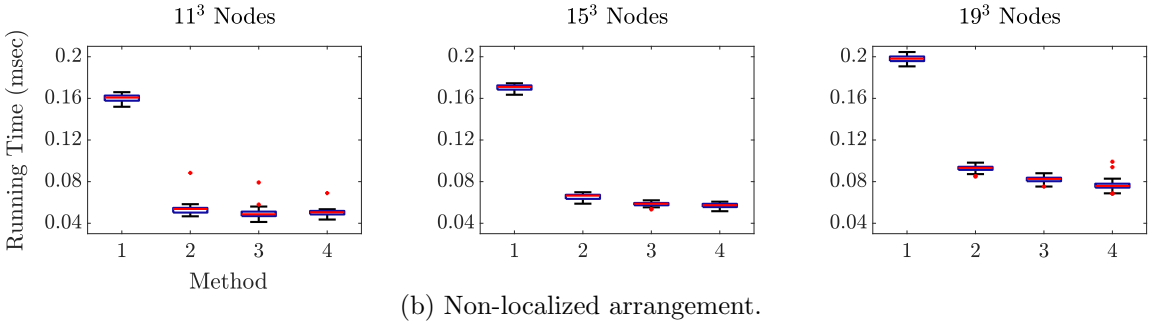


(b) Non-localized arrangement.

Figure 6.8: Comparison of running time of Av for 2D problems.

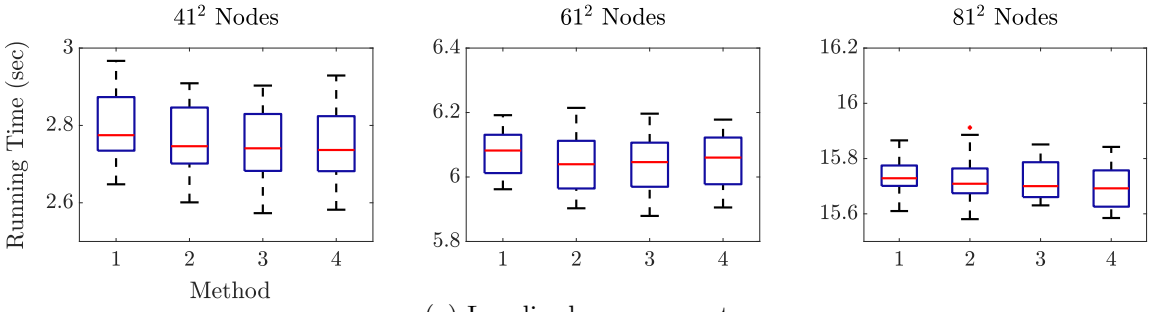


(a) Localized arrangement.

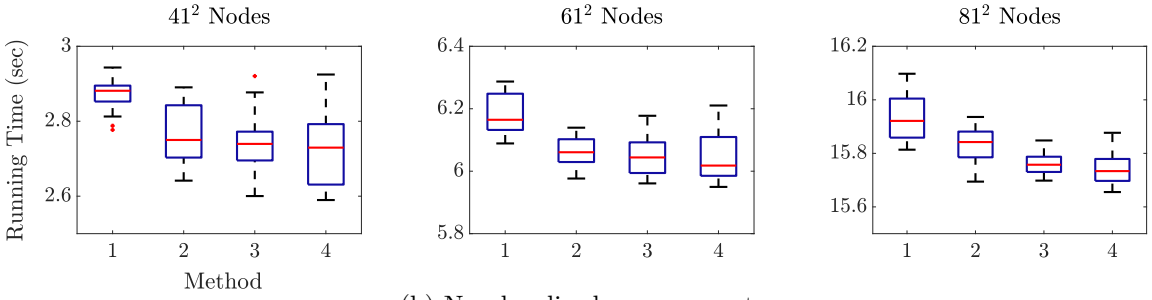


(b) Non-localized arrangement.

Figure 6.9: Comparison of running time of Av for 3D problems.

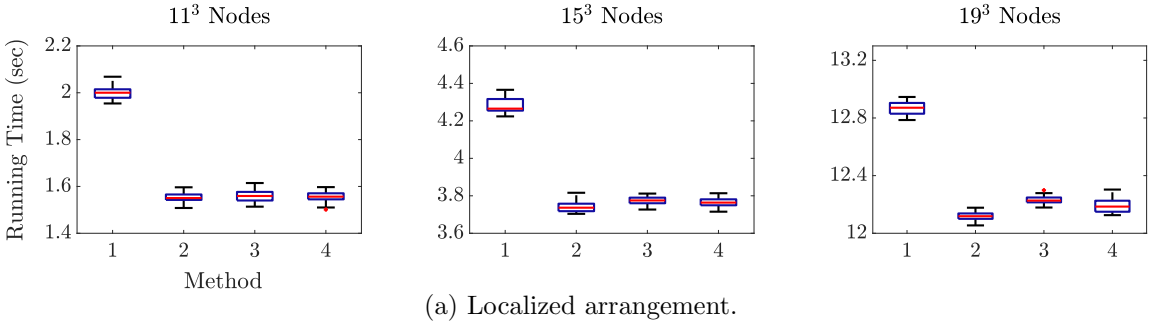


(a) Localized arrangement.

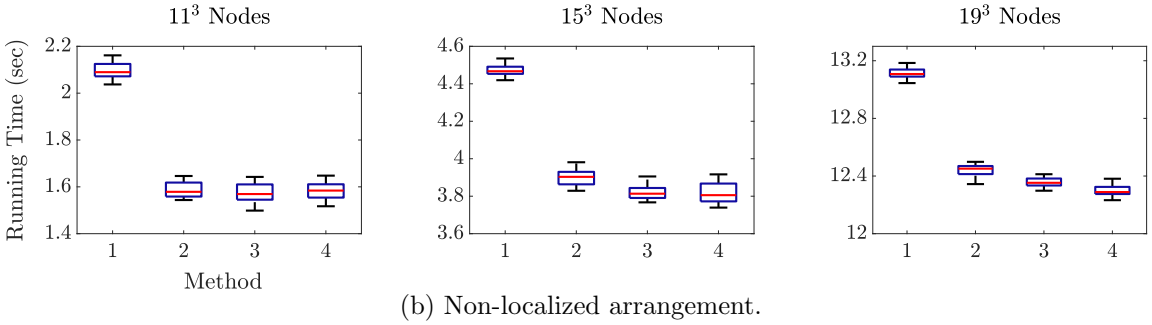


(b) Non-localized arrangement.

Figure 6.10: Comparison of running time of objective function evaluation for 2D problems.



(a) Localized arrangement.



(b) Non-localized arrangement.

Figure 6.11: Comparison of running time of objective function evaluation for 3D problems.

stands for the specialized PCG introduced in Section 6.2.5. The CPU version was run on a 2-socket system with Intel Xeon E5-2670 v3 (12 cores, 2.30GHz) with 48 OpenMP threads. In addition, we used multithreaded BLAS and LAPACK libraries `mwblas` and `mwlapack` provided by MATLAB. On the other hand, we ran our GPU version using cuBLAS and cuSOLVER on NVIDIA TITAN V. The data were stored in the row major format for the CPU version and in the column major format for the GPU version, and we simply replaced a CUDA kernel by `omp parallel for` for comparison; our OpenMP code was not NUMA aware. All operations were performed in double precision. We acknowledge that the speedup is problem dependent. The speedup in Tables 6.6 and 6.7 is meant to provide the reader an idea of potential acceleration.

Table 6.6: Speedup of objective function evaluation of piecewise-rigid motion.

Number of Nodes	Running Time (sec)		Speedup	Number of Nodes	Running Time (sec)		Speedup
	OpenMP	CUDA			OpenMP	CUDA	
41^2	3.068	0.364	8.4	11^3	1.162	0.219	5.3
61^2	14.915	0.990	15.1	15^3	6.109	0.712	8.6
81^2	45.350	2.430	18.7	19^3	33.545	2.285	14.7
91^2	70.687	3.769	18.8	20^3	45.067	3.059	14.7
101^2	109.173	5.809	18.8	22^3	81.409	5.846	13.9
111^2	163.613	9.051	18.1	23^3	110.443	8.070	13.7

Table 6.7: Speedup of objective function evaluation of the atrophy model.

Number of Nodes	Running Time (sec)			Speedup
	OpenMP + PCG	OpenMP + PCG*	CUDA + PCG*	
41^2	44.792	6.889	2.730	16.4
61^2	366.210	41.945	6.018	60.9
81^2	1227.783	131.735	15.734	78.0
91^2	1997.998	209.534	26.024	76.8
101^2	3102.931	330.595	39.289	79.0
111^2	4603.193	504.160	59.822	76.9
11^3	19.958	5.353	1.584	12.6
15^3	186.558	26.289	3.806	49.0
19^3	860.830	105.921	12.289	70.0
20^3	1157.632	148.062	16.493	70.2
22^3	2104.303	262.582	28.792	73.1
23^3	2804.688	359.998	39.647	70.7

Chapter 7

Conclusion

In the previous chapters, we have covered examples, theorems, and computation of the framework

$$\min_{\theta \in \Theta} \left(\int_0^T \Lambda(\varphi(t), \theta(t))\, dt + \rho(\varphi(T, \Omega_0), \Omega_{\text{targ}}) \right) \tag{2.10}$$

subject to

$$\varphi(t, x) = x + \int_0^t \mathcal{M}(\varphi_{|[0,\, s]}, \theta(s))(\varphi(s, x))\, ds \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d.$$

Our two examples demonstrate that this framework enables us to concentrate on the modeling of deformation vector fields. If the developed model satisfies the sufficient conditions of our theorems, diffeomorphic shape evolutions and diffeomorphic shape registrations will follow automatically. However, more sophisticated models could be far more complicated than the two examples we presented, especially when we take into account the interaction between the dynamics of shapes and their surrounding environment. For example, cell migration in fluid involves the interaction between deformable cell membranes and the fluid flow; the motion of heart valves is intertwined with the heart muscle and the blood flow. These topics reside in the regime of PDE-constrained optimization problems [15], which are more technically demanding. To tackle these problems numerically, more advanced methods like parallel algorithms for PDE-constrained optimization problems [3] and the immersed finite element method [107] are certainly required. Despite all the challenges, we envision the model-based direction as an attempt to partially answer the meaning of a sensible registration.

