Variational Convergence
and
Discrete Minimal Surfaces

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen
Doktorgrades
Doctor rerum naturalium
der Georg-August-Universität Göttingen

im Promotionsstudiengang *Mathematical Sciences*
der Georg-August University School of Science (GAUSS)

vorgelegt von
Henrik Schumacher
aus Ostercappeln

Göttingen 2015

Betreuungsausschuss

Prof. Dr. Max Wardetzky	Institut für Numerische und Angewandte Mathematik
Prof. Dr. Russell Luke	Institut für Numerische und Angewandte Mathematik

Mitglieder der Prüfungskommission

Referent:	Prof. Dr. Max Wardetzky	Institut für Numerische und Angewandte Mathematik
Korreferent:	Prof. Dr. Russell Luke	Institut für Numerische und Angewandte Mathematik
Externer Gutachter:	Prof. Dr. Samir Adly	Université de Limoges

Weitere Mitglieder der Prüfungskommission

Prof. Dr. Dorothea Bahns	Mathematisches Institut
Prof. Dr. Thorsten Hohage	Institut für Numerische und Angewandte Mathematik
Prof. Dr. Viktor Pidstrygach	Mathematisches Institut
Prof. Dr. Karl-Henning Rehren	Institut für Theoretische Physik

Tag der mündlichen Prüfung: 9. 12. 2014

Danksagung

Vielen lieben Menschen, die mir stets mehr zugetraut haben als ich mir selber, bin ich zu tiefem Dank verpflichtet.

Zuvorderst natürlich meinem Betreuer Max Wardetzky: Lieber Max, vier Jahre lang hast Du mich auf meinem Weg zur Promotion angeleitet, mir aber auch in großem Umfang die Freiheit gelassen, meinen eigenen Ideen nachzugehen. Mit einem riesigen Vertrauensvorschuss hast Du mir ein hervorragendes Arbeitsumfeld geboten, mir immer wieder Denkanstöße gegeben und mich mit interessanten Leuten zusammen gebracht. Schließlich hast Du mich mit sanftem Druck zur Fertigstellung und Abgabe der Arbeit gebracht. Danke, Max! Das werde ich Dir nie vergessen.

Russell Luke, meinem zweiten Betreuer, danke ich ganz herzlich für die Betreuung und die Einblicke in Variationelle Konvergenz und Komplexe Optimierung, diesen mir so fremden Sachgebieten.

Ein ganz besonders großer Dank geht an Andy Furnas, der die Freundlichkeit besessen hat, dieser Arbeit mit Engelsgeduld eine gewisse sprachliche Qualität zu verleihen. (Alle noch enthaltenen Fehler gehen natürlich auf meinen Deckel.) Andy, Du bist ein sehr, sehr toller Kollege! Dank gilt auch meinem langjährigen Freund Carsten Thiel, der ebenfalls einen Blick auf ein Manuskript dieser Arbeit geworfen hat.

Einiges in dieser Dissertation geht auf Überlegungen zurück, die in der Zusammenarbeit mit Sebastian Scholtes über die diskrete Bernoulli-Energie entstanden. Sebastian war auch derjenige, der mich zuerst auf Γ-Konvergenz aufmerksam gemacht und mich damit auf die richtige Spur gebracht hat. Vielen Dank dafür!

Gert Lube danke ich für das in mich gesetzte Vertrauen und für die Unterweisung in der Nichtlinearen Funktionalanalysis. Heiko von der Mosel möchte ich für die Diskussion über die Regularität von Minimalflächen und seine Gastfreundschaft danken. Thomas Schick danke ich für einige hilfreiche Tipps aus der Knotentheorie. Laurant Bartholdi gebührt Dank, weil er so freundlich war, mir die Knotenmodelle aus der Mathematischen Sammlung und deren digitale Rohdaten für Experimente zur Verfügung zu stellen.

Für die vielen Diskussionen und Anregungen danke ich meinen lieben Freunden und Kollegen Daniel Arndt, Ulrich Bauer, Friedrich Bös, Keenan Crane, Rasmus Fuhse, Knut Heidemann, Julia Plehnert, Alessio Quaglino, Lars Röhe und Clarisse Weischedel. Ohne Euch wäre die Zeit nicht halb so schön gewesen!

Meinen Eltern danke ich für das immerwährend in mich gesetzte Vertrauen, ihre Geduld und ihre Unterstützung.

Ich wäre heute nicht da, wo ich jetzt bin, ohne den Halt und Beistand meiner geliebten Katharina. Ich bin so froh, dass es Dich gibt!
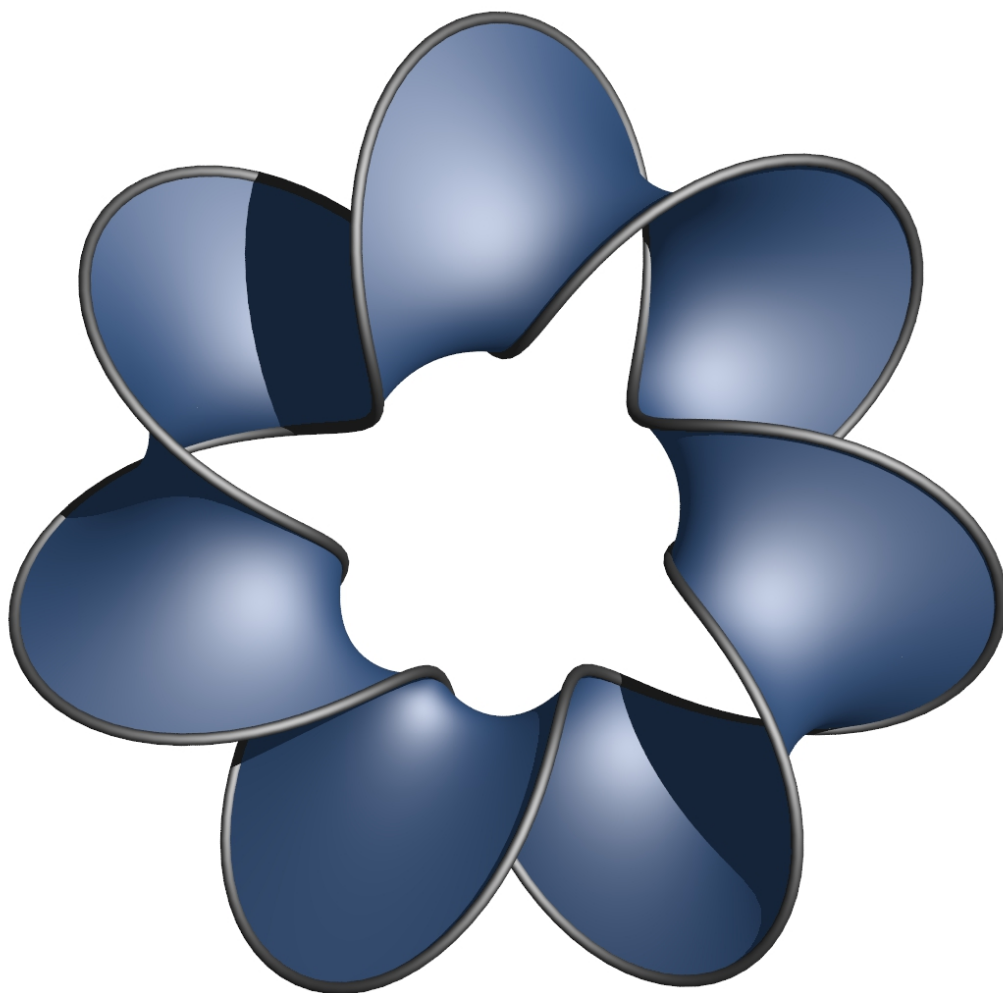
Figure 0.1: Minimal surface spanning a septfoil knot, demonstrating spontaneous
symmetry breaking and nonuniqueness of the solutions to the least area
problem: While the boundary curve has sevenfold rotation symmetry,
the presented surface lacks this symmetry; rotation by integer multiples
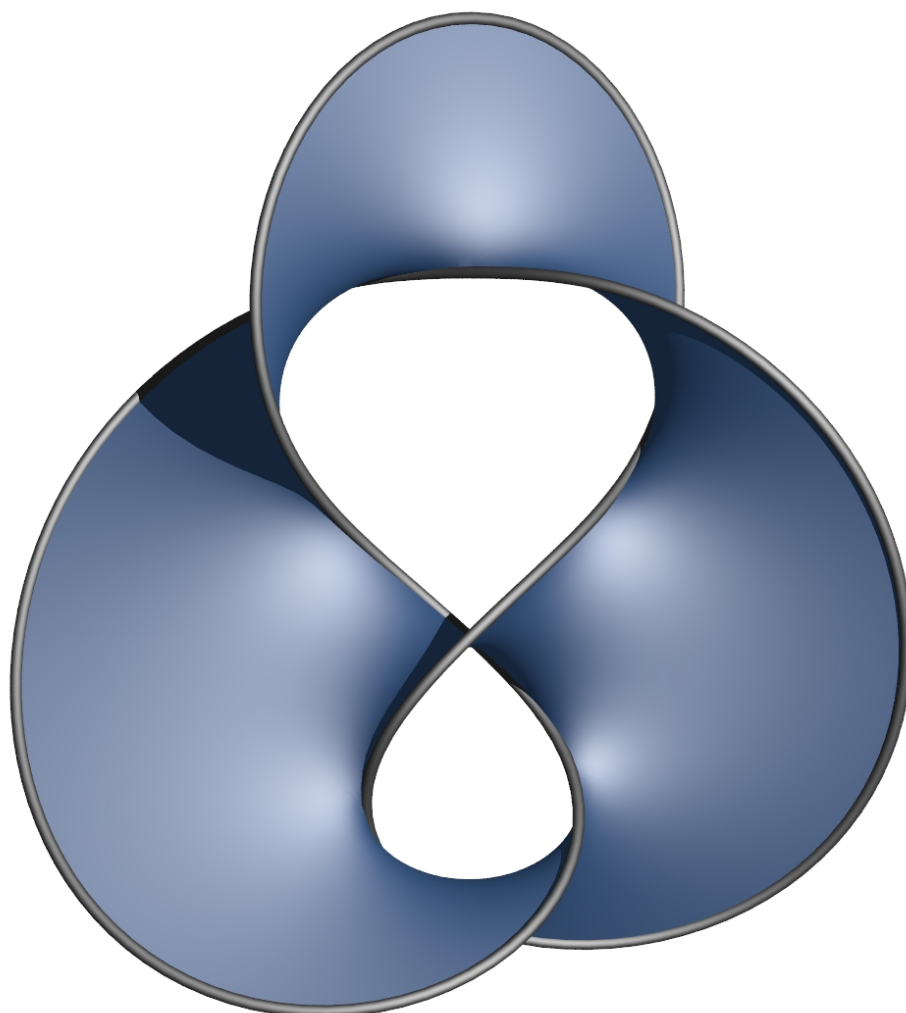of $\frac{2\pi}{7}$ produces further minimal surfaces.

Figure 0.2: Minimal surface spanning a figure-eight knot. (A 3D-print of the bound-
ary curve is part of the *Göttingen Collection of Mathematical Models
and Instruments*.)

Introduction

This thesis is concerned with the convergence behavior of the solutions to parametric variational problems. An emphasis is put on sequences of variational problems that arise as discretizations of either infinite-dimensional optimization problems or infinite-dimensional operator problems.

Apart from Chapter 1, which is of purely preparatory nature, the present work consists of essentially three parts, each of interest on its own but still depending on each other:

The first part consists of Chapters 2 and 3. In particular, Chapter 2 is the theoretical backbone of the whole work. We discuss parametric minimization problems in their most general form. In particular, the chapter can be read as a blueprint for conducting convergence analysis in practice (and we will use it this way in the third part of this work). We introduce a language that will sound familiar to workers in the area of variational analysis as well as to those in the finite elements community. This language is based on Strang's second lemma and provides sufficient conditions for variational convergence of sequences of minimization problems. These conditions are divided into three groups: *consistency*, *proximity*, and *stability*. Consistency encodes the ability to approximate the values of the functionals being minimized; proximity encodes the ability to approximate the solutions; and stability encodes the growth rates of the objective functions away from their minimizers. Consistency and proximity are closely related to epigraphical convergence. Often, they can be verified by standard methods from approximation theory. However, stability is more difficult to investigate. We introduce two notions of stability: the first, which we term *topological stability*, implies Kuratowski convergence (and in certain cases also Hausdorff convergence) of minimizers and is tightly related to lower semi-continuity; the second, *quantitative stability*, allows for convergence rates to be easily deduced, but is not readily verifiable in concrete applications.

In Chapter 3 we investigate the consequences of Chapter 2 on parametric operator problems. In particular, we obtain nonlinear generalizations of the main convergence theorems in the theory of generalized Ritz-Galerkin schemes, namely Cea's lemma and the two lemmata of Strang (Section 3.3).

We point out that this theory is capable of treating nonlinear variational problems with nonunique solutions; it encompasses the existing convergence theory for convex minimization problems and monotone operator problems, but it is not limited to these cases; and it was particularly designed to be easily applicable to *nonconforming* discretization schemes.

Although we focus on convergence of discretization schemes, the results of Chapters 2 and 3 may also be applicable to other limiting problems such as thin shell and rod limits in elasticity, continuum limits of particle systems, homogenization problems or for analyzing regularization methods.

One of the hardest tasks in dealing with geometric variational problems is to find a definition of a configuration space together with a technically feasible and geometrically meaningful notion of distance: e.g., even in the classical, seemingly simple, least area problem of surfaces, it is not obvious what a *surface* should be. The second part of this thesis is devoted to this task. We develop a suitable configuration space for the least area problem. It turns out that this configuration space is also of use for problems in nonlinear elasticity. A central building block for this configuration space and its metrics is a certain Riemannian metric on the space of inner products on a finite-dimensional real vector space. This will be introduced in Chapter 4 where we also investigate its invariance properties. In Chapter 5 we introduce a space of Lipschitz immersions from an abstract, compact smooth manifold into $\mathbb{R}^m$, equipped with metrics that are invariant under the action of the diffeomorphism group. Being invariant, these metrics descend to metrics on *shape space*, the quotient of the space of Lipschitz immersions by the group of diffeomorphisms.

In the third and final part we apply the theories of parametric optimization and of Lipschitz immersions to two concrete examples: the approximation of minimal surfaces (Chapter 7); and the approximation of static solutions of a full-dimensional, nonlinear elasticity model (Chapter 8), both by finite element discretizations. The configuration spaces of both problems consist of immersions of an abstract smooth manifold into Euclidean space and we discretize these spaces by considering immersed simplicial complexes, i.e., triangle meshes and tetrahedral meshes, as discrete configurations (Chapter 6). To some extent, the discretization can be described as continuous, piecewise linear finite elements. As mentioned above, Chapter 2 will serve us as a guideline in the proof of variational convergence of these discretizations: Under certain assumptions on existence and regularity of solutions, we show consistency, proximity and topological stability of the discretization schemes in terms of the metrics introduced in Chapter 5. The theory developed in Chapter 2 then implies Kuratowski convergence of discrete (almost-)minimizers to smooth minimizers.

The approximation theory needed for showing consistency and proximity will be developed in Chapter 6.

The discretization and the detailed convergence analysis for minimal surfaces will be treated in Chapter 7. While the used discretization has been successfully applied several times in numerical computations[1] there has been no notable convergence analysis for it so far.

We also discuss another discretization scheme for which a partial convergence analysis was conducted in [20] (Section 7.3).

In Section 7.7 we briefly discuss two area decreasing flows on the configuration space of immersion. Suitable discretizations of these flows can be used to numerically approximate minimizers of the discrete area functional.

Chapter 8 is devoted to a nonlinear elastic energy which involves the Hencky strain tensor. To our surprise, the Hencky strain tensor arose naturally as a certain gradient vector field on the space of inner products on a finite-dimensional vector space (Section 4.4). Since we belief that this perspective is new, we spend some effort to outline the relationship between the Riemannian geometry of the space of inner products on the one hand; and the role of the Hencky strain tensor in nonlinear elasticity on the other hand (Sections 8.1 and 8.2). Afterwards, we discretize the smooth elasticity model by using tetrahedral meshes immersed into $\mathbb{R}^3$ and apply it to the problem of finding static solutions of a material under the influence of a potential and under partial Dirichlet boundary conditions. Assuming existence and $W^{2,\infty}$-regularity of solutions, we show consistency, proximity, and topological stability with respect to the $W^{1,\infty}$-norm, thus Kuratowski convergence of the discrete solutions to the solutions of the smooth problem.

---

[1]The earliest account we could find is by Wagner [37] and goes back to 1977.

Contents

1 Preliminaries

In this text, the symbol C usually refers to a "generic constant" that may be increased during the course of a proof. Sometimes, we mark the dependence on certain quantities by attaching them in parentheses. For example, $C(s)$ depends on the entity s. In this context, s is usually assumed to be fixed as well.

1.1 Functional Analysis

If not otherwise stated, vector spaces are assumed to be over the reals. For normed spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_X)$, we denote the space of continuous linear maps from X to Y by $L(X; Y)$. For $A \in L(X; Y)$, we denote by $\|A\|$ the operator norm. The continuous *dual space* of X is denoted by X' and if not otherwise stated, $\langle \cdot, \cdot \rangle$ refers to the bilinear pairing

$$\langle \cdot, \cdot \rangle \colon X' \times X \to \mathbb{R}, \qquad \langle \xi, x \rangle = \xi(x).$$

The dual map of A will be written as $A' \in L(Y'; X')$. Differentiability is always meant in the sense of Fréchet differentiability.

1.2 Differential Geometry

We will discuss smooth manifolds, smooth manifolds with boundary, and smooth manifolds with corners. It belongs to the curiosities of mathematical language that every smooth manifold is also a manifold with boundary; and that every manifold with boundary is also a manifold with corners. In particular, even a smooth manifold M without boundary (i.e., with $\partial M = \emptyset$) is a smooth manifold with boundary—and even a smooth manifold with corners.

Let M, M_0 be smooth manifolds with corners and $f \colon M \to M_0$ a differentiable map. We write TM for the *tangent bundle* of M and $T'M$ for the *cotangent bundle* (the *continuous* linear forms on TM). The *tangent map* or *(total) differential* of f is denoted by $Tf \colon TM \to TM_0$. In the case that $M_0 = Y$ is a Banach space, the tangent bundle $TM_0 \cong Y \times Y$ is trivial and we may write $\mathrm{d}f$ for the fiber component of the

tangent map Tf. This leads to the identity $T_x f = (f(x), \mathrm{d}_x f)$ for all $x \in M$. More generally, we use d for the *exterior derivative* of differential forms. If $M \subset X$ is an open subset of a Banach space X, we may write $T_x f\, u = (f(x), Df(x)\,u)$, with $x \in U$, $u \in T_x U \cong X$, and $Df\colon U \to L(X; Y)$ the total derivative (the Jacobi "matrix").

In general, we denote the *pullback* of tensors along $f\colon M \to M_0$ by $f^\#$ instead of the often used notation f^*. We will often write $|_x$ (read as "at x") for point evaluation at x, e.g., $\mathrm{d}f|_x = \mathrm{d}_x f$, in particular, if other notation would be confusing.

Let $\pi\colon E \to M_0$ be a (locally trivial) fiber bundle. We denote the space of *sections* by $\Gamma(M_0; E) := \{\, \varphi \in C^0(M_0; E) \mid \pi \circ \varphi = \mathrm{id}_{M_0} \,\}$. For a continuous map $f\colon M \to M_0$, we denote by

$$f^\# E := \{\, (x, e) \in M \times E \mid f(x) = \pi(e) \,\}$$

the *pullback of the fiber bundle E along f*.

Let g and g_0 be Riemannian metrics on M and M_0, respectively. The induced inner products for tensors on M are denoted by $\langle \cdot, \cdot \rangle_g$, the induced inner product norm by $|\cdot|_g$; operator norms of tensors are simply denoted by $\|\cdot\|$. For example, if M and M_0 are finite dimensional and $A \in L(T_x M; T_y M_0)$ is an endomorphism of tangent spaces, one may identify $L(TM; TM_0) \cong TM_0 \otimes T'M$ and consider both the operator norm and the Frobenius norm

$$\|A\| = \sup_{\substack{u \in T_x M \\ u \neq 0}} \frac{|A\,u|_{g_0}}{|u|_g} \quad \text{and} \quad |A|_{g,g_0} = (\mathrm{tr}(A^{*_{g,g_0}} A))^{\frac{1}{2}},$$

where $A^{*_{g,g_0}}$ denotes the adjoint of A with respect to the metrics $g|_x$ and $g_0|_y$.

In this work, the symbol ∇ is used for *covariant derivatives* exclusively, not for gradients. Depending on context, the *Hessian* $\mathrm{Hess}(f)$ may refer to $\nabla\nabla f$ (if $f \in \Gamma(M; E)$ is the section of a vector bundle), $\nabla \mathrm{d}f$ (if $f\colon M \to X$ is a vector-valued mapping) or ∇Tf (if $f\colon (M, g) \to (M_0, g_0)$ is a mapping between Riemannian manifolds). In any case it is the second derivative with respect to the covariant derivative ∇. We may write ∇^g, Hess^g or ∇^{g,g_0}, Hess^{g,g_0} if they are meant with respect to the Levi-Civita connections of the Riemannian metrics g and g_0. For example, let $f\colon (M, g) \to (M_0, g_0)$ be a smooth mapping. Then Tf can be interpreted as a section $Tf \in \Gamma(M; L(TM; f^\# TM_0))$. The Levi-Civita connections of g and g_0 induce a covariant derivative ∇^{g,g_0} on the vector bundle $L(TM; f^\# TM_0) \to M$: For smooth vector fields $u, v \in \Gamma(M; f^\# TM_0)$ along f, one has

$$\mathrm{Hess}^{g,g_0}(f)(u, v) = \nabla^{g_0} Tf(u, v) = \nabla^{g_0}_u (Tf\,v) - Tf\,(\nabla^g_u v).$$

Here, ∇^{g_0} denotes the pullback connection along f.

1.3 Sobolev Spaces

Let $m \geq k$ be positive integers and let Σ be a compact, k-dimensional smooth manifold with corners. Fix a smooth Riemannian metric g on Σ and denote the Euclidean metric on $\mathbb{R}^m$ by g_0. For $l \in \mathbb{N} \cup \{0\}$ and $1 \leq p \leq \infty$, denote by $W_g^{l,p}(\Sigma; \mathbb{R}^m)$ the Sobolev space of l-times weakly differentiable mappings from Σ to $\mathbb{R}^m$ with weak derivatives up to order l in L_g^p. These vector spaces are independent of the choice of g. However, their norms $\|\cdot\|_{W_g^{l,p}}$ *do* depend on g:

$$\|u\|_{W_g^{l,p}}^p := \sum_{i=0}^{l} \int_{\Sigma} |\nabla^i u|_{g,g_0}^p \, \text{vol}_g, \qquad \text{for } u \in W_g^{l,p}(\Sigma; \mathbb{R}^m), \, 1 \leq p < \infty,$$

$$\|u\|_{W_g^{l,\infty}} := \max_{i \in \{0,1,\dots,l\}} \operatorname*{ess\,sup}_{x \in \Sigma} |\nabla^i u|_x|_{g,g_0}, \quad \text{for } u \in W_g^{l,\infty}(\Sigma; \mathbb{R}^m).$$

Here, ∇^i denotes i-fold covariant differentiation and $|\cdot|_{g,g_0}$ denotes the inner product norm induced by g, g_0 on the tensor bundles $T'\Sigma \otimes \dots \otimes T'\Sigma \otimes u^\# T\mathbb{R}^m$.

Moreover, we denote by $C_0^\infty(\Sigma; \mathbb{R}^m)$ the vector space of smooth mappings $\Sigma \to \mathbb{R}^m$ whose support is compact and contained in $\Sigma \setminus \partial\Sigma$. When $l \geq 1$, one may define

$$W_{0,g}^{l,p}(\Sigma; \mathbb{R}^m) := \overline{C_0^\infty(\Sigma; \mathbb{R}^m)}^{W_g^{l,p}}, \quad \text{for } 1 \leq p < \infty,$$

$$W_{0,g}^{l,\infty}(\Sigma; \mathbb{R}^m) := \{ u \in W_g^{l,\infty}(\Sigma; \mathbb{R}^m) \mid \text{for } i = 0, 1, \dots, l-1: \ (\nabla^i u)|_{\partial\Sigma} = 0 \},$$

where the restriction mapping is given by

$$|_{\partial\Sigma} \colon C^0(\Sigma; T'\Sigma \otimes \dots \otimes T'\Sigma \otimes u^\# T\mathbb{R}^m) \to C^0(\partial\Sigma; (T'\Sigma \otimes \dots \otimes T'\Sigma \otimes u^\# T\mathbb{R}^m)|_{\partial\Sigma}).$$

More generally, we write the trace operator as

$$|_{\partial\Sigma} \colon W_g^{l,p}(\Sigma; \mathbb{R}^m) \to W_g^{l-1,p}(\partial\Sigma; \mathbb{R}^m).$$

For later use, we note the Poincaré inequality (see [21, Section 5.6, Theorem 3] for a proof) and an import implication of it:

Lemma 1.1 (Poincaré inequality)
Let (Σ, g) be a connected, compact, smooth Riemannian manifold with non-empty boundary. For each $p \in [1, \infty]$ there is a constant $C_p \geq 0$ such that the following holds:

$$\|u\|_{L^p} \leq C_p \|du\|_{L^p} \quad \text{for all } u \in W_{0,g}^{1,p}(\Sigma; \mathbb{R}^m).$$

The *Hölder conjugate q* of $p \in [1, \infty]$ is defined by

$$q = \begin{cases} \infty, & p = 1, \\ \frac{1}{1-p^{-1}}, & p \in \,]1, \infty[, \\ 1, & p = \infty. \end{cases}$$

Lemma 1.2 *Let (Σ, g) be a connected, compact, smooth Riemannian manifold with non-empty boundary. Let $1 < p \le \infty$ and let $1 \le q < \infty$ the Hölder conjugate of p. Then the* Laplacian

$$\bar{A} \colon W^{1,p}_{0,g}(\Sigma; \mathbb{R}^m) \to (W^{1,q}_{0,g}(\Sigma; \mathbb{R}^m))', \qquad u \mapsto (v \mapsto \int_\Sigma \langle du, dv \rangle_g \, \mathrm{vol}_g)$$

is an isomorphism of Banach spaces and one has $\|\bar{A}^{-1}\| \le (1 + C_p)(1 + C_q)$.

Proof. Note that the mapping

$$\kappa \colon L^p_g(\Sigma; T'\Sigma \otimes \mathbb{R}^m) \to (L^q_g(\Sigma; T'\Sigma \otimes \mathbb{R}^m))', \quad \alpha \mapsto (\beta \mapsto \int_\Sigma \langle \alpha, \beta \rangle_g \, \mathrm{vol}_g)$$

is an isometric isomorphism for all $1 < p \le \infty$. Since $\langle \bar{A}(u), v \rangle = \langle \kappa(du), dv \rangle$, $\bar{A}$ is well-defined and continuous with $\|\bar{A}\| \le 1$. From the Poincaré inequality (Lemma 1.1), we obtain

$$\begin{aligned} \|u\|_{W^{1,p}} &\le (1 + C_p)\|du\|_{L^p} &&\text{for all } u \in W^{1,p}_{0,g}(\Sigma; \mathbb{R}^m), \\ \|v\|_{W^{1,q}} &\le (1 + C_q)\|dv\|_{L^q}, &&\text{for all } v \in W^{1,q}_{0,g}(\Sigma; \mathbb{R}^m). \end{aligned}$$

Now, we may estimate for $u \in W^{1,p}_{0,g}(\Sigma; \mathbb{R}^m)$:

$$\begin{aligned} \|\bar{A}(u)\|_{(W^{1,q}_{0,g})'} &= \sup_{\substack{v \in W^{1,q}_{0,g} \\ v \ne 0}} \frac{\langle \kappa(du), dv \rangle}{\|v\|_{W^{1,q}_g}} \ge \frac{1}{1 + C_q} \sup_{\substack{v \in W^{1,q}_{0,g} \\ v \ne 0}} \frac{\langle \kappa(du), dv \rangle}{\|dv\|_{L^q_g}} \\ &= \frac{1}{1 + C_q}\|\kappa(du)\|_{(L^q_g)'} = \frac{1}{1 + C_q}\|du\|_{L^p_g} \\ &\ge \frac{1}{(1 + C_p)(1 + C_q)}\|u\|_{W^{1,p}_{0,g}}. \end{aligned}$$

This shows that $\bar{A}$ is a continuous, open, injective and linear map between Banach spaces. Thus $\bar{A}$ is an isomorphism and the operator norm of its inverse is readily estimated by $(1 + C_p)(1 + C_q)$. $\qquad\square$

1.4 Variational Analysis

Although hardly recognizable, most of the topics in this section are covered in detail by chapters 4 and 7 of [34]. Our aim here is to give a brief summary without introducing too much non-standard notation.

For a function $f\colon X \to \,]-\infty, \infty]$, we define the *epigraph* as

$$\mathrm{epi}(f) := \{\,(x, t) \in X \times \,]-\infty, \infty] \mid t \geq f(x)\,\}.$$

We use $\inf(f) = \inf\{\,f(x) \mid x \in X\,\}$ for the *infimal value* of f and

$$\arg\min(f) := \{\,x \in X \mid \forall y \in X\colon\ f(x) \leq f(y)\,\} = \{\,x \in X \mid f(x) = \inf(f)\,\}$$

for the set of *minimizers* of f. Moreover, we define for $\delta \in [0, \infty]$ the set of *δ-minimizers* by

$$\arg\min{}^{\delta}(f) := \{\,x \in X \mid \forall y \in X\colon\ f(x) \leq f(y) + \delta\,\}.$$

In particular, one has $\arg\min^{0}(f) = \arg\min(f)$ and

$$\arg\min{}^{\delta}(f) = \begin{cases} X, & \delta = \infty, \\ \emptyset, & \inf(f) = -\infty \text{ and } \delta \in [0, \infty[, \\ \{\,x \in X \mid f(x) \leq \inf(f) + \delta\,\}, & \inf(f) > -\infty \text{ and } \delta \in [0, \infty[. \end{cases} \quad (1.1)$$

When X is a topological space, a function $f\colon X \to \,]-\infty, \infty]$ is called *lower semi-continuous*, if all lower level sets $f^{-1}(]-\infty, t])$, $t \in \,]-\infty, \infty]$ are closed or, equivalently, if the epigraph $\mathrm{epi}(f) \subset X \times \,]-\infty, \infty]$ is a closed set (with respect to the product topology).