Appendix A

Bagatelles

A.1 Approximation of $x^\nu K_\nu(x)$, $\nu \in \mathbb{N}$

The modified Bessel function of the second kind K_ν has the recurrence relation [75, 51:5:2]

$$x^\nu\, K_\nu(x) = 2(\nu-1)\left(x^{\nu-1} K_{\nu-1}(x)\right) + x^2\left(x^{\nu-2}\, K_{\nu-2}(x)\right),$$

so we only need to approximate $x^2\, K_0(x)$ and $x\, K_1(x)$ to obtain $x^\nu K_\nu(x)$, $\nu \in \mathbb{N}$. The following approximations of $x^2\, K_0(x)$ and $x\, K_1(x)$ are modified from [29], whose Fortran code for K_0 and K_1 can be downloaded from `http://www.netlib.org/specfun/k0` and `http://www.netlib.org/specfun/k1`. The approximations of K_0 and K_1 from [29] have at least 18 significant decimal digits theoretically. The coefficients of polynomials $P_0^{(i)}$, $Q_0^{(i)}$, $P_1^{(i)}$, $Q_1^{(i)}$ are listed in Tables A.1–A.6, and $\varepsilon_{\text{mach}}$ stands for the machine epsilon. Note that the upper bound of the numerical support of K_0 and K_1 (705.342 and 705.343) could be machine dependent; we simply used 700 in our implementation. In addition, we can use the formula [75, 51:10:4]

$$\frac{d}{dx}(x^\nu\, K_\nu(x)) = -x^\nu\, K_{\nu-1}(x)$$

to further obtain approximate derivatives.

$$x^2\, K_0(x) \approx \begin{cases} 0, & \text{if } 0 \le x < \varepsilon_{\text{mach}}; \\ \dfrac{P_0^{(1)}(x^2)}{Q_0^{(1)}(x^2)}\, x^2 - \dfrac{P_0^{(2)}(x^2)}{Q_0^{(2)}(x^2)}\, x^4 \log x - x^2 \log x, & \text{if } \varepsilon_{\text{mach}} \le x \le 1; \\ \dfrac{P_0^{(3)}(1/x)}{Q_0^{(3)}(1/x)}\, x^{3/2}\, e^{-x}, & \text{if } 1 < x \le 705.342; \\ 0, & \text{if } x > 705.342. \end{cases}$$

$$x\,K_1(x) \approx \begin{cases} 1, & \text{if } 0 \le x < \varepsilon_{\text{mach}}; \\ \dfrac{P_1^{(1)}(x^2)}{Q_1^{(1)}(x^2)} + \dfrac{P_1^{(2)}(x^2)}{Q_1^{(2)}(x^2)}\, x^2 \log x, & \text{if } \varepsilon_{\text{mach}} \le x \le 1; \\ \dfrac{P_1^{(3)}(1/x)}{Q_1^{(3)}(1/x)}\, x^{1/2}\, e^{-x}, & \text{if } 1 < x \le 705.343; \\ 0, & \text{if } x > 705.343. \end{cases}$$

Table A.1: Coefficients of polynomials $P_0^{(1)}$ and $Q_0^{(1)}$.

	$P_0^{(1)}$	$Q_0^{(1)}$
1	2.4708152720399552679 e+03	2.1312714303849120380 e+04
x	5.9169059852270512312 e+03	−2.4994418972832303646 e+02
x^2	4.6850901201934832188 e+02	1
x^3	1.1999463724910714109 e+01	0
x^4	1.3166052564989571850 e−01	0
x^5	5.8599221412826100000 e−04	0

Table A.2: Coefficients of polynomials $P_0^{(2)}$ and $Q_0^{(2)}$.

	$P_0^{(2)}$	$Q_0^{(2)}$
1	−4.0320340761145482298 e+05	−1.6128136304458193998 e+06
x	−1.7733784684952985886 e+04	2.9865713163054025489 e+04
x^2	−2.9601657892958843866 e+02	−2.5064972445877992730 e+02
x^3	−1.6414452837299064100 e+00	1

A.2 Supplementary Proofs

Lemma A.2.1. *For all a, $b \in \mathbb{R}$ and $\varepsilon > 0$, we have*

$$ab \le \frac{1}{2\varepsilon}\, a^2 + \frac{\varepsilon}{2}\, b^2.$$

Proof. Note that

$$\left(\frac{1}{\sqrt{2\varepsilon}}\, a - \sqrt{\frac{\varepsilon}{2}}\, b \right)^2 \ge 0.$$

Rearranging the terms gives the result. □

Table A.3: Coefficients of polynomials $P_0^{(3)}$ and $Q_0^{(3)}$.

	$P_0^{(3)}$	$Q_0^{(3)}$
1	1.1600249425076035558 e+02	9.2556599177304839811 e+01
x	2.3444738764199315021 e+03	1.8821890840982713696 e+03
x^2	1.8321525870183537725 e+04	1.4847228371802360957 e+04
x^3	7.1557062783764037541 e+04	5.8824616785857027752 e+04
x^4	1.5097646353289914539 e+05	1.2689839587977598727 e+05
x^5	1.7398867902565686251 e+05	1.5144644673520157801 e+05
x^6	1.0577068948034021957 e+05	9.7418829762268075784 e+04
x^7	3.1075408980684392399 e+04	3.1474655750295278825 e+04
x^8	3.6832589957340267940 e+03	4.4329628889746408858 e+03
x^9	1.1394980557384778174 e+02	2.0013443064949242491 e+02
x^{10}	0	1

Table A.4: Coefficients of polynomials $P_1^{(1)}$ and $Q_1^{(1)}$.

	$P_1^{(1)}$	$Q_1^{(1)}$
1	−2.2149374878243304548 e+06	−2.2149374878243304548 e+06
x	7.1938920065420586101 e+05	3.7264298672067697862 e+04
x^2	1.7733324035147015630 e+05	−2.8143915754538725829 e+02
x^3	7.1885382604084798576 e+03	1
x^4	9.9991373567429309922 e+01	0
x^5	4.8127070456878442310 e−01	0

Table A.5: Coefficients of polynomials $P_1^{(2)}$ and $Q_1^{(2)}$.

	$P_1^{(2)}$	$Q_1^{(2)}$
1	−1.3531161492785421328 e+06	−2.7062322985570842656 e+06
x	−1.4758069205414222471 e+05	4.3117653211351080007 e+04
x^2	−4.5051623763436087023 e+03	−3.0507151578787595807 e+02
x^3	−5.3103913335180275253 e+01	1
x^4	−2.2795590826955002390 e−01	0

Table A.6: Coefficients of polynomials $P_1^{(3)}$ and $Q_1^{(3)}$.

	$P_1^{(3)}$	$Q_1^{(3)}$
1	2.2196792496874548962 e+00	1.7710478032601086579 e+00
x	4.4137176114230414036 e+01	3.4552228452758912848 e+01
x^2	3.4122953486801312910 e+02	2.5951223655579051357 e+02
x^3	1.3319486433183221990 e+03	9.6929165726802648634 e+02
x^4	2.8590657697910288226 e+03	1.9448440788918006154 e+03
x^5	3.4540675585544584407 e+03	2.1181000487171943810 e+03
x^6	2.3123742209168871550 e+03	1.2082692316002348638 e+03
x^7	8.1094256146537402173 e+02	3.3031020088765390854 e+02
x^8	1.3182609918569941308 e+02	3.6001069306861518855 e+01
x^9	7.5584584631176030810 e+00	1
x^{10}	6.4257745859173138767 e−02	0

Proposition A.2.2. *The layered elasticity tensor can be expressed as*

$$\begin{aligned}\mathcal{E}_\Phi(\varepsilon_u,\varepsilon_u) &= \lambda_{\tan}\big(\zeta_{11}+\zeta_{22}\big)^2 + \mu_{\tan}\big(\zeta_{11}^2+\zeta_{22}^2+2\,\zeta_{12}^2\big) + \mu_{\mathrm{tsv}}\,\zeta_{33}^2 + 2\mu_{\mathrm{ang}}\big(\zeta_{13}^2+\zeta_{23}^2\big)\\ &= \lambda_{\tan}\left(\mathrm{tr}(\varepsilon_u) - N^\top\varepsilon_u\,N\right)^2 + \mu_{\tan}\left(\mathrm{tr}(\varepsilon_u^2) - 2\,|\varepsilon_u\,N|^2 + (N^\top\varepsilon_u\,N)^2\right)\\ &\quad + \mu_{\mathrm{tsv}}\,(S^\top\varepsilon_u\,S)^2 + 2\mu_{\mathrm{ang}}\left(|\varepsilon_u S|^2 - (N^\top\varepsilon_u\,S)^2\right),\end{aligned}$$

where ζ_{ij} is the ij-th element of $\zeta_u = F^\top\varepsilon_u\,F$ and $F = \left[\begin{array}{ccc} T_1, & T_2, & S \end{array}\right]$.

Proof. We examine the four terms separately. First we notice that $\zeta_{33} = S^\top\varepsilon_u\,S$, which gives the third term. Since $\{T_1, T_2, N\}$ is an orthonormal basis, we have

$$|\varepsilon_u S|^2 \;=\; (T_1^\top\varepsilon_u\,S)^2 + (T_2^\top\varepsilon_u\,S)^2 + (N^\top\varepsilon_u\,S)^2 \;=\; \zeta_{13}^2+\zeta_{23}^2 + (N^\top\varepsilon_u\,S)^2,$$

which implies the fourth term

$$\zeta_{13}^2+\zeta_{23}^2 = |\varepsilon_u S|^2 - (N^\top\varepsilon_u\,S)^2.$$

To derive the first and second terms, we define a projection matrix by

$$P := \left[\begin{array}{ccc} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0 \end{array}\right].$$

For the first term, we observe that

$$\zeta_{11}+\zeta_{22} = \mathrm{tr}(P\,\zeta_u\,P) = \mathrm{tr}\big(P\,(F^\top\varepsilon_u\,F)\,P\big) = \mathrm{tr}\big(\varepsilon_u\,(F\,P\,F^\top)\big),$$

and we can write the second term as

$$\zeta_{11}^2 + \zeta_{22}^2 + 2\,\zeta_{12}^2 = \operatorname{tr}\Big((P\,\zeta_u\,P)^2\Big) = \operatorname{tr}\Big(\big(P\,(F^\top \varepsilon_u\,F)\,P\big)\big(P\,(F^\top \varepsilon_u\,F)\,P\big)\Big)$$
$$= \operatorname{tr}\Big(\big(\varepsilon_u\,(F\,P\,F^\top)\big)^2\Big).$$

Since $F\,P\,F^\top = I_3 - NN^\top$, we obtain the first term

$$(\zeta_{11} + \zeta_{22})^2 = \Big(\operatorname{tr}\big(\varepsilon_u\,(I_3 - NN^\top)\big)\Big)^2 = \Big(\operatorname{tr}(\varepsilon_u) - N^\top \varepsilon_u\,N\Big)^2$$

and the second term

$$\zeta_{11}^2 + \zeta_{22}^2 + 2\,\zeta_{12}^2 = \operatorname{tr}\Big(\big(\varepsilon_u\,(I_3 - NN^\top)\big)^2\Big) = \operatorname{tr}(\varepsilon_u^2) - 2\,N^\top \varepsilon_u^2\,N + (N^\top \varepsilon_u\,N)^2,$$

which completes the proof. □

Theorem A.2.3. *Let $v \in L^1([0,T], C_0^{p+1}(\mathbb{R}^d, \mathbb{R}^d))$. Then the initial value problem*

$$\varphi(t) = id + \int_0^t v(s) \circ \varphi(s)\,ds$$

has a unique solution in $C([0,T], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$.

Proof. Suppose that we have obtained a unique solution $\varphi_{t_0} \in C([0,t_0], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$ up to time t_0, which is true right at the beginning with $\varphi_0(0) = id$. Denote $\xi_{t_0} := \varphi_{t_0}(t_0)$. If we can show that there exist a fixed $\eta > 0$ and a unique $\varphi \in C([t_0, t_0+\eta], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$ satisfying

$$\varphi(t) = id + \int_{t_0}^t v(s) \circ \varphi(s)\,ds, \tag{A.1}$$

then since $w \mapsto w \circ \xi_{t_0}$ is in $\mathscr{L}(C_0^p(\mathbb{R}^d, \mathbb{R}^d), C_0^p(\mathbb{R}^d, \mathbb{R}^d))$, the extension $\varphi_{t_0} \oplus \varphi \in C([0, t_0 + \eta], \mathrm{Diff}_{id}^p(\mathbb{R}^d))$ defined by

$$(\varphi_{t_0} \oplus \varphi)(t) = \begin{cases} \varphi_{t_0}(t) & \text{if } t \in [0, t_0] \\ \varphi(t) \circ \xi_{t_0} & \text{if } t \in (t_0, t_0 + \eta] \end{cases}$$

will be the unique solution on $[0, t_0+\eta]$. Our proof will be complete by extending the unique solution repeatedly to $[0,T]$.

Since $\mathrm{Diff}_{id}^p(\mathbb{R}^d)$ is open in $id + C_0^p(\mathbb{R}^d, \mathbb{R}^d)$, there exists a fixed $r > 0$ such that $\overline{B}(id, r) := \{\xi' \in id + C_0^p(\mathbb{R}^d, \mathbb{R}^d) : \|\xi' - id\|_{p,\infty} \le r\}$ is contained in $\mathrm{Diff}_{id}^p(\mathbb{R}^d)$. With t_0 and r fixed, we define the iterate mapping $\Gamma : C([t_0, t_0+\eta], \overline{B}(id, r)) \to C([t_0, t_0+\eta], id + C_0^p(\mathbb{R}^d, \mathbb{R}^d))$ by

$$\Gamma(\varphi)(t) = id + \int_{t_0}^t v(s) \circ \varphi(s)\,ds.$$

The domain $C([t_0, t_0+\eta], \overline{B}(id, r))$ and the codomain $C([t_0, t_0+\eta], id + C_0^p(\mathbb{R}^d, \mathbb{R}^d))$ are both equipped with the metric

$$d(\varphi, \psi) := \sup_{t \in [t_0, t_0+\eta]} \|\varphi(t) - \psi(t)\|_{p,\infty},$$

which makes the domain a complete metric space. We are going to choose a fixed $\eta > 0$ such that Γ is a well-defined contraction. The Banach fixed point theorem will then imply the existence of a unique $\varphi \in C([t_0, t_0+\eta], \overline{B}(id, r)) \subset C([t_0, t_0+\eta], \mathrm{Diff}^p_{id}(\mathbb{R}^d))$ that satisfies (A.1). Since $s \mapsto v(s) \circ \varphi(s)$ is strongly measurable and

$$\int_{t_0}^{t} \|v(s) \circ \varphi(s)\|_{p,\infty}\, ds \le \int_{t_0}^{t} C_r \|v(s)\|_{p,\infty}\, ds \le C_r \int_{t_0}^{t_0+\eta} \|v(s)\|_{p,\infty}\, ds < \infty,$$

we know that the Bochner integral is well defined. Moreover, the same inequality also gives

$$\|\Gamma(\varphi)(t) - id\|_{p,\infty} \le C_r \int_{t_0}^{t_0+\eta} \|v(s)\|_{p,\infty}\, ds.$$

On the other hand, we have

$$d(\Gamma(\varphi), \Gamma(\psi)) \le \left(C_r' \int_{t_0}^{t_0+\eta} \|v(s)\|_{p+1,\infty}\, ds \right) d(\varphi, \psi).$$

We conclude that a unique fixed point $\varphi \in C([t_0, t_0+\eta], \overline{B}(id, r))$ of Γ exists as long as $\eta > 0$ satisfies

$$\int_{t_0}^{t_0+\eta} \|v(s)\|_{p+1,\infty}\, ds < \min\left\{ \frac{r}{C_r},\ \frac{1}{C_r'} \right\}.$$

Since the right-hand side is a fixed quantity, a fixed η can be chosen. □

Theorem A.2.4 (Grönwall's lemma)**.** *Let α, β, and u be measurable functions on $[a, b]$.*

(i) Suppose that α is nonnegative, β is nonnegative and integrable, and u is essentially bounded on $[a, b]$. If

$$u(t) \le \alpha(t) + \int_a^t \beta(s)\, u(s)\, ds \quad \text{for all } t \in [a, b],$$

then

$$u(t) \le \alpha(t) + \int_a^t \alpha(s)\, \beta(s)\, \exp\left(\int_s^t \beta(s')\, ds' \right) ds \quad \text{for all } t \in [a, b].$$

The bound could be infinity.

(ii) Suppose that β is nonnegative and integrable and u is essentially bounded on $[a, b]$. If

$$u(t) \le C + \int_a^t \beta(s)\, u(s)\, ds \quad \text{for all } t \in [a, b],$$

where $C \in \mathbb{R}$, then

$$u(t) \leq C \, \exp\left(\int_a^t \beta(s) \, ds \right) \qquad \textit{for all } t \in [a, b].$$

Proof. (i) The assumption is

$$u(t) \leq \alpha(t) + \int_a^t \beta(s) \, u(s) \, ds.$$

Since β is nonnegative, we plug the above inequality into its right-hand side and obtain

$$\begin{aligned} u(t) &\leq \alpha(t) + \int_a^t \beta(s_0) \, \left(\alpha(s_0) + \int_a^{s_0} \beta(s_1) \, u(s_1) \, ds_1 \right) ds_0 \\ &= \alpha(t) + \int_a^t \beta(s_0) \, \alpha(s_0) \, ds_0 + \int_a^t \int_a^{s_0} \beta(s_0) \, \beta(s_1) \, u(s_1) \, ds_1 \, ds_0. \end{aligned}$$