1.4.1 Metric spaces

Let (X, d) be a metric space. For $r \geq 0$ and $x \in X$, we denote by $B(x, r) := \{\,y \in X \mid d(x, y) < r\,\}$ the *open ball* and by $\bar{B}(x, r) := \{\,y \in X \mid d(x, y) \leq r\,\}$ the *closed ball* of radius r around x.

For a set $A \subset X$, we write

$$\mathrm{dist}(x, A) := \inf\{\,d(x, a) \mid a \in A\,\}.$$

Moreover, we define the *open and the closed r-thickening* of a set $A \subset X$ by

$$B(A, r) := \{\,x \in X \mid \mathrm{dist}(x, A) < r\,\} \quad\text{and}\quad \bar{B}(A, r) := \{\,x \in X \mid \mathrm{dist}(x, A) \leq r\,\}.$$

Remark 1.3 These definitions seem to be quite simple. However, one has to keep in mind some subtleties:

1. The closed ball $\bar{B}(x, r)$ needs not to coincide with the closure $\overline{B(x, r)}$ of the open ball $B(x, r)$. As a counterexample, consider the discrete metric space $X = \{0, 1\}$ with $d(x, y) = |x - y|$ and $r = 1$: One has $\bar{B}(0, 1) = \{0, 1\}$, but $B(0, 1) = \{0\}$. Since singletons are closed in metric spaces, one has $\overline{B(0, 1)} = \{0\}$.

2. While one has $\bigcup_{a \in A} B(a, r) = B(A, r)$ and $\bigcup_{a \in A} \bar{B}(a, r) \subset \bar{B}(A, r)$, equality in the latter does not hold in general. A counterexample is given by $X = \mathbb{R}$ with $d(x, y) := |x - y|$ and $A =]0, 1]$: One has $\bar{B}(A, 1) = [-1, 2]$, but $\bigcup_{a \in A} \bar{B}(a, r) =]-1, 2]$.

3. With the convention $\inf(\emptyset) = \infty$, one has $B(\emptyset, r) = \bar{B}(\emptyset, r) = \emptyset$ for finite $r \in [0, \infty[$, but $B(\emptyset, \infty) = \emptyset$, $\bar{B}(\emptyset, \infty) = X$.

Let $f \colon (X, d_X) \to (Y, d_Y)$ be a Lipschitz continuous mapping between metric spaces. We denote by

$$\mathrm{Lip}(f) := \sup_{\substack{x_1, x_2 \in X \\ x_1 \neq x_2}} \frac{d_Y(f(x_1), f(x_2))}{d_X(x_1, x_2)}$$

the optimal Lischitz constant of f.

1.4.2 Hausdorff convergence

Definition 1.4 Let (X, d) be a metric space and $A_1, A_2 \subset X$ subsets. We define the *Hausdorff distance* between A_1 and A_2 by

$$\mathrm{dist}(A_1, A_2) := \inf \{ r > 0 \mid A_1 \subset \bar{B}(A_2, r) \text{ and } A_2 \subset \bar{B}(A_1, r) \},$$

with the convention $\inf \emptyset = \infty$.

The subtleties mentioned in Remark 1.3 could cause one to believe that one has to be very cautious with the definition of Hausdorff distance.[1] Fortunately, this is not the case. The following lemma lists frequently used definitions of Hausdorff distance and shows that they are equivalent:

Lemma 1.5 *Let (X, d) be a metric space and let A_1, $A_2 \subset X$ be some sets. The following numbers are all equal:*

---

[1] For example, one could replace the closed thickening in the definition by open thickenings.

1. $m_1 := \inf\{\, r \geq 0 \mid A_1 \subset \bar{B}(A_2, r) \text{ and } A_2 \subset \bar{B}(A_1, r) \,\}$.

2. $m_2 := \inf\{\, r \geq 0 \mid A_1 \subset \bigcup_{a \in A_2} \bar{B}(a, r) \text{ and } A_2 \subset \bigcup_{a \in A_1} \bar{B}(a, r) \,\}$.

3. $m_3 := \inf\{\, r \geq 0 \mid A_1 \subset B(A_2, r) \text{ and } A_2 \subset B(A_1, r) \,\}$.

4. $m_4 := \max\left(\sup_{a \in A_1} \inf_{b \in A_2} d(a, b), \sup_{b \in A_2} \inf_{a \in A_1} d(a, b)\right)$.

PROOF. For each $\delta > 0$ and an arbitrary set $A \subset X$, one has

$$B(A, r) \subset \bigcup_{a \in A} \bar{B}(a, r) \subset \bar{B}(A, r) \subset B(A, r + \delta),$$

thus $m_1 \leq m_2 \leq m_3 \leq m_1 + \delta$.

$$\begin{aligned}
m_4 &= \max\left(\sup_{a \in A_1} \operatorname{dist}(a, A_2), \sup_{b \in A_2} \operatorname{dist}(b, A_1)\right) \\
&= \inf\left\{\, r \geq 0 \;\middle|\; \begin{array}{l} \text{for all } a \in A_1\colon \operatorname{dist}(a, A_2) \leq r \\ \text{for all } b \in A_2\colon \operatorname{dist}(b, A_1) \leq r \end{array} \right\} \\
&= \inf\{\, r \geq 0 \mid A_1 \subset \bar{B}(A_2, r) \text{ and } A_2 \subset \bar{B}(A_1, r) \,\} = m_1. \qquad \square
\end{aligned}$$

Remark 1.6 Note that two sets $A_1, A_2 \subset X$ have Hausdorff distance 0 if and only if their closures coincide: $\bar{A}_1 = \bar{A}_2$. Moreover, the Hausdorff distance between A_1 and A_2 may be infinite. Thus, the Hausdorff distance is not a metric on the power set $\mathfrak{P}(X)$: In general it is only an extended semi-metric in the sense that it is a function $\operatorname{dist}\colon \mathfrak{P}(X) \times \mathfrak{P}(X) \to [0, \infty]$ with:

1. $\operatorname{dist}(A, A) = 0$ for each $A \in \mathfrak{P}(X)$.

2. $\operatorname{dist}(A_1, A_2) = \operatorname{dist}(A_2, A_1)$ for all $A_1, A_2 \in \mathfrak{P}(X)$.

3. $\operatorname{dist}(A_1, A_3) \leq \operatorname{dist}(A_1, A_2) + \operatorname{dist}(A_2, A_3)$ for all $A_1, A_2, A_3 \in \mathfrak{P}(X)$.

However, the restriction of dist to the set $\mathfrak{B}(X)$ of all non-empty, closed and bounded sets leads to a metric.

Definition 1.7 Let (X, d) be a metric space and $A_n, A \subset X$ subsets, $n \in \mathbb{N}$. We say, *A_n converges to A uniformly* or *A_n Hausdorff converges to A*, if for every $\varepsilon > 0$ there is an $n \in \mathbb{N}$ such that for every integer $k \geq n$:

$$A \subset \bar{B}(A_k, \varepsilon) \quad \text{and} \quad A_k \subset \bar{B}(A, \varepsilon).$$

Remark 1.8 Let $A_n \neq \emptyset$ for infinitely many $n \in \mathbb{N}$. Then Hausdorff convergence of A_n to A as $n \to \infty$ implies $A \neq \emptyset$. Moreover, Hausdorff convergence of A_n to A is equivalent to $\lim_{n\to\infty} \text{dist}(A, A_n) = 0$.

The metric space $(\mathfrak{B}(X), \text{dist})$ inherits some nice properties from (X, d):

Theorem 1.9 *Let (X, d) be a metric space.*

1. *$(\mathfrak{B}(X), \text{dist})$ is complete if and only if (X, d) is complete.*

2. *$(\mathfrak{B}(X), \text{dist})$ is compact if and only if (X, d) is compact.*

Follow the instructions in [28, p. 280] or see [23] for a proof. The second statement is sometimes called *Blaschke selection theorem* (see [3, § 18]).

1.4.3 Kuratowski convergence

Definition 1.10 Let X be a topological space and for $x \in X$, denote by $\mathfrak{U}(x)$ the set of all open neighborhoods of x. For a sequence of sets $(A_n)_{n\in\mathbb{N}}$ in X one defines the *limit inferior* or *inner limit* $\text{Li}_{n\to\infty} A_n$ and the *limit superior* or *outer limit* $\text{Ls}_{n\to\infty} A_n$, respectively, in the following way.

$$\operatorname*{Li}_{n\to\infty} A_n := \{\, x \in X \mid \forall U \in \mathfrak{U}(x) \, \exists n \in \mathbb{N} \, \forall k \geq n\colon \ U \cap A_k \neq \emptyset \,\},$$
$$\operatorname*{Ls}_{n\to\infty} A_n := \{\, x \in X \mid \forall U \in \mathfrak{U}(x) \, \forall n \in \mathbb{N} \, \exists k \geq n\colon \ U \cap A_k \neq \emptyset \,\}$$

If $A := \text{Ls}_{n\to\infty} A_n = \text{Li}_{n\to\infty} A_n$ agree, one says that A_n *converges to A in the sense of Kuratowski* and writes $\text{Lt}_{n\to\infty} A_n = A$.

Remark 1.11 Both the lower and the upper limit are closed sets and one has

$$\operatorname*{Li}_{n\to\infty} A_n \subset \operatorname*{Ls}_{n\to\infty} A_n.$$

One often refers to $\text{Ls}_{n\to\infty} A_n$ as the *set of cluster points* since x is an element of $\text{Ls}_{n\to\infty} A_n$ if and only if there is a sequence of elements $x_n \in A_n$ that has x as a cluster point, i.e., there is a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ that converges to x as $k \to \infty$.

A very useful identity which we use frequently is given by the following lemma.

Lemma 1.12 *Let X be a topological space and $A_n, A \subset X$, $n \in \mathbb{N}$. Then one has*

$$\operatorname*{Ls}_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} \overline{\bigcup_{k\geq n} A_k}.$$

PROOF. "⊂": Fix $x \in \mathrm{Ls}_{n \to \infty} A_n$ and let $n \in \mathbb{N}$ be arbitrary. For each $U \in \mathfrak{U}(x)$ one finds a $k \geq n$ with $U \cap A_k \neq \emptyset$, hence one has $x \in \overline{\bigcup_{k \geq n} A_k}$. Because this holds for each $n \in \mathbb{N}$, we obtain $x \in \bigcap_{n \in \mathbb{N}} \overline{\bigcup_{k \geq n} A_k}$.

"⊃": Now, fix $x \in \bigcap_{n \in \mathbb{N}} \overline{\bigcup_{k \geq n} A_k}$ and let $U \in \mathfrak{U}(x)$ be arbitrary. For every $n \in \mathbb{N}$, the intersection $U \cap \bigcup_{k \geq n} A_k$ has to be nonempty because x lies in the closure of $\bigcup_{k \geq n} A_k$. Thus, there is some $k \geq n$ with $U \cap A_k \neq \emptyset$, which shows $x \in \mathrm{Ls}_{n \to \infty} A_n$. □

The lower and upper limits are monotone:

Lemma 1.13 *Let X be a topological space and let $A_n \subset B_n \subset X$, $n \in \mathbb{N}$. Then one has*

$$\mathop{\mathrm{Li}}_{n \to \infty} A_n \subset \mathop{\mathrm{Li}}_{n \to \infty} B_n \quad \text{and} \quad \mathop{\mathrm{Ls}}_{n \to \infty} A_n \subset \mathop{\mathrm{Ls}}_{n \to \infty} B_n.$$

For later use, we point out that the Kuratowski limits are robust under metric thickening:

Lemma 1.14 *Let (X, d) be a metric space, $r_n \geq 0$ with $\lim_{n \to \infty} r_n = 0$ and $A_n \subset X$. Then one has*

$$\mathop{\mathrm{Li}}_{n \to \infty} \bar{B}(A_n, r_n) = \mathop{\mathrm{Li}}_{n \to \infty} A_n \quad \text{and} \quad \mathop{\mathrm{Ls}}_{n \to \infty} \bar{B}(A_n, r_n) = \mathop{\mathrm{Ls}}_{n \to \infty} A_n.$$

PROOF. Monotonicity of the Kuratowski limits leads to

$$\mathop{\mathrm{Li}}_{n \to \infty} A_n \subset \mathop{\mathrm{Li}}_{n \to \infty} \bar{B}(A_n, r_n), \quad \text{and} \quad \mathop{\mathrm{Ls}}_{n \to \infty} A_n \subset \mathop{\mathrm{Ls}}_{n \to \infty} \bar{B}(A_n, r_n).$$

Fix arbitrary $a \in \mathrm{Ls}_{n \to \infty} \bar{B}(A_n, r_n)$, $\varepsilon > 0$, and $N \in \mathbb{N}$ with $r_k < \frac{\varepsilon}{3}$ for all $k \geq N$. By definition of the Kuratowski limit superior, for each $n \geq N$ there is a $k \geq n$ and an $a_k \in B(a, \frac{\varepsilon}{3}) \cap \bar{B}(A_k, r_k)$. Now, choose $x_k \in A_k \cap B(a_k, 2r_k)$ and observe $d(a, x_k) \leq d(a, a_k) + d(a_k, x_k) < \frac{\varepsilon}{3} + 2r_k \leq \varepsilon$. Thus, we have $x_k \in B(a, \varepsilon) \cap A_k \neq \emptyset$ for all such k. This shows $a \in \mathrm{Ls}_{n \to \infty} A_n$.

Analogously, one shows $\mathrm{Li}_{n \to \infty} A_n \supset \mathrm{Li}_{n \to \infty} \bar{B}(A_n, r_n)$. □

In metric spaces, Hausdorff convergence implies Kuratowski convergence:

Lemma 1.15 *Let (X, d) be a metric space. Let $A_n \subset X$ be Hausdorff convergent to the set $A \subset X$. Then A_n converges to $\bar{A}$ in the sense of Kuratowski.*

9

PROOF. Let $(s_n)_{n \in \mathbb{N}}$ be a sequence of non-negative real numbers with $s_n \to 0$ as $n \to \infty$ such that

$$A_n \subset \bar{B}(A, s_n) \quad \text{and} \quad A \subset \bar{B}(A_n, s_n)$$

holds for each $n \in \mathbb{N}$. Put $r_n := \sup_{k \geq n} s_k$ and observe $r_n \to 0$ as $n \to \infty$. Now monotonicity of the Kuratowski limits and Lemma 1.14 lead to

$$\operatorname*{Ls}_{n \to \infty} A_n \subset \operatorname*{Ls}_{n \to \infty} \bar{B}(A, r_n) = \bar{A} = \operatorname*{Li}_{n \to \infty} \bar{B}(A, r_n) \subset \operatorname*{Li}_{n \to \infty} \bar{B}(A_n, 2r_n) = \operatorname*{Li}_{n \to \infty} A_n. \qquad \square$$

For non-empty sets in compact metric spaces, the notions of Kuratowski convergence and Hausdorff convergence coincide; in non-compact metric spaces, Hausdorff convergence is a stronger notion:

Lemma 1.16 *Let (X, d) be a metric space. The following are equivalent:*

1. *X is compact.*

2. *Every Kuratowski convergent sequence $(A_n)_{n \in \mathbb{N}}$ of sets $A_n \subset X$ with $\mathrm{Lt}_{n \to \infty} A_n \neq \emptyset$ is also Hausdorff convergent.*

Proof. "1. $\Rightarrow$ 2.": Let X be compact and let $(A_n)_{n \in \mathbb{N}}$ be a sequence of sets in X with $A = \mathrm{Lt}_{n \to \infty} A_n \neq \emptyset$. For $\varepsilon > 0$ and $n \in \mathbb{N}$ define the sets

$$X_\varepsilon := X \setminus B(A, \varepsilon) \quad \text{and} \quad U_n := X \setminus \overline{\bigcup_{k \geq n} A_k}.$$

Observe that X_ε is compact and that $(U_n)_{n \in \mathbb{N}}$ is an increasing sequence of open sets. Moreover, $(U_n)_{n \in \mathbb{N}}$ is a covering of U:

$$\bigcup_{n \in \mathbb{N}} U_n = X \setminus \left(\bigcap_{n \in \mathbb{N}} \overline{\bigcup_{k \geq n} A_k} \right) = X \setminus A \supset X \setminus B(A, \varepsilon) = X_\varepsilon.$$

Thus, there has to be a finite $n \in \mathbb{N}$ with $X_\varepsilon \subset U_n$. By construction of X_ε and U_n, we have $A_k \subset B(A, \varepsilon)$ for all $k \geq n$.

The next argument is by contradiction: *Assume* that there exists an $\varepsilon > 0$ such that for each $n \in \mathbb{N}$ there is a $k(n) \geq n$ with $A \not\subset B(A_{k(n)}, \varepsilon)$. Then we may choose a point $x_n \in A$ with $d(x_n, A_{k(n)}) > \varepsilon$. Because A is compact, there is a subsequence $(n_m)_{m \in \mathbb{N}}$ and some $x \in A$ with $x_{n_m} \to x$ as $m \to \infty$. Now, x is an element of the lower limit, hence there is some $N \in \mathbb{N}$ such that $d(x, A_n) < \frac{\varepsilon}{2}$ holds for all $n \geq N$. Additionally, by increasing N if necessary, we may assume $d(x, x_{n_m}) < \frac{\varepsilon}{2}$ whenever $n_m \geq N$. For $n_m \geq N$, we have $k(n_m) \geq n_m \geq N$ and the triangle inequality implies the *contradiction*

$$\varepsilon < d(x_{n_m}, A_{k(n_m)}) \leq d(x_{n_m}, x) + d(x, A_{k(n_m)}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon.$$

This shows for each $\varepsilon > 0$ that

$$A \subset B(A_n, \varepsilon) \quad \text{and} \quad A_n \subset B(A, \varepsilon)$$

hold for sufficiently large $n \in \mathbb{N}$, which is Hausdorff convergence.

"2. $\Rightarrow$ 1.": Let X be non-compact. We are going to show the existence of a sequence of non-empty sets $(A_n)_{n \in \mathbb{N}}$ with non-empty Kuratowski limit but without Hausdorff limit. Because X is non-compact, there is a sequence $(x_n)_{n \in \mathbb{N}}$ without any cluster points. In particular, there is some $\varepsilon > 0$ with $d(x_1, x_n) \geq \varepsilon$ for all $n \in \mathbb{N}$. Now define $A_n := \{x_1, x_n\}$, $n \in \mathbb{N}$. We have $A = \mathrm{Lt}_{n \to \infty} A_n = \{x_1\} \neq \emptyset$. If $(A_n)_{n \in \mathbb{N}}$ were Hausdorff convergent, its limit would have to coincide with A by Lemma 1.15. But for all $n > 1$, one has $A_n \setminus \bar{B}(A, \varepsilon) = \{x_n\} \neq \emptyset$, thus $(A_n)_{n \in \mathbb{N}}$ cannot Hausdorff converge. $\square$

1.4.4 Gamma convergence

Definition 1.17 Let X be a metric space and let $F, F_n\colon X \to \ ]-\infty, \infty]$, $n \in \mathbb{N}$ be some functions. One says, F_n *Γ-converges to F*, if the two following conditions are fulfilled:

1. lim inf-condition: *Every* sequence $(x_n)_{n\in\mathbb{N}}$ in X with $x_n \xrightarrow{n\to\infty} x$ satisfies

$$F(x) \leq \liminf_{n\to\infty} F_n(x_n).$$

2. lim sup-condition: For every $x \in X$ there *exists* a sequence $x_n \xrightarrow{n\to\infty} x$ with

$$F(x) \geq \limsup_{n\to\infty} F_n(x_n).$$

Such a sequence $(x_n)_{n\in\mathbb{N}}$ is called a *recovery sequence* of F at x.

Some authors refer to Γ-convergence as *epi-convergence* because of the following lemma (see [34, 7.2 Proposition] for details).

Lemma 1.18 *Let (X, d) be a metric space and $F, F_n\colon X \to \ ]-\infty, \infty]$, $n \in \mathbb{N}$ some functions. The following statements are equivalent:*

1. *F_n Γ-converges to F.*

2. *epi(F_n) converges to epi(F) in the sense of Kuratowski.*

As an immediate corollary, we see that the epigraph of a Γ-limit also is a Kuratowski limit, thus closed. Hence we have the following necessary conditions for a function to be a Γ-limit:

Corollary 1.19 *Let F_n Γ-converge to F. Then F is lower semi-continuous. In particular,* $\arg\min(F)$ *is closed.*

A basic link between Γ-convergence and convergence of minimizers is given by the following lemma:

Lemma 1.20 *Let F_n Γ-converge to F, assume* $\inf(F) < \infty$ *and let $(\varrho_n)_{n\in\mathbb{N}}$ be a sequence in $[0, \infty[$ with $\varrho_n \xrightarrow{n\to\infty} 0$. Then one has*

$$\limsup_{n\to\infty} \inf(F_n) \leq \inf(F) \quad and \quad \operatorname*{Ls}_{n\to\infty} \arg\min{}^{\varrho_n}(F_n) \subset \arg\min(F).$$

Proof. For $\varrho > 0$, let $x \in \arg\min^{\varrho}(F) \neq \emptyset$. By the $\lim\sup$-condition, there is a sequence $x_n \xrightarrow{n\to\infty} x$ with

$$\inf(F) + \varrho \geq F(x) \geq \limsup_{n\to\infty} F_n(x_n) \geq \limsup_{n\to\infty} \inf(F_n).$$

Because $\varrho > 0$ can be chosen arbitrarily, one obtains the first statement.

Let $x \in \mathrm{Ls}_{n\to\infty} \arg\min^{\varrho_n}(F_n)$. There is a sequence $x_{n_k} \in \arg\min^{\varrho_{n_k}}(F_{n_k})$ with $x_{n_k} \xrightarrow{k\to\infty} x$. By repeating elements of this sequence where necessary, we may construct a sequence $(x_n)_{n\in\mathbb{N}}$ containing $(x_{n_k})_{k\in\mathbb{N}}$ with $x_n \xrightarrow{n\to\infty} x$. Now, the $\lim\inf$-condition applies:

$$\begin{aligned}
F(x) &\leq \liminf_{n\to\infty} F_n(x_n) \leq \liminf_{k\to\infty} F_{n_k}(x_{n_k}) \\
&\leq \liminf_{k\to\infty} (\inf(F_{n_k}) + \varrho_{n_k}) = \liminf_{k\to\infty} \inf(F_{n_k}) \\
&\leq \limsup_{k\to\infty} \inf(F_{n_k}) \leq \limsup_{n\to\infty} \inf(F_n) \leq \inf(F).
\end{aligned}$$

This implies $x \in \arg\min(F)$, thus $\mathrm{Ls}_{n\to\infty} \arg\min^{\varrho_n}(F_n) \subset \arg\min(F)$. $\square$

In particular, the preceding lemma yields

$$\mathop{\mathrm{Ls}}_{n\to\infty} \arg\min(F_n) \subset \arg\min(F),$$

thus the cluster points of minimizers of $(F_n)_{n\in\mathbb{N}}$ are minimizers of F. This can be further improved by the following result (see [11, Theorem 7.19] for details):

Theorem 1.21 *Let F_n Γ-converge to F with $-\infty < \inf F < \infty$. Then one has*

$$\mathop{\mathrm{Ls}}_{n\to\infty} \arg\min(F_n) \subset \arg\min(F) = \bigcap_{\varrho > 0} \mathop{\mathrm{Li}}_{n\to\infty} \arg\min^{\varrho}(F_n).$$

In finite dimensional Euclidean space $X = \mathbb{R}^m$ (and presumably in all metric spaces with the Heine-Borel property), this result can be strengthened to (see [34, Theorem 7.31] for a proof):

Theorem 1.22 *Let $F, F_n : \mathbb{R}^m \to]-\infty, \infty]$, $n \in \mathbb{N}$ such that F_n Γ-converges to F with $-\infty < \inf F < \infty$ and assume $\arg\min(F) \neq \emptyset$. Then the following statements are equivalent:*

1. *$\inf(F_n) \xrightarrow{n\to\infty} \inf(F)$.*

2. *There is a monotonically decreasing sequence $\varrho_n \searrow 0$ such that*

$$\mathop{\mathrm{Lt}}_{n\to\infty} \arg\min^{\varrho_n}(F_n) = \arg\min(F).$$

1.4.5 Epigraph distances

A way to quantify Γ-convergence is by measuring the Hausdorff distance of truncated epigraphs. This idea was introduced in [1] and [2]. See also [34, Sections 7.I and 7.J] for a detailed treatment.

Definition 1.23 Let X be some set and $F\colon X \to\ ]-\infty, \infty]$ some functions. For a set $K \subset X$ and a number $r \in [-\infty, \infty]$ define the *truncated epigraph* $\mathrm{epi}_K^r(F)$ *of* F by

$$\mathrm{epi}_K^r(F) := \{\, (x, t) \in K \times\ ]-\infty, \infty] \mid F(x) \le t \le r \,\}.$$

Note the identities $\mathrm{epi}_K^{-\infty}(F) = \emptyset$ and $\mathrm{epi}_K^\infty(F) = K \times\ ]-\infty, \infty]$.

Using the box metric on $X \times \mathbb{R}$ given by

$$d((x, s), (y, t)) := \max\left(|y - x|, |t - s|\right), \quad \text{for all } x, y \in X, \ s, t \in \mathbb{R},$$

the central result of the theory of epigraph distances can be formulated as:

Theorem 1.24 *Let (X, d_X) be a metric space, let $F, G\colon X \to\ ]-\infty, \infty]$ be some functions, $K \subset X$ some subset, and $r \in \mathbb{R}$ such that the following hold:*

- $-\infty < \inf(F) \le r$ *and* $-\infty < \inf(G) \le r$.

- $\emptyset \neq \arg\min(F) \subset K$ *and* $\emptyset \neq \arg\min(G) \subset K$.

Then for each $\delta > 0$ with

$$\delta > \mathrm{dist}\left(\mathrm{epi}_K^r(F), \mathrm{epi}_K^r(G)\right) \quad \text{and} \quad \delta > r - \min(\inf(F), \inf(G)),$$

one has:

1. $|\inf(F) - \inf(G)| \le \mathrm{dist}\left(\mathrm{epi}_K^r(F), \mathrm{epi}_K^r(G)\right)$,

2. $\arg\min(F) \subset \bar{B}(\arg\min^{2\delta}(G), \delta)$,

3. $\arg\min(G) \subset \bar{B}(\arg\min^{2\delta}(F), \delta)$.

PROOF. Let $\mathrm{pr}_X\colon X \times \mathbb{R} \to X$ and $\mathrm{pr}_\mathbb{R}\colon X \times \mathbb{R} \to \mathbb{R}$ be the projections onto the factors. One has for each $\varepsilon > \mathrm{dist}\left(\mathrm{epi}_K^r(F)\right), \mathrm{epi}_K^r(G))$:

$$]\inf(F), r] \subset \mathrm{pr}_\mathbb{R}\left(\mathrm{epi}_K^r(F)\right) \subset \mathrm{pr}_\mathbb{R}\left(\bar{B}(\mathrm{epi}_K^r(G), \varepsilon)\right) \subset [\inf(G) - \varepsilon, r + \varepsilon],$$

thus $\inf(F) \geq \inf(G) - \varepsilon$. Vice versa, one obtains $]\inf(G), r] \subset [\inf(F) - \varepsilon, r + \varepsilon]$ and $\inf(G) \geq \inf(F) - \varepsilon$, which leads to the first statement. Similarly, one has for each δ as above:

$$
\begin{aligned}
\arg\min(F) \subset \arg\min^{r-\inf(F)}(F) &= \operatorname{pr}_X\left(\operatorname{epi}_K^r(F)\right) \\
&\subset \operatorname{pr}_X\left(\bar{B}(\operatorname{epi}_K^r(G), \delta)\right) \\
&\subset \bar{B}(\arg\min^{r-\inf(G)+\delta}(G), \delta) \\
&\subset \bar{B}(\arg\min^{2\delta}(G), \delta),
\end{aligned}
$$

which is the second statement. Analogously, one obtains the third statement. $\qquad\square$

If the functions F and G are well-behaved and δ is small, then $\arg\min^{2\delta}(F)$ and $\arg\min^{2\delta}(G)$ are not too far away from $\arg\min(F)$ and $\arg\min(G)$, respectively. What "well-behaved" means will be made precise later in Section 2.3.3.

2 Parametric Optimization Problems

In this chapter we present the theoretical core of our considerations on the convergence of nonconforming Ritz-Galerkin methods. The central aim is to break the overwhelming task of convergence analysis into smaller, manageable subtasks: showing consistency, proximity, and stability.

We keep the presentation as broad as possible. Concrete applications will be given later in Section 3.3, Chapter 7, and Chapter 8.

Let C be a topological space and let $\mathcal{F}: C \to \mathbb{R}$ be a function. We denote the set of minimizers by

$$\mathcal{M} := \arg\min(\mathcal{F}) = \{ x \in C \mid \mathcal{F}(x) = \inf(\mathcal{F}) \}$$

and sets of δ-minimizers by

$$\mathcal{M}^\delta := \arg\min{}^\delta(\mathcal{F}) = \{ x \in C \mid \forall y \in C: \ \mathcal{F}(x) \le \mathcal{F}(y) + \delta \}, \quad \text{for } \delta \in [0, \infty].$$

Moreover, let topological spaces C_n and functions $\mathcal{F}_n: C_n \to \mathbb{R}$ with minimizers $\mathcal{M}_n := \arg\min(\mathcal{F}_n)$ and δ-minimizers $\mathcal{M}_n^\delta := \arg\min{}^\delta(\mathcal{F}_n)$ be given for each $n \in \mathbb{N}$. One may think of $\mathcal{F}_n$ as small perturbations of $\mathcal{F}$ or—in the context of Ritz-Galerkin methods—as a discretization of $\mathcal{F}$. We are interested in the behavior of the sets $\mathcal{M}_n$ as $n \to \infty$.[1] Ideally, $\mathcal{M}_n$ should approximate $\mathcal{M}$ "in some way". Therefore, we need a method to relate these sets. A rather general way is letting C and C_n communicate with each other via some mappings

$$\mathcal{S}_n: \ \mathrm{dom}(\mathcal{S}_n) \subset C \to C_n \quad \text{and} \quad \mathcal{R}_n: \ \mathrm{dom}(\mathcal{R}_n) \subset C_n \to C.$$

We are going to refer to $\mathcal{S}_n$ as the *sampling operator* and to $\mathcal{R}_n$ as the *reconstruction operator*. If one has $\mathcal{M}_n \subset \mathrm{dom}(\mathcal{R}_n)$ and $\mathcal{M} \subset \mathrm{dom}(\mathcal{S}_n)$, the pairs of sets $(\mathcal{M}, \mathcal{R}_n(\mathcal{M}_n))$ and $(\mathcal{M}_n, \mathcal{S}_n(\mathcal{M}))$ lie in common spaces C and C_n, respectively. Thus, they can be compared.