We then plug the first inequality into the right-hand side of the second inequality:

$$\begin{aligned} u(t) &\leq \alpha(t) + \int_a^t \beta(s_0) \, \alpha(s_0) \, ds_0 \\ &\quad + \int_a^t \int_a^{s_0} \beta(s_0) \, \beta(s_1) \, \left(\alpha(s_1) + \int_a^{s_1} \beta(s_2) \, u(s_2) \, ds_2 \right) ds_1 \, ds_0 \\ &= \alpha(t) + \int_a^t \beta(s_0) \, \alpha(s_0) \, ds_0 + \int_a^t \int_a^{s_0} \beta(s_0) \, \beta(s_1) \, \alpha(s_1) \, ds_1 \, ds_0 \\ &\quad + \int_a^t \int_a^{s_0} \int_a^{s_1} \beta(s_0) \, \beta(s_1) \, \beta(s_2) \, u(s_2) \, ds_2 \, ds_1 \, ds_0. \end{aligned}$$

Thus, at the n-th iterate, we have

$$\begin{aligned} u(t) \leq \alpha(t) &+ \int_a^t \beta(s_0) \, \alpha(s_0) \, ds_0 + \int_a^t \int_a^{s_0} \beta(s_0) \, \beta(s_1) \, \alpha(s_1) \, ds_1 \, ds_0 + \cdots \\ &+ \int_a^t \int_a^{s_0} \cdots \int_a^{s_{n-1}} \beta(s_0) \cdots \beta(s_{n-1}) \, \alpha(s_{n-1}) \, ds_{n-1} \cdots ds_0 \\ &+ \int_a^t \int_a^{s_0} \cdots \int_a^{s_n} \beta(s_0) \cdots \beta(s_n) \, u(s_n) \, ds_n \cdots ds_0. \end{aligned} \tag{A.2}$$

Since α, β are nonnegative and u is essentially bounded on $[a, b]$, Tonelli's theorem yields

$$\begin{aligned} u(t) \leq \alpha(t) &+ \int_a^t \beta(s_0) \, \alpha(s_0) \, ds_0 + \int_a^t \beta(s_1) \, \alpha(s_1) \left(\int_{s_1}^t \beta(s_0) ds_0 \right) ds_1 + \cdots \\ &+ \int_a^t \beta(s_{n-1}) \, \alpha(s_{n-1}) \left(\int_{s_{n-1}}^t \cdots \int_{s_1}^t \beta(s_0) \cdots \beta(s_{n-2}) \, ds_0 \cdots ds_{n-2} \right) ds_{n-1} \\ &+ \left(\operatorname*{ess\,sup}_{s_n \in [a,b]} |u(s_n)| \right) \int_a^t \int_{s_n}^t \cdots \int_{s_1}^t \beta(s_0) \cdots \beta(s_n) \, ds_0 \cdots ds_n. \end{aligned}$$

Denote by μ the measure with density β with respect to the Lebesgue measure. Note that

$\int_x^t \int_{s_k}^t \cdots \int_{s_1}^t \beta(s_0)\cdots\beta(s_k)\,ds_0\cdots ds_k$ is the μ-measure of one of the $(k+1)!$ simplexes that equally split the μ-measure of the cube $[x,t]^{k+1}$, thus

$$\begin{aligned}\int_x^t \int_{s_k}^t \cdots \int_{s_1}^t \beta(s_0)\cdots\beta(s_k)\,ds_0\cdots ds_k &= \frac{1}{(k+1)!}\int_{[x,\,t]^{k+1}} \beta(s_0)\cdots\beta(s_k)\,ds_0\cdots ds_k \\ &= \frac{1}{(k+1)!}\left(\int_x^t \beta(s')\,ds'\right)^{k+1}.\end{aligned}$$

We continue the above inequality and write

$$\begin{aligned}u(t) \le \alpha(t) &+ \int_a^t \beta(s_0)\,\alpha(s_0)\,ds_0 + \int_a^t \beta(s_1)\,\alpha(s_1)\;\frac{1}{1!}\left(\int_{s_1}^t \beta(s')\,ds'\right)ds_1 + \cdots \\ &+ \int_a^t \beta(s_{n-1})\,\alpha(s_{n-1})\;\frac{1}{(n-1)!}\left(\int_{s_{n-1}}^t \beta(s')\,ds'\right)^{n-1} ds_{n-1} \\ &+ \Big(\operatorname*{ess\,sup}_{s\,\in\,[a,\,b]} |u(s)|\Big)\;\frac{1}{(n+1)!}\left(\int_a^t \beta(s')\,ds'\right)^{n+1} \\ = \alpha(t) &+ \int_a^t \beta(s)\,\alpha(s)\left(\sum_{k=0}^{n-1}\frac{1}{k!}\left(\int_s^t \beta(s')\,ds'\right)^{k}\right)ds \\ &+ \frac{\operatorname{ess\,sup}|u(s)|}{(n+1)!}\left(\int_a^t \beta(s')\,ds'\right)^{n+1}.\end{aligned}$$

When $n\to\infty$, the last term goes to zero since β is integrable. Hence we conclude that

$$u(t) \le \alpha(t) + \int_a^t \alpha(s)\,\beta(s)\,\exp\!\left(\int_s^t \beta(s')\,ds'\right)ds.$$

(ii) Observe that we only use the nonnegativeness of α to invoke Tonelli's theorem. Thus we can replace α by a constant $C\in\mathbb{R}$ in (A.2), apply the Tonelli's theorem thanks to the nonnegativeness of β, and then arrive at

$$u(t) \le C\sum_{k=0}^{n}\frac{1}{k!}\left(\int_a^t \beta(s')\,ds'\right)^{k} + \frac{\operatorname{ess\,sup}|u(s)|}{(n+1)!}\left(\int_a^t \beta(s')\,ds'\right)^{n+1}.$$

Letting $n\to\infty$ leads to

$$u(t) \le C\,\exp\!\left(\int_a^t \beta(s)\,ds\right).$$

□

Proposition A.2.5. *Let $f\in L^2([0,T],H^1(\Omega))$. Then $\partial_t f = g \in L^2([0,T],(H^1(\Omega))^*)$, where the time derivative is in the sense of distribution, if and only if*

$$\frac{d}{dt}\langle f(\cdot),\psi\rangle_{L^2} = (g(\cdot)\mid\psi)_{(H^1)^*,\,H^1}$$

for almost every $t \in [0,T]$ *and all* $\psi \in H^1(\Omega)$.

Proof. ($\Rightarrow$) By definition, the distributional time derivative $\partial_t f = g$ means that

$$-\int_0^T f(t)\,\dot{\varphi}(t)\,dt = \int_0^T g(t)\,\varphi(t)\,dt \;\in (H^1(\Omega))^* \quad \text{for all } \varphi \in \mathscr{D}\big((0,T)\big).$$

Evaluating the above equation at an arbitrary $\psi \in H^1(\Omega)$ gives

$$\Big(-\int_0^T f(t)\,\dot{\varphi}(t)\,dt \,\Big|\, \psi\Big)_{(H^1)^*,\,H^1} = \Big(\int_0^T g(t)\,\varphi(t)\,dt \,\Big|\, \psi\Big)_{(H^1)^*,\,H^1} \qquad \text{for all } \varphi \in \mathscr{D}\big((0,T)\big),$$

which, from Theorem 2.2.8, implies that

$$-\int_0^T \langle f(t),\psi\rangle_{L^2}\,\dot{\varphi}(t)\;dt = \int_0^T (g(t)\mid\psi)_{(H^1)^*,\,H^1}\,\varphi(t)\;dt \qquad \text{for all } \varphi \in \mathscr{D}\big((0,T)\big).$$

In other words, the weak derivative of the real-valued function $\langle f(\cdot),\psi\rangle_{L^2}$ is $\big(g(\cdot)\mid\psi\big)_{(H^1)^*,\,H^1}$, that is, $\frac{d}{dt}\langle f(\cdot),\psi\rangle_{L^2} = (g(\cdot)\mid\psi)_{(H^1)^*,\,H^1}$ for almost every $t \in [0,T]$.

($\Leftarrow$) Note that we can reverse the arguments in the above proof. □

A.3 Reaction and Atrophy Functions

In our numerical experiments of the atrophy model, we have used the reaction and atrophy functions shown in Figure 3.16. Both reaction and atrophy functions are shifted and scaled from C^2 functions compactly supported on $[-1,1]$ with maximum 1, that is,

$$R(\tau) := R_{\max}\,\widetilde{R}\Big(-1+2\,\frac{\tau-\tau_{\min}}{\tau_{\max}-\tau_{\min}}\Big)$$

and

$$\alpha(\tau) := \alpha_{\max}\,\widetilde{\alpha}\Big(-1+2\,\frac{\tau-\tau_{\min}}{\tau_{\max}-\tau_{\min}}\Big).$$

The C^2 function $\widetilde{R}:\mathbb{R}\to\mathbb{R}$ is defined by

$$\widetilde{R}(x) = \begin{cases} b_1\,(x+1)^3\,(x-b_2), & \text{if } x \in [-1,-1+\frac{d}{2});\\ c_1\,x^3 + c_2\,x^2 + c_3\,x + c_4, & \text{if } x \in [-1+\frac{d}{2},-1+d);\\ -a\,x^2+1, & \text{if } x \in [-1+d,1-d);\\ -c_1\,x^3 + c_2\,x^2 - c_3\,x + c_4, & \text{if } x \in [1-d,1-\frac{d}{2});\\ b_1\,(-x+1)^3\,(-x-b_2), & \text{if } x \in [1-\frac{d}{2},1];\\ 0, & \text{otherwise,}\end{cases}$$

where

$$\begin{cases} b_1 = \dfrac{8\,ad^2 - 48\,ad + 48\,(a-1)}{3\,d^4} \\ b_2 = \dfrac{29\,ad^3 - 156\,ad^2 + (264\,a - 120)\,d - 144\,(a-1)}{24\,ad^2 - 144\,ad + 144\,(a-1)} \\ c_1 = \dfrac{-ad^2 + 12\,ad - 24\,(a-1)}{9\,d^3} \\ c_2 = \dfrac{-2\,ad^3 - 13\,ad^2 + (36\,a - 24)\,d - 24\,(a-1)}{3\,d^3} \\ c_3 = -\dfrac{(1-d)^2\,(ad^2 - 12\,ad + 24\,(a-1))}{3\,d^3} \\ c_4 = \dfrac{ad^5 - 15\,ad^4 + (63\,a - 15)\,d^3 - (109\,a - 72)\,d^2 + (84\,a - 72)\,d - 24\,(a-1)}{9\,d^3} \end{cases}.$$

We used $a = 0.3$ and $d = 0.5$ in our numerical experiments. The C^2 function $\widetilde{\alpha} : \mathbb{R} \to \mathbb{R}$ is defined by

$$\widetilde{\alpha}(x) = \begin{cases} b_1\,(x+1)^3\,(x - b_2), & \text{if } x \in [-1, -1+d); \\ -a\,x^2 + 1, & \text{if } x \in [-1+d, 1-d); \\ b_1\,(-x+1)^3\,(-x - b_2), & \text{if } x \in [1-d, 1]; \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\begin{cases} a = \dfrac{6}{d^2 - 6\,d + 6} \\ b_1 = \dfrac{3\,(d-2)}{d^3\,(d^2 - 6\,d + 6)} \\ b_2 = \dfrac{8\,d^2 - 15\,d + 6}{3\,(d-2)} \end{cases}.$$

We used $d = 0.5$ in our numerical experiments.

A.4 Gradient Computation of Piecewise Rigid Motion

To simplify our presentation, we consider the case when the shape has only one connected component. Given a discretized initial shape $\{q_{j,0}\}_{j=1}^{m}$, a rotation center c_0, and a discretized target shape $\{q_{j,\mathrm{targ}}\}_{j=1}^{m'}$, the objective function is given by

$$J(\theta) := \frac{1}{2}\int_0^T \Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t) + |\theta(t)|^2\Big)\,dt + \rho\Big(\{q_j(T)\}_j, \{q_{j,\mathrm{targ}}\}_j\Big)$$

subject to

$$\begin{cases} c(t) = c_0 + \displaystyle\int_0^t u(s)\,ds \\ \alpha(t) = \displaystyle\int_0^t \omega(s)\,ds \\ q_j(t) = \mathcal{P}\big(c(t), \alpha(t), q_{j,0}, c_0\big) \\ v_j(t) = \dot{\mathcal{P}}\big(u(t), \omega(t), \alpha(t), q_{j,0}, c_0\big) \end{cases},$$

where $\theta(t) = \big(u(t), \omega(t)\big)$. We let $y := (c, \alpha)$ be our state variable. We then introduce the costate $p : [0, T] \to \mathbb{R}^{\frac{d(d+1)}{2}}$ and form the Lagrangian

$$\begin{aligned} L(y, p, \theta) := \frac{1}{2}\int_0^T \Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t) + |\theta(t)|^2\Big)\,dt + \rho\Big(\{q_j(T)\}_j, \{q_{j,\mathrm{targ}}\}_j\Big) \\ + \int_0^T p(t)^\top\big(\dot{y}(t) - \theta(t)\big)\,dt \end{aligned}$$

subject to

$$\begin{cases} q_j(t) = \mathcal{P}\big(c(t), \alpha(t), q_{j,0}, c_0\big) \\ v_j(t) = \dot{\mathcal{P}}\big(u(t), \omega(t), \alpha(t), q_{j,0}, c_0\big) \end{cases}.$$

For each θ, we let

$$y_\theta(t) = y_0 + \int_0^t \theta(s)\,ds$$

and look for p_θ such that

$$\begin{cases} \partial_y L(y_\theta, p_\theta, \theta) = 0 \\ \partial_p L(y_\theta, p_\theta, \theta) = 0 \end{cases}.$$

With such chosen y_θ and p_θ, we deduce that

$$dJ(\theta) = \partial_\theta L(y_\theta, p_\theta, \theta) = \partial_{\theta'} L(y_\theta, p_\theta, \theta')\big|_{\theta'=\theta},$$

thus $\partial_{\theta'} L(y_\theta, p_\theta, \theta')|_{\theta'=\theta}$ is the gradient of the objective function J. We now show how to obtain p_θ. Derivatives of L with respect to y and p are given by

$$\begin{cases} (\partial_y L(y, p, \theta) \mid \delta y) = \dfrac{1}{2}\displaystyle\int_0^T \partial_{y(t)}\Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\Big)^\top \delta y(t)\,dt \\ \qquad\qquad + \Big(\partial_{y(T)}\,\rho\Big(\{q_j(T)\}_j, \{q_{j,\mathrm{targ}}\}_j\Big)\Big)^\top \delta y(T) \\ \qquad\qquad + p(T)^\top \delta y(T) - \displaystyle\int_0^T \dot{p}(t)^\top \delta y(t)\,dt \\ (\partial_p L(q, p, \theta) \mid \delta p) = \displaystyle\int_0^T \delta p(t)^\top(\dot{y}(t) - \theta(t))\,dt \end{cases}.$$

Note that $\partial_y L(y,p,\theta)=0$ is equivalent to

$$\begin{cases} \dot{p}(t) = \frac{1}{2}\,\partial_{y(t)}\Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\Big) & \text{for almost every } t\in[0,T] \\ p(T) = -\partial_{y(T)}\,\rho\Big(\{q_j(T)\}_j,\{q_{j,\mathrm{targ}}\}_j\Big) \end{cases}.$$

In addition, $\partial_p L(y,p,\theta)=0$ is equivalent to $\dot{y}(t)=\theta(t)$ for almost every $t\in[0,T]$. Hence we can compute the gradient at θ as follows. First we compute

$$y_\theta(t) = y_0 + \int_0^t \theta(s)\,ds,$$

which satisfies $\partial_p L(y_\theta,p,\theta)=0$ for all p. Plugging $y_\theta=(c_\theta,\alpha_\theta)$ into the dynamics of p, recall that $\theta=(u,\omega)$, next we compute

$$\begin{cases} p_\theta(t) = -\partial_{y(T)}\,\rho\Big(\{q_j(T)\}_j,\{q_{j,\mathrm{targ}}\}_j\Big) + \displaystyle\int_T^t \frac{1}{2}\,\partial_{y(t)}\Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\Big)\,dt \\ q_j(t) = \mathcal{P}\big(c_\theta(t),\alpha_\theta(t),q_{j,0},c_0\big) \\ v_j(t) = \dot{\mathcal{P}}\big(u(t),\omega(t),\alpha_\theta(t),q_{j,0},c_0\big) \end{cases},$$

where we have

$$\begin{aligned} &\frac{1}{2}\,\partial_{y(t)}\Big(\boldsymbol{v}(t)^\top K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)\Big) \\ &= \Big(\partial_{y(t)}\boldsymbol{v}(t)\Big)\,K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t) - \frac{1}{2}\,\Big(\partial_{y(t)}\boldsymbol{q}(t)\Big)\;\partial_{\boldsymbol{q}(t)}\Big(\boldsymbol{w}^\top K_V(\boldsymbol{q}(t))\,\boldsymbol{w}\Big)\Big|_{\boldsymbol{w}=K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t)}. \end{aligned}$$

Since y_θ and p_θ satisfy the requirements, we can obtain the gradient of J as

$$\Big(\partial_{\theta'} L(y_\theta,p_\theta,\theta')\big|_{\theta'=\theta}\Big)(t) = \Big(\partial_{\theta(t)}\boldsymbol{v}(t)\Big)\,K_V(\boldsymbol{q}(t))^{-1}\,\boldsymbol{v}(t) + \theta(t) - p_\theta(t).$$

A.5 Derivation of Mass and Stiffness Matrices

We aim to approximate

$$\int_{\mathcal{T}_0} \psi_{i,0}\,\psi_{j,0}\,\det D\varphi(t)\;dx, \tag{A.3}$$

$$\int_{\mathcal{T}_0} \nabla\psi_{i,0}^\top\,W_{\varphi(t)}\,\nabla\psi_{j,0}\,\det D\varphi(t)\;dx, \tag{A.4}$$

and

$$\int_{\mathcal{T}_0} R\Big(\sum_{k=1}^n \tau_k(t)\,\psi_{k,0}\Big)\psi_{i,0}\,\det D\varphi(t)\;dx. \tag{A.5}$$

Since the only difference in the derivation between $d=2$ and $d=3$ is the piecewise linear basis functions, which lead to different coefficients, our discussion will be based on $d=2$

for simplicity. We prepare some notations. Denote the positions of three nodes at vertices of $\mathcal{T}_0$ by $q_{0,0}$, $q_{1,0}$, and $q_{2,0}$, and denote their evolution by $t \mapsto q_i(t)$. We also define

$$Q_0 := \begin{bmatrix} q_{1,0} - q_{0,0}, & q_{2,0} - q_{0,0} \end{bmatrix} \quad \text{and} \quad Q(t) := \begin{bmatrix} q_1(t) - q_0(t), & q_2(t) - q_0(t) \end{bmatrix}.$$

In addition, we denote by $\mathcal{T}^P$ the triangular parent element formed by vertices $e_0 := (0,0)$, $e_1 := (1,0)$, and $e_2 := (0,1)$. The linear basis functions on the parent domain are given by

$$\begin{cases} \psi_0^P(x,y) = -x - y + 1 \\ \psi_1^P(x,y) = x \\ \psi_2^P(x,y) = y \end{cases}.$$

By relabeling if necessary, we assume that the global indices of the three nodes, $q_0(t)$, $q_1(t)$, $q_2(t)$, are also equal to the local indices, $0, 1, 2$, to avoid an additional index mapping between the parent and the physical domain.