---

[1] We point out that this framework also covers the more general problem of analyzing $\lim_{\lambda \to \lambda_\infty} \mathcal{M}_\lambda$, where $(\mathcal{F}_\lambda: C_\lambda \to \mathbb{R})_{\lambda \in \Lambda}$ is a family of functions parameterized by a sequential topological space Λ. For the sake of brevity, we only discuss $\Lambda = \mathbb{N} \cup \{\infty\}$ with the topology $\mathfrak{T} = \{\emptyset\} \cup \{ U \subset \mathbb{N} \cup \{\infty\} \mid \infty \in U \text{ and } \mathrm{card}(U) = \infty \}$.

Example 2.1 In the finite element method, C_n is often a finite-dimensional affine subspace of a Banach space C, $\mathcal{R}_n$ is the canonical embedding and $\mathcal{S}_n$ is an interpolation operator such that $\mathcal{R}_n \circ \mathcal{S}_n$ is a projection onto the ansatz space $\mathcal{R}_n(C_n)$. Then, the Hausdorff distance between $\mathcal{M}$ and $\mathcal{R}_n(\mathcal{M}_n)$ with respect to d_C yields a canonical measure of approximation.

More generally, one may consider mappings $\Psi \colon C \to \mathcal{X}$, $\Psi_n \colon C_n \to \mathcal{X}$ to some metric space $(\mathcal{X}, d_{\mathcal{X}})$ and analyze the Hausdorff distance between the sets $\Psi(\mathcal{M})$ and $\Psi_n(\mathcal{M}_n)$ therein. For example, Ψ, Ψ_n could be embeddings into a space $\mathcal{X}$ whose metric topology is weaker than those of C, C_n. But Ψ, Ψ_n need not to be injective at all, yielding only partial information: They could also represent restriction or trace mappings, truncations in infinite decompositions (e.g. Fourier or modal representations, projections on subspaces etc.), state variables in physical systems, (locally) averaged quantities or even quotient mappings. We propose to view Ψ, Ψ_n as nonlinear variants of *test functions*.

In practice, one might profit considerably from using a priori information on minimizers (such as higher regularity or energy bounds) in order to achieve quantitative approximation results. We are going to incoorporate a priori information in the form of subsets $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ such that $\Psi(\mathcal{A} \cap \mathcal{M})$, $\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n)$ contain $\Psi(\mathcal{M})$, $\Psi_n(\mathcal{M}_n)$ respectively.[2]

We summarize the information given so far in the following (not necessarily commutative) diagrams

$$(2.1)$$

$$(2.2)$$

Assume for the moment $\emptyset \neq \mathcal{M} \subset \mathcal{A}$, $\emptyset \neq \mathcal{M}_n \subset \mathcal{A}_n$. If (2.1) were commutative, one would have the following implications

1. For each $x \in \mathcal{M}$ and $y \in C_n$:

$$\mathcal{F}_n \circ \mathcal{S}_n(x) = \mathcal{F}(x) \leq \mathcal{F} \circ \mathcal{R}_n(y) = \mathcal{F}_n(y), \quad \text{thus} \quad \mathcal{S}_n(\mathcal{M}) \subset \mathcal{M}_n.$$

2. For each $y \in \mathcal{M}_n$ and $x \in C$:

$$\mathcal{F} \circ \mathcal{R}_n(y) = \mathcal{F}_n(y) \leq \mathcal{F}_n \circ \mathcal{S}_n(x) = \mathcal{F}(x), \quad \text{thus} \quad \mathcal{R}_n(\mathcal{M}_n) \subset \mathcal{M}.$$

If, in addition, (2.2) were commutative, this would lead to

3. $\Psi(\mathcal{M}) = \Psi_n \circ \mathcal{S}_n(\mathcal{M}) \subset \Psi_n(\mathcal{M}_n),$

4. $\Psi_n(\mathcal{M}_n) = \Psi \circ \mathcal{R}_n(\mathcal{M}_n) \subset \Psi(\mathcal{M}),$

hence $\Psi(\mathcal{M}) = \Psi_n(\mathcal{M}_n)$.

Alas, in practice, these diagrams rarely commute. But one may hope that they *almost* commute, i.e., they commute up to some errors that can be uniformly bounded, at least on the sets of a priori information.

In Sections 2.1 and 2.2, we will name these non-commutativity errors and analyze what information can be deduced if these errors are sufficiently small. Afterwards, we will single out additional conditions that guarantee convergence of minimizers (Section 2.3).

2.1 Consistency

2.1.1 General theory

We start with the first diagram (2.1). For non-empty sets $\mathcal{A} \subset \text{dom}(\mathcal{S}_n)$, $\mathcal{A}_n \subset \text{dom}(\mathcal{R}_n)$, there are two squares of interest:

$$\tag{2.3}$$

Each square is equipped with its own non-cummutativity error:

Definition 2.2 (Consistency) For non-empty sets $\mathcal{A} \subset \mathrm{dom}(\mathcal{S}_n)$, $\mathcal{A}_n \subset \mathrm{dom}(\mathcal{R}_n)$, define

1. the *sampling consistency error*

$$\delta_n^{\mathcal{S}} := \delta(\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{A}) := \sup_{a \in \mathcal{A}} (\mathcal{F}_n \circ \mathcal{S}_n(a) - \mathcal{F}(a))^+, \qquad (2.4)$$

2. the *reconstruction consistency error*

$$\delta_n^{\mathcal{R}} := \delta(\mathcal{F}, \mathcal{F}_n, \mathcal{R}_n, \mathcal{A}_n) := \sup_{a \in \mathcal{A}_n} (\mathcal{F} \circ \mathcal{R}_n(a) - \mathcal{F}_n(a))^+, \qquad (2.5)$$

3. the *total consistency error*

$$\delta_n := \delta(\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n, \mathcal{A}, \mathcal{A}_n) := \delta_n^{\mathcal{S}} + \delta_n^{\mathcal{R}}, \qquad (2.6)$$

where $t^+ := \max\{t, 0\}$ denotes the non-negative part of $t \in \mathbb{R}$. We say, the sequence $((\mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ is *consistent with respect to $\mathcal{F}$ on $((\mathcal{A}, \mathcal{A}_n))_{n \in \mathbb{N}}$*, if its consistency error $\delta(\mathcal{S}_n, \mathcal{R}_n, \mathcal{A}, \mathcal{A}_n)$ converges to 0 for $n \to \infty$. In that case, we also say that $((\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ is *consistent on $((\mathcal{A}, \mathcal{A}_n))_{n \in \mathbb{N}}$*.

Remark 2.3 A stronger notion of (total) consistency error (but also one which would be harder to verify) would be

$$\sup_{a \in \mathcal{A}} |\mathcal{F}_n \circ \mathcal{S}_n(a) - \mathcal{F}(a)| + \sup_{a \in \mathcal{A}_n} |\mathcal{F} \circ \mathcal{R}_n(a) - \mathcal{F}_n(a)|.$$

In light of the latter expression, our definition of consistency error could be termed *upper consistency error*. However, our definition is sufficient for our needs and we omit "upper" for the sake of brevity. Of course, one may also define a *lower consistency error*, which would be the notion of choice for maximization problems.

Definition 2.4 We call $\mathcal{A} \subset C$ *valid with respect to the pair $(\mathcal{F}, \mathcal{S}_n)$*, if $\emptyset \neq \mathcal{A} \subset \mathrm{dom}(\mathcal{S}_n)$ and $\inf(\mathcal{F}) = \inf(\mathcal{F}|_{\mathcal{A}})$ hold. Analogously, we call $\mathcal{A}_n \subset C_n$ *valid with respect to the pair $(\mathcal{F}_n, \mathcal{R}_n)$*, if $\emptyset \neq \mathcal{A}_n \subset \mathrm{dom}(\mathcal{R}_n)$ and $\inf(\mathcal{F}_n) = \inf(\mathcal{F}_n|_{\mathcal{A}_n})$ hold. For the sake of brevity, we will simply say that $\mathcal{A}, \mathcal{A}_n$ are *valid* whenever $(\mathcal{F}, \mathcal{S}_n)$ and $(\mathcal{F}_n, \mathcal{R}_n)$ can be deduced from the context.

This definition is a bit subtle, but crucial: On the one hand, validity allows us to consider sampling and reconstruction errors by demanding that valid sets are contained in the domains of the respective operators. Hence, this part of the definition merely serves as a short-hand notation. On the other hand, $\inf(\mathcal{F}) = \inf(\mathcal{F}|_{\mathcal{A}})$ and $\inf(\mathcal{F}_n) = \inf(\mathcal{F}_n|_{\mathcal{A}_n})$ are vital requirements for the following momentous lemma:

Lemma 2.5 *Let $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ be valid with respect to $(\mathcal{F}, \mathcal{S}_n)$ and $(\mathcal{F}_n, \mathcal{R}_n)$, respectively. Assume that both the sampling consistency error $\delta_n^{\mathcal{S}}$ and the reconstruction consistency error $\delta_n^{\mathcal{R}}$ are finite. Then one has*

$$\inf(\mathcal{F}_n) \leq \inf(\mathcal{F}) + \delta_n^{\mathcal{S}} \quad \text{and} \quad \inf(\mathcal{F}) \leq \inf(\mathcal{F}_n) + \delta_n^{\mathcal{R}}.$$

Hence, one has either $\inf(\mathcal{F}_n) = \inf(\mathcal{F}) = -\infty$ or both $\inf(\mathcal{F}_n)$ and $\inf(\mathcal{F})$ are finite with

$$|\inf(\mathcal{F}_n) - \inf(\mathcal{F})| \leq \max(\delta_n^{\mathcal{S}}, \delta_n^{\mathcal{R}}).$$

PROOF. Choose a minimizing sequence $(x_m)_{m \in \mathbb{N}}$ in $\mathcal{A}$ for $\mathcal{F}$ and a minimizing sequence $(y_m)_{m \in \mathbb{N}}$ in $\mathcal{A}_n$ for $\mathcal{F}_n$, i.e.,

$$\inf(\mathcal{F}) = \lim_{m \to \infty} \mathcal{F}(x_m) \quad \text{and} \quad \inf(\mathcal{F}_n) = \lim_{m \to \infty} \mathcal{F}_n(y_m).$$

Then (2.4) and (2.5) imply

$$\inf(\mathcal{F}_n) \leq \mathcal{F}_n(\mathcal{S}_n(x_m)) \leq \mathcal{F}(x_m) + \delta_n^{\mathcal{S}} \xrightarrow{m \to \infty} \inf(\mathcal{F}) + \delta_n^{\mathcal{S}},$$

$$\inf(\mathcal{F}) \leq \mathcal{F}(\mathcal{R}_n(y_m)) \leq \mathcal{F}_n(y_m) + \delta_n^{\mathcal{R}} \xrightarrow{m \to \infty} \inf(\mathcal{F}_n) + \delta_n^{\mathcal{R}}. \qquad \square$$

Knowing the total consistency error puts one into the position to compare ϱ-minimizers:

Lemma 2.6 *Let $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ be valid sets. Denote by δ_n the total consistency error. Then one has for $\varrho \in [0, \infty]$*

$$\mathcal{S}_n(\mathcal{A} \cap \mathcal{M}^{\varrho}) \subset \mathcal{S}_n(\mathcal{A}) \cap \mathcal{M}_n^{\varrho + \delta_n} \quad \text{and} \quad \mathcal{R}_n(\mathcal{A}_n \cap \mathcal{M}_n^{\varrho}) \subset \mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{\varrho + \delta_n}.$$

PROOF. **Case 1:** $\varrho = \infty$ or $\delta_n = \infty$. The inclusions hold because of $\mathcal{M}_n^{\varrho + \delta_n} = \mathcal{M}_n^{\infty} = C_n$ and $\mathcal{M}^{\varrho + \delta_n} = \mathcal{M}^{\infty} = C_n$.[3]
Case 2: Both ϱ and δ_n are finite. Then, by Lemma 2.5, either both $\inf(\mathcal{F})$ and $\inf(\mathcal{F}_n)$ equal $-\infty$ or both of them are finite.
Case 2.a: $\inf(\mathcal{F}) = \inf(\mathcal{F}_n) = -\infty$. All the sets $\mathcal{M}^{\varrho}$, $\mathcal{M}^{\varrho + \delta_n}$, $\mathcal{M}_n^{\varrho}$, $\mathcal{M}_n^{\varrho + \delta_n}$ are empty such that the inclusions hold trivially.
Case 2.b: $\inf(\mathcal{F}), \inf(\mathcal{F}_n) > -\infty$. We discuss only the first inclusion for the second one follows analogously. In the case that $\mathcal{A} \cap \mathcal{M}^{\varrho}$ is empty, nothing is to show. Otherwise, let $x \in \mathcal{A} \cap \mathcal{M}^{\varrho}$. We apply (1.1) and Lemma 2.5 in order to estimate

$$\mathcal{F}_n \circ \mathcal{S}_n(x) \leq \mathcal{F}(x) + \delta_n^{\mathcal{S}} \leq \inf(\mathcal{F}) + \varrho + \delta_n^{\mathcal{S}} \leq \inf(\mathcal{F}_n) + \delta_n^{\mathcal{R}} + \varrho + \delta_n^{\mathcal{S}}$$

This leads to $\mathcal{S}_n(x) \in \mathcal{M}_n^{\varrho + \delta_n}$ which shows the first inclusion. $\qquad \square$

---

[3] Admittedly, this case is of little practical relevance.

Remark 2.7 For the moment, it may appear as a superfluous burden to drag along $\mathcal{S}_n(\mathcal{A})$, $\mathcal{R}_n(\mathcal{A}_n)$ on the right hand side of the previous lemma's conclusions. However, this may be crucial when treating optimization problems with *non-compact* lower level sets as we will see in Corollary 2.30. The area functional of immersed surfaces as discussed in Chapter 7 is such an example. For a demonstration of the non-compactness of lower level sets see Figure 7.3.

Corollary 2.8 *Let $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ be valid, let $\mathcal{F} \colon C \to \mathbb{R}$ be lower semi-continuous on the set $\mathcal{B} := \mathrm{Ls}_{n\to\infty} \mathcal{R}_n(\mathcal{A}_n)$, and let $((\mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n))_{n\in\mathbb{N}}$ be consistent on $((\mathcal{A}, \mathcal{A}_n))_{n\in\mathbb{N}}$. Then one has for $\varrho \in [0, \infty]$:*

$$\mathrm{Ls}_{n\to\infty} (\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}_n^{\varrho}) \subset \mathcal{B} \cap \mathcal{M}^{\varrho}.$$

Proof. In the case that $\varrho = \infty$, this is obviously true. Hence, let us assume that ϱ is finite. Denote by δ_n the total consistency error, $\mathcal{K}_n^{\varrho} := \mathcal{A}_n \cap \mathcal{M}_n^{\varrho}$ and let $\eta_n := \sup\{\delta_k \mid k \geq n\}$. Observe that $\eta_n \searrow 0$ as $n \nearrow \infty$ by consistency. By Lemma 2.6, we have for all $m, n \in \mathbb{N}$ with $m \geq n$:

$$\mathcal{R}_m(\mathcal{K}_m^{\varrho}) \subset (\textstyle\bigcup_{k\geq m} \mathcal{R}_k(\mathcal{A}_k)) \cap \mathcal{M}^{\varrho+\eta_n}.$$

Taking closures and intersections leads to

$$\bigcap_{n\in\mathbb{N}} \overline{\bigcup_{k\geq n} \mathcal{R}_k(\mathcal{K}_k^{\varrho})} = \bigcap_{n\in\mathbb{N}} \bigcap_{m\geq n} \overline{\bigcup_{k\geq m} \mathcal{R}_k(\mathcal{K}_k^{\varrho})}$$

$$\subset \bigcap_{n\in\mathbb{N}} \left(\left(\bigcap_{m\geq n} \overline{\bigcup_{k\geq m} \mathcal{R}_k(\mathcal{A}_k)} \right) \cap \overline{\mathcal{M}^{\varrho+\eta_n}} \right) = \bigcap_{n\in\mathbb{N}} \left(\mathcal{B} \cap \overline{\mathcal{M}^{\varrho+\eta_n}} \right).$$

Because $\mathcal{B}$ is closed and $\mathcal{F}|_{\mathcal{B}}$ is lower-semicontinuous, one has

$$\mathcal{B} \cap \overline{\mathcal{M}^{\varrho+\eta_n}} = \mathcal{B} \cap \mathcal{M}^{\varrho+\eta_n}.$$

Finally, $\bigcap_{n\in\mathbb{N}} \mathcal{M}^{\varrho+\eta_n} = \mathcal{M}^{\varrho}$ completes the proof. $\qquad\square$

Using $\varrho = 0$, this leads immediately to

Corollary 2.9 *Let $\mathcal{A} \supset \mathcal{M}$, $\mathcal{A}_n \supset \mathcal{M}_n$ be valid. Denote by δ_n the total consistency error. Then one has*

$$\mathcal{S}_n(\mathcal{M}) \subset \mathcal{S}_n(\mathcal{A}) \cap \mathcal{M}_n^{\delta_n} \quad and \quad \mathcal{R}_n(\mathcal{M}_n) \subset \mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{\delta_n}.$$

Assuming consistency of $((\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n))_{n\in\mathbb{N}}$ on $((\mathcal{A}, \mathcal{A}_n))_{n\in\mathbb{N}}$ and lower semi-continuity of $\mathcal{F}$, cluster points of minimizers of $(\mathcal{F}_n)_{n\in\mathbb{N}}$ are minimizers of $\mathcal{F}$ in the sense that

$$\mathrm{Ls}_{n\to\infty} \mathcal{R}_n(\mathcal{M}_n) \subset \mathcal{M}.$$

Remark 2.10 Note that the preceding corollary can be used to prove the existence of minimizers of $\mathcal{F}$ if one manages to verify that cluster points of minimizers of $(\mathcal{F}_n)_{n \in \mathbb{N}}$ actually exist.

2.1.2 Coupling

Consistency errors behave well under coupling of several functionals. We simply state the results as their proofs are straightforward.

Lemma 2.11 *Let non-empty sets $\mathcal{A}$, $\mathcal{B} \subset \mathrm{dom}(\mathcal{S}_n)$, $\mathcal{A}_n$, $\mathcal{B}_n \subset \mathrm{dom}(\mathcal{R}_n)$ and functions $\mathcal{F}$, $\mathcal{G} \colon C \to \mathbb{R}$, $\mathcal{F}_n$, $\mathcal{G}_n \colon C_n \to \mathbb{R}$ be given. For α, $\beta \in [0, \infty[$ define $\mathcal{H} := \alpha \mathcal{F} + \beta \mathcal{G}$ and $\mathcal{H}_n := \alpha \mathcal{F}_n + \beta \mathcal{G}_n$. If $\mathcal{A} \cap \mathcal{B}$ and $\mathcal{A}_n \cap \mathcal{B}_n$ are non-empty, one has the following estimates for the consistency errors:*

$$\delta(\mathcal{H}, \mathcal{H}_n, \mathcal{S}_n, \mathcal{A} \cap \mathcal{B}) \leq \alpha \, \delta(\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{A}) + \beta \, \delta(\mathcal{G}, \mathcal{G}_n, \mathcal{S}_n, \mathcal{B}),$$
$$\delta(\mathcal{H}, \mathcal{H}_n, \mathcal{R}_n, \mathcal{A}_n \cap \mathcal{B}_n) \leq \alpha \, \delta(\mathcal{F}, \mathcal{F}_n, \mathcal{R}_n, \mathcal{A}_n) + \beta \, \delta(\mathcal{G}, \mathcal{G}_n, \mathcal{R}_n, \mathcal{B}_n).$$

Corollary 2.12 *The set of sequences $((\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ as in (2.3) that are consistent on $((\mathcal{A}, \mathcal{A}_n))_{n \in \mathbb{N}}$ is a convex cone.*

Remark 2.13 In the same way, one could say that the set of lower and upper consistent sequences $((\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ is a vector space over $\mathbb{R}$ (see also Remark 2.3).

2.2 Proximity

2.2.1 General theory

Corollary 2.9 is quite similar to the inclusion $\mathrm{Ls}_{n \to \infty} \arg \min(F_n) \subset \arg \min(F)$ from Theorem 1.21. We currently have no analogue for the inequality $\arg \min(F) \subset \mathrm{Li}_{n \to \infty} \arg \min^\varrho(F_n)$ for $\varrho > 0$, but only

$$\mathcal{S}_n(\mathcal{A} \cap \mathcal{M}) \subset \mathcal{S}_n(\mathcal{A}) \cap \mathcal{M}_n^{\delta_n}.$$

If $\mathcal{R}_n \circ \mathcal{S}_n$ were the identity on C and assuming consistency, one would obtain for given $\varrho > 0$ and for all sufficiently large n:

$$\mathcal{A} \cap \mathcal{M} = \mathcal{R}_n \circ \mathcal{S}_n(\mathcal{A} \cap \mathcal{M}) \subset \mathcal{A} \cap \mathcal{R}_n(\mathcal{M}_n^{\delta_n}) \subset \mathcal{A} \cap \mathcal{R}_n(\mathcal{M}_n^\varrho).$$

This would yield the desired result

$$\mathcal{A} \cap \mathcal{M} \subset \mathcal{A} \cap \bigcap_{\varrho > 0} \mathrm{Li}_{n \to \infty} \mathcal{R}_n(\mathcal{M}_n^\varrho).$$

In the case that C is an infinite-dimensional Banach space and C_n a finite-dimensional Banach space, $\mathcal{R}_n \circ \mathcal{S}_n = \mathrm{id}_C$ cannot occur. Even worse: In this context, $\mathcal{R}_n \circ \mathcal{S}_n$ may be a compact operator so it cannot converge uniformly to id_C. Hence we have to establish a sufficiently weak notion for $\mathcal{R}_n \circ \mathcal{S}_n$ being "sufficiently close to id_C", a notion that does not imply uniform approximation.

We do this in a slightly more general way by discussing diagram (2.2). At times, it may be instructive for the reader to substitute $X = C$, $\Psi = \mathrm{id}_C$ and $\Psi_n = \mathcal{R}_n$. Again, for non-empty sets $\mathcal{A} \subset \mathrm{dom}(\mathcal{S}_n)$, $\mathcal{A}_n \subset \mathrm{dom}(\mathcal{R}_n)$, there are two squares of interest:

$$\begin{array}{ccc} & C & \\ \mathcal{A} \nearrow & \searrow \Psi & \\ & & X\,, \\ \mathcal{S}_n \searrow & \nearrow \Psi_n & \\ & C_n & \end{array} \qquad \begin{array}{ccc} & C & \\ \mathcal{R}_n \nearrow & \searrow \Psi & \\ \mathcal{A}_n & & X\,. \\ \searrow & \nearrow \Psi_n & \\ & C_n & \end{array} \tag{2.7}$$

Note that for our purposes, we do not require Ψ, Ψ_n to be defined on all of C, C_n but at least on the sets $\mathcal{A} \cup \mathcal{R}_n(\mathcal{A}_n)$, $\mathcal{A}_n \cup \mathcal{S}_n(\mathcal{A})$, respectively. Each of the two diagrams has its own non-commutativity error:

Definition 2.14 (Proximity) For non-empty sets $\mathcal{A} \subset \mathrm{dom}(\mathcal{S}_n)$, $\mathcal{A}_n \subset \mathrm{dom}(\mathcal{R}_n)$, define

1. the *sampling proximity error*

$$\varepsilon_n^{\mathcal{S}} := \varepsilon(\Psi, \Psi_n, \mathcal{S}_n, \mathcal{A}) := \sup_{a \in \mathcal{A}} d_X(\Psi_n \circ \mathcal{S}_n(a), \Psi(a)), \tag{2.8}$$

2. the *reconstruction proximity error*

$$\varepsilon_n^{\mathcal{R}} := \varepsilon(\Psi, \Psi_n, \mathcal{R}_n, \mathcal{A}_n) := \sup_{a \in \mathcal{A}_n} d_X(\Psi \circ \mathcal{R}_n(a), \Psi_n(a)), \tag{2.9}$$

3. the *proximity error*

$$\varepsilon_n := \varepsilon(\Psi, \Psi_n, \mathcal{S}_n, \mathcal{R}_n, \mathcal{A}, \mathcal{A}_n) := \max\left(\varepsilon_n^{\mathcal{S}}, \varepsilon_n^{\mathcal{R}}\right). \tag{2.10}$$

We say that the sequence $((\mathcal{S}_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ is *proximate with respect to* $((\Psi, \Psi_n))_{n \in \mathbb{N}}$ *on* $((\mathcal{A}, \mathcal{A}_n))_{n \in \mathbb{N}}$, if its proximity error ε_n converges to 0 as $n \to \infty$.

Lemma 2.15 *Let $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ be valid. Denote by δ_n the total consistency error. Then one has for $\varrho \in [0, \infty]$*

$$\Psi(\mathcal{A} \cap \mathcal{M}^\varrho) \subset \bar{B}(\Psi_n(\mathcal{S}_n(\mathcal{A}) \cap \mathcal{M}_n^{\varrho+\delta_n}), \varepsilon_n^\mathcal{S}),$$
$$\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^\varrho) \subset \bar{B}(\Psi(\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{\varrho+\delta_n}), \varepsilon_n^\mathcal{R}).$$

Proof. For $x \in \Psi(\mathcal{A} \cap \mathcal{M}^\varrho)$ (if existent), fix an $a \in \mathcal{A} \cap \mathcal{M}^\varrho$ with $x = \Psi(a)$. According to Lemma 2.6, we have that $\mathcal{S}_n(a) \in \mathcal{S}_n(\mathcal{A}) \cap \mathcal{M}_n^{\varrho+\delta_n}$. Now, the definition of the sampling proximity error implies

$$d_X(\Psi(a), \Psi_n(\mathcal{S}_n(a))) \le \varepsilon_n^\mathcal{S},$$

thus $\Psi(a) \in \bar{B}(\Psi_n(\mathcal{S}_n(\mathcal{A}) \cap \mathcal{M}_n^{\varrho+\delta_n}), \varepsilon_n^\mathcal{S})$. The proof of the second statement is now straightforward. $\square$

Lemma 2.16 *In addition to the previous lemma, assume $\mathcal{S}_n(\mathcal{A}) \subset \mathcal{A}_n$ and proximity, i.e., $\varepsilon_n \xrightarrow{n\to\infty} 0$. Then one has*

$$\Psi(\mathcal{A} \cap \mathcal{M}) \subset \underset{n\to\infty}{\mathrm{Li}}\ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n})$$
$$\subset \underset{n\to\infty}{\mathrm{Ls}}\ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}) \subset \underset{n\to\infty}{\mathrm{Ls}}\ \Psi(\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{2\delta_n}).$$

Proof. By applying the previous lemma twice—once with $\varrho = 0$, once with $\varrho = \delta_n$—and by the triangle inequality, one has for all $n \in \mathbb{N}$:

$$\Psi(\mathcal{A} \cap \mathcal{M}) \subset \bar{B}(\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}), \varepsilon_n^\mathcal{S})$$
$$\subset \bar{B}(\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}), \varepsilon_n^\mathcal{S})$$
$$\subset \bar{B}(\Psi(\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}_n^{2\delta_n}), \varepsilon_n^\mathcal{S} + \varepsilon_n^\mathcal{R}).$$

Because of the monotonicity properties of Li and Ls, we may apply Li, Ls, Ls to the three terms on the right hand side, respectively, without invalidating the inclusions. The statement then follows from thickening robustness (Lemma 1.14). $\square$

Assume for a moment that $\mathcal{A}$, $\mathcal{A}_n$ are sets of true a priori information, i.e., $\mathcal{A} \supset \mathcal{M}$, $\mathcal{A}_n \supset \mathcal{M}_n$. Lemma 2.15 tells us that—up to Ψ, Ψ_n—minimizers are at least close to δ_n-minimizers:

$$\Psi(\mathcal{M}) \subset \bar{B}(\Psi_n(\mathcal{M}_n^{\delta_n}), \varepsilon_n^\mathcal{S}) \quad \text{and} \quad \Psi_n(\mathcal{M}_n) \subset \bar{B}(\Psi(\mathcal{M}^{\delta_n}), \varepsilon_n^\mathcal{R}). \tag{2.11}$$

If $\mathcal{F}$, $\mathcal{F}_n$ are "not too shallow", one may expect existence of a (small) $r_n \ge 0$ with

$$\Psi_n(\mathcal{M}_n^{\delta_n}) \subset \bar{B}(\Psi(\mathcal{M}), r_n) \quad \text{and} \quad \Psi(\mathcal{M}^{\delta_n}) \subset \bar{B}(\Psi_n(\mathcal{M}_n), r_n). \tag{2.12}$$

25

Under these circumstances, the triangle inequality would yield:

$$\Psi(\mathcal{M}) \subset \bar{B}(\Psi_n(\mathcal{M}_n), r_n + \varepsilon_n^{\mathcal{S}}) \quad \text{and} \quad \Psi_n(\mathcal{M}_n) \subset \bar{B}(\Psi(\mathcal{M}), r_n + \varepsilon_n^{\mathcal{R}}).$$

This idea will guide us in our attempt to deduce quantitative convergence rates in section 2.3.3. The major task will be to find a way to express what "not too shallow" actually means. This approach will depend crucially on some very detailed information on $\mathcal{F}$, $\mathcal{F}_n$—information that may be prohibitively hard to obtain in practice.

Therefore, we establish less restrictive conditions that allow us to deduce Kuratowksi convergence (or even Hausdorff convergence) from (2.11) *without giving a precise convergence rate*. This will be the focus of section 2.3.1, where Lemma 2.16 will be used.

For both approaches, we will have to transport variational information of $\mathcal{F}$ forward to $\mathcal{X}$ along Ψ. This is why we introduce the (variational) pushforward first.

2.2.2 Pushforward

Definition 2.17 Let $F \colon Y \to \;]-\infty, \infty]$ be a function and $\psi \colon Y \to X$ a mapping to a topological space X. With the convention $\inf(\emptyset) = \infty$, define the *(variational) pushforward of F along ψ* by

$$(\psi_{\#}F)(x) := \inf \{\, F(y) \mid y \in Y \colon \psi(y) = x \,\} = \inf_{y \in \psi^{-1}(x)} F(y).$$

Example 2.18 For injections, the pushforward reduces to the well-known and frequently used extension by infinity: Assume that $\Psi \colon C \hookrightarrow X$ and $\Psi_n \colon C_n \hookrightarrow X$ are injections. Then $\Psi_{\#}\mathcal{F}$ and $(\Psi_n)_{\#}\mathcal{F}_n$ are given by

$$(\Psi_{\#}\mathcal{F})(x) = \begin{cases} \mathcal{F}(x), & x \in C \\ \infty, & \text{else} \end{cases}, \qquad ((\Psi_n)_{\#}\mathcal{F}_n)(x) = \begin{cases} \mathcal{F}_n(x), & x \in C_n \\ \infty, & \text{else} \end{cases}.$$

This allows one to treat the optimization problems for $\mathcal{F}$ and $\mathcal{F}_n$ on a common space.