To approximate (A.3), first we note that

$$\det D\varphi(t,x) \approx \frac{|\det Q(t)|}{|\det Q_0|} \quad \text{for } x \in \mathcal{T}_0.$$

We make a change of variables $x \mapsto q_{0,0} + Q_0\, x$ which maps $\mathcal{T}^P$ to $\mathcal{T}_0$ and obtain

$$\begin{aligned} \int_{\mathcal{T}_0} \psi_{i,0}\, \psi_{j,0}\, \det D\varphi(t)\; dx &\approx \int_{\mathcal{T}^P} \psi_i^P(x)\, \psi_j^P(x)\, \frac{|\det Q(t)|}{|\det Q_0|}\, |\det Q_0|\; dx \\ &= \begin{cases} \dfrac{1}{12}\, |\det Q(t)|, & \text{if } i,j \in \mathcal{T}_0 \text{ and } i = j\,; \\ \dfrac{1}{24}\, |\det Q(t)|, & \text{if } i,j \in \mathcal{T}_0 \text{ and } i \neq j\,; \\ 0, & \text{otherwise}\,. \end{cases} \end{aligned}$$

For the second integral (A.4), we recall from (3.13) that

$$W_{\varphi(t)} = \big(D\varphi(t)\big)^{-1} \big(U_{\varphi(t)} \circ \varphi(t)\big)\; D\varphi(t)^{-\top}.$$

It follows that

$$\begin{aligned} &\int_{\mathcal{T}_0} \nabla\psi_{i,0}^\top\, W_{\varphi(t)}\, \nabla\psi_{j,0}\, \det D\varphi(t)\; dx \\ &= \int_{\mathcal{T}_0} \Big(D\varphi(t)^{-\top} \nabla\psi_{i,0}\Big)^\top \big(U_{\varphi(t)} \circ \varphi(t)\big)\, \Big(D\varphi(t)^{-\top} \nabla\psi_{j,0}\Big)\, \det D\varphi(t)\; dx. \end{aligned}$$

We approximate $U_{\varphi(t)} \circ \varphi(t)$ on $\mathcal{T}_0$ by a constant matrix $U_{\boldsymbol{q}(t)}$ and approximate $D\varphi(t)$ by $Q(t)\, Q_0^{-1}$. Moreover, since $\psi_i^P(x) = \psi_{i,0}(q_{0,0} + Q_0\, x)$, we have $\nabla\psi_i^P(x) = Q_0^\top\, \nabla\psi_{i,0}(q_{0,0} +$

$Q_0\,x$). Thus a change of variables $x \mapsto q_{0,0} + Q_0\,x$ will change

$$D\varphi(t)^{-\top} \nabla\psi_{i,0} \approx \big(Q(t)\,Q_0^{-1}\big)^{-\top}\,\nabla\psi_{i,0}$$

to

$$\big(Q(t)\,Q_0^{-1}\big)^{-\top}\,\big(Q_0^{-\top}\,\nabla\psi_i^P\big) = Q(t)^{-\top}\,\nabla\psi_i^P. \tag{A.6}$$

Using the identity (A.6), we make a change of variables and approximate the integral by

$$\begin{aligned}
&\int_{\mathcal{T}_0} \nabla\psi_{i,0}^\top\, W_{\varphi(t)}\,\nabla\psi_{j,0}\,\det D\varphi(t)\;dx \\
&\approx \int_{\mathcal{T}^P} \big(Q(t)^{-\top}\,\nabla\psi_i^P\big)^\top U_{\boldsymbol{q}(t)}\,\big(Q(t)^{-\top}\,\nabla\psi_j^P\big)\;\frac{|\det Q(t)|}{|\det Q_0|}\,|\det Q_0|\;dx \\
&= \begin{cases} \dfrac{1}{2}\,\big(\nabla\psi_i(t)\big)^\top_{\mathcal{T}(t)}\,U_{\boldsymbol{q}(t)}\,\big(\nabla\psi_j(t)\big)_{\mathcal{T}(t)}\,|\det Q(t)|, & \text{if } i,j \in \mathcal{T}_0\,; \\ 0, & \text{otherwise}\,. \end{cases}
\end{aligned}$$

In the third integral (A.5), we approximate the nonlinear function $R\Big(\sum_{k=1}^n \tau_k(t)\,\psi_{k,0}\Big)$ by linear interpolation:

$$R\Big(\sum_{k=1}^n \tau_k(t)\,\psi_{k,0}\Big) \approx R(\tau_i(t))\,\psi_{i,0} + R(\tau_j(t))\,\psi_{j,0} + R(\tau_k(t))\,\psi_{k,0},$$

which leads to

$$\begin{aligned}
&\int_{\mathcal{T}_0} R\Big(\sum_{k=1}^n \tau_k(t)\,\psi_{k,0}\Big)\psi_{i,0}\,\det D\varphi(t)\;dx \\
&\approx \int_{\mathcal{T}_0} \Big(R(\tau_i(t))\,\psi_{i,0}(x) + R(\tau_j(t))\,\psi_{j,0}(x) + R(\tau_k(t))\,\psi_{k,0}(x)\Big)\,\psi_{i,0}(x)\;\frac{|\det Q(t)|}{|\det Q_0|}\;dx \\
&= \int_{\mathcal{T}^P} \Big(R(\tau_i(t))\,\psi_i^P(x) + R(\tau_j(t))\,\psi_j^P(x) + R(\tau_k(t))\,\psi_k^P(x)\Big)\,\psi_i^P(x)\;\frac{|\det Q(t)|}{|\det Q_0|}\,|\det Q_0|\;dx \\
&= \frac{1}{24}\Big(2R(\tau_i(t)) + R(\tau_j(t)) + R(\tau_k(t))\Big)\,|\det Q(t)|.
\end{aligned}$$

Bibliography

[1] Achdou, Y., Franchi, B., Marcello, N., and Tesi, M. C. A qualitative model for aggregation and diffusion of β-amyloid in Alzheimer's disease. *Journal of Mathematical Biology 67* (2013), 1369–1392.

[2] Adams, R., and Fournier, J. *Sobolev Spaces*, 2 ed., vol. 140 of *Pure and Applied Mathematics*. Academic Press, 2003.

[3] Akçelik, V., Biros, G., Ghattas, O., Hilla, J., Keyes, D., and van Bloemen Waanders, B. Parallel algorithms for PDE-constrained optimization. In *Parallel Processing for Scientific Computing*. SIAM, 2006, ch. 16, pp. 291–322.

[4] Arguillère, S., Trélat, E., Trouvé, A., and Younes, L. Shape deformation analysis from the optimal control viewpoint. *Journal de Mathématiques Pures et Appliquées 104*, 1 (2015), 139–178.

[5] Aronszajn, N. Theory of reproducing kernels. *Transactions of the American Mathematical Society 68*, 3 (1950), 337–404.

[6] Arsigny, V., Pennec, X., and Ayache, N. Polyrigid and polyaffine transformations: A new class of diffeomorphisms for locally rigid or affine registration. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003* (Berlin, Heidelberg, 2003), R. E. Ellis and T. M. Peters, Eds., Springer Berlin Heidelberg, pp. 829–837.

[7] Baden, A., Crane, K., and Kazhdan, M. Möbius registration. *Computer Graphics Forum 37*, 5 (2018), 211–220.

[8] Baiker, M., Milles, J., Vossepoel, A. M., Que, I., Kaijzel, E. L., Lowik, C. W. G. M., Reiber, J. H. C., Dijkstra, J., and Lelieveldt, B. P. F. Fully automated whole-body registration in mice using an articulated skeleton atlas. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2007), pp. 728–731.

[9] Bari, M., Stoltzfus, L., Lin, P.-H., Liao, C., Emani, M., and Chapman, B. Is data placement optimization still relevant on newer GPUs? In *2018 IEEE/ACM Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS)* (11 2018), pp. 83–96.

[10] Beg, M. F., and Khan, A. Computing an average anatomical atlas using LDDMM and geodesic shooting. In *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.* (2006), pp. 1116–1119.

[11] Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision 61*, 2 (2005), 139–157.

[12] Ben Amar, M., and Bordner, A. Mimicking cortex convolutions through the wrinkling of growing soft bilayers. *Journal of Elasticity 129* (2017), 213–238.

[13] Benseghir, T., Malandain, G., and Vaillant, R. Iterative closest curve: A framework for curvilinear structure registration application to 2D/3D coronary arteries registration. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013* (Berlin, Heidelberg, 2013), K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds., Springer Berlin Heidelberg, pp. 179–186.

[14] Bertsch, M., Franchi, B., Meschini, V., Tesi, M. C., and Tosin, A. A sensitivity analysis of a mathematical model for the synergistic interplay of amyloid beta and tau on the dynamics of Alzheimer's disease. *Brain Multiphysics 2* (2021), 100020.

[15] Biegler, L. T., Ghattas, O., Heinkenschloss, M., and van Bloemen Waanders, B., Eds. *Large-Scale PDE-Constrained Optimization*, vol. 30 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag Berlin Heidelberg, 2003.

[16] Bochev, P., and Lehoucq, R. Energy principles and finite element methods for pure traction linear elasticity. *Computational Methods in Applied Mathematics 5*, 10 (2011), 1–15.

[17] Bolz, J., Farmer, I., Grinspun, E., and Schröder, P. Sparse matrix solvers on the GPU: Conjugate gradients and multigrid. In *ACM SIGGRAPH 2003 Papers* (2003), Association for Computing Machinery, p. 917–924.

[18] Bookstein, F. L. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence 11*, 6 (1989), 567–585.

[19] Breen, D. E., and Whitaker, R. T. A level-set approach for the metamorphosis of solid models. *IEEE Transactions on Visualization and Computer Graphics 7*, 2 (2001), 173–192.

[20] Bressan, A., and Lewicka, M. A model of controlled growth. *Archive for Rational Mechanics and Analysis 227*, 1223–1266 (2018).

[21] Camara, O., Schweiger, M., Scahill, R. I., Crum, W. R., Sneller, B. I., Schnabel, J. A., Ridgway, G. R., Cash, D. M., Hill, D. L. G., and Fox, N. C. Phenomenological model of diffuse global and regional atrophy using finite-element methods. *IEEE Transactions on Medical Imaging 25*, 11 (2006), 1417–1430.

[22] Carbonell, F., Iturria-Medina, Y., and Evans, A. C. Mathematical modeling of protein misfolding mechanisms in neurological diseases: A historical overview. *Frontiers in Neurology 9*, 37 (2018).

[23] Carmen Calzada, M., Camacho, G., Fernández-Cara, E., and Marín, M. Fictitious domains and level sets for moving boundary problems. applications to the

numerical simulation of tumor growth. *Journal of Computational Physics 230*, 4 (2011), 1335–1358.

[24] Cech, P., Andronache, A., Wang, L., Székely, G., and Cattin, P. Piecewise rigid multimodal spine registration. In *Bildverarbeitung für die Medizin* (01 2006), pp. 211–215.

[25] Cecka, C., Lew, A. J., and Darve, E. Assembly of finite element methods on graphics processors. *International Journal for Numerical Methods in Engineering 85*, 5 (2011), 640–669.

[26] Charlier, B., Charon, N., and Trouvé, A. The fshape framework for the variability analysis of functional shapes. *Foundations of Computational Mathematics 17*, 2 (2017), 287–357.

[27] Charon, N., and Trouvé, A. The varifold representation of nonoriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences 6*, 4 (2013), 2547–2580.

[28] Christensen, G. E., Rabbitt, R. D., and Miller, M. I. Deformable templates using large deformation kinematics. *IEEE transactions on image processing 5*, 10 (1996), 1435–1447.

[29] Cody, W. J. Algorithm 715: SPECFUN—a portable FORTRAN package of special function routines and test drivers. *ACM Transactions on Mathematical Software 19*, 1 (1993), 22—32.

[30] Cohen-Or, D., Solomovic, A., and Levin, D. Three-dimensional distance field metamorphosis. *ACM transactions on graphics 17*, 2 (1998), 116–141.

[31] Cristini, V., Lowengrub, J., and Nie, Q. Nonlinear simulation of tumor growth. *Journal of Mathematical Biology 46* (2003), 191–224.

[32] Csapo, I., Davis, B., Shi, Y., Sanchez, M., Styner, M., and Niethammer, M. Longitudinal image registration with temporally-dependent image similarity measure. *IEEE Transactions on Medical Imaging 32*, 10 (2013), 1939–1951.

[33] Danuser, G., Allard, J., and Mogilner, A. Mathematical modeling of eukaryotic cell migration: Insights beyond experiments. *Annual Review of Cell and Developmental Biology 29*, 1 (2013), 501–528.

[34] Diestel, J., and J. J. Uhl, J. *Vector Measures*, vol. 15 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1977.

[35] Dompierre, J., Labbé, P., Vallet, M., and Camarero, R. How to subdivide pyramids, prisms, and hexahedra into tetrahedra. In *Proceedings of the 8th International Meshing Roundtable* (1999), pp. 195–204.

[36] Dou, Y., Fan, Y., Zhao, J., and Gregersen, H. Longitudinal residual strain and stress-strain relationship in rat small intestine. *Biomedical engineering online 5*, 37 (2006).

[37] du Bois d'Aische, A., De Craene, M., Geets, X., Grégoire, V., Macq, B., and Warfield, S. K. Estimation of the deformations induced by articulated bodies: Registration of the spinal column. *Biomedical Signal Processing and Control 2*, 1 (2007), 16–24.

[38] Dziekonski, A., Sypek, P., Lamecki, A., and Mrozowski, M. Finite element matrix generation on a GPU. *Progress In Electromagnetics Research 128* (2012), 249–265.

[39] Eisenberger, M., Lähner, Z., and Cremers, D. Divergence-free shape correspondence by deformation. *Computer Graphics Forum 38*, 5 (2019), 1–12.

[40] Escher, J. Classical solutions to a moving boundary problem for an elliptic-parabolic system. *Interfaces and Free Boundaries 6* (2004), 175–193.

[41] Evans, L. C. *Partial Differential Equations*, 2 ed., vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 2010.

[42] Farhat, C., and Crivelli, L. A general approach to nonlinear FE computations on shared-memory multiprocessors. *Computer Methods in Applied Mechanics and Engineering 72*, 2 (1989), 153–171.

[43] Fornefett, M., Rohr, K., and Stiehl, H. S. Elastic registration of medical images using radial basis functions with compact support. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)* (1999), vol. 1, pp. 402–407.

[44] Fung, Y. C. What are the residual stresses doing in our blood vessels? *Annals of Biomedical Engineering 19* (1991), 237–249.

[45] Grenander, U., Srivastava, A., and Saini, S. A pattern-theoretic characterization of biological growth. *IEEE Transactions on Medical Imaging 26*, 5 (2007), 648–659.

[46] Gris, B. Incorporation of a deformation prior in image reconstruction. *Journal of Mathematical Imaging and Vision 61*, 5 (2019), 691–709.

[47] Gris, B., Durrleman, S., and Trouvé, A. A sub-Riemannian modular framework for diffeomorphism-based analysis of shape ensembles. *SIAM Journal on Imaging Sciences 11*, 1 (2018), 802–833.

[48] Gu, X., Wang, Y., Chan, T. F., Thompson, P. M., and Yau, S.-T. Genus zero surface conformal mapping and its application to brain surface mapping. In *Information Processing in Medical Imaging* (Berlin, Heidelberg, 2003), C. Taylor and J. A. Noble, Eds., Springer Berlin Heidelberg, pp. 172–184.

[49] Gurtin, M. E. *An Introduction to Continuum Mechanics*, vol. 158 of *Mathematics in Science and Engineering*. Academic Press, 1981.

[50] Haber, E., and Modersitzki, J. Numerical methods for volume preserving image registration. *Inverse problems 20* (2004), 1621–1638.

[51] Hogea, C., Davatzikos, C., and Biros, G. Brain–tumor interaction biophysical models for medical image registration. *SIAM Journal on Scientific Computing 30*, 6 (2008), 3050–3072.

[52] Hsieh, D.-N., and Younes, L. Piecewise rigid motion in diffeomorphism groups with strong right-invariant metrics. In *Mathematics of Shapes and Applications*. World Scientific, 2019, pp. 97–114.

[53] Huang, Q.-X., Adams, B., Wicke, M., and Guibas, L. J. Non-rigid registration under isometric deformations. In *Proceedings of the Symposium on Geometry Processing* (2008), SGP '08, Eurographics Association, pp. 1449—1457.

[54] Huesman, R. H., Klein, G. J., Kimdon, J. A., Kuo, C., and Majumdar, S. Deformable registration of multi-modal data including rigid structures. In *2002 IEEE Nuclear Science Symposium Conference Record* (2002), vol. 3, pp. 1879–1882.

[55] Johnson, W. P. The curious history of Faà di Bruno's formula. *The American Mathematical Monthly 109*, 3 (2002), 217–234.

[56] Joshi, S. C., and Miller, M. I. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing 9*, 8 (2000), 1357–1370.

[57] Khanal, B., Lorenzi, M., Ayache, N., and Pennec, X. A biophysical model of brain deformation to simulate and analyze longitudinal MRIs of patients with Alzheimer's disease. *NeuroImage 134* (2016), 35–52.

[58] Khmelinskii, A., Baiker, M., Kaijzel, E. L., Chen, J., Reiber, J. H. C., and Lelieveldt, B. P. F. Articulated whole-body atlases for small animal image analysis: Construction and applications. *Molecular imaging and biology 13* (2011), 898–910.

[59] Kiran, U., Sharma, D., and Gautam, S. S. GPU-warp based finite element matrices generation and assembly using coloring method. *Journal of Computational Design and Engineering 6*, 4 (2019), 705–718.

[60] Kulason, S., Miller, M. I., Trouvé, A., and Alzheimer's Disease Neuroimaging Initiative. Reaction-diffusion model of cortical atrophy spread during early stages of Alzheimer's disease. *bioRxiv:2020.11.02.362855* (2020).

[61] Laga, H., Guo, Y., Tabia, H., Fisher, R. B., and Bennamoun, M. *3D Shape Analysis: Fundamentals, Theory, and Applications*. Wiley, 2019.

[62] Lions, J. L., and Magenes, E. *Non-Homogeneous Boundary Value Problems and Applications Vol. 1*, vol. 181 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 1972.