We list some elementary properties of the pushforward:

Lemma 2.19 *Let $F \colon Y \to \;]-\infty, \infty]$ be a function with $\inf(F) < \infty$, $\psi \colon Y \to X$ some mapping and $\varrho \in [0, \infty[$. Then one has $\inf(\psi_{\#}F) = \inf(F)$ and*

$$\psi(\arg\min^{\varrho}(F)) \subset \arg\min^{\varrho}(\psi_{\#}F) = \bigcap_{\sigma > \varrho} \psi(\arg\min^{\sigma}(F)).$$

Equality holds, e.g., if for each $x \in \arg\min^{\varrho}(\psi_{\#}F)$, the function $F|_{\psi^{-1}(x)}$ attains its infimum.

Proof. First, note that $(\psi_\# F)(\psi(y)) = \inf_{a \in \psi^{-1}(\psi(y))} F(a) \le F(y)$ holds for all $y \in Y$. This leads to $\inf(\psi_\# F) \le \inf(F)$ and

$$\psi(\arg \min^\varrho(F)) \subset \arg \min^\varrho(\psi_\# F) \quad \text{for all } \varrho \in [0, \infty]. \tag{2.13}$$

Because of $\inf(\psi_\# F) \le \inf(F) < \infty$, there exists a sequence $(x_n)_{n \in \mathbb{N}}$ with $(\psi_\# F)(x_n) < \infty$ for all $n \in \mathbb{N}$ and $\liminf_{n \to \infty}(\psi_\# F)(x_n) = \inf(\psi_\# F)$. For each $n \in \mathbb{N}$, x_n has to be in the image of ψ since $(\psi_\# F)(x_n)$ is finite. So, we may choose $y_n \in \psi^{-1}(x_n)$ with $F(y_n) \le (\psi_\# F)(x_n) + \frac{1}{n}$. This leads to

$$\inf(F) \le \liminf_{n \to \infty} F(y_n) \le \liminf_{n \to \infty} \left((\psi_\# F)(x_n) + \tfrac{1}{n} \right) = \inf(\psi_\# F),$$

thus $\inf(F) = \inf(\psi_\# F)$. The case $\inf(F) = -\infty$ is also included.

From now on, let $\varrho \in [0, \infty[$ be finite. We are going to show

$$\arg \min^\varrho(\psi_\# F) \subset \psi(\arg \min^\sigma(F)) \quad \text{for each } \sigma > \varrho. \tag{2.14}$$

Therefore, let $x \in \arg \min^\varrho(\psi_\# F)$. Since one has $(\psi_\# F)(x) \le \inf(\psi_\# F) + \varrho = \inf(F) + \varrho < \infty$, there is a minimizing sequence $(y_n)_{n \in \mathbb{N}}$ of $F|_{\psi^{-1}(x)}$, i.e., $y_n \in \psi^{-1}(x)$ and

$$\begin{aligned} F(y_n) &\le \inf_{a \in \psi^{-1}(x)} F(a) + \tfrac{1}{n} = (\psi_\# F)(x) + \tfrac{1}{n} \\ &\le \inf(\psi_\# F) + \varrho + \tfrac{1}{n} = \inf(F) + \varrho + \tfrac{1}{n}. \end{aligned}$$

For $n > \frac{1}{\sigma - \varrho}$, one has $y_n \in \arg \min^\sigma(F)$ and $x = \psi(y_n) \in \psi(\arg \min^\sigma(F))$. This shows $\arg \min^\varrho(\psi_\# F) \subset \psi(\arg \min^\sigma(F))$ for each $\sigma > \varrho$. Now, (2.13) and (2.14) together yield

$$\arg \min^\varrho(\psi_\# F) \subset \bigcap_{\sigma > \varrho} \psi(\arg \min^\sigma(F)) \subset \bigcap_{\sigma > \varrho} \arg \min^\sigma(\psi_\# F) = \arg \min^\varrho(\psi_\# F).$$

If for $x \in \arg \min^\varrho(\psi_\# F)$ the function $F|_{\psi^{-1}(x)}$ attains its infimum, say at $y \in Y$, one has $F(y) = \inf_{a \in \psi^{-1}(x)} F(a) = (\psi_\# F)(x) \le \varrho$, showing that $x \in \psi(\arg \min^\varrho(F))$. $\qquad \square$

Remark 2.20 As a curiosity with few practical value, we would like to mention that $\arg \min^\varrho(\psi_\# F) = \bigcap_{\sigma > \varrho} \psi(\arg \min^\sigma(F))$ also holds true for $\varrho = \infty$ if one interprets $\bigcap_{\sigma > \varrho}$ as an "intersection in the space X", i.e. if one uses the definition

$$\bigcap_{\sigma > \varrho} \psi(\arg \min^\sigma(F)) := \{ x \in X \mid \forall \sigma > \varrho \colon x \in \psi(\arg \min^\sigma(F)) \}.$$

Since the set $\{ \sigma \in [-\infty, \infty] \mid \sigma > \infty \}$ is empty, one has

$$\bigcap_{\sigma > \infty} \psi(\arg \min^\sigma(F)) = X = \arg \min^\infty(\psi_\# F).$$

Remark 2.21 Colloquially, the sufficient condition for equality in the preceding lemma can be restated as: Non-empty ψ-slices $\psi^{-1}(x)$, $x \in X$ are "small" enough to allow $F|_{\psi^{-1}(x)}$ to be minimizable.

For example, this sufficient condition is met if there is a topology on Y such that for each $\varrho \in [0, \infty[$ and each $x \in X$, the intersection $\arg\min^{\varrho}(F) \cap \psi^{-1}(x)$ is closed and countably compact.

The pushforward allows us to draw a connection to epigraph distances of functions. The following lemma may be seen as an epigraphical variant of Lemma 2.15.

Lemma 2.22 *Let $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ be valid sets and assume $\mathcal{S}_n(\mathcal{A}) \subset \mathcal{A}_n$, $\mathcal{R}_n(\mathcal{A}_n) \subset \mathcal{A}$, and $-\infty < \inf(\mathcal{F})$, $\inf(\mathcal{F}_n) < \infty$. Put $\mathcal{H} := (\Psi|_{\mathcal{A}})_{\#}(\mathcal{F}|_{\mathcal{A}})$, $\mathcal{H}_n := (\Psi_n|_{\mathcal{A}_n})_{\#}(\mathcal{F}_n|_{\mathcal{A}_n})$, $\mathcal{N}^{\varrho} := \arg\min^{\varrho}(\mathcal{H})$, and $\mathcal{N}_n^{\varrho} := \arg\min^{\varrho}(\mathcal{H}_n)$. Denote by ε_n the proximity error of $(\mathcal{S}_n, \mathcal{R}_n)$ on $(\mathcal{A}, \mathcal{A}_n)$ and by $\delta_n^{\mathcal{S}}$, $\delta_n^{\mathcal{R}}$, δ_n the consistency errors of $(\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n)$ on $(\mathcal{A}, \mathcal{A}_n)$. With $m := \inf(\mathcal{H})$, $m_n := \inf(\mathcal{H}_n)$, and $r_n := \max\{\varepsilon_n, \delta_n\}$, one has for $\infty > \sigma > \varrho \geq 0$:*

1. $\mathrm{epi}^{m+\varrho}(\mathcal{H}) \subset \bar{B}(\mathrm{epi}^{m+\sigma+\delta_n^{\mathcal{S}}}(\mathcal{H}_n), r_n)$,

2. $\mathrm{epi}^{m_n+\varrho}(\mathcal{H}_n) \subset \bar{B}(\mathrm{epi}^{m_n+\sigma+\delta_n^{\mathcal{R}}}(\mathcal{H}), r_n)$,

3. $\mathcal{N}^{\varrho} \subset \bar{B}(\mathcal{N}_n^{\sigma+2\delta_n}, r_n)$,

4. $\mathcal{N}_n^{\varrho} \subset \bar{B}(\mathcal{N}^{\sigma+2\delta_n}, r_n)$.

PROOF. *Claim 1:* Let $\infty > \sigma > \varrho \geq 0$. If $\mathrm{epi}^{m+\varrho}(\mathcal{H})$ is empty, we are done. Otherwise, let $(x, t) \in \mathrm{epi}^{m+\varrho}(\mathcal{H})$. Since $\mathcal{H}(x) \leq t \leq m + \varrho < \infty$ is finite, there is an $a \in \mathcal{A}$ with $\Psi(a) = x$ and $\mathcal{F}(a) \leq \mathcal{H}(x) + \sigma - \varrho$. Put $y := \Psi_n \circ \mathcal{S}_n(a) \in \Psi_n(\mathcal{A}_n)$. By consistency, one has

$$\mathcal{H}_n(y) \leq \mathcal{F}_n(\mathcal{S}_n(a)) \leq \mathcal{F}(a) + \delta_n^{\mathcal{S}} \leq t + \sigma - \varrho + \delta_n^{\mathcal{S}} \leq \sigma + \delta_n^{\mathcal{S}}, \tag{2.15}$$

thus $(y, t + \delta_n^{\mathcal{S}}) \in \mathrm{epi}^{m+\sigma+\delta_n^{\mathcal{S}}}(\mathcal{H}_n)$. Using the definition of the proximity error leads to

$$d_X(x, y) = d_X(\Psi(a), \Psi_n(\mathcal{S}_n(a))) \leq \varepsilon_n,$$
$$\mathrm{dist}_{X \times \mathbb{R}}((x, t), \mathrm{epi}^{m+\sigma+\delta_n}(\mathcal{H}_n)) \leq d_{X \times \mathbb{R}}((x, t), (y, t + \delta_n^{\mathcal{S}})) \leq \max\{\varepsilon_n, \delta_n\} = r_n,$$

which shows the first claim.

Claim 3: Because $\mathcal{A}$, $\mathcal{A}_n$ are valid, one has for the infimal values $m = \inf(\mathcal{H}) = \inf(\mathcal{F})$, $m_n = \inf(\mathcal{H}_n) = \inf(\mathcal{F}_n)$ and by Lemma 2.5 $|m - m_n| \leq \delta_n$. Denote by $\mathrm{pr}_X \colon X \times]-\infty, \infty[\to X$ the canonical projection. With the box metric on $X \times]-\infty, \infty[$,

one has $\text{pr}_X(\bar{B}(\mathcal{U}, r)) \subset \bar{B}(\text{pr}_X(\mathcal{U}), r)$ for any set $\mathcal{U} \subset X \times \,]\!-\!\infty, \infty[$. Combining this with Claim 1 and using Lemma 2.5 leads to

$$
\begin{aligned}
\mathcal{N}^{\varrho} = P(\text{epi}^{m+\sigma}(\mathcal{H})) &\subset P\big(\bar{B}(\text{epi}^{m+\sigma+\delta_n^S}(\mathcal{H}_n), r_n)\big) \\
&\subset \bar{B}\big(P(\text{epi}^{m+\sigma+\delta_n^S}(\mathcal{H}_n)), r_n\big) \\
&= \bar{B}\big(\mathcal{N}_n^{m-m_n+\sigma+\delta_n^S}, r_n\big) \subset \bar{B}(\mathcal{N}_n^{\sigma+2\delta_n}, r_n).
\end{aligned}
$$

The proofs of Claims 2 and 4 are analogous. $\qquad\square$

Remark 2.23 In general, the sets $\mathcal{N}^0$, $\mathcal{N}_n^0$ need not to coincide with $\Psi(\mathcal{A} \cap \mathcal{M})$ and $\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n)$. Moreover, handling both conditions $\mathcal{S}_n(\mathcal{A}) \subset \mathcal{A}_n$ and $\mathcal{R}_n(\mathcal{A}_n) \subset \mathcal{A}$ at once may be quite difficult in practice.

Remark 2.24 It is instructive to apply the preceding lemma in the setting of Example 2.18 together with a priori information $\emptyset \neq \mathcal{M} \subset \mathcal{A}$ and $\emptyset \neq \mathcal{M}_n \subset \mathcal{A}_n$. Then one has $\text{epi}^{\varrho}(\mathcal{H}) = \text{epi}_{\mathcal{A}}^{\varrho}(\mathcal{F})$ and $\text{epi}^{\varrho}(\mathcal{H}_n) = \text{epi}_{\mathcal{A}_n}^{\varrho}(\mathcal{F}_n)$ for $\varrho \in \mathbb{R}$. The first two results of the lemma can be simplified to

$$
\text{epi}_{\mathcal{A}}^{m+\varrho}(\mathcal{F}) \subset \bar{B}(\text{epi}_{\mathcal{A}_n}^{m+\sigma+\delta_n^S}(\mathcal{F}_n), r_n) \quad \text{and} \quad \text{epi}_{\mathcal{A}_n}^{m_n+\varrho}(\mathcal{F}_n) \subset \bar{B}(\text{epi}_{\mathcal{A}}^{m_n+\sigma+\delta_n^{\mathcal{R}}}(\mathcal{F}), r_n).
$$

For $\varrho = 0$, the third and fourth statements lead to

$$
\mathcal{M} \subset \bar{B}(\mathcal{M}_n^{\sigma+2\delta_n}, r_n), \quad \text{and} \quad \mathcal{M}_n \subset \bar{B}(\mathcal{M}^{\sigma+2\delta_n}, r_n) \quad \text{for all } \sigma > 0.
$$

2.3 Stability

In order to deduce set convergence of $\Psi_n(\mathcal{M}_n)$ to $\Psi(\mathcal{M})$ from Lemma 2.15 or Lemma 2.16, one requires a reasonable interplay between $\mathcal{F}$ and Ψ (and probably an analogous interplay between $\mathcal{F}_n$ and Ψ_n). We term the presence of such an interplay as *stability*. First, we give a rather weak, purely qualitative condition and point out its relation to the concept of lower semi-continuity. Afterwards, we will discuss other conditions that are more suitable for quantitative results.

2.3.1 Topological stability

Definition 2.25 Let $F: Y \rightarrow \,]\!-\!\infty, \infty]$ be a function, $\psi: Y \rightarrow X$ a mapping to a topological space X, $K \subset X$ a closed set. We call F *topologically stable along ψ over K*, if

$$
K \cap \psi(\arg\min(F)) = \bigcap_{\varrho>0} \overline{K \cap \psi(\arg\min^{\varrho}(F))}.
$$

The notion of topological stability is a generalization of lower semi-continuity in the context of test mappings:

Example 2.26 Let X be a topological space, $K \subset X$ a closed set, and $F \colon X \to {]}{-}\infty, \infty]$ lower semi-continuous on K, i.e., the lower level sets of $F|_K$ are closed in K (and thus in X because K is closed). Then F is topologically stable along id_X over K.

Example 2.27 Let X be a topological space, $Y \subset X$ a closed set, and $F \colon Y \to {]}{-}\infty, \infty]$ lower semi-continuous on Y. Denote by $\psi \colon Y \hookrightarrow X$ the inclusion mapping. Since $\psi_{\#}F$ is the extension by infinity (see Example 2.18), it is lower semi-continuous, thus topologically stable on X.

We arrive at the first main theorem of this chapter.

Theorem 2.28 (Kuratowksi convergence of minimizers)
Let $\mathcal{A} \subset C$ and $\mathcal{A}_n \subset C_n$ be valid sets and let $\mathcal{K} \subset X$ be a closed set such that $\Psi(\mathcal{R}_n(\mathcal{A}_n)) \subset \mathcal{K}$ holds for all sufficiently large n. Assume consistency and proximity, i.e., $\delta_n \overset{n\to\infty}{\longrightarrow} 0$ and $\varepsilon_n \overset{n\to\infty}{\longrightarrow} 0$, and topological stability of $\mathcal{F}$. Then one has

$$\underset{n\to\infty}{\mathrm{Ls}}\ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n) \subset \mathcal{K} \cap \Psi(\mathcal{M}).$$

If $\Psi(\mathcal{M}) \subset \Psi(\mathcal{A} \cap \mathcal{M}) \cap \mathcal{K}$ and $\mathcal{S}_n(\mathcal{A}) \subset \mathcal{A}_n$ hold for all sufficiently large n, then one has Kuratowski convergence

$$\Psi(\mathcal{M}) = \underset{n\to\infty}{\mathrm{Lt}}\ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}).$$

Proof. From the second statement of Lemma 2.15 with $\varrho = 0$, we have for sufficiently large $n \in \mathbb{N}$:

$$\underset{n\to\infty}{\mathrm{Ls}}\ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n) \subset \underset{n\to\infty}{\mathrm{Ls}}\ \bar{B}\big(\Psi(\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{\delta_n}), \varepsilon_n^{\mathcal{R}}\big)$$

$$\subset \underset{n\to\infty}{\mathrm{Ls}}\ \bar{B}(\mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n}), \varepsilon_n^{\mathcal{R}}).$$

Using Lemma 1.14 with $A_n := \mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n})$ and $r_n := \varepsilon_n^{\mathcal{R}}$, we obtain

$$\underset{n\to\infty}{\mathrm{Ls}}\ \bar{B}(\mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n}), \varepsilon_n^{\mathcal{R}}) = \underset{n\to\infty}{\mathrm{Ls}}\ \mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n}) = \bigcap_{n\in\mathbb{N}} \overline{\mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n})}.$$

Now, topological stability of $\mathcal{F}$ along Ψ over $\mathcal{K}$ leads to the first statement. In the same way, one shows

$$\underset{n\to\infty}{\mathrm{Ls}}\ \Psi(\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{2\delta_n}) \subset \underset{n\to\infty}{\mathrm{Ls}}\ \mathcal{K} \cap \Psi(\mathcal{M}^{2\delta_n}) \subset \mathcal{K} \cap \Psi(\mathcal{M}).$$

The condition $\mathcal{S}_n(\mathcal{A}) \subset \mathcal{A}_n$ allows us to use Lemma 2.16, leading to

$$\Psi(\mathcal{A} \cap \mathcal{M}) \subset \underset{n\to\infty}{\mathrm{Li}} \ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n})$$

$$\subset \underset{n\to\infty}{\mathrm{Ls}} \ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}) \subset \underset{n\to\infty}{\mathrm{Ls}} \ \Psi(\mathcal{R}_n(\mathcal{A}_n) \cap \mathcal{M}^{2\delta_n}) \subset \mathcal{K} \cap \Psi(\mathcal{M}).$$

If both $\Psi(\mathcal{A} \cap \mathcal{M})$ and $\mathcal{K}$ contain all images of minimizers $\Psi(\mathcal{M})$, the above chain of inclusions is closed. In particular, $\mathrm{Lt}_{n\to\infty} \ \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n})$ exists and coincides with $\Psi(\mathcal{M})$. $\qquad\square$

Remark 2.29 We point out that this theorem holds as well if $\delta_n \geq 0$ is any upper bound for the consistency errors with $\delta_n \xrightarrow{n\to\infty} 0$. Thus, this theorem does not only show the *existence* of some $\varrho_n \xrightarrow{n\to\infty} 0$ with $\arg\min(\mathcal{F}) = \mathrm{Lt}_{n\to\infty} \arg\min^{\varrho_n}(\mathcal{F}_n)$, as Theorem 1.22 does in the finite dimensional case, but it also tells us *how* to obtain such a sequence.

Finally, the equivalence of Kuratowski and Hausdorff convergence in compact metric spaces (Lemma 1.16) yields:

Corollary 2.30 *In addition to all the conditions in Theorem 2.28, assume that the sets $\mathcal{M}$ and $\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}$ are non-empty for all sufficiently large n and that the set $\mathcal{K} \subset \mathcal{X}$ is compact. Then one has Hausdorff convergence, i.e.,*

$$\lim_{n\to\infty} \mathrm{dist}_{\mathcal{X}} \left(\Psi(\mathcal{M}), \Psi_n(\mathcal{A}_n \cap \mathcal{M}_n^{\delta_n}) \right) = 0.$$

2.3.2 Examples

The notion of topological stability seems quite artificial. Therefore, some further examples are in order.

Lemma 2.31 *Let Y, X be topological spaces, $F \colon Y \to \ ]-\infty, \infty]$, $\psi \colon Y \to X$, and $K \subset X$ a closed set. Assume $\inf(F) < \infty$, $K \cap \psi(\arg\min^{\varrho}(F)) = K \cap \arg\min^{\varrho}(\psi_\# F)$ for all $\varrho \in [0, \infty[$, and that $\psi_\# F$ is lower semi-continuous on K.*
  Then F is topologically stable along ψ over K.

PROOF. Note that the set $\arg\min^{\varrho}(\psi_\# F)$ is closed for all $\varrho \in [0, \infty[$ because $\psi_\# F$ is lower semi-continuous. One has by Lemma 2.19

$$K \cap \psi(\arg\min(F)) = K \cap \arg\min(\psi_\# F) = \bigcap_{\varrho>0} K \cap \arg\min^{\varrho}(\psi_\# F)$$

$$= \bigcap_{\varrho>0} \overline{K \cap \arg\min^{\varrho}(\psi_\# F)} = \bigcap_{\varrho>0} \overline{K \cap \psi(\arg\min^{\varrho}(F))}. \qquad\square$$

Example 2.32 Let X be a Banach space, Y a reflexive Banach space, $\psi\colon Y \to X$ a compact linear map, and $F\colon Y \to \mathbb{R}$ a lower semi-continuous map with convex and bounded lower level sets. The lower level set $\arg\min^\varrho(F)$ is closed (Hahn-Banach theorem) and sequentially compact (Banach-Alaoglu theorem) in the weak topology of Y. Because ψ is a compact operator, it maps $\arg\min^\varrho(F)$ onto a compact, thus closed set in X. Moreover, for each $x \in X$, $\arg\min^\varrho(F) \cap \psi^{-1}(x)$ is closed and compact in the weak topology. By Lemma 2.19, we have $\psi(\arg\min^\varrho(F)) = \arg\min^\varrho(\psi_\# F)$, which is why $\psi_\# F$ is lower semi-continuous. Now, Lemma 2.31 implies that F is topologically stable along ψ over X.

A prominent example is given by the canonical embedding $\psi\colon W_0^{1,2}(\Omega) \to L^2(\Omega)$ and the Dirichlet energy $F(u) := \frac{1}{2} \int_\Omega |\mathrm{d}u|^2 \, \mathrm{d}\mu$, $u \in W_0^{1,2}(\Omega)$ for a bounded domain $\Omega \subset \mathbb{R}^d$.

Example 2.33 Let Y be a topological space, $\sim \,\subset Y \times Y$ an equivalence relation and $F\colon Y \to \mathbb{R}$ a continuous function that is invariant on equivalence classes, i.e., $y_1 \sim y_2$ implies $F(y_1) = F(y_2)$. Let $X := Y/\!\sim$ be the quotient space, equipped with the quotient topology and denote by $\psi\colon Y \to X := Y/\!\sim$ the quotient mapping. Then $\psi_\# F$ is continuous and F factors through the quotient:

$$
\begin{array}{ccc}
Y & \xrightarrow{\ F\ } & \mathbb{R} \;. \\
{\scriptstyle \psi}\big\downarrow & \nearrow {\scriptstyle \psi_\# F} & \\
X & &
\end{array}
$$

Again, Lemma 2.31 shows that F is topologically stable along ψ over X.

Example 2.34 In particular, the previous setting is powerful, if $\sim$ is induced by the orbits of an action $m\colon G \times Y \to Y$ of a non-compact topological group G that leaves F invariant. Even if $\arg\min(F)$ is non-compact, the *moduli space*

$$
\arg\min(\psi_\# F) = \arg\min(F)/G
$$

may be compact. Thus, F is stable along ψ over every compact set K containing $\psi(\arg\min(F))$ and there is a chance to put Corollary 2.30 to use.

For example, this setup may occur in gauge theory where the Lagrangian is invariant under the action of the group of gauge transformations. Other examples can be found in geometric optimization problems, where the objective function is invariant under the isometry group or under some other subgroup of the diffeomorphisms of a manifold—as will be the case in our treatment of minimal surfaces in Chapter 7.

2.3.3 Quantitative stability

Still, the result of Theorem 2.28 is a bit dissatisfactory, because it does not provide any answers to the following questions:

1. *Are* there any cluster points of minimizers of the $\mathcal{F}_n$ at all, i.e.,

$$\operatorname*{Ls}_{n\to\infty} (\mathcal{A}_n \cap \Psi_n(\mathcal{M}_n)) \neq \emptyset?$$

Note that together with Claim 1 of Theorem 2.28, a positive answer would provide a proof for the existence of minimizers of $\mathcal{F}$.

2. Do the discrete minimizers *converge* in the sense of Kuratowski, i.e.,

$$\operatorname*{Li}_{n\to\infty} (\mathcal{A}_n \cap \Psi_n(\mathcal{M}_n)) = \operatorname*{Ls}_{n\to\infty} (\mathcal{A}_n \cap \Psi_n(\mathcal{M}_n))?$$

3. Can *all* smooth minimizers be approximated by discrete *minimizers*, i.e.,

$$\Psi(\mathcal{M}) = \operatorname*{Lt}_{n\to\infty} (\mathcal{A}_n \cap \Psi_n(\mathcal{M}_n))?$$

4. If yes, do the discrete minimizers converge uniformly, i.e.,

$$\lim_{n\to\infty} \operatorname{dist} (\Psi(\mathcal{M}), \Psi_n(\mathcal{M}_n)) = 0?$$

5. If so, what is the uniform convergence rate?

These questions can be addressed by appropriate notions of conditioning and stability. The essential ideas are not new for they can be found, e.g., in [2], [34], [5], and [4]. We adapted them with small changes in order to make them applicable in the presence of test functionals.

Definition 2.35 We call the pair (F, ψ) *faithful*, if the pushforward along ψ does not introduce "new" minimizers in the sense that $\psi(\arg\min(F)) = \arg\min(\psi_\# F)$.

Definition 2.36 (Conditioning) Let $f\colon [0,\infty] \to [0,\infty]$ be a nondecreasing function with $f(0) = 0$, let (X, d) be a metric space, $H\colon X \to \,]-\infty,\infty]$ be a function, and $K \subset X$ a set. We say H is *f-conditioned on K* if $\arg\min(H)$ is non-empty and if one has

$$H(x) \geq \inf(H) + f(\operatorname{dist}(x, \arg\min(H))) \quad \text{for all } x \in K.$$

For a function $F\colon Y \to \,]-\infty,\infty]$ and a mapping $\psi\colon Y \to X$, we say *F is f-conditioned along ψ over K*, if (F, ψ) is faithful and if $\psi_\# F$ is f-conditioned on K.

Remark 2.37 Note that our notion of conditioning also includes *well-posedness* in the sense that $\emptyset \neq \arg\min(\psi_\# F) = \psi(\arg\min(F))$.

The usefulness of conditioning lies in the properties of the quantile function:

Lemma 2.38 *Let $f\colon [0,\infty] \to [0,\infty]$ be a nondecreasing function with $f(0) = 0$ and let $s, t \in [0,\infty]$. Then $f(t) \leq s$ implies $t \leq f^\dagger(s)$, where*

$$f^\dagger \colon [0,\infty] \to [0,\infty], \qquad s \mapsto \inf\{\, t \in [0,\infty] \mid s \leq f(t) \,\}$$

is the quantile function *of f.*

Corollary 2.39 *Let $H\colon X \to \,]{-}\infty,\infty]$ be f-conditioned on $K \subset X$. Then one has*

$$K \cap \arg\min{}^\varrho(H) \subset \bar{B}(\arg\min(H), f^\dagger(\varrho)).$$

Corollary 2.40 *Let $F\colon Y \to \,]{-}\infty,\infty]$ be f-conditioned along $\psi\colon Y \to X$ over $K \subset X$. Then one has $K \cap \psi(\arg\min{}^\varrho(F)) \subset \bar{B}(\psi(\arg\min(F)), f^\dagger(\varrho))$ for each $\varrho \in [0,\infty[$.*

Proof. By Lemma 2.19 and Corollary 2.39, we have

$$K \cap \psi(\arg\min{}^\varrho(F)) \subset K \cap \arg\min{}^\varrho(\psi_\# F) \subset \bar{B}(\arg\min(\psi_\# F), f^\dagger(\varrho)).$$

Using the faithfulness of (F, ψ) finishes the proof. $\qquad\qquad\square$

The second main theorem of this chapter is:

Theorem 2.41 *Let $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ be valid sets and let $\mathcal{K}$, $\mathcal{K}_n \subset X$ be sets with $\Psi \circ \mathcal{R}_n(\mathcal{A}_n) \subset \mathcal{K}$ and $\Psi_n \circ \mathcal{S}_n(\mathcal{A}) \subset \mathcal{K}_n$. Let $f, f_n\colon [0,\infty] \to [0,\infty]$ be nondecreasing functions with $f(0) = f_n(0) = 0$. Let δ_n be the total consistency error and $\varepsilon_n^{\mathcal{S}}$, $\varepsilon_n^{\mathcal{R}}$ be the sampling and reconstruction proximity errors, respectively. Then:*

1. *If $\mathcal{F}$ is f-conditioned along Ψ over $\mathcal{K}$, one has*

$$\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n) \subset \bar{B}(\Psi(\mathcal{M}), r_n^{\mathcal{R}}) \quad \text{with} \quad r_n^{\mathcal{R}} := \varepsilon_n^{\mathcal{R}} + f^\dagger(\delta_n).$$

2. *If $\mathcal{F}_n$ is f_n-conditioned along Ψ_n over $\mathcal{K}_n$, one has*

$$\Psi(\mathcal{A} \cap \mathcal{M}) \subset \bar{B}(\Psi_n(\mathcal{M}_n), r_n^{\mathcal{S}}) \quad \text{with} \quad r_n^{\mathcal{S}} := \varepsilon_n^{\mathcal{S}} + f_n^\dagger(\delta_n).$$

If both conditions are satisfied and if one has $\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n) \supset \Psi_n(\mathcal{M}_n)$ and $\Psi(\mathcal{A} \cap \mathcal{M}) \supset \Psi(\mathcal{M})$, then the Hausdorff distance between $\Psi_n(\mathcal{M}_n)$ and $\Psi(\mathcal{M})$ is bounded by $\max(r_n^{\mathcal{S}}, r_n^{\mathcal{R}})$.

PROOF. *Claim 1:* By Lemma 2.15 and $\Psi(\mathcal{R}_n(\mathcal{A}_n)) \subset \mathcal{K}$, we have

$$\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n) \subset \bar{B}(\mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n}), \varepsilon_n^{\mathcal{R}}).$$

The conditioning of $\mathcal{F}$ allows us to use Corollary 2.40 in order to obtain

$$\mathcal{K} \cap \Psi(\mathcal{M}^{\delta_n}) \subset \bar{B}(\Psi(\mathcal{M}), f^{\dagger}(\delta_n)).$$

Now, the triangle inequality leads to

$$\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n) \subset \bar{B}(\bar{B}(\Psi(\mathcal{M}), f^{\dagger}(\delta_n)), \varepsilon_n^{\mathcal{R}}) \subset \bar{B}(\Psi(\mathcal{M}), \varepsilon_n^{\mathcal{R}} + f^{\dagger}(\delta_n)).$$

Claim 2: Analogously, we obtain from Lemma 2.15, $\Psi_n(\mathcal{S}_n(\mathcal{A})) \subset \mathcal{K}_n$, Corollary 2.40, and the triangle inequality (in that order) that:

$$\begin{aligned}
\Psi(\mathcal{A} \cap \mathcal{M}) &\subset \bar{B}(\Psi_n \circ \mathcal{S}_n(\mathcal{A}) \cap \Psi_n(\mathcal{M}_n^{\delta_n}), \varepsilon_n^{\mathcal{S}}) \\
&\subset \bar{B}(\mathcal{K}_n \cap \Psi_n(\mathcal{M}_n^{\delta_n}), \varepsilon_n^{\mathcal{S}}) \\
&\subset \bar{B}(\bar{B}(\Psi_n(\mathcal{M}_n), f_n^{\dagger}(\delta_n)), \varepsilon_n^{\mathcal{S}}) \\
&\subset \bar{B}(\Psi_n(\mathcal{M}_n), f_n^{\dagger}(\delta_n) + \varepsilon_n^{\mathcal{S}}). \qquad \square
\end{aligned}$$

Remark 2.42 Note that the previous theorem can also be applied to the functions $\mathcal{F}_n$ and $\mathcal{F}_m$ for large $n, m \in \mathbb{N}$ as a technique to show that $\overline{(\Psi_n(\mathcal{A}_n \cap \mathcal{M}_n))}_{n \in \mathbb{N}}$ is a Cauchy sequence in $(\mathfrak{B}(\mathcal{X}), \mathrm{dist}_{\mathcal{X}})$. Successfully applied, this can be used together with Theorem 1.9 and Theorem 2.28 to show *existence* of minimizers of $\mathcal{F}$.

In light of Theorem 2.41, it is desirable to consider a particular class of conditioning functions:

Definition 2.43 A nondecreasing function $f \colon [0, \infty] \to [0, \infty]$ with $f(0) = 0$ is called a *modulus of stability* if its quantile function satisfies $f^{\dagger}(s) \xrightarrow{s \to 0} 0$.

Notice that even if each $\mathcal{F}_n$ is f_n-conditioned with modulus of stability f_n, this does not guarantee $f_n^{\dagger}(\delta_n) \xrightarrow{n \to \infty} 0$. This is why we introduce the following notion.

Definition 2.44 (Stability) Let $\mathcal{K}, \mathcal{K}_n \subset \mathcal{X}$ be sets. We call $(\mathcal{F}, \mathcal{F}_n)$ *f-stable along* (Ψ, Ψ_n) *over* $(\mathcal{K}, \mathcal{K}_n)$, if there is a modulus of stability f such that

1. $\mathcal{F}$ is f-conditioned along Ψ over $\mathcal{K}$.

2. For each $n \in \mathbb{N}$, $\mathcal{F}_n$ is f-conditioned along Ψ_n over $\mathcal{K}_n$.

We formulated our definitions such that—under mild additional assumptions—our results read much like the "fundamental theorem of numerical analysis". This becomes even clearer in the case $X = C$, $\Psi = \mathrm{id}_C$ and $\Psi_n = \mathcal{R}_n$:

Theorem 2.45 *Assume $\emptyset \neq \mathcal{M} \subset \mathcal{A}$, $\emptyset \neq \mathcal{M}_n \subset \mathcal{A}_n$, $\mathcal{R}_n(\mathcal{A}_n) \subset \mathcal{K}$, $(\mathcal{R}_n \circ \mathcal{S}_n)(\mathcal{A}) \subset \mathcal{K}_n$ and $\varepsilon_n := \sup_{a \in \mathcal{A}} d_C(\mathcal{R}_n \circ \mathcal{S}_n(a), a) \overset{n \to \infty}{\longrightarrow} 0$.*
Then consistency[4] and f-stability[5] imply convergence

$$\mathcal{R}_n(\mathcal{M}_n) \overset{n \to \infty}{\longrightarrow} \mathcal{M}$$

in Hausdorff distance with respect to d_C with convergence rate $\varepsilon_n + f^\dagger(\delta_n)$.

PROOF. Observe that $\varepsilon_n^{\mathcal{R}} = 0$ so ε_n is precisely the proximity error of $(\mathcal{S}_n, \mathcal{R}_n)$ with respect to $(\mathrm{id}_C, \mathcal{R}_n)$ on $(\mathcal{A}, \mathcal{A}_n)$. Hence, Theorem 2.41 is directly applicable. $\qquad \square$

2.3.4 Examples

In general, quantitative conditioning of a given function may be quite hard to show. However, there are some straightforward and well-known examples:

Example 2.46 Coercivity of bilinear forms is related to conditioning:

Let $(X, \langle \cdot, \cdot \rangle)$ be a Hilbert space, $A : X \to X$ a (not necessarily continuous) self-adjoint operator, and $v \in X$. Consider the function $F(x) := \langle A x, x \rangle + \langle v, x \rangle$ and observe that $\arg \min(F) \neq \emptyset$ holds if and only if A is positive semi-definite and v is contained in $\mathrm{im}(A) = \ker(A)^\perp$. Assume that $v \in \mathrm{im}(A)$ and that A is positive semi-definite with positive spectral gap $\lambda = \inf(\sigma(A) \setminus \{0\}) > 0$, where $\sigma(A)$ denotes the spectrum of A. In this case, one has $\arg \min(F) = A^\dagger v + \ker(A)$ with the Moore-Penrose pseudoinverse $A^\dagger$ of A. Hence F is f-conditioned with $f(t) = \lambda t^2$.

One calls the bilinear form $b : (x, y) \mapsto \langle Ax, y \rangle$ coercive with coercivity constant λ, if and only if A is positive definite with $\lambda = \inf(\sigma(A)) > 0$. The Lax-Milgram theorem implies that A is continuously invertible, thus $v \in \mathrm{im}(A)$ is readily fulfilled. This shows that the coercivity of b implies the f-conditioning of F.

---

[4]of $(\mathcal{F}, \mathcal{F}_n, \mathcal{S}_n, \mathcal{R}_n)$ on $(\mathcal{A}, \mathcal{A}_n)$
[5]of $(\mathcal{F}, \mathcal{F}_n)$ along $(\mathrm{id}_C, \mathcal{R}_n)$ over $(\mathcal{K}, \mathcal{K}_n)$

Example 2.47 Let (X, d) be a convex length space in the sense that every pair of points can be joined by a length minimizing geodesic. Let $F \colon X \to \mathbb{R}$ be a uniformly convex function in the sense

$$F(\gamma(t)) \leq (1 - t)F(\gamma(0)) + t\,F(\gamma(1)) - t(1 - t)f(d(\gamma(0), \gamma(1))), \quad t \in [0, 1],$$

for an arbitrary length minimizing geodesic $\gamma \colon [0, 1] \to X$. Here, $f \colon [0, \infty[\to [0, \infty[$ is the *modulus of convexity* of F, i.e., a nondecreasing function with $f(0) = 0$. Note that $\arg\min(F)$ is convex. Assume $\arg\min(F) \neq \emptyset$. For arbitrary $x \in X$, $z \in \arg\min(F)$, let $\gamma \colon [0, 1] \to X$ be a length minimizing geodesic from z to x. Then one has $F(\gamma(t)) \geq F(\gamma(0))$ for all $t \in [0, 1]$, hence

$$
\begin{aligned}
F(x) = F(\gamma(1))) &\geq \sup_{t \in [0,1]} \tfrac{1}{t}(F(\gamma(t)) - (1 - t)\,F(\gamma(0)) + t(1 - t)\,f(d(x, z))) \\
&\geq \sup_{t \in [0,1]} \tfrac{1}{t}(t\,F(\gamma(0)) + t(1 - t)\,f(d(x, z))) \\
&= \inf(F) + f(d(x, z)).
\end{aligned}
$$

Thus F is f-conditioned. See also [15, Chapter 1].

Both going over to a weaker metric and extension by infinity (see Example 2.18) do not essentially destroy conditioning as the following lemma shows:

Lemma 2.48 *Let (X, d_X), (Y, d_Y) be metric spaces, let $F \colon Y \to \mathbb{R}$ be f-conditioned over Y, and $\psi \colon Y \to X$ be Lipschitz continuous such that (F, ψ) is faithful. Then $\psi_\# F$ is h-conditioned along ψ over X with $h(t) = f(\mathrm{Lip}(\psi)^{-1}\,t)$.*

Proof. Note that $\arg\min(F)$ is necessarily non-empty, hence

$$\arg\min(\psi_\# F) = \psi(\arg\min(F))$$

is also non-empty by faithfulness. For $x \in X \setminus \psi(Y)$, one has $(\psi_\# F)(x) = \infty$. Hence, we have to check conditioning for points in $\psi(Y)$, only. For $x \in \psi(Y)$, one has

$$\mathrm{dist}_X(x, \psi(\arg\min(F))) \leq \mathrm{Lip}(\psi)\,\mathrm{dist}_Y(y, \arg\min(F)) \quad \text{for all } y \in \psi^{-1}(x).$$

Together with faithfulness, this leads to

$$
\begin{aligned}
(\psi_\# F)(x) = \inf\{\,F(y) \mid y \in \psi^{-1}(x)\,\} \\
&\geq \inf\{\,\inf(F) + f(\mathrm{dist}_Y(y, \arg\min(F))) \mid y \in \psi^{-1}(x)\,\} \\
&\geq \inf(F) + f(\,\mathrm{Lip}(\psi)^{-1}\,\mathrm{dist}_X(x, \psi(\arg\min(F)))) \\
&= \inf(F) + h(\,\mathrm{dist}_X(x, \arg\min(\psi_\# F))). \qquad \square
\end{aligned}
$$

2.4 Constraints

To conclude this chapter, we briefly demonstrate how sampling and reconstruction operators can be obtained in the presence of constraints.

Let $\tilde{C}$, $\tilde{C}_n$, $\mathcal{Y}$, $\mathcal{Y}_n$ be metric spaces and let the configuration spaces C, C_n be given by constraints of the form

$$C := \{\, x \in \tilde{C} \mid \Phi(x) \in \mathcal{B} \,\} = \Phi^{-1}(\mathcal{B}),$$
$$C_n := \{\, x \in \tilde{C}_n \mid \Phi_n(x) \in \mathcal{B}_n \,\} = \Phi_n^{-1}(\mathcal{B}_n),$$

where $\Phi \in C^0(\tilde{C}; \mathcal{Y})$, $\Phi_n \in C^0(\tilde{C}_n; \mathcal{Y}_n)$ are continuous mappings and $\mathcal{B} \subset \mathcal{Y}$, $\mathcal{B}_n \subset \mathcal{Y}_n$ are non-empty, closed subsets.

Moreover, assume that the domains of the mappings $\mathcal{F}$, Ψ and $\mathcal{F}_n$, Ψ_n include $\tilde{C}$ and $\tilde{C}_n$, respectively. In practice, it may be difficult to construct exact sampling and reconstruction operators

$$\mathcal{S}_n \colon \mathcal{A} \to C_n \quad \text{and} \quad \mathcal{R}_n \colon \mathcal{A}_n \to C$$

explicitly. But often, operators $\tilde{\mathcal{S}}_n \colon \mathcal{A} \to \tilde{C}_n$ and $\tilde{\mathcal{R}}_n \colon \mathcal{A}_n \to \tilde{C}$ can be constructed such that the following errors are "small":

$$\delta_n^{\tilde{\mathcal{S}}} := \sup_{a \in \mathcal{A}} (\mathcal{F}_n \circ \tilde{\mathcal{S}}_n(a) - \mathcal{F}(a))^+, \quad \delta_n^{\tilde{\mathcal{R}}} := \sup_{a \in \mathcal{A}_n} (\mathcal{F} \circ \tilde{\mathcal{R}}_n(a) - \mathcal{F}_n(a))^+,$$

$$\varepsilon_n^{\tilde{\mathcal{S}}} := \sup_{a \in \mathcal{A}} d_X(\Psi_n \circ \tilde{\mathcal{S}}_n(a), \Psi_n(a)), \quad \varepsilon_n^{\tilde{\mathcal{R}}} := \sup_{a \in \mathcal{A}_n} d_X(\Psi \circ \tilde{\mathcal{R}}_n(a), \Psi_n(a)),$$

$$\eta_n^{\tilde{\mathcal{S}}} := \sup_{a \in \mathcal{A}} \operatorname{dist}_{\mathcal{Y}_n}(\Phi_n \circ \tilde{\mathcal{S}}_n(a), \mathcal{B}), \quad \eta_n^{\tilde{\mathcal{R}}} := \sup_{a \in \mathcal{A}_n} \operatorname{dist}_{\mathcal{Y}}(\Phi \circ \tilde{\mathcal{R}}_n(a), \mathcal{B}).$$

As for the convergence analysis, we need merely the *existence* of sampling and reconstruction operators. Thus, an implicit argument for their existence in the vicinities of $\tilde{\mathcal{S}}_n$ and $\tilde{\mathcal{R}}_n$ suffices. Such implicit arguments can be provided by suitable "openness" conditions on Φ and Φ_n. We discuss the setting for reconstruction operators; the approach for sampling operators is analogous.

Definition 2.49 Let (X, d_X), (Y, d_Y) be metric spaces, $r > 0$, $\lambda > 0$ be constants, and $U \subset X$ a set. We call a mapping $f \colon X \to Y$ (ϑ, σ)-*open on* U if $\bar{B}(f(a), \varepsilon) \subset f(\bar{B}(a, \sigma \varepsilon))$ holds for all $a \in U$ and all $0 \le \varepsilon < \vartheta$.

For a more detailed treatment of openness and the related notion of metric regularity conditions we refer to [34, Section 9.G]. At the moment, we are satisfied with the following simple implications:

Lemma 2.50 *Assume $\eta_n^{\tilde{R}} \le \vartheta_0 < \vartheta$ and that Φ is (ϑ, σ)-open on $\tilde{R}_n(\mathcal{A}_n)$. Then for each $\tau \in]\sigma, \frac{\vartheta}{\vartheta_0}\sigma[$, there is a reconstruction operator $\mathcal{R}_n \colon \mathcal{A}_n \to C$ with*

$$\sup_{a \in \mathcal{A}_n} d_{\tilde{C}}(\tilde{R}_n(a), \mathcal{R}_n(a)) \le \tau \eta_n^{\tilde{R}}.$$

Proof. For $b \in \mathcal{A}_n$ let $a = \tilde{R}_n(b)$. Because one has $\mathrm{dist}_{\mathcal{Y}}(\Phi(a), \mathcal{B}) \le \eta_n^{\tilde{R}} \le \vartheta_0$ and $\frac{\tau}{\sigma}\eta_n^{\tilde{R}} < \vartheta$, there is a $y \in \mathcal{B}$ with $y \in \bar{B}(\Phi(a), \frac{\tau}{\sigma}\eta_n^{\tilde{R}})$. The (σ, ϑ)-openness of Φ implies

$$y \in \bar{B}(\Phi(a), \tfrac{\tau}{\sigma}\eta_n^{\tilde{R}}) \subset \Phi(\bar{B}(a, \tfrac{\tau}{\sigma}\sigma \eta_n^{\tilde{R}})) = \Phi(\bar{B}(a, \tau \eta_n^{\tilde{R}})).$$

Hence, there is an $x \in \bar{B}(a, \tau \eta_n^{\tilde{R}})$ with $\Phi(x) = y \in \mathcal{B}$, thus $x \in C$. Now define $\mathcal{R}_n(b) := x$ and observe $d_{\tilde{C}}(\tilde{R}_n(b), \mathcal{R}_n(b)) = d_{\tilde{C}}(a, x) \le \tau \eta_n^{\tilde{R}}$. $\qquad\square$

Corollary 2.51 *In addition to the condition of the previous lemma, assume that $\mathcal{F}$ and Ψ are Lipschitz continuous on $\bar{B}(\tilde{R}_n(\mathcal{A}_n), \tau \vartheta)$. Then one has the following estimates for the consistency and proximity errors $\delta_n^{\mathcal{R}}$, $\varepsilon_n^{\mathcal{R}}$:*

$$\delta_n^{\mathcal{R}} \le \delta_n^{\tilde{R}} + \mathrm{Lip}(\mathcal{F})\,\tau \eta_n^{\tilde{R}}, \quad and \quad \varepsilon^{\mathcal{R}} \le \varepsilon^{\tilde{R}} + \mathrm{Lip}(\Psi)\,\tau \eta_n^{\tilde{R}}$$

The situation that we actually have in mind is the following:

Lemma 2.52 *Let $R > 0$ and $\mathcal{U} \subset \tilde{C}$ be an open set with $\bar{B}(\tilde{R}_n(\mathcal{A}_n), R) \subset \mathcal{U}$. Moreover, assume that there are $C \ge 0$, $\Lambda > 0$ such that the following conditions are fulfilled:*

1. *$\Phi \in C^{1,1}(\mathcal{U}; \mathcal{Y})$ with $\mathrm{Lip}(D\Phi) \le C$.*

2. *For each $a \in \Psi_n(\mathcal{A}_n)$, there is a closed vector space $\mathcal{Z}_a \subset \tilde{C}$ such that $\tilde{C} = \ker(D\Phi(a)) \oplus \mathcal{Z}_a$ and $\|D\Phi(a)\,u\|_{\mathcal{Y}} \ge \Lambda \|u\|_{\tilde{C}}$ holds for each $u \in \mathcal{Z}$.*

Let $\vartheta = \frac{\Lambda}{2}\min\left(\frac{\Lambda}{C}, R\right)$ and $\sigma = \frac{2}{\Lambda}$. Then Φ is (ϑ, σ)-open on $\tilde{R}_n(\mathcal{A}_n)$.

Proof. Fix $b \in \mathcal{A}_n$ and put $a = \Psi_n(a)$. Let $X := \mathcal{Z}_a$ be the aforementioned closed vector space and let $\varepsilon \ge 0$ with $\varepsilon < \vartheta$ be arbitrary. With $r = \frac{2}{\Lambda}\varepsilon < \min\left(\frac{\Lambda}{C}, R\right)$ and $U := (\bar{B}(0, r) \cap \mathcal{Z}_a)$ define the mapping

$$f \colon U \to Y, \quad f(x) := \Phi(a + x).$$

For all $x \in U$ and $u \in X$, one has

$$a + x \in \mathcal{U} \quad and \quad Df(x)\,u = D\Phi(a + x)\,u,$$

thus the conditions $\|Df(0)\,u\| \ge \Lambda\|u\|$ and $\mathrm{Lip}(Df) \le C$ of the quantitative implicit function theorem (see Theorem 2.53 below) are fulfilled; one obtains

$$\bar{B}(\Phi(a), \varepsilon) = \bar{B}(f(0), \varepsilon) \subset f(\bar{B}(0, \tfrac{2}{\Lambda}\varepsilon)) = \Phi(\bar{B}(a, \tfrac{2}{\Lambda}\varepsilon)). \qquad\square$$

Theorem 2.53 (Quantitative inverse mapping theorem)

Let X, Y be Banach spaces, $U \subset X$ be an open, convex set, and $a \in U$ a point. Let $f \in C^{1,1}(U; Y)$ such that there are $C \geq 0$ and $\Lambda > 0$ with:

1. *For all $v \in X$: $\|Df(a)\,v\| \geq \Lambda\|v\|$, and*

2. *$\mathrm{Lip}(Df) \leq C$.*

Fix $r < \min(\frac{\Lambda}{C}, \mathrm{dist}(a, X \setminus U))$. Then the restriction $f|_{\bar{B}(a,r)}$ is a $C^{1,1}$-diffeomorphism onto its image and one has the inclusion

$$\bar{B}(f(a), \varepsilon) \subset f(\bar{B}(a, 2\Lambda^{-1}\varepsilon)) \quad \text{for each } 0 \leq \varepsilon \leq \tfrac{\Lambda}{2} r.$$

Proof. Let $\Lambda' := \Lambda - Cr > 0$. For $x \in \bar{B}(a, r)$ and $v \in X$, Lipschitz continuity of Df yields

$$\|Df(x)\,v\| \geq \|Df(a)\,v\| - \|(Df(x) - Df(a))\,v\|$$
$$\geq (\Lambda - C\|x - a\|)\|v\| \geq (\Lambda - Cr)\|v\| = \Lambda'\|v\|. \tag{2.16}$$

Let $0 < r' < \frac{\Lambda'}{C}$, $r' \leq r$ and $\Lambda'' := \Lambda' - Cr' > 0$. For $x, y \in \bar{B}(a, r')$, Taylor's theorem implies

$$\|f(y) - f(x)\| \geq \|Df(x)\,(y - x)\| - \tfrac{C}{2}\|y - x\|^2$$
$$\geq (\Lambda' - Cr')\|y - x\| = \Lambda''\|y - x\|.$$

This shows that $f|_{\bar{B}(a,r)}$ is injective. From

$$\|Df(a)^{-1}Df(x) - \mathrm{Id}\| \leq \|Df(a)^{-1}\|\,\|Df(x) - Df(a)\| \leq \tfrac{C}{\Lambda}\|x - a\| < 1$$

and $A^{-1} = -((\mathrm{Id} - A) - \mathrm{Id})^{-1} = -\sum_{k=0}^{\infty}(\mathrm{Id} - A)^k$ it follows that $Df(x)$ is invertible for all $x \in \bar{B}(a, r)$. The inverse function theorem [40, Corollary 4.37, p. 172] states that

$$h := f|_{\bar{B}(a,r)} : \bar{B}(a, r) \to f(\bar{B}(a, r))$$

is a $C^{1,1}$-diffeomorphism. Furthermore, we have for all $x \in \bar{B}(a, r)$

$$\|h(x) - h(a)\| = \|f(x) - f(a)\| \geq \|Df(a)\,(x - a)\| - \tfrac{C}{2}\|x - a\|^2$$
$$\geq (\Lambda - \tfrac{C}{2} r)\|x - a\| \geq \tfrac{\Lambda}{2}\|x - a\|.$$

Let $\varepsilon \in \,]0, \frac{\Lambda}{2} r]$. Since h is a diffeomorphism, the sets $W := h(B(a, 2\Lambda^{-1}\varepsilon))$ and $U := B(h(a), \varepsilon) \cap W$ are open. Moreover, we have for each $y \in \partial W$ that

$$\|y - h(a)\| = \|h(h^{-1}(y)) - h(a)\| \geq \tfrac{\Lambda}{2}\|h^{-1}(y) - a\| = \tfrac{\Lambda}{2}\varepsilon.$$

Hence $V := B(h(a), \varepsilon) \setminus W = B(h(a), \varepsilon) \setminus \overline{W}$ is also an open set. Thus, U, V is an open covering of the connected set $B(h(a), \varepsilon)$. Since U contains $h(a)$, V has to be empty which leads to $B(h(a), \varepsilon) \subset h(B(a, 2\Lambda^{-1}\varepsilon))$. Taking closures yields

$$
\begin{aligned}
\bar{B}(f(a), \varepsilon) = \bar{B}(h(a), \varepsilon) = \overline{B(h(a), \varepsilon)} &\subset \overline{h(B(a, 2\Lambda^{-1}\varepsilon))} \\
&= h\left(\overline{B(a, 2\Lambda^{-1}\varepsilon)}\right) = h(\bar{B}(a, 2\Lambda^{-1}\varepsilon)) = f(\bar{B}(a, 2\Lambda^{-1}\varepsilon)),
\end{aligned}
$$

where we used that $\overline{B(a, \delta)} = \bar{B}(a, \delta)$ holds in Banach spaces for all $\delta > 0$ and that h is a homeomorphism. $\qquad\square$
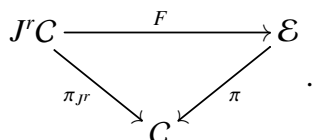
3 Parametric Operator Problems

To some extent, the theory of parametric optimization can be directly applied to operator problems. This shall be demonstrated in Section 3.1. We will be led to a notion of conditioning for operators which we explore by discussing several examples, including the classical notions of coercive linear operators and monotone operators (Section 3.2). In particular, we focus our attention on those operators that can be derived in a certain way from 1-forms (Section 3.2.2). We do this by having in mind that virtually all theories of modern physics are based on the principle of stationary action: The constituting equation of a physical system is $\mathrm{d}\mathcal{L} = F$, where $\mathcal{L}$ is the Lagrangian of the system and F encodes non-conservative forces such as friction. If the configuration space consists of fields (as in elasticity, electrodynamics, or fluid dynamics), the constituting equations are usually partial differential equations and generalized Ritz-Galerkin methods may be applied in order to approximate their solutions. Therefore, we demonstrate the consequences of our considerations for generalized Ritz-Galerkin methods in Section 3.3.

3.1 General Theory

From now on, let C be a topological space and $\pi\colon \mathcal{E} \to C$ a continuous, locally trivial fiber bundle of metric spaces with fiber metric $\mathrm{d}_\mathcal{E}$. Let $\alpha, \beta \in \Gamma(C; \mathcal{E})$ be given sections, i.e., $\pi \circ \alpha = \pi \circ \beta = \mathrm{id}_C$. We are going to consider α as the "operator" and β as the "right-hand side" of the *operator problem*:

$$\text{Find the } \textit{cut locus } \mathcal{N} := \{\, x \in C \mid \alpha(x) = \beta(x) \,\}. \tag{3.1}$$

Example 3.1 Differential equations can be formulated as such operator problems by considering C as the function space and putting $\alpha(x) = F \circ J^r(x)$, where J^r is the r-jet of x and F is a morphism of fiber bundles over C:

$$
\begin{array}{ccc}
J^r C & \xrightarrow{\quad F \quad} & \mathcal{E} \\
{\scriptstyle \pi_{J^r}} \searrow & & \swarrow {\scriptstyle \pi} \\
& C &
\end{array}
\;.
$$

Additionally, consider the family of operator problems

$$\text{Find the } \textit{cut locus } \mathcal{N}_n := \{\, x \in C_n \mid \alpha_n(x) = \beta_n(x) \,\}, \tag{3.2}$$

with topological spaces C_n, continuous, locally trivial fiber bundles $\pi_n \colon \mathcal{E}_n \to C_n$ with fiber metrics $d_{\mathcal{E}_n}$, and sections $\alpha_n, \beta_n \in \Gamma(C_n, \mathcal{E}_n)$.

As before, we also assume the existence of communication $S_n \colon \operatorname{dom}(S_n) \subset C \to C_n$ and $\mathcal{R}_n \colon \operatorname{dom}(\mathcal{R}_n) \subset C_n \to C$ and test mappings $\Psi \colon C \to X$, $\Psi_n \colon C_n \to X$ to some metric space (X, d_X). Approximation of $\mathcal{N}$ by $\mathcal{N}_n$ will be formulated by Hausdorff convergence of $\Psi_n(\mathcal{N}_n)$ to $\Psi(\mathcal{N})$ in the metric space (X, f_X).

It is insightful to define the real-valued functions

$$\begin{aligned} \mathcal{G} \colon \; & C \;\to\; \mathbb{R}, \quad \mathcal{G}(x) \;:=\; d_{\mathcal{E}}(\alpha(x), \beta(x)), \\ \mathcal{G}_n \colon \; & C_n \to \mathbb{R}, \quad \mathcal{G}_n(x) \;:=\; d_{\mathcal{E}_n}(\alpha_n(x), \beta_n(x)), \end{aligned}$$

since they allow us to use the theory of parametric optimization on $\mathcal{G}$ and $\mathcal{G}_n$ in order to obtain convergence results for $\mathcal{N}_n$. Note that if $\mathcal{N}$ and $\mathcal{N}_n$ are non-empty, they coincide precisely with the global minimizers of $\mathcal{G}$ and $\mathcal{G}_n$, respectively. This leads us to the following notions of consistency and conditioning.

Definition 3.2 (Consistency) Let $\mathcal{A} \subset \operatorname{dom}(S_n)$, $\mathcal{A}_n \subset \operatorname{dom}(\mathcal{R}_n)$ be non-empty sets. By definition, the consistency errors of (α_n, β_n) with respect to (α, β) coincide with the consistency errors of $\mathcal{G}_n$ with respect to $\mathcal{G}$. More explicitly, we define:

1. the *sampling consistency error*

$$\eta_n^{\mathcal{S}} := \sup_{a \in \mathcal{A}} \left(d_{\mathcal{E}_n}(\alpha_n \circ S_n(a), \beta_n \circ S_n(a)) - d_{\mathcal{E}}(\alpha(a), \beta(a)) \right)^+,$$

2. the *reconstruction consistency error*

$$\eta_n^{\mathcal{R}} := \sup_{a \in \mathcal{A}_n} \left(d_{\mathcal{E}}(\alpha \circ \mathcal{R}_n(a), \beta \circ \mathcal{R}_n(a)) - d_{\mathcal{E}_n}(\alpha_n(a), \beta_n(a)) \right)^+,$$

3. the *total consistency error*

$$\eta_n := \eta_n^{\mathcal{S}} + \eta_n^{\mathcal{R}}.$$

We say the sequence $((\alpha_n, \beta_n, S_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ is *consistent with respect to* (α, β) *on* $((\mathcal{A}, \mathcal{A}_n))_{n \in \mathbb{N}}$ if its consistency error η_n converges to 0 as $n \to \infty$.

Definition 3.3 Let $g \colon [0, \infty] \to [0, \infty]$ be a nondecreasing function with $g(0) = 0$ and let $\mathcal{K} \subset X$ be some set. We say (α, β) is *g-conditioned along* Ψ *over* $\mathcal{K}$ if $\mathcal{G}$ is.

The main theorem of this section is thus readily deduced from Theorem 2.41:

Theorem 3.4 *Assume that $\mathcal{A} \cap \mathcal{N} \neq \emptyset$ and $\mathcal{A}_n \cap \mathcal{N}_n \neq \emptyset$. Let $\mathcal{K}, \mathcal{K}_n \subset X$ be sets with $\Psi \circ \mathcal{R}_n(\mathcal{A}_n) \subset \mathcal{K}$ and $\Psi_n \circ \mathcal{S}_n(\mathcal{A}) \subset \mathcal{K}_n$ and let $g, g_n : [0, \infty] \rightarrow [0, \infty]$ be nondecreasing with $g(0) = g_n(0) = 0$. Denote by $\varepsilon_n^{\mathcal{S}}$, $\varepsilon_n^{\mathcal{R}}$ the proximity errors of (Ψ, Ψ_n) on $(\mathcal{A}, \mathcal{A}_n)$ and by $\eta_n^{\mathcal{S}}$, $\eta_n^{\mathcal{R}}$ the consistency errors of (α_n, β_n) with respect to (α, β) on $(\mathcal{A}, \mathcal{A}_n)$. Then:*

1. *If (α, β) is g-conditioned along Ψ over $\mathcal{K}$, one has*

$$\Psi_n(\mathcal{A}_n \cap \mathcal{N}_n) \subset \bar{B}(\Psi(\mathcal{N}), r_n^{\mathcal{R}}) \quad \text{with } r_n^{\mathcal{R}} := \varepsilon_n^{\mathcal{R}} + g_n^{\dagger}(\eta_n^{\mathcal{R}}).$$

2. *If (α_n, β_n) is g_n-conditioned along Ψ_n over $\mathcal{K}_n$, one has*

$$\Psi(\mathcal{A} \cap \mathcal{N}) \subset \bar{B}(\Psi_n(\mathcal{N}_n), r_n^{\mathcal{S}}) \quad \text{with } r_n^{\mathcal{S}} := \varepsilon_n^{\mathcal{S}} + g^{\dagger}(\eta_n^{\mathcal{S}}).$$

If both conditions are fulfilled and if one has $\Psi(\mathcal{N}) \supset \Psi(\mathcal{A} \cap \mathcal{N})$ and $\Psi_n(\mathcal{N}_n) \supset \Psi_n(\mathcal{A}_n \cap \mathcal{N}_n)$ then the Hausdorff distance between the sets $\Psi(\mathcal{N})$ and $\Psi(\mathcal{N}_n)$ is bounded by $\max(r_n^{\mathcal{S}}, r_n^{\mathcal{R}})$.

PROOF. The assumptions $\mathcal{A} \cap \mathcal{N} \neq \emptyset$ and $\mathcal{A}_n \cap \mathcal{N}_n \neq \emptyset$ guarantee that $\mathcal{A}$ and $\mathcal{A}_n$ are valid with respect to $\mathcal{G}$ and $\mathcal{G}_n$, respectively. We could apply Theorem 2.41 directly in order to obtain essentially the same result. But since we have the further information $\inf(\mathcal{G}) = \inf(\mathcal{G}_n) = 0$, we may improve that a little in the following way:

Let $z \in \mathcal{A} \cap \mathcal{N}$. One has

$$g_n(\text{dist}_{C_n}(\mathcal{S}_n(z), \mathcal{N}_n)) \leq \mathcal{G}_n \circ \mathcal{S}_n(z) = (\mathcal{G}_n \circ \mathcal{S}_n(z) - \mathcal{G}(z))^+ \leq \eta_n^{\mathcal{S}},$$

thus

$$\text{dist}_X(\Psi(z), \Psi_n(\mathcal{N}_n)) \leq \text{dist}_X(\Psi(z), \Psi_n(\mathcal{S}_n(z))) + \text{dist}_X(\Psi_n(\mathcal{S}_n(z)), \Psi_n(\mathcal{N}_n))$$
$$\leq \varepsilon_n^{\mathcal{S}} + g_n^{\dagger}(\eta_n^{\mathcal{S}}).$$

Analogously, one obtains $\text{dist}_X(\Psi_n(z), \Psi(\mathcal{N})) \leq \varepsilon_n^{\mathcal{R}} + g^{\dagger}(\eta_n^{\mathcal{R}})$ for all $z \in \mathcal{A}_n \cap \mathcal{N}_n$. $\square$

3.2 Examples

Theorem 3.4 would be of little value, if there were no reasonable examples of well-conditioned operators. In the following, we present a collection of examples together with some useful properties.

Example 3.5 If $\pi\colon \mathcal{E} \to C$ happens to be a locally trivial bundle of normed vector spaces, one may equivalently consider $\omega := \alpha - \beta$. Then $\mathcal{N}$ coincides with the set of zeroes of ω and one has $\mathcal{G}(x) = \|\omega(x)\|_{\mathcal{E}}$.

Example 3.6 When $\mathcal{E} \cong C \times Y$ is a trivial vector bundle with the normed vector space Y as fiber, one may write ω as $\omega(x) = (x, A(x))$, $x \in C$ with a mapping $A\colon C \to Y$. Then one has $\mathcal{G}(x) = \|A(x)\|_Y$ and $\mathcal{N} = \{\, x \in C \mid A(x) = 0 \,\}$ is the set of zeroes of A.

For simplicity, we assume from now on that $\mathcal{E}$ is a vector bundle.

3.2.1 Conditional cones

In view of Theorem 3.4, it is desirable that $\mathcal{S}_n(\mathcal{A})$ and $\mathcal{R}_n(\mathcal{A}_n)$ are contained in some sets $\mathcal{K}_n$, $\mathcal{K} \subset \mathcal{X}$ on which the operators $\omega_n = \alpha_n - \beta_n$ and $\omega = \alpha - \beta$ are well-conditioned. It is by no means necessary that these sets $\mathcal{K}_n$, $\mathcal{K}$ are neighborhoods of $\Psi_n(\mathcal{N}_n)$, $\Psi(\mathcal{N})$. In particular, this shifts the perspective from the question if ω is well-conditioned (everywhere) to the question *where* is it well-conditioned. We demonstrate in the following that $\mathcal{K}$ may be, e.g., a union of conditional cones.

Definition 3.7 Let X and Y be normed vector spaces and let $A \in L(X; Y)$ be a continuous linear operator. For $\lambda > 0$, define the (not necessarily convex) *conditional λ-cone of A* by

$$\mathrm{Cone}_\lambda(A) := \{\, u \in X \mid \|Au\| \geq \lambda\|u\| \,\}.$$

Lemma 3.8 *Let X, Y be Banach spaces, $U \subset X$ an open convex set, and $A \in C^{1,1}(U; Y)$. Let $\lambda > 0$ and $0 < \vartheta < 1$. Then for any two points $x, y \in U$ with*

$$y - x \in \mathrm{Cone}_\lambda(DA(x)) \quad \text{and} \quad \|y - x\| \leq \vartheta\tfrac{2\lambda}{\mathrm{Lip}(DA)},$$

the following estimate holds

$$(1 - \vartheta)\lambda\,\|y - x\| \leq \|A(y) - A(x)\|.$$

Proof. By Taylor's theorem, one has

$$\|A(y) - A(x) - DA(x)(y - x)\| \le \tfrac{1}{2} \operatorname{Lip}(DA) \|x - y\|^2,$$

leading to

$$\|A(y) - A(x)\| \ge \|DA(x)(y - x)\| - \tfrac{1}{2} \operatorname{Lip}(DA)\|x - y\|^2.$$

The cone condition and $\|y - x\| \le \vartheta \frac{2\lambda}{\operatorname{Lip}(DA)}$ imply

$$\|A(y) - A(x)\| \ge \lambda\|x - y\| - \tfrac{1}{2} \operatorname{Lip}(DA)\|x - y\|^2 \ge (1 - \vartheta)\lambda\|x - y\|. \qquad \square$$

Lemma 3.9 *Let X, Y be Banach spaces, $U \subset X$ an open convex set, and let $A \in C^{1,1}(U; Y)$ be a mapping with $C := \operatorname{Lip}(DA)$ and $\mathcal{N} := \{\, x \in U \mid A(x) = 0 \,\} \ne \emptyset$. Let $\lambda > 0$ and $0 < \vartheta < 1$. Then A is f-conditioned along id_U over the closure of the set*

$$W := \bigcup_{x \in \mathcal{N}} \bar{B}(x, \vartheta\tfrac{2\lambda}{C}) \cap (x + \operatorname{Cone}_\lambda(DA(x))),$$

with the function $f(t) = (1 - \vartheta)\lambda\, t$.

Proof. Let $z \in \overline{W}$. For an arbitrary $\varepsilon > 0$ there exists a $y \in B(z, \varepsilon) \cap W$. By the construction of the set W, there is some $x \in \mathcal{N}$ such that $\|y - x\| \le \vartheta\frac{2\lambda}{C}$ and $y - x \in \operatorname{Cone}_\lambda(DA(x))$. By Lemma 3.8, we obtain

$$
\begin{aligned}
\|A(z)\| = \|A(z) - A(x)\| \\
&\ge \|A(y) - A(x)\| - \|A(z) - A(y)\| \\
&\ge (1 - \vartheta)\lambda\|y - x\| - C\varepsilon \\
&\ge (1 - \vartheta)\lambda(\|z - x\| - \|z - y\|) - C\varepsilon \\
&\ge (1 - \vartheta)\lambda\|z - x\| - ((1 - \vartheta)\lambda + C)\varepsilon \\
&\ge (1 - \vartheta)\lambda \operatorname{dist}_X(z, \mathcal{N}) - ((1 - \vartheta)\lambda + C)\varepsilon.
\end{aligned}
$$

Note that we used the Lipschitz continuity of A with Lipschitz constant C in order to get from the second to the third line. Taking the supremum over all $\varepsilon > 0$ yields $\|A(z)\| \ge f(\operatorname{dist}_X(z, \mathcal{N}))$. $\qquad \square$

Example 3.10 Let X be a Hilbert space, $F\colon X \to \mathbb{R}$, $F(x) = |x|^4 - 2|x|^2$ be a double well potential. Consider $A\colon X \to X'$ given by the differential of F:

$$\langle A(x), u \rangle = \langle dF|_x, u \rangle = 4|x|^2\langle x, u \rangle - 4\langle x, u \rangle \quad \text{for all } u \in X.$$

The set $\mathcal{N}$ of zeroes of A consists of the origin and the unit sphere:

$$\mathcal{N} = \{0\} \cup \mathbb{S}, \quad \text{where} \quad \mathbb{S} := \{\, x \in X \mid |x| = 1 \,\}.$$

The differential $DA \colon X \to L(X; X')$ of A (and hence the Hessian of F) is given by

$$\langle DA(x)\, u, v \rangle = 4|x|^2 \langle u, v \rangle + 8\langle x, u \rangle \langle x, v \rangle - 4\langle u, v \rangle \quad \text{for all } u, v \in X.$$

Identifying $X' \cong X$ by the Riesz isomorphism, we may treat DA as a mapping $DA \colon X \to L(X; X)$. While $DA(0) = -4\, \mathrm{id}_X$ is a well-conditioned linear operator in the classical sense, we have $\ker(DA(x)) = x^\perp \neq \{0\}$ for all $x \in \mathbb{S} = \mathcal{N} \setminus \{0\}$. However, we have at least $x \in \mathrm{Cone}_8(DA(x))$ for each $x \in \mathbb{S}$. Observe that

$$\langle D^2 A(x)(u, v), w \rangle = 8\langle x, u \rangle \langle v, w \rangle + 8\langle x, v \rangle \langle w, u \rangle + 8\langle x, w \rangle \langle u, v \rangle$$

leads to $\mathrm{Lip}(DA(x)) \leq 24|x|$. Thus, for $r > 0$ and $U := B(0, 1 + r)$, we have $C(r) := \mathrm{Lip}(DA|_U) \leq 24(1 + r)$. Choose r such that it fulfills $r = \vartheta \frac{2\lambda}{C(r)} = \frac{\vartheta}{3(1+r)}$. Now, Lemma 3.9 tells us for $0 < \vartheta < 1$ and $\lambda = 4$ that A is f-conditioned on $\overline{W}$ with $f(t) = 4(1 - \vartheta)\, t$, where

$$W = \bar{B}(\mathcal{N}, r) = \bar{B}(0, r) \cup \bar{B}(0, 1 + r) \setminus \bar{B}(0, 1 - r).$$

Moreover, one readily verifies that $\|A(x)\| \geq 4(1 - \vartheta) \min(\|x\|, \|\|x\| - 1\|)$ holds for all $x \in X \setminus W$. Hence, A is globally f-conditioned.

For applications, it may be very helpful to know that conditional cones have certain continuity properties:

Lemma 3.11 *Let X, Y be Banach spaces, $\lambda > 0$, and A, $B \in L(X; Y)$ with $\|A - B\| < \lambda$. Then one has $\mathrm{Cone}_\lambda(A) \subset \mathrm{Cone}_{\lambda - \|A - B\|}(B)$.*

Proof. For $u \in \mathrm{Cone}_\lambda(A)$ one computes

$$\|Bu\| \geq \|Au\| - \|(B - A)u\| \geq \lambda \|u\| - \|B - A\| \|u\| \geq (\lambda - \|A - B\|) \|u\|,$$

which shows that $u \in \mathrm{Cone}_{\lambda - \|A - B\|}(B)$. $\qquad \square$

3.2.2 Operators induced by differential 1-forms

As explained in the introduction to this chapter, many important examples are covered by the following situation: C is a Banach manifold, $T'C$ is the continuous cotangent bundle, and $\omega \in \Gamma(C; T'C)$ is a (not necessarily differentiable) differential 1-form.

Example 3.12 Let $C = X$ be a Banach space. Then one may identify $T'C$ with $X \times X'$ and 1-forms on C coincide with operators $A \colon X \to X'$ via $\omega(x) = (x, A(x))$.

It is now tempting to apply Theorem 3.4 directly to ω by using the function $\mathcal{G}(x) := \|\omega(x)\|_{T'_x C}$. But this may be suboptimal for stability considerations (see Example 3.13 below). In the following, we outline a more promising approach.

Frequently, the linear functional $\omega(x) \colon T_x C \to \mathbb{R}$ is also continuous with respect to a weaker norm $\|\cdot\|_{\mathcal{E}_x}$ on $T_x C$. Let $(\mathcal{E}|_x, \|\cdot\|_{\mathcal{E}_x})$ be the completion of $T_x C$ with respect to $\|\cdot\|_{\mathcal{E}_x}$ and assume that the family $(\mathcal{E}_x)_{x \in C}$ gives rise to a locally trivial bundle $\pi_{\mathcal{E}} \colon \mathcal{E} \to C$ of Banach spaces. Let $\bar{\omega}(x) \in \mathcal{E}'_x$ be the unique continuous linear extension of $\omega(x) \in T'_x C$ and let $\iota_x \colon T_x C \hookrightarrow \mathcal{E}_x$ be the canonical inclusion. This leads to vector bundle mappings

$$
\begin{array}{ccc}
TC \xrightarrow{\quad \iota \quad} \mathcal{E} & & \mathcal{E}' \xrightarrow{\quad \iota' \quad} T'C \\
\pi_{TC} \searrow \quad \swarrow \pi_{\mathcal{E}} & \text{and} & \pi_{\mathcal{E}'} \searrow \quad \swarrow \pi_{T'C} \\
C & & C
\end{array}
$$

Note that one has $\omega = \iota' \circ \bar{\omega}$ and that the image of ι_x in $\mathcal{E}_x$ is dense, thus ι'_x is injective. Hence, the zeroes of ω and $\bar{\omega}$ coincide:

$$
\mathcal{N} = \{ x \in C \mid \omega(x) = 0 \} = \{ x \in C \mid \bar{\omega}(x) = 0 \} .
$$

Define the function $\bar{\mathcal{G}} \colon C \to \mathbb{R}$, $\bar{\mathcal{G}}(x) := \|\bar{\omega}(x)\|_{\mathcal{E}'_x}$. If $x \mapsto \|\iota_x\|$ is uniformly bounded by some $C \geq 0$, one has

$$
\begin{aligned}
\bar{\mathcal{G}}(x) = \|\bar{\omega}(x)\|_{\mathcal{E}'_x} &= \sup_{\substack{u \in \mathcal{E}_x \\ u \neq 0}} \frac{|\langle \bar{\omega}(x), u \rangle|}{\|u\|_{\mathcal{E}_x}} = \sup_{\substack{v \in T_x C \\ v \neq 0}} \frac{|\langle \bar{\omega}(x), \iota_x v \rangle|}{\|\iota_x v\|_{\mathcal{E}_x}} \\
&\geq \sup_{\substack{v \in T_x C \\ v \neq 0}} \frac{|\langle \iota'_x \bar{\omega}(x), v \rangle|}{\|\iota_x\| \, \|v\|_{\mathcal{T}_x C}} \geq C^{-1} \mathcal{G}(x).
\end{aligned}
$$

This tells us that—up to a constant—$\bar{\mathcal{G}}$ is never worse-conditioned than $\mathcal{G}$. Even more, $\bar{\mathcal{G}}$ may be much better conditioned than $\mathcal{G}$, as the following example shall illustrate:

Example 3.13 Let (Σ, g) be a compact, connected, smooth Riemannian manifold with non-empty boundary. For $p \in [2, \infty]$ and its Hölder conjugate $q \in [1, 2]$, put $X = W^{1,p}_{0,g}(\Sigma; \mathbb{R}^m)$ and $Y = W^{1,q}_{0,g}(\Sigma; \mathbb{R}^m)$. The canonical inclusion $J \colon X \hookrightarrow Y$ has dense image in Y, thus J' is injective.

Consider the Laplacian $A\colon X \to X'$ given by $\langle A(f), u \rangle = \int_\Sigma \langle \mathrm{d}f, \mathrm{d}u \rangle_g \, \mathrm{vol}_g$. As it turns out, $A(f)$ is continuously extendable onto Y, producing an element $\bar{A}(f) \in Y'$. We may write $\bar{A}\colon X \to Y'$ with $A = J' \circ \bar{A}$. Let $\eta \in Y'$ be arbitrary and define the operators $\omega(f) = (f, A(f) - \eta)$ and $\bar{\omega}(f) = (f, \bar{A}(f) - \eta)$. Let $f_0 \in \mathcal{N}$ be a solution to $A(f_0) = \eta$ and let $f \in X$ be arbitrary. Together with Lemma 1.2, one obtains

$$\|\bar{\omega}(f)\|_{Y'} = \|\bar{\omega}(f) - \bar{\omega}(f_0)\|_{Y'} = \|\bar{A}(f - f_0)\|$$
$$\geq \tfrac{1}{(1+C_p)(1+C_q)} \|f - f_0\|_{W^{1,p}_{0,g}} \geq \tfrac{1}{(1+C_p)(1+C_q)} \, \mathrm{dist}_X(f, \mathcal{N}).$$

As a consequence, the operator $\bar{\omega}$ is g-conditioned (along id_X over X) with the linear modulus of stability $g(t) = \frac{t}{(1+C_p)(1+C_q)}$.

Note that when $p > 2$, the operator $\omega\colon X \to T'X$, $\omega(f) := (f, A(f) - \eta)$ *cannot* be g-conditioned with a *linear* modulus of stability: Otherwise, $A\colon X \to X'$ would be an isomorphism of Banach spaces and the continuous *symmetric* bilinear form $b\colon X \times X \to \mathbb{R}$ defined by $b(u, v) := \langle A(u), Jv \rangle$, $u, v \in X$ would be coercive, rendering $(W^{1,p}_{0,g}(\Sigma; \mathbb{R}^m), b)$ into a Hilbert space.

Remark 3.14 The reasoning of the preceding theorem shows that for a smooth function $F\colon X \to \mathbb{R}$ on a Banach space X, the Hessian of F, interpreted as a linear operator $A|_x := \mathrm{Hess}(F)|_x\colon X \to X'$ can only be invertible if X is a Hilbert space. Hence, in general, the Newton method is not at disposal. However, if $\mathrm{d}F$ can be interpreted as a mapping $\mathrm{d}F\colon X \to Y'$ with some Banach space $Y \supset X$, there is a chance that $\bar{A}|_x\colon X \to Y'$ is invertible and one obtains a Newton-like vector field $-\bar{A}^{-1}|_x(\mathrm{d}F|_x)$. We will use the same idea in Section 7.7.2 in order to introduce a gradient-like vector field.

3.2.3 Monotone operators

An important class of operators with conditioning properties is given by monotone operators:

Definition 3.15 Let X, Y be Banach spaces, $J\colon X \hookrightarrow Y$ be an injective, continuous linear map with dense image, and $A\colon X \to X'$ be an operator that factors through J', i.e., $A = J' \circ \bar{A}$ with some operator $\bar{A}\colon X \to Y'$. Moreover, let $\psi\colon X \to Z$ be a mapping to the metric space (Z, d_Z), $K \subset Z$ some set and let $g\colon [0, \infty] \to [0, \infty]$ be a nondecreasing function with $g(0) = 0$.

We say A is *g-monotone along (ψ, J) over K* if

$$\langle A(x_2) - A(x_1), x_2 - x_1 \rangle \geq g(d_Z(\psi(x_2), \psi(x_1))) \cdot \|Jx_2 - Jx_1\|_Y$$

holds for all $x_1, x_2 \in \psi^{-1}(K)$.

Example 3.16 When $K = Z = Y = X$, $J = \psi = \mathrm{id}_X$ and $g(t) = c\,t$ is a linear function, the notion of g-monotonicity along (J, ψ) over K reduces to the usual notion of *strong monotonicity*:

$$\langle A(x_2) - A(x_1), x_2 - x_1 \rangle \geq c\|x_2 - x_1\|_X^2 \quad \text{for all } x_1, x_2 \in X.$$

Lemma 3.17 *Let* $A\colon X \to X'$ *be* g-*monotone along* (ψ, J) *over* K. *Let* $\bar{G}\colon X \to \mathbb{R}$ *be given by* $\bar{G}(x) := \|\bar{A}(x)\|_{Y'}$.

Assume that $\mathcal{N} := \{\, x \in X \mid A(x) = 0 \,\}$ *is non-empty and that* $(\bar{G}, \psi)$ *is faithful. Then* $\bar{A}$ *is* g-*conditioned along* ψ *over* K.

Proof. Observe that injectivity of J' implies

$$\{\, x \in X \mid \bar{A}(x) = 0 \,\} = \{\, x \in X \mid A(x) = 0 \,\} = \mathcal{N} \neq \emptyset.$$

One immediately deduces for $x_0 \in \mathcal{N}$ and arbitrary $x \in \psi^{-1}(K)$:

$$\begin{aligned}
\|\bar{A}(x)\|_{Y'}\|Jx - Jx_0\|_Y &= \|\bar{A}(x) - \bar{A}(x_0)\|_{Y'}\|Jx - Jx_0\|_Y \\
&\geq \langle \bar{A}(x) - \bar{A}(x_0), Jx - Jx_0 \rangle \\
&= \langle A(x) - A(x_0), x - x_0 \rangle \\
&\geq g(d_Z(\psi(x), \psi(x_0))) \cdot \|Jx - Jx_0\|_Y.
\end{aligned}$$

Division by $\|Jx - Jx_0\|_Y$ and taking the infimum over all $x_0 \in \mathcal{N} = \arg\min(\bar{G})$ yields

$$\bar{G}(x) = \|\bar{A}(x)\|_{Y'} \geq g(\mathrm{dist}_Z(\psi(x), \psi(\arg\min(\bar{G})))).$$

Because of faithfulness, one has $\arg\min(\psi_\#\bar{G}) = \psi(\arg\min(\bar{G}))$. Let $z \in Z$. Taking the infimum over all $x \in \psi^{-1}(z)$, one obtains

$$(\psi_\#\bar{G})(z) \geq g(\mathrm{dist}(z, \arg\min(\psi_\#\bar{G}))). \qquad \square$$

3.2.4 Symmetric conditional cones

Let X, Y be normed spaces and let $J \in L(X; Y)$. For $A \in L(X; Y')$ and $\lambda > 0$ define the *symmetric conditional* λ-*cone*

$$\mathrm{SymCone}_\lambda(A, J) := \{\, u \in X \mid |\langle Au, Ju \rangle| \geq \lambda \, \|u\|_X \|Ju\|_Y \,\}.$$

In particular, we have the Hessian $A = \mathrm{Hess}(F)|_x\colon X \to X'$ (and its induced operator $\bar{A}\colon X \to Y'$) of a twice differentiable function $F\colon X \to \mathbb{R}$ in mind (see also Remark 3.14).

Lemma 3.18 *Let X, Y be normed spaces and let $J \in L(X; Y)$ be injective. Then one has* $\mathrm{SymCone}_\lambda(A, J) \subset \mathrm{Cone}_\lambda(A)$

Proof. Let $u \in \mathrm{SymCone}_\lambda(A, J)$, $u \neq 0$.

$$\|Au\|_{Y'} = \sup_{\substack{w \in Y \\ w \neq 0}} \frac{|\langle Au, w \rangle|}{\|w\|_Y} \geq \sup_{\substack{v \in X \\ Jv \neq 0}} \frac{|\langle Au, Jv \rangle|}{\|Jv\|_Y} \geq \frac{|\langle Au, Ju \rangle|}{\|Ju\|_Y} \geq \lambda \|u\|_X. \qquad \square$$

The reason for introducing $\mathrm{SymCone}_\lambda(A, J)$ is that $\mathrm{Cone}_\lambda(A)$ for arbitrary linear operators does not behave well under restriction to linear subspaces:

Example 3.19 Let the linear operator $A \in L(\mathbb{R}^2; \mathbb{R}^2)$ given by counter-clockwise rotation about the origin. Observe $\mathrm{Cone}_\lambda(A) = \mathbb{R}^2$ for all $0 < \lambda \leq 1$. Let $u \in \mathbb{R}^2 \setminus \{0\}$ be an arbitrary vector and let $I \colon \mathbb{R}\,u \hookrightarrow \mathbb{R}^2$ be the canonical inclusion. Then $I'AI = 0$ and thus $\mathrm{Cone}_\lambda(I'AI) = \{0\}$ for all $\lambda > 0$.

In contrast, we have:

Lemma 3.20 *Let X, Y, Z be normed spaces and $A \in L(X; Y')$, $B \in L(Z; X)$, $J \in L(X; Y)$ continuous linear operators. With $I = JB$, one has for $\lambda, \mu > 0$*

$$\mathrm{Cone}_\mu(B) \cap B^{-1}(\mathrm{SymCone}_{\lambda/\mu}(A, J)) \subset \mathrm{SymCone}_\lambda(AB, I),$$
$$\mathrm{SymCone}_\lambda(AB, I) \subset B^{-1}(\mathrm{SymCone}_{\lambda/\|B\|}(A, J)).$$

Proof. For $u \in Z$, one has $|\langle ABu, Iu \rangle| = |\langle ABu, JBu \rangle|$. Moreover, note that we only have to consider the case $B \neq 0$.

On the one hand, if $u \in \mathrm{Cone}_\mu(B) \cap B^{-1}\mathrm{SymCone}_{\lambda/\mu}(A, J)$, one has

$$|\langle ABu, Iu \rangle| = |\langle ABu, JBu \rangle| \geq \tfrac{\lambda}{\mu} \|Bu\|_X \|JBu\|_Y \geq \lambda \|u\|_Z \|Iu\|_Y,$$

hence $u \in \mathrm{SymCone}_\lambda(AB, I)$.

On the other hand, let $u \in \mathrm{SymCone}_\lambda(AB, I)$. Then $Bu \in \mathrm{SymCone}_{\lambda/\|B\|}(A, J)$ follows from the estimate

$$|\langle ABu, JBu \rangle| = |\langle ABu, Iu \rangle| \geq \lambda \|u\|_Z \|Iu\|_Y \geq \tfrac{\lambda}{\|B\|} \|Bu\|_X \|JBu\|_Y. \qquad \square$$

Corollary 3.21 *Let X, Y be normed spaces, $J \in L(X; Y)$, $X_n \subset X$ a linear subspace, $Y_n := \overline{J(X_n)}$, and $J_n = J|_{X_n}$. Let $A_n \in L(X_n; Y_n')$ be the* Ritz-Galerkin discretization *of A given by*

$$\langle A_n u, J_n v \rangle := \langle Au, Jv \rangle \quad \text{for all } u, v \in X_n.$$

Then one has for each $\lambda > 0$:

$$\mathrm{SymCone}_\lambda(A_n, J_n) = X_n \cap \mathrm{SymCone}_\lambda(A, J).$$

PROOF. Apply the previous lemma with $Z = X_n$ to the embedding $B\colon X_n \hookrightarrow X$ and observe that B is isometric. $\qquad\Box$

Moreover, symmetric conditional cones have essentially the same continuity properties as general conditional cones:

Lemma 3.22 *Let X, Y be Banach spaces, $J \in L(X; Y)$, $\lambda > 0$, and A, $B \in L(X; Y')$ with $\|A - B\| \leq \lambda$. Then one has $\mathrm{SymCone}_\lambda(A, J) \subset \mathrm{SymCone}_{\lambda - \|A-B\|}(B, J)$.*

PROOF. Let $u \in \mathrm{SymCone}_\lambda(A, J)$. Then one has

$$|\langle Bu, Ju\rangle| \geq |\langle Au, Ju\rangle| - |\langle (B-A)u, Ju\rangle|$$
$$\geq \lambda \|u\|_X \|Ju\|_Y - \|B - A\| \|u\|_X \|Ju\|_Y \geq (\lambda - \|B - A\|) \|u\|_X \|Ju\|_Y,$$

showing that $u \in \mathrm{SymCone}_{\lambda - \|A-B\|}(B, J)$. $\qquad\Box$

3.3 Application to Generalized Ritz-Galerkin Methods

Besides demonstrating the applicability of the presented theory, this section has a second motivation: Ritz-Galerkin methods formed the starting point of our considerations. Cea's and Strang's lemmata were prototypical results to aim for. Moreover, Strang's second lemma was the major motivation for introducing test mappings and the notions of consistency and proximity errors. We would like to emphasize that in this very situation, it really paid off to give symbols to inclusion mappings which are often treated rather stepmotherly.

Within this section, let $C = X = X$ be a separable Banach space and let $A\colon X \to X'$ be an operator with non-empty set of zeroes $\mathcal{N} = \{\, x \in X \mid A(x) = 0 \,\}$.

3.3.1 Conforming Ritz-Galerkin method

Definition 3.23 A *conforming Ritz-Galerkin scheme* is a sequence $(X_n)_{n\in\mathbb{N}}$ of finite-dimensional subspaces of X with $\mathrm{Ls}_{n\to\infty} X_n = \bigcap_{n\in\mathbb{N}} \overline{\bigcup_{k\geq n} X_n} = X$. Let $\mathcal{R}_n\colon X_n \hookrightarrow X$ be the canonical inclusion. The operator

$$A_n\colon X_n \to X_n', \quad \langle A_n(x), u\rangle := \langle (A \circ \mathcal{R}_n)(x), \mathcal{R}_n(u)\rangle \quad x, u \in X_n$$

is the *Ritz-Galerkin discretization* of the operator A and one refers to

$$\mathcal{N}_n := \{\, x \in X_n \mid A_n(x) = 0 \,\}$$

as the set of *discrete solutions*.

Let $(X_n)_{n \in \mathbb{N}}$ be a conforming Ritz-Galerkin scheme and let A_n be the Ritz-Galerkin discretization of $A : X \to X'$. Choose a continuous, linear and surjective mapping $S_n : X \to X_n$ such that $\mathcal{R}_n \circ S_n$ is a projector. Moreover let $\Psi = \mathrm{id}_X$ and $\Psi_n = \mathcal{R}_n$.

When $\mathcal{N}$ is non-empty in the above setting, the sampling proximity error over $\mathcal{N}$ is the classical approximation error:

$$\varepsilon_n^{\mathcal{S}} = \sup_{a \in \mathcal{N}} \|(\Psi_n \circ S_n)(a) - \Psi(a)\|_X = \sup_{a \in \mathcal{N}} \|(\mathcal{R}_n \circ S_n)(a) - a\|_X$$
$$= \sup_{a \in \mathcal{N}} \inf_{x \in X_n} \|x - a\|_X = \sup_{a \in \mathcal{N}} \mathrm{dist}(a, X_n).$$

Assuming that existence has already been shown, Theorem 3.4 leads to the following generalization of Cea's lemma (see e.g., [6, Chapter II, Lemma 4.2]).

Lemma 3.24
Assume that $\mathcal{N}$ and $\mathcal{N}_n$ are non-empty, that A_n is g-conditioned, and that A is uniformly continuous with modulus of continuity h, i.e.,

$$\|A_n(x)\| \geq g(\mathrm{dist}_X(x, \mathcal{N})) \quad \text{for all } x \in X_n,$$
$$\|A(x) - A(y)\| \leq h(\|x - y\|) \qquad \text{for all } x, y \in X.$$

Then one has $\mathcal{N} \subset \bar{B}(\mathcal{N}_n, \varepsilon_n^{\mathcal{S}} + g^\dagger(h(\varepsilon_n^{\mathcal{S}})))$, where $\varepsilon_n^{\mathcal{S}}$ denotes the sampling proximity error on $\mathcal{N}$.

PROOF. The sampling consistency error on $\mathcal{N}$ can be estimated by

$$\eta_n^{\mathcal{S}} = \sup_{a \in \mathcal{N}} (\|A_n \circ S_n(a)\| - \|A(a)\|)^+$$
$$= \sup_{a \in \mathcal{N}} \|\mathcal{R}_n{}' \circ A \circ \mathcal{R}_n \circ S_n(a) - \mathcal{R}_n{}' \circ A(a)\|$$
$$\leq \sup_{a \in \mathcal{N}} \|\mathcal{R}_n{}'\| \, \|A \circ \mathcal{R}_n \circ S_n(a) - A(a)\|$$
$$\leq \sup_{a \in \mathcal{N}} h(\|\mathcal{R}_n \circ S_n(a) - a\|) = h(\varepsilon_n^{\mathcal{S}}).$$

Thus, the stated result follows from the second statement of Theorem 3.4. □

Remark 3.25 Assume that $\mathcal{N} = \{ x \}$ and $\mathcal{N}_n = \{ x_n \}$ both consist of precisely one element.[1] Then Cea's lemma in the above form gives the convergence rate

$$\|x - x_n\| \leq \varepsilon_n^{\mathcal{S}} + g^\dagger(h(\varepsilon_n^{\mathcal{S}})).$$

---

[1]Note that this is the case, e.g., with strictly monotone operators.

If A is a strongly monotone operator then $g(t) = c\, t$ can be chosen with some $c > 0$, (see Example 3.16 and Lemma 3.17). If A is additionally Lipschitz continuous (i.e., $h(t) = C\, t$ with some $C \geq 0$), one obtains a quasi-optimal convergence rate as in Cea's lemma:

$$\|x - x_n\| \leq \left(1 + \frac{C}{c}\right) \varepsilon_n^{\mathcal{S}}.$$

Remark 3.26 If solutions are not unique, one may additionally use the first statement of Theorem 3.4. Note that the reconstruction error $\varepsilon_n^{\mathcal{R}}$ on X_n vanishes in this setting:

$$\varepsilon_n^{\mathcal{R}} = \sup_{x \in X_n} \|\Psi \circ \mathcal{R}_n(x) - \Psi_n(x)\|_X = \sup_{x \in X_n} \|\mathcal{R}_n(x) - \mathcal{R}_n(x)\|_X = 0.$$

Hence, the remaining ingredients would be estimates on the reconstruction consistency error and on the conditioning of A.

3.3.2 Strang's first lemma

In practice, due to rounding errors, one has to use *approximations $A_n \colon X_n \to X_n'$* of A on X_n. Moreover, one may reduce the computational costs by using numerical methods (e.g. quadrature rules) whose accuracy is adjusted to the expected error level: There is no point in performing expensive calculations with very high precision if the discretization error is magnitudes higher. Classically, Strang's first lemma addresses these issues by giving estimates on the overall error of solutions of the discretized problem. Theorem 3.4 induces a variant of Strang's first lemma, alas without existence and uniqueness statements:

Lemma 3.27
Assume that $\mathcal{N}$ and $\mathcal{N}_n$ are non-empty and that

$$\begin{aligned}
\|A_n(x)\|_{X_n'} &\geq g(\operatorname{dist}_{X_n}(x, \mathcal{N}_n)) \quad \text{for all } x \in X_n, \\
\|A(x) - A(y)\|_{X'} &\leq h(\|x - y\|) \qquad\qquad \text{for all } x, y \in X.
\end{aligned}$$

Then one has

$$\mathcal{N} \subset \bar{B}(\mathcal{N}_n, \varepsilon_n^{\mathcal{S}} + g^{\dagger}(\chi_n + h(\varepsilon_n^{\mathcal{S}}))),$$

where $\varepsilon_n^{\mathcal{S}}$ is the sampling proximity error on $\mathcal{N}$ and

$$\chi_n := \sup_{b \in \mathcal{S}_n(\mathcal{N})} \|A_n(b) - (\mathcal{R}_n' \circ A \circ \mathcal{R}_n)(b)\|_{X_n'}.$$

Proof. The sampling consistency error $\eta_n^{\mathcal{S}}$ on $\mathcal{N}$ can be estimated by

$$
\begin{aligned}
\eta_n^{\mathcal{S}} &= \sup_{a \in \mathcal{N}} \left(\|(A_n \circ \mathcal{S}_n)(a)\|_{X_n'} - \|A(a)\|_{X'} \right)^+ \\
&\leq \sup_{a \in \mathcal{N}} \left(\left((\|A_n \circ \mathcal{S}_n)(a)\|_{X_n'} - \|(A \circ \mathcal{R}_n \circ \mathcal{S}_n)(a)\|_{X'} \right)^+ + \|(A \circ \mathcal{R}_n \circ \mathcal{S}_n)(a) - A(a)\|_{X'} \right) \\
&\leq \sup_{b \in \mathcal{S}_n(\mathcal{N})} \left(\|A_n(b)\|_{X_n'} - \|(\mathcal{R}_n' \circ A \circ \mathcal{R}_n)(b)\|_{X_n'} \right)^+ + \sup_{a \in \mathcal{A}} h(\|(\mathcal{R}_n \circ \mathcal{S}_n)(a) - a\|_X) \\
&\leq \sup_{b \in \mathcal{S}_n(\mathcal{N})} \|A_n(b) - (\mathcal{R}_n' \circ A \circ \mathcal{R}_n)(b)\|_{X_n'} + h(\varepsilon_n^{\mathcal{S}}) = \chi_n + h(\varepsilon_n^{\mathcal{S}}).
\end{aligned}
$$

We emphasize that we used here that h is nondecreasing. Now, the statement follows immediately from Theorem 3.4. □

Example 3.28 Note that χ_n measures in some way the deviation of A_n from the Ritz-Galerkin discretization. This is traditionally termed consistency error. The connection to Strang's first lemma becomes even clearer when analyzing the classical setting:

Let $\xi \in X'$, $\xi_n \in X_n'$ be continuous linear forms and let $B \colon X \times X \to \mathbb{R}$ and $B_n \colon X_n \times X_n \to \mathbb{R}$ be bilinear forms such that B is continuous and B_n is coercive, i.e., there are constants $C \geq 0$, $c > 0$ with

$$
|B(u,v)| \leq C\|u\|_X \|v\|_X \quad \text{and} \quad B_n(w,w) \geq c\,\|w\|_X^2 \quad \text{for all } u, v \in X, \text{ and all } w \in X_n.
$$

Consider the operators

$$
\begin{aligned}
A &\colon X \to X', \quad \langle A(x), u \rangle := B(x,u) - \langle \xi, u \rangle, \\
A_n &\colon X_n \to X_n', \quad \langle A_n(y), v \rangle := B_n(y,v) - \langle \xi_n, v \rangle.
\end{aligned}
$$

By the Lax-Milgram theorem, $\mathcal{N}_n = \{\, x_n \,\}$ is a singleton. Using the preceding lemmawith the functions $g(s) = c\,s$ and $h(s) = C\,s$, one obtains

$$
\|x - x_n\| \leq \left(1 + \tfrac{C}{c} \right) \varepsilon_n + \tfrac{1}{c} \chi_n \quad \text{for each } x \in \mathcal{N}.
$$

Note that χ_n can be estimated by

$$
\chi_n \leq \sup_{b \in \mathcal{S}_n(\mathcal{N})} \sup_{w \in X_n \setminus \{0\}} \left(\frac{|B(b,w) - B_n(b,w)|}{\|w\|_X} + \frac{|\langle \xi - \xi_n, w \rangle|}{\|w\|_X} \right).
$$

This is exactly the consistency error of the classical Strang lemma (see [6, Chapter III, Lemma 1.1]).

3.3.3 Nonconforming Ritz-Galerkin method

Let a sequence X_n of finite-dimensional Banach spaces and operators $A_n \colon X_n \to X_n'$, $n \in \mathbb{N}$ be given. Note that from now on, we do not assume that X_n is a subspace of X. The frequent setting is that X, X_n are continuously injected into a larger Banach space $\mathcal{X} := Y$ in a canonical way. Denote these injections by $\Psi \colon X \hookrightarrow Y$, $\Psi_n \colon X_n \hookrightarrow Y$. If $\Psi_n(X_n) \not\subset \Psi(X)$, one calls (X_n, A_n) a *nonconforming Ritz-Galerkin scheme*. In practice, there are essentially two reasons why a Ritz-Galerkin scheme is nonconforming:

1. The elements of X_n may violate certain constraints on X. In particular, boundary conditions may be an issue: The elements of the function space X_n may satisfy boundary conditions only on a restricted class of boundary shapes, e.g., polygonal lines, simplicial manifolds, or spline surfaces.

2. The (differential) operator $A \colon X \to X'$ cannot be extended to X_n because the elements of X_n fail to have the necessary smoothness, e.g., they are discontinuous.

Traditionally, one calls the quantity

$$\sup_{a \in \mathcal{N}} \inf_{w \in X_n} \|\Psi(a) - \Psi_n(w)\|_Y,$$

the *approximation error*. Note that the approximation error can be bounded by the sampling proximity error $\varepsilon_n^{\mathcal{S}}$ on $\mathcal{N}$:

$$\sup_{a \in \mathcal{N}} \inf_{w \in X_n} \|\Psi(a) - \Psi_n(w)\|_Y \leq \sup_{a \in \mathcal{N}} \|\Psi(a) - (\Psi_n \circ \mathcal{S}_n)(a)\|_Y = \varepsilon_n^{\mathcal{S}}$$

In contrast to the classical Strang lemma, we may circumvent the need to extend A_n to Y by using $A_n \circ \mathcal{S}_n$ in order to define what is traditionally called the *consistency error*:

$$\sup_{a \in \mathcal{N}} \inf_{w \in X_n} \frac{|\langle (A_n \circ \mathcal{S}_n)(a), w \rangle|}{\|w\|_{X_n}}.$$

Note that this is precisely the sampling consistency error $\delta_n^{\mathcal{S}}$ on $\mathcal{N}$:

$$\sup_{a \in \mathcal{N}} \inf_{w \in X_n} \frac{|\langle (A_n \circ \mathcal{S}_n)(a), w \rangle|}{\|w\|_{X_n}} = \sup_{a \in \mathcal{N}} \|(A_n \circ \mathcal{S}_n)(a)\|_{X_n'}$$
$$= \sup_{a \in \mathcal{N}} \|(A_n \circ \mathcal{S}_n)(a) - A(a)\|_{X_n'} = \delta_n^{\mathcal{S}}.$$

It turns out that Theorem 3.4 implies the variant of Strang's second lemma that Braess briefly mentions in a side remark (see [6, Chapter III, Remark 1.3]):

Lemma 3.29 *Assume that $\mathcal{N}$ and $\mathcal{N}_n$ are non-empty and that*

$$g(\mathrm{dist}_Y(\Psi_n(x), \Psi_n(\mathcal{N}_n))) \leq \|A_n(x)\|_{X_n} \qquad \text{for all } x \in X_n,$$
$$\|A(x) - A(y)\|_X \leq h(\mathrm{dist}_Y(\Psi(x), \Psi(y))) \quad \text{for all } x, y \in X.$$

Then one has

$$\Psi(\mathcal{N}) \subset \bar{B}(\Psi_n(\mathcal{N}_n), \varepsilon_n^{\mathcal{S}} + g^\dagger(\delta_n^{\mathcal{S}} + h(\varepsilon_n^{\mathcal{S}}))).$$

3.4 Openness, Existence, and Convergence

We conclude this chapter with a remark on the relationship between openness of the operators A, A_n (see Definition 2.49) and the convergence behavior of their solution sets.

Theorem 3.30 *Let operators $A\colon C \to Y$ and $A_n\colon C_n \to Y_n$ be given, where C, C_n are metric spaces and Y, Y_n are normed vector spaces. For some non-empty sets $\mathcal{A} \subset C$, $\mathcal{A}_n \subset C_n$ denote by by $\varepsilon_n^{\mathcal{S}}$, $\varepsilon_n^{\mathcal{R}}$ the proximity errors of (Ψ, Ψ_n) on $(\mathcal{A}, \mathcal{A}_n)$ and by $\eta_n^{\mathcal{S}}$, $\eta_n^{\mathcal{R}}$ the consistency errors of A_n with respect to A on $(\mathcal{A}, \mathcal{A}_n)$. Moreover, let ϑ, σ be positive real numbers.*

1. *Assume that A_n is (ϑ, σ)-open on $\mathcal{S}_n(\mathcal{A} \cap \mathcal{N})$ and that $\delta_n^{\mathcal{S}} < \vartheta$. If $\mathcal{A} \cap \mathcal{N}$ is non-empty, then also $\mathcal{N}_n$ is non-empty and one has*

$$\Psi(\mathcal{A} \cap \mathcal{N}) \subset \bar{B}(\Psi_n(\mathcal{N}_n), r_n^{\mathcal{S}}) \quad \text{with } r_n^{\mathcal{S}} = \varepsilon_n^{\mathcal{S}} + \mathrm{Lip}(\Psi_n)\,\sigma\,\delta_n^{\mathcal{S}}.$$

2. *Assume that A is (ϑ, σ)-open on $\mathcal{R}_n(\mathcal{A}_n \cap \mathcal{N}_n)$ and that $\delta_n^{\mathcal{R}} < \vartheta$. If $\mathcal{A}_n \cap \mathcal{N}_n$ is non-empty, then also $\mathcal{N}$ is non-empty and one has*

$$\Psi_n(\mathcal{A}_n \cap \mathcal{N}_n) \subset \bar{B}(\Psi(\mathcal{N}), r_n^{\mathcal{R}}) \quad \text{with } r_n^{\mathcal{R}} = \varepsilon_n^{\mathcal{R}} + \mathrm{Lip}(\Psi)\,\sigma\,\delta_n^{\mathcal{R}}.$$

PROOF. Let $x \in \mathcal{A} \cap \mathcal{N}$, i.e., $\mathcal{G}(x) = \|A(x)\|_Y = 0$. One has

$$\|(A \circ \mathcal{S}_n)(x)\|_{Y_n} = (\mathcal{G}_n \circ \mathcal{S}_n)(x) \leq \mathcal{G}(x) + \delta_n^{\mathcal{S}} < \vartheta.$$

Now, (ϑ, σ)-openness implies

$$0 \in \bar{B}((A_n \circ \mathcal{S}_n)(x), \delta_n^{\mathcal{S}}) \subset A_n(\bar{B}(\mathcal{S}_n(x), \sigma\,\delta_n^{\mathcal{S}})).$$

Hence, there is a $y \in \mathcal{N}_n$ with $d_{C_n}(\mathcal{S}_n(x), y) \leq \sigma\,\delta_n^{\mathcal{S}}$ and one obtains

$$d_X(\Psi(x), \Psi_n(y)) \leq d_X(\Psi(x), (\Psi_n \circ \mathcal{S}_n)(x)) + d_X((\Psi_n \circ \mathcal{S}_n)(x), \Psi_n(y))$$
$$\leq \varepsilon_n^{\mathcal{S}} + \mathrm{Lip}(\Psi_n)\,\sigma\,\delta_n^{\mathcal{S}} = r_n^{\mathcal{S}},$$

which shows the first claim. The proof of the second claim is analogous. $\square$

Note that in the setting of an operator A of class $C^{1,1}$ between Banach spaces (or more generally: Banach manifolds), openness of A at $x \in C$ can be shown with the quantitative inverse function theorem (Theorem 2.53), even in the case that $DA(x)$ has a kernel, as long as it has a sufficiently transversal closed complement (see the proof of Lemma 2.52).

Interpreting $\mathcal{N}_n$ as the discrete problem, the second claim of Theorem 3.30 can be used for *a posteriori* estimates: Having found $x \in \mathcal{N}_n$ one may sometimes be able to estimate the openness of A at $\mathcal{R}_n(x)$:

Example 3.31 Let C, C_n, Y, Y_n be Banach spaces and let both A and A_n be of class $C^{1,1}$. Let $x \in \mathcal{N}_n$. Assume that $DA_n(x)$ satifies $\|DA_n(x)\,u\| \geq \lambda_n\|u\|$ for all $u \in T_x C_n$. Moreover assume that one can show $\|DA(\mathcal{R}_n(x))\,u\| \geq \lambda\|u\|$ for all $u \in T_{\mathcal{R}_n(x)}C$ with some $\lambda \geq \lambda_n - c\,\varepsilon_n > 0$. Then again, the quantitative inverse function theorem would imply a certain openness of A and Theorem 3.30 yields existence of a smooth solution and an error estimate.

4 The Space of Inner Products

In this chapter we summarize some facts about the Riemannian manifold of inner product of a finite-dimensional real vector space, i.e., of symmetric, positive definite bilinear forms. This space and its Riemannian distance will be crucial in defining the metric space of Lipschitz immersions (see Chapter 5). The latter will be used as configuration space in our treatment of discrete minimal surfaces (Chapter 7), as well as in our discretization of Hencky elasticity (Chapter 8).

This space is traditionally discussed as the homogeneous space $\mathrm{GL}_k(\mathbb{R})/\mathrm{O}(k)$ by differential geometers. However, the representation as a quotient may not be convenient if one aims at numerical computations. Fortunately, the manifold of symmetric, positive definite matrices and its Riemannian structures have recently caught the attention of applied mathematicians so there is also a concise theory in terms of matrices (see e.g., [27] and references therein). We try to be self-contained and to give proofs for the relevant results, although these may be found elsewhere, too.

In the course of this chapter, we will also be led in a natural way to a certain vector field (on the manifold of inner products), which reappears as the Hencky strain tensor in elasticity theory.

4.1 Basic Definitions

Let V be a k-dimensional real vector space with $k \in \mathbb{N}$ and let $P(V)$ denote the space of symmetric, positive definite bilinear forms on V. The group $\mathrm{GL}(V)$ acts from the right on $P(V)$ via pullback:

$$\mathrm{GL}(V) \times P(V) \to P(V), \quad (A, b) \mapsto A^{\#}b = b(A\cdot, A\cdot).$$

As an open set in the vector space $\mathrm{Sym}(V) = V' \odot V' \subset V' \otimes V'$ of symmetric bilinear forms on V, the space $P(V)$ is a smooth manifold with tangent bundle given by $T_b P(V) := \mathrm{Sym}(V)$.

We equip $P(V)$ with a Riemannian structure g_P given by

$$g_P|_b(X, Y) := \langle X, Y \rangle_b \quad \text{for all } X, Y \in T_b P(V) = \mathrm{Sym}(V),$$

where $\langle \cdot, \cdot \rangle_b$ denotes the inner product on $V' \otimes V'$ that is induced by b.

We use the so-called *musical isomorphisms*

$$\flat_b \colon V \to V', \quad u \mapsto b(u, \cdot) = (v \mapsto b(u, v)) \quad \text{and} \quad \sharp_b := \flat_b^{-1} \colon V' \to V.$$

In finite-dimensional vector spaces, one may identify $\text{End}(V) \cong V \otimes V'$ and

$$\text{id}_V \otimes \sharp_b \colon \text{Sym}(V) \to \{ A \in \text{End}(V) \mid A^{*_b} = A \}$$

identifies $\text{Sym}(V)$ isometrically with the b-self-adjoint endomorphisms:

$$\langle X, Y \rangle_b = \langle (\text{id}_V \otimes \sharp_b) X, (\text{id}_V \otimes \sharp_b) Y \rangle_b = \text{tr} \left(((\text{id}_V \otimes \sharp_b) X)^{*_b} (\text{id}_V \otimes \sharp_b) Y \right),$$

where $X, Y \in \text{Sym}(V)$.

Theoretically, one could deduce all the result of this chapter in terms of this identification. However, it proves less cumbersome to perform computations in terms of Gram matrices.

Definition 4.1 For a basis $e = (e_1, \ldots, e_k)$ of V, define the *Gram mapping* which maps a bilinear form to its Gram matrix:

$$\mathbf{G}_e \colon V' \otimes V' \to \text{Mat}_{k \times k}(\mathbb{R}), \quad X \mapsto (X(e_i, e_j))_{1 \le i, j \le k}.$$

Remark 4.2 In terms of the Gram mapping, one has the following representation that we will use throughout our discussion:

$$g_P(X, Y)|_b = \langle X, Y \rangle_b = \text{tr}(\mathbf{G}_e(b)^{-1} \mathbf{G}_e(X)^\mathsf{T} \mathbf{G}_e(b)^{-1} \mathbf{G}_e(Y)), \tag{4.1}$$

where $b \in P(V)$ and $X, Y \in T_b P(V)$.

Remark 4.3 Whenever a basis e of V is given, the dual basis $\eta = (\eta_1, \ldots, \eta_k)$ of V' can be written as $\eta_i = \sum_{j=1}^k (\mathbf{G}_e(b)^{-1})_{ij} \flat_b e_j$. A basis $(\eta_{ij})_{1 \le i \le j \le k}$ for $\text{Sym}(V)$ is induced by e via $\eta_{ij} = \eta_i \otimes \eta_j + \eta_j \otimes \eta_i$. Moreover, since $\mathbf{G}_e$ is a chart, we have globally defined coordinate vector fields $X_{ij} \in \mathfrak{X}(P(V))$ on $P(V)$ by:

$$\begin{aligned}
X_{ij}|_b &:= \eta_i \otimes \eta_j + \eta_j \otimes \eta_i \\
&= \sum_{\alpha, \beta = 1}^k (\mathbf{G}_e(b)^{-1})_{i\alpha} (\mathbf{G}_e(b)^{-1})_{j\beta} (\flat_b e_i \otimes \flat_b e_j + \flat_b e_j \otimes \flat_b e_i).
\end{aligned}$$

Because the vector fields X_{ij} are coordinate vector fields, their commutators vanish:

$$[X_{\alpha\beta}, X_{\gamma\delta}] = 0 \quad \text{for all } 1 \le \alpha, \beta, \gamma, \delta \le k.$$

We even have $\text{d}(\mathbf{G}_e(X_{ij})) = 0$. Note that every smooth vector field $Y \in \mathfrak{X}(P(V))$ can be written as $Y = \sum_{1 \le \alpha \le \beta \le k} \varphi_{\alpha\beta} X_{\alpha\beta}$ with appropriately chosen $\varphi_{\alpha\beta} \in C^\infty(P(V); \mathbb{R})$.

4.2 The Levi-Civita Connection

In order to compute distances, we are interested in the geodesics in $(P(V), g_P)$. Therefore, we deduce a representation of the Levi-Civita connection in terms of its covariant derivative

$$\nabla \colon \mathfrak{X}(P(V)) \times \mathfrak{X}(P(V)) \to \mathfrak{X}(P(V)).$$

Lemma 4.4 *Let X, $Y \in \mathfrak{X}(P(V))$ be smooth vector fields and let e be a basis of V. The Levi-Civita connection with respect to g_P can be written as*

$$\mathbf{G}_e(\nabla_X Y|_b) = \langle \mathrm{d}(\mathbf{G}_e(Y))|_b, X \rangle - \sigma(\mathbf{G}_e(X)\mathbf{G}_e(b)^{-1}\mathbf{G}_e(Y)),$$

where $\sigma \colon \mathrm{Mat}_{k \times k}(\mathbb{R}) \to \mathrm{Mat}_{k \times k}(\mathbb{R})$, $\mathbf{A} \mapsto \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)$ denotes symmetrization.

PROOF. According to Remark 4.3, there are vector fields $X_1, \ldots, X_m \in \mathfrak{X}(P(V))$, $m = \frac{1}{2}k(k+1)$ with $\mathrm{d}(\mathbf{G}_e(X_\alpha)) = 0$ for all $1 \le \alpha \le m$ such that each vector field $Y \in \mathfrak{X}(P(V))$ can be written as $Y = \sum_{\alpha=1}^m \varphi_\alpha X_\alpha$, with appropriately chosen $\varphi_\alpha \in C^\infty(P(V); \mathbb{R})$. Note that this also implies $[X_\alpha, X_\beta] = 0$ for all $1 \le \alpha, \beta \le m$. The Koszul formula (see e.g., [9, Chapter 2, Equation 9]) tells us:

$$g_P(X_\alpha, \nabla_{X_\beta} X_\gamma) = \tfrac{1}{2}(X_\beta\, g_P(X_\gamma, X_\alpha) + X_\gamma\, g_P(X_\alpha, X_\beta) - X_\alpha\, g_P(X_\beta, X_\gamma)).$$

Abbreviating $\mathbf{B} := \mathbf{G}_e(b)$ and $\mathbf{X}_\alpha := \mathbf{G}_e(X_\alpha)$, we have from (4.1) that

$$g_P(X_\gamma, X_\alpha) = \mathrm{tr}\,(\mathbf{B}^{-1}\mathbf{X}_\gamma \mathbf{B}^{-1}\mathbf{X}_\alpha).$$

Using the well-known rules for differentiating products and inverses, we obtain

$$
\begin{aligned}
X_\beta\, g_P(X_\gamma, X_\alpha)|_b \\
&= -\mathrm{tr}\,(\mathbf{B}^{-1}\mathbf{X}_\beta \mathbf{B}^{-1}\mathbf{X}_\gamma \mathbf{B}^{-1}\mathbf{X}_\alpha) - \mathrm{tr}\,(\mathbf{B}^{-1}\mathbf{X}_\gamma \mathbf{B}^{-1}\mathbf{X}_\beta \mathbf{B}^{-1}\mathbf{X}_\alpha) \\
&= -2\,\mathrm{tr}\,(\mathbf{B}^{-1}\sigma(\mathbf{X}_\beta \mathbf{B}^{-1}\mathbf{X}_\gamma)\mathbf{B}^{-1}\mathbf{X}_\alpha) \\
&= -2 g_P\big(\mathbf{G}_e^{-1}(\sigma(\mathbf{X}_\beta \mathbf{B}^{-1}\mathbf{X}_\gamma)), X_\alpha|_b\big).
\end{aligned}
$$

Symmetry of all occurring matrices and conjugation invariance of the trace lead to

$$X_\beta\, g_P(X_\gamma, X_\alpha) = X_\gamma\, g_P(X_\alpha, X_\beta) = X_\alpha\, g_P(X_\beta, X_\gamma),$$

hence to $\nabla_{X_\beta} X_\gamma|_b = -\mathbf{G}_e^{-1}(\sigma(\mathbf{X}_\beta \mathbf{B}^{-1}\mathbf{X}_\gamma))$ and

$$\mathbf{G}_e(\nabla_{X_\beta} X_\gamma|_b) = \langle \mathrm{d}\mathbf{G}_e(X_\beta), X_\gamma \rangle - \sigma(\mathbf{X}_\beta \mathbf{B}^{-1}\mathbf{X}_\gamma).$$

For arbitrary vector fields X and Y, write $Y = \sum_{\beta=1}^m \varphi_\beta X_\beta$. The statement now follows from the Leibniz rule for covariant differentiation. $\qquad\square$

Corollary 4.5 *The vector field $Z \in \mathfrak{X}(P(V))$ given by $Z|_b = b$ is covariantly constant.*

PROOF. Let $X \in \mathfrak{X}(P(V))$ be an arbitrary vector field and e some basis of V. By the preceding lemma, we may compute

$$\mathbf{G}_e(\nabla_X Z|_b) = \langle d\mathbf{G}_e(b), X \rangle - \sigma(\mathbf{G}_e(X)\mathbf{G}_e(b)^{-1}\mathbf{G}_e(b)) = \mathbf{G}_e(X) - \mathbf{G}_e(X) = 0. \quad \square$$

4.3 Geodesics

Lemma 4.6 *Let V be a finite-dimensional real vector space, $b \in P(V)$, and $X \in T_b P(V)$. The geodesic $\gamma \colon]{-\varepsilon, \varepsilon}[\to P(V)$ starting from b in direction $X \in T_b P(V)$ with respect to the Riemannian metric g_P is given by*

$$\gamma(t) = \mathbf{G}_e^{-1}\left(\mathbf{L}^{\mathsf{T}} \exp\left(t\,\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(X)\mathbf{L}^{-1}\right)\mathbf{L}\right)$$

for every $\mathbf{L} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$ with $\mathbf{L}^{\mathsf{T}}\mathbf{L} = \mathbf{G}_e(b)$.

PROOF. Let γ be the geodesic with $\gamma(0) = b$ and $\dot{\gamma}(0) = X$. Define $\mathbf{B}(t) := \mathbf{G}_e(\gamma(t))$. Since $\mathbf{G}_e(\dot{\gamma}(t)) = \dot{\mathbf{B}}(t)$, the geodesic equations can be written as

$$\begin{aligned}
0 = \mathbf{G}_e(\nabla_{\dot{\gamma}}\dot{\gamma})(t) &= \langle d\mathbf{G}_e(\dot{\gamma}(t)), \dot{\gamma}(t) \rangle - \sigma(\mathbf{G}_e(\dot{\gamma}(t))\mathbf{G}_e(\gamma(t))^{-1}\mathbf{G}_e(\dot{\gamma}(t))) \\
&= \ddot{\mathbf{B}}(t) - \dot{\mathbf{B}}(t)\,\mathbf{B}^{-1}(t)\,\dot{\mathbf{B}}(t).
\end{aligned}$$

We use the ansatz $\mathbf{B}(t) = \mathbf{L}^{\mathsf{T}}\exp(t\,\mathbf{C})\mathbf{L}$ with a matrix $\mathbf{L} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$ and a symmetric matrix $\mathbf{C} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$. One computes

$$\dot{\mathbf{B}}(t) = \mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{L} \quad \text{and} \quad \ddot{\mathbf{B}}(t) = \mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{C}\mathbf{L}$$

and checks that substituting our ansatz solves the geodesic equation:

$$\begin{aligned}
\mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{C}\mathbf{L} &- \mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{L}\mathbf{L}^{-1}\exp(-t\,\mathbf{C})\mathbf{L}^{-\mathsf{T}}\mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{L} \\
&= \mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{C}\mathbf{L} - \mathbf{L}^{\mathsf{T}}\mathbf{C}\exp(t\,\mathbf{C})\,\mathbf{C}\mathbf{L} = 0.
\end{aligned}$$

For satisfying the initial conditions, one has to find $\mathbf{L}$ and $\mathbf{C}$ such that

$$\mathbf{B}(0) = \mathbf{L}^{\mathsf{T}}\mathbf{L} = \mathbf{G}_e(b) \quad \text{and} \quad \dot{\mathbf{B}}(0) = \mathbf{L}^{\mathsf{T}}\mathbf{C}\mathbf{L} = \mathbf{G}_e(X)$$

hold. Note that by the uniqueness of the solutions for second-order ODEs, γ does not depend on the actual choice of $\mathbf{L}$. $\quad \square$

Example 4.7 When choosing $\mathbf{L} = \mathbf{G}_e(b)^{\frac{1}{2}}$, one has

$$\mathbf{G}_e(\gamma(t)) = \mathbf{G}_e(b)^{\frac{1}{2}} \exp\left(t \cdot \mathbf{G}_e(b)^{-\frac{1}{2}} \mathbf{G}_e(X) \mathbf{G}_e(b)^{-\frac{1}{2}}\right) \mathbf{G}_e(b)^{\frac{1}{2}}.$$

Note that $\gamma(t)$ exists for all times $t \in \mathbb{R}$ and that $\gamma(t)$ is always positive definite, which leads us to:

Corollary 4.8 *The Riemannian manifold $(P(V), g_P)$ is geodesically complete, i.e., for each $b \in P(V)$ and each $X \in T_b P(V)$, there is a geodesic $\gamma \colon \mathbb{R} \to P(V)$ with $\gamma(0) = b$ and $\dot{\gamma}(0) = X$.*

For every pair of points in a geodesically complete space, there is always a length minimizing geodesic connecting them. In $(P(V), g_P)$, there is exactly one geodesic between any pair of points and this geodesic can be directly expressed in terms of Gram matrices.

Lemma 4.9 *For any two points b_0, $b_1 \in (P(V), g_P)$, there is a* unique *geodesic $\gamma \colon [0, 1] \to P(V)$ with $\gamma(0) = b_0$ and $\gamma(1) = b_1$.*

PROOF. Let b_0, $b_1 \in P(V)$. Choose a basis e of V and $\mathbf{L} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$ such that $\mathbf{G}_e(b_0) = \mathbf{L}^{\mathsf{T}} \mathbf{L}$.

Existence: One easily verifies with Lemma 4.6 that the geodesic γ starting at b_0 in direction $X = \mathbf{G}_e^{-1}(\mathbf{L}^{\mathsf{T}} \log(\mathbf{L}^{-\mathsf{T}} \mathbf{G}_e(b_1) \mathbf{L}^{-1}) \mathbf{L})$ satisfies $\gamma(1) = b_1$.

Uniqueness: Let $\gamma \colon [0, 1] \to P(V)$ with $\gamma(0) = b_0$ and $\gamma(1) = b_1$. Put $X := \dot{\gamma}(0)$. By Lemma 4.6, one has $b_1 = \gamma(1) = \mathbf{G}_e^{-1}(\mathbf{L}^{\mathsf{T}} \exp(\mathbf{L}^{-\mathsf{T}} \mathbf{G}_e(X) \mathbf{L}^{-1}) \mathbf{L})$. Since the matrix exponential is a diffeomorphism from the symmetric matrices onto the symmetric, positive definite matrices, this equation can be solved for X, delivering the same X as in the existence proof. Note that X does not depend on the particular choice of $\mathbf{L}$: Any other choice can be written as $\mathbf{UL}$ with a orthonormal matrix $\mathbf{U} \in O(\mathbb{R}^n)$. By the rules of the functional calculus of self-adjoint operators we find:

$$
\begin{aligned}
(\mathbf{UL})^{\mathsf{T}} &\log\left((\mathbf{UL})^{-\mathsf{T}} \mathbf{G}_e(b_1)(\mathbf{UL})^{-1}\right)(\mathbf{UL}) \\
&= \mathbf{L}^{\mathsf{T}} \mathbf{U}^{\mathsf{T}} \log\left(\mathbf{U} \mathbf{L}^{-\mathsf{T}} \mathbf{G}_e(b_1) \mathbf{L}^{-1} \mathbf{U}^{-1}\right) \mathbf{U} \mathbf{L} \\
&= \mathbf{L}^{\mathsf{T}} \mathbf{U}^{\mathsf{T}} \mathbf{U} \log\left(\mathbf{L}^{-\mathsf{T}} \mathbf{G}_e(b_1) \mathbf{L}^{-1}\right) \mathbf{U}^{-1} \mathbf{U} \mathbf{L} \\
&= \mathbf{L}^{\mathsf{T}} \log\left(\mathbf{L}^{-\mathsf{T}} \mathbf{G}_e(b_1) \mathbf{L}^{-1}\right) \mathbf{L}. \qquad \square
\end{aligned}
$$

4.4 The Hencky Field

Motivated by the preceding result and because of its central role in nonlinear elasticity (see Chapter 8), we introduce the Hencky field, a vector field on $P(V)$:

Definition 4.10 Fix $c \in P(V)$. We define the *Hencky field* $H_c \in \mathfrak{X}(P(V))$ by

$$H_c|_b := (\mathrm{id} \otimes \flat_b)(\log_b ((\mathrm{id} \otimes \sharp_b)c)).$$

Here, $\log_b$ is the logarithm of the functional calculus of b-self-adjoint operators.

Remark 4.11 In terms of a basis e of V and a matrix $\mathbf{L} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$ with $\mathbf{G}_e(b) = \mathbf{L}^\mathsf{T}\mathbf{L}$, one may write

$$H_c|_b = \mathbf{G}_e^{-1}(\mathbf{L}^\mathsf{T} \log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1}) \mathbf{L}).$$

Thus, by Lemma 4.9, the Hencky field X_c always points to c in the sense that

$$\exp_b^P (H_c|_b) = c,$$

where $\exp_b^P \colon T_b P(V) \to P(V)$ is the Riemannian exponential map with respect to g_P.

Corollary 4.12 *The geodesic distance of $(P(V), g_P)$ is given by*

$$d_P(b, c) = |H_c|_b|_b = |\log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1})| = \left(\textstyle\sum_{i=1}^k \log(\lambda_i)^2\right)^{\frac{1}{2}}, \quad b, c \in P(V).$$

Here, $|\cdot|$ denotes the Frobenius norm of matrices, $\mathbf{L} \in \mathrm{Mat}_{k \times k}(V)$ is a matrix with $\mathbf{L}^\mathsf{T}\mathbf{L} = \mathbf{G}_e(b)$, and $\lambda_1, \dots, \lambda_k$ are the eigenvalues of $\mathbf{G}_e(g)$ with respect to $\mathbf{G}_e(b)$.

PROOF. While $d_P(b, c) = |H_c|_b|_b$ follows from the fact that geodesics have constant speed, the second equality follows from (4.1):

$$
\begin{aligned}
|H_c|_b|_b^2 &= \mathrm{tr}\left((\mathbf{L}^\mathsf{T}\mathbf{L})^{-1}\mathbf{G}_e(H_c|_b)(\mathbf{L}^\mathsf{T}\mathbf{L})^{-1}\mathbf{G}_e(H_c|_b)\right) \\
&= \mathrm{tr}\left((\mathbf{L}^\mathsf{T}\mathbf{L})^{-1}\mathbf{L}^\mathsf{T} \log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1}) \mathbf{L}(\mathbf{L}^\mathsf{T}\mathbf{L})^{-1}\mathbf{L}^\mathsf{T} \log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1}) \mathbf{L}\right) \\
&= \mathrm{tr}\left(\mathbf{L}^{-1} \log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1})^2 \mathbf{L}\right) = \mathrm{tr}\left(\log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1})^2\right) \\
&= |\log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1})|^2.
\end{aligned}
$$

One may choose e as an orthonormal basis of b such that $\mathbf{G}_e(c) = \mathbf{diag}(\lambda_1, \dots, \lambda_k)$ is a diagonal matrix. With $\mathbf{L} = \mathbf{I}$, one obtains

$$|\log (\mathbf{L}^{-\mathsf{T}}\mathbf{G}_e(c)\mathbf{L}^{-1})|^2 = \textstyle\sum_{i=1}^k \log(\lambda_i)^2. \qquad \square$$

Corollary 4.13 *Fix $c \in P(V)$ and define $f_c \colon P(V) \to \mathbb{R}$, $f_c(b) = \frac{1}{2}d_P^2(b, c)$. This function is smooth and the downward gradient coincides with the Hencky field:*

$$\mathrm{grad}^{g_P} f_c = -H_c.$$

4.5 Densities

We summarize some definitions and facts about densities. For a more detailed introduction to this topic see, e.g., [26, pp. 375–382].

Definition 4.14 Let V be a k-dimensional real vector space. A *density* on V is a function $\varrho \colon \prod_{i=1}^{k} V \to \mathbb{R}$ with the properties:

1. For all $v_1, \ldots, v_k \in V$ and all $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ the following holds:

$$\varrho(\lambda_1 v_1 \ldots, \lambda_n v_k) = |\lambda_1 \cdots \lambda_k| \varrho(v_1 \ldots, v_k).$$

2. For all $v_1, \ldots, v_k \in V$, $\lambda \in \mathbb{R}$ and $i \neq j$ the following holds:

$$\varrho(v_1 \ldots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \ldots, v_k) = \varrho(v_1 \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k).$$

We define $\Omega(V)$ to be the space of densities on V.

The space $\Omega(V)$ is a one-dimensional vector space over $\mathbb{R}$. A linear map $A \colon V \to V$ induces the *pullback* $A^{\#} \colon \Omega(V) \to \Omega(V)$, a linear operator defined by

$$(A^{\#}\varrho)(v_1, \ldots, v_k) = \varrho(A v_1, \ldots, A v_k) \quad \text{for } v_1, \ldots, v_k \in V.$$

One may define the *absolute value of the determinant of A* by $|\det(A)| \varrho = A^{\#}\varrho$ for all $\varrho \in \Omega(V)$. Thus, densities have a transformation behavior under pullback that is compatible with the transformation formula of integrals. This is why one can define the integral of a section of the density bundle $\Omega(\Sigma) := \coprod_{x \in \Sigma} \Omega(T_x \Sigma)$, leading to a linear functional $\int_{\Sigma} \colon \Omega(\Sigma) \to \mathbb{R}$.

A density $\varrho \in \Omega(V)$ is called *positive*, if $\varrho(e_1, \ldots, e_k) > 0$ holds for all bases $e = (e_1, \ldots, e_k)$ of V. Denote the space of positive densities on V by $\mathrm{Vol}(V)$. Note that the multiplicative group $(\mathbb{R}_{>0}, \cdot)$ acts transitively on $\mathrm{Vol}(V)$. For two densities ϱ and σ denote the unique positive number $t \in \mathbb{R}_{>0}$ with $\sigma = t \varrho$ by $\frac{\sigma}{\varrho}$. We define the distance d_{Vol} by

$$d_{\mathrm{Vol}}(\varrho, \sigma) := |\log\left(\tfrac{\sigma}{\varrho}\right)| \quad \text{for all } \varrho, \sigma \in \mathrm{Vol}(V).$$

Every $g \in P(V)$ induces a unique density vol_g on V fulfilling $\mathrm{vol}_g(e_1, \ldots, e_k) = 1$ for any g-orthonormal basis $e_1, \ldots, e_k$ of V. Thus, one has

$$\frac{\mathrm{vol}_b}{\mathrm{vol}_g} = \frac{\det(\mathbf{G}_e(b))}{\det(\mathbf{G}_e(g))} \quad \text{and} \quad d_{\mathrm{Vol}}(\mathrm{vol}_b, \mathrm{vol}_g) = |\log\left(\det(\mathbf{G}_e(b))\right) - \log\left(\det(\mathbf{G}_e(g))\right)|,$$

for any two elements $g, b \in P(V)$ and *any* basis e of V. For an injective linear map $A \colon (V_1, g_1) \to (V_2, g_2)$ between finite-dimensional inner product spaces, we point out the formula

$$\mathrm{vol}_{A^\# g_2} = |\det(A^* A)|^{\frac{1}{2}} \, \mathrm{vol}_{g_1}, \tag{4.2}$$

where A^* denotes the adjoint of A with respect to g_1 and g_2.

Lemma 4.15 *The mapping*

$$\mathrm{vol} \colon P(V) \to \mathrm{Vol}(V), \quad g \mapsto \mathrm{vol}_g$$

is Lipschitz-continous with Lipschitz constant $\dim(V)^{\frac{1}{2}}$.

PROOF. Let $g, b \in P(V)$ and choose an orthonormal basis $e = (e_1, \ldots, e_k)$ of g that diagonalizes $\mathbf{G}_e(b)$. Let $\lambda_1, \ldots, \lambda_k$ be the eigenvalues of $\mathbf{G}_e(b)$. One has $\mathrm{vol}_b = \lambda_1 \cdots \lambda_k \, \mathrm{vol}_g$, thus

$$d_{\mathrm{Vol}}(\mathrm{vol}_g, \mathrm{vol}_b) = |\log(\lambda_1 \cdots \lambda_k)| \leq \sum_{i=1}^{k} |\log(\lambda_i)| \leq \sqrt{k} \, d_P(g, b). \tag{4.3}$$

$\square$

Remark 4.16 Note that when $\dim(V) = 1$, the mapping $\mathrm{vol} \colon P(V) \to \mathrm{Vol}(V)$ is an isometric diffeomorphism. Hence, $\mathrm{Vol}(V)$ is also a Riemannian manifold and we may denote the induced Riemannian metric by g_{Vol}. The diagram

$$
\begin{array}{ccc}
TP(V) & \xrightarrow{\exp^P} & P(V) \\
{\scriptstyle T\,\mathrm{vol}} \downarrow & & \downarrow {\scriptstyle \mathrm{vol}} \\
T\,\mathrm{Vol}(V) & \xrightarrow{\exp^{\mathrm{Vol}}} & \mathrm{Vol}(V)
\end{array}
$$

is commutative because of the identities $\exp(\mathrm{tr}(\mathbf{X})) = \det(\exp(\mathbf{X}))$ for $\mathbf{X} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$ and $(T_b \, \mathrm{vol})X = g_P(b, X) \, \mathrm{vol}_b$ for $X \in T_b P(V)$.

4.6 Invariant Metrics

Here, we give a characterization of all GL(V)-invariant Riemannian metrics on $P(V)$ which highlights the outstanding roles of the Riemannian metrics g_P and g_{Vol} on the one hand; and of the Hencky field on the other. This also shows that Hencky's elastic energy (see Chapter 8) is a quite natural deformation measure (at least, for full-dimensional domains).

Lemma 4.17 *Let $A\colon V \to W$ be a linear isomorphism between two finite-dimensional real vector spaces. Then the pullback along A induces an isometric diffeomorphism*

$$A^{\#}\colon (P(W), g_P) \to (P(V), g_P), \qquad b \mapsto b(A\cdot, A\cdot).$$

PROOF. Fix $c \in P(W)$ and $b = A^{\#}c$. Let $e = (e_1, \ldots, e_k)$ and $f = (f_1, \ldots, f_k)$ be bases of V and W, respectively. Denote by ξ the dual basis of f and define the matrix $\mathbf{A} \in \mathrm{Mat}_{k\times k}(\mathbb{R})$ by $\mathbf{A}_{ij} := \langle \xi_i, A e_j \rangle$, $1 \le i, j \le k$. Now, let $Y \in T_c P(W)$ and $X = (T_c A^{\#}) Y \in T_b P(V)$ and write $\mathbf{Y} := \mathbf{G}_f(Y)$ and $\mathbf{X} := \mathbf{G}_e(X)$. The identities

$$\mathbf{G}_e(b) = \mathbf{A}^{\mathsf{T}} \mathbf{G}_f(c) \, \mathbf{A} \quad \text{and} \quad \mathbf{X} = \mathbf{A}^{\mathsf{T}} \mathbf{Y} \, \mathbf{A}$$

together with (4.1) lead to

$$
\begin{aligned}
g_P(X, X)|_b &= \mathrm{tr}\left(\mathbf{G}_e(b)^{-1}\mathbf{X}^{\mathsf{T}}\,\mathbf{G}_e(b)^{-1}\mathbf{X}\right) \\
&= \mathrm{tr}\left((\mathbf{A}^{-1}\mathbf{G}_f(c)^{-1}\mathbf{A}^{-\mathsf{T}})(\mathbf{A}^{\mathsf{T}}\mathbf{Y}^{\mathsf{T}}\,\mathbf{A})(\mathbf{A}^{-1}\mathbf{G}_f(c)^{-1}\mathbf{A}^{-\mathsf{T}})(\mathbf{A}^{\mathsf{T}}\mathbf{Y}\,\mathbf{A})\right) \\
&= \mathrm{tr}\left(\mathbf{A}^{-1}(\mathbf{G}_f(c)^{-1}\mathbf{Y}^{\mathsf{T}}\mathbf{G}_f(c)^{-1}\mathbf{Y})\,\mathbf{A}\right) \\
&= \mathrm{tr}\left(\mathbf{G}_f(c)^{-1}\mathbf{Y}^{\mathsf{T}}\mathbf{G}_f(c)^{-1}\mathbf{Y}\right) = g_P(Y, Y)|_c. \qquad \square
\end{aligned}
$$

In particular, the group GL(V) acts smoothly from the right on $P(V)$ via pullback and we obtain:

Corollary 4.18 *The metric g_P is invariant under the action of GL(V) on $P(V)$.*

Theorem 4.19 *Every GL(V)-invariant Riemannian metric on $P(V)$ can be written as*

$$g_{\mu,\lambda}(X, Y)|_b := 2\mu\, g_P(X, Y)|_b + \lambda\,(\mathrm{vol}^{\#} g_{\mathrm{Vol}})(X, Y)|_b,$$

with some parameters $\mu > 0$ and $\lambda > -\frac{2\mu}{\dim(V)}$. Here, $X, Y \in T_b P(V)$ are tangent vectors at $b \in P(V)$ and $\mathrm{vol}\colon P(V) \to \mathrm{Vol}(V)$ is the Riemannian density operator (see Lemma 4.15).

PROOF. That $g_{\mu,\lambda}$ is GL(V)-invariant follows from Lemma 4.17. We show that $g_{\mu,\lambda}$ is positive definite for $\mu > 0$, $\lambda > -\frac{2\mu}{k}$, where $k = \dim(V)$:

For $X \in T_b P(V)$ with $g_P(X, X) \leq 0$ choose a basis e of V such that $\mathbf{G}_e(b) = \mathbf{I}$ is the identity matrix and $\mathbf{G}_e(X) = \mathbf{diag}(\lambda_1, \ldots, \lambda_k)$ is diagonal. Then define the function $f \colon \mathbb{R}^k \to \mathbb{R}$ by

$$f(x_1, \ldots, x_n) := \mu \sum_{i=1}^{k} x_i^2 + \tfrac{\lambda}{2}\Big(\sum_{i=1}^{k} x_i\Big)^2$$

and observe $f(\lambda_1, \ldots, \lambda_n) = \frac{1}{2} g_{\mu,\lambda}(X, X)$. Note that f is a quadratic functional with $f(0, \ldots, 0) = 0$, $Df(0, \ldots, 0) = 0$ and its Hessian is given by

$$\mathrm{Hess}_0(f) = \begin{pmatrix} 2\mu + \lambda & \lambda & \cdots & \lambda \\ \lambda & \ddots & & \lambda \\ \vdots & & \ddots & \vdots \\ \lambda & \cdots & \lambda & 2\mu + \lambda \end{pmatrix}.$$

The m-th principal minor of $\mathrm{Hess}_0(f)$ is $(m\lambda + 2\mu)\mu^{k-1}$. Thus by Sylvester's criterion, $\mathrm{Hess}_0(f)$ is positive definite (and f is strictly convex) if and only if $\mu > 0$ and $\lambda > -\frac{2\mu}{k} > -\frac{2\mu}{k-1} > \cdots > -\frac{2\mu}{1}$. Hence $f(\lambda_1, \ldots, \lambda_n) \leq 0$ is only possible for $\lambda_1 = \cdots = \lambda_k = 0$. This implies $X = 0$.

Now, let g be an arbitrary GL(V)-invariant metric on $P(V)$. Fix $b \in P(V)$. Then $g|_b$ has to be invariant under the stabilizer O(V, b) of b. Choose an orthonormal basis e of V with respect to b such that $\mathbf{G}_e(b) = \mathbf{I}$ is the identity matrix. One has for all symmetric $\mathbf{X}, \mathbf{Y} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$:

$$g_P|_b(\mathbf{G}_e^{-1}(\mathbf{X}), \mathbf{G}_e^{-1}(\mathbf{Y})) = \langle \mathbf{X}, \mathbf{Y} \rangle_{\mathrm{Frob}}.$$

In the same vein, define

$$G(\mathbf{X}, \mathbf{Y}) := g|_b(\mathbf{G}_e^{-1}(\mathbf{X}), \mathbf{G}_e^{-1}(\mathbf{Y}))$$

for all symmetric $\mathbf{X}, \mathbf{Y} \in \mathrm{Mat}_{k \times k}(\mathbb{R})$. Since $g|_b$ is O(V, b)-invariant, G is O(n)-invariant. By diagonalizing the symmetric matrix $\mathbf{X}$, one realizes that $G(\mathbf{X}, \mathbf{X})$ is a symmetric quadratic polynomial in the eigenvalues of $\mathbf{X}$. Thus, there are $\mu, \lambda \in \mathbb{R}$ with $G(\mathbf{X}, \mathbf{X}) = 2\mu \, \mathrm{tr}(\mathbf{X}\mathbf{X}) + \lambda \, \mathrm{tr}(\mathbf{X}) \, \mathrm{tr}(\mathbf{X})$. The polarization formula implies

$$G(\mathbf{X}, \mathbf{Y}) = 2\mu \langle \mathbf{X}, \mathbf{Y} \rangle_{\mathrm{Frob}} + \lambda \langle \mathbf{X}, \mathbf{I} \rangle_{\mathrm{Frob}} \langle \mathbf{I}, \mathbf{Y} \rangle_{\mathrm{Frob}}.$$

By pullback along $\mathbf{G}_e$, we obtain

$$\begin{aligned} g(X, Y)|_b &= 2\mu \, g_P(X, Y)|_b + \lambda \, g_P(X, b) \, g_P(b, Y) \\ &= 2\mu \, g_P(X, Y)|_b + \lambda \, (\mathrm{vol}^\# \, g_{\mathrm{Vol}})(X, Y)|_b \end{aligned}$$

The inequality $\lambda > -\frac{2\mu}{k}$ follows from the fact that $g|_b$ is positive definite and from the considerations at the beginning of this proof. □

Remark 4.20 As we will see in Remark 8.7, the parameters μ, λ can be interpreted as Lamé coefficients.

Lemma 4.21 *The Levi-Civita connection of $g_{\mu,\lambda}$ is given by that of g_P.*

PROOF. Let ∇ be the covariant derivative of g_P. From Corollary 4.5 we have that the vector field $W \in \mathfrak{X}(P(V))$, $W|_b = b$ is covariantly constant. For $X, Y, Z \in \mathfrak{X}(P(V))$, we compute

$$
\begin{aligned}
Xg_{\mu,\lambda}(Y,Z) &= 2\mu\, Xg_P(Y,Z) + \lambda\, X(g_P(Y,W)g_P(W,Z)) \\
&= 2\mu\, g_P(\nabla_X Y, Z) + \mu\, g_P(Y, \nabla_X Z) \\
&\quad + \lambda\, (g_P(\nabla_X Y, W)g_P(W,Z)) + \lambda\, (g_P(Y,W)g_P(W,\nabla_X Z)) \\
&= g_{\mu,\lambda}(\nabla_X Y, Z) + g_{\mu,\lambda}(Y, \nabla_X Z).
\end{aligned}
$$

This shows $\nabla g_{\mu,\lambda} = 0$, which suffices to prove the statement. □

Corollary 4.22 *The geodesics with respect to g_P and $g_{\mu,\lambda}$ coincide.*

Lemma 4.23 *Let $b, c \in P(V)$. Denote the geodesic distance with respect to $g_{\mu,\lambda}$ by $d_{\mu,\lambda}$. Then one has*

$$
d_{\mu,\lambda}^2(b,c) = 2\mu\, d_P^2(b,c) + \lambda\, d_{\mathrm{Vol}}^2(\mathrm{vol}_b, \mathrm{vol}_g),
$$

where vol_b, vol_c denote the densities defined by b, c respectively.

PROOF. Let $\gamma\colon [0,1] \to P(V)$ be the unique geodesic from b to c. According to Remark 4.11, we have $\dot\gamma(0) = -H_c|_b$ and

$$
d_{\mu,\lambda}^2(b,c) = |H_c|_b|_{g_{\mu,\lambda}}^2 = 2\mu\, g_P(H_c|_b, H_c|_b) + \lambda\, g_P(H_c|_b, b)^2.
$$

Choosing a b-orthonormal basis e of V with $\mathbf{G}_e(c) = \mathbf{diag}(\lambda_1, \ldots, \lambda_k)$, one obtains

$$
\begin{aligned}
|g_P(H_c|_b, b)| &= |\mathrm{tr}\log(\mathbf{diag}(\lambda_1, \ldots, \lambda_k))| = |\log(\det \mathbf{diag}(\lambda_1, \ldots, \lambda_k))| \\
&= \left|\log\left(\tfrac{\mathrm{vol}_c}{\mathrm{vol}_b}\right)\right| = d_{\mathrm{Vol}}(\mathrm{vol}_b, \mathrm{vol}_c).
\end{aligned}
$$
□

5 Lipschitz Immersions

We need a suitable shape space of immersed manifolds for our applications to elasticity theory and minimal surfaces theory. Therefore, we define a space of parameterized (locally) Lipschitz immersions and equip it with a reparameterization-invariant distance. This distance is in terms of zeroth and first derivatives of immersions. It descends to a distance on the *shape space*, i.e., the quotient space of unparameterized, immersed manifolds.[1]

5.1 Basic Definitions

Throughout this chapter Σ will be a compact, k-dimensional smooth manifold with boundary. Let g be a smooth Riemannian metric on Σ. With a slight abuse of notation, we denote with vol_g not only the Riemannian density induced by g, but also the complete measure induced by it. Define the locally trivial fiber bundle $\pi\colon P_\Sigma \to \Sigma$ by $P_\Sigma|_x := P(T_x\Sigma)$ for all $x \in \Sigma$ and equip the fibers with the distance function d_P. Define

$$\mathcal{P}(\Sigma) := \left\{ b\colon \Sigma \to P_\Sigma \ \mathrm{vol}_g\text{-measurable.} \ \middle| \ \begin{matrix} \pi \circ b = \mathrm{id}_\Sigma \ \mathrm{vol}_g\text{-a. e.,} \\ \mathrm{ess\,sup}_{x\in\Sigma} d_P(b|_x, g|_x) < \infty \end{matrix} \right\} / \sim$$

with the equivalence relation $b_1 \sim b_2$ if $b_1 = b_2$ holds vol_g-almost everywhere (vol_g-a. e.). We introduce the distance $d_{\mathcal{P}}$ on $\mathcal{P}(\Sigma)$ by

$$d_{\mathcal{P}}(b_1, b_2) := \mathrm{ess\,sup}_{x\in\Sigma} d_P(b_1|_x, b_2|_x).$$

Note that neither the space $\mathcal{P}(\Sigma)$ nor the distance $d_{\mathcal{P}}$ depend on the choice of g.

Definition 5.1 For $f \in W^{1,\infty}(\Sigma; \mathbb{R}^m)$, the derivative Tf exists at almost every point $x \in \Sigma$ and we obtain an almost everywhere defined $f^\# g_0$. We use this construction to define the *space of Lipschitz immersions* as

$$\mathrm{Imm}(\Sigma; \mathbb{R}^m) := \{ f \in W^{1,\infty}(\Sigma; \mathbb{R}^m) \mid f^\# g_0 \in \mathcal{P}(\Sigma) \}$$

---

[1]Contrary to the meaning we associate to it, the term "shape space" is frequently used for certain classes of subsets of $\mathbb{R}^3$ *modulo the action of the Euclidean group.*

and equip it with the distance

$$d_{\mathrm{Imm}}(f_1, f_2) := \|f_1 - f_2\|_{L^\infty} + d_{\mathcal{P}}(f_1^\# g_0, f_2^\# g_0) \quad \text{for } f_1, f_2 \in \mathrm{Imm}(\Sigma; \mathbb{R}^m).$$

Remark 5.2 Our notion of Lipschitz immersions may seem quite similar to the one established by Rivière in [33]. Note however that Rivière's notion depends on the boundedness of $\mathrm{vol}_{f^\# g}$ only, while ours depends on the boundedness of the full distortion tensor $f^\# g$.

The distance $d_{\mathcal{P}}$ involves only little first order information about f_1 and f_2 in the case $k < m$. Even in the case $k = m$, this distance is insensitive to the orientation of the tangent spaces. Hence we add a distance between oriented tangent planes of f_1 and f_2. Since we also would like to treat non-orientable manifolds Σ, we use the following construction:

Let $\hat{\pi} \colon \hat{\Sigma} \to \Sigma$ be the *orientation covering* of Σ (see, e.g., [26, p. 330]). Note that $\hat