[63] Little, J., Hill, D., and Hawkes, D. Deformations incorporating rigid structures. *Computer Vision and Image Understanding 66*, 2 (1997), 223–232.

[64] Loeckx, D., Maes, F., Vandermeulen, D., and Suetens, P. Nonrigid image registration using free-form deformations with a local rigidity constraint. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004* (2004),

C. Barillot, D. R. Haynor, and P. Hellier, Eds., Springer Berlin Heidelberg, pp. 639–646.

[65] Ma, J., Miller, M. I., and Younes, L. A Bayesian generative model for surface template estimation. *International journal of biomedical imaging 2010* (2010).

[66] Mang, A., Gholami, A., Davatzikos, C., and Biros, G. PDE-constrained optimization in medical image analysis. *Optimization and Engineering 19*, 3 (2018), 765–812.

[67] Manyuhina, O. V., Mayett, D., and Schwarz, J. M. Elastic instabilities in a layered cerebral cortex: a revised axonal tension model for cortex folding. *New Journal of Physics 16*, 123058 (2014).

[68] Markall, G. R., Slemmer, A., Ham, D. A., Kelly, P. H. J., Cantwell, C. D., and Sherwin, S. J. Finite element assembly strategies on multi-core and many-core architectures. *International Journal for Numerical Methods in Fluids 71*, 1 (2013), 80–97.

[69] Martin-Fernandez, M. A., Munyoz-Moreno, E., Martin-Fernandez, M., and Alberola-Lopez, C. Articulated registration: elastic registration based on a wire-model. In *Medical Imaging 2005: Image Processing* (2005), J. M. Fitzpatrick and J. M. Reinhardt, Eds., vol. 5747, pp. 182–191.

[70] Miller, M. I., Younes, L., Ratnanather, J. T., Brown, T., Trinh, H., Lee, D. S., Tward, D., Mahon, P. B., Mori, S., Albert, M., et al. Amygdalar atrophy in symptomatic Alzheimer's disease based on diffeomorphometry: the BIOCARD cohort. *Neurobiology of Aging 36* (2015), S3–S10.

[71] Natarajan, R. Finite element applications on a shared-memory multiprocessor: algorithms and experimental results. *Journal of Computational Physics 94*, 2 (1991), 352–381.

[72] Nelson, D. Experimental methods for determining residual stresses and strains in various biological structures. *Experimental Mechanics 54* (2014), 695–708.

[73] NVIDIA Corporation. CUDA C++ best practices guide. `https://docs.nvidia.com/cuda/cuda-c-best-practices-guide/index.html`, 2020.

[74] NVIDIA Corporation. CUDA C++ programming guide. `https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html`, 2020.

[75] Oldham, K., Myland, J., and Spanier, J. *An Atlas of Functions*, 2 ed. Springer, 2008.

[76] Omens, J. H., and Fung, Y.-C. Residual strain in rat left ventricle. *Circulation Research 66*, 1 (1990), 37–45.

[77] Papademetris, X., Dione, D. P., Dobrucki, L. W., Staib, L. H., and Sinusas, A. J. Articulated rigid registration for serial lower-limb mouse imaging. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005* (2005), J. S. Duncan and G. Gerig, Eds., Springer Berlin Heidelberg, pp. 919–926.

[78] Paulsen, V. I., and Raghupathi, M. *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, vol. 152 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2016.

[79] Persson, P.-O., and Strang, G. A simple mesh generator in MATLAB. *SIAM Review 46*, 2 (2004), 329–345.

[80] Pitiot, A., Bardinet, E., Thompson, P. M., and Malandain, G. Piecewise affine registration of biological images for volume reconstruction. *Medical Image Analysis 10*, 3 (2006), 465–483. Special Issue on The Second International Workshop on Biomedical Image Registration (WBIR'03).

[81] Qiu, A., Albert, M., Younes, L., and Miller, M. I. Time sequence diffeomorphic metric mapping and parallel transport track time-dependent shape changes. *NeuroImage 45*, 1, Supplement 1 (2009), S51–S60. Mathematics in Brain Imaging.

[82] Ratnanather, J. T., Arguillère, S., Kutten, K. S., Hubka, P., Kral, A., and Younes, L. 3D normal coordinate systems for cortical areas. In *Mathematics of Shapes and Applications*. World Scientific, 2019, pp. 167–179.

[83] Richman, D. P., Stewart, R. M., Hutchinson, J. W., and Caviness, V. S. Mechanical model of brain convolutional development. *Science 189*, 4196 (1975), 18–21.

[84] Rivest-Hénault, D., Sundar, H., and Cheriet, M. Nonrigid 2D/3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. *IEEE Transactions on Medical Imaging 31*, 8 (2012), 1557–72.

[85] Rohlfing, T., Maurer, C. R., Bluemke, D. A., and Jacobs, M. A. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE transactions on medical imaging 22*, 6 (2003), 730–741.

[86] Ruan, D., Fessler, J. A., Roberson, M., Balter, J., and Kessler, M. Nonrigid registration using regularization that accommodates local tissue rigidity. In *Medical Imaging 2006: Image Processing* (2006), J. M. Reinhardt and J. P. W. Pluim, Eds., vol. 6144, SPIE, pp. 346–354.

[87] Russon, A. E., and Blair, J. M. Rational function minimax approximations for the modified Bessel functions $K_0(x)$ and $K_1(x)$. *Report AECL-3461, Atomic energy of Canada limited, Chalk River, Ontario* (1969).

[88] Schonsheck, S. C., Bronstein, M. M., and Lai, R. Nonisometric surface registration via conformal Laplace–Beltrami basis pursuit. *Journal of Scientific Computing 86*, 3 (2021), 1–24.

[89] Segel, L. A., and Slemrod, M. The quasi-steady-state assumption: A case study in perturbation. *SIAM Review 31*, 3 (1989), 446–477.

[90] Shuttleworth, R., and Trucu, D. Two-scale moving boundary dynamics of cancer invasion: Heterotypic cell populations' evolution in heterogeneous ECM. In *Cell Movement*. Birkhäuser, Cham, 2018, pp. 1–26.

[91] Srivastava, A., Samir, C., Joshi, S. H., and Daoudi, M. Elastic shape models for face analysis using curvilinear coordinates. *Journal of Mathematical Imaging and Vision 33* (2009), 253–265.

[92] Staring, M., Klein, S., and Pluim, J. P. W. A rigidity penalty term for nonrigid registration. *Medical Physics 34*, 11 (2007), 4098–4108.

[93] Subbaroyan, J., Martin, D. C., and Kipke, D. R. A finite-element model of the mechanical effects of implantable microelectrodes in the cerebral cortex. *Journal of Neural Engineering 2* (2005), 103–113.

[94] Sundar, H., Davatzikos, C., and Biros, G. Biomechanically-constrained 4D estimation of myocardial motion. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2009), Springer, pp. 257–265.

[95] Sỳkora, D., Dingliana, J., and Collins, S. As-rigid-as-possible image registration for hand-drawn cartoon animations. In *Proceedings of the 7th International Symposium on Non-photorealistic Animation and Rendering* (2009), pp. 25–33.

[96] Tallinen, T., Chung, J. Y., Rousseau, F., Girard, N., Lefèvre, J., and Mahadevan, L. On the growth and form of cortical convolutions. *Nature Physics 12* (2016), 588–593.

[97] Toga, A. W., and Thompson, P. M. The role of image registration in brain mapping. *Image and Vision Computing 19*, 1–2 (2001), 3–24.

[98] Van Kaick, O., Zhang, H., Hamarneh, G., and Cohen-Or, D. A survey on shape correspondence. *Computer graphics forum 30*, 6 (2011), 1681–1707.

[99] Werner, R., Ehrhardt, J., Schmidt-Richberg, A., and Handels, H. Validation and comparison of a biophysical modeling approach and non-linear registration for estimation of lung motion fields in thoracic 4D CT data. In *Medical Imaging 2009: Image Processing* (2009), J. P. W. Pluim and B. M. Dawant, Eds., vol. 7259, International Society for Optics and Photonics, SPIE, pp. 293–300.

[100] Wolgemuth, C. W., and Zajac, M. The moving boundary node method: A level set-based, finite volume algorithm with applications to cell motility. *Journal of Computational Physics 229*, 19 (2010), 7287–7308.

[101] Xu, G., Bayly, P. V., and Taber, L. A. Residual stress in the adult mouse brain. *Biomech Model Mechanobiol 8*, 4 (2009), 253–262.

[102] Yip, S., Perk, T., and Jeraj, R. Development and evaluation of an articulated registration algorithm for human skeleton registration. *Physics in Medicine and Biology 59*, 6 (2014), 1485–1499.

[103] Yosida, K. *Functional Analysis*, 6 ed. Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 1995.

[104] Younes, L. *Shapes and Diffeomorphisms*, 2 ed., vol. 171 of *Applied Mathematical Sciences*. Springer-Verlag Berlin Heidelberg, 2019.

[105] Younes, L., Kutten, K. S., and Ratnanather, J. T. Normal and equivolumetric coordinate systems for cortical areas. *arXiv:1911.07999* (2019).

[106] Zhang, J., and Shen, D. GPU-based implementation of finite element method for elasticity using CUDA. In *2013 IEEE 10th International Conference on High Performance Computing and Communications and 2013 IEEE International Conference on Embedded and Ubiquitous Computing* (2013), pp. 1003–1008.

[107] Zhang, L., Gerstenberger, A., Wang, X., and Liu, W. K. Immersed finite element method. *Computer Methods in Applied Mechanics and Engineering 193*, 21–22 (2004), 2051–2067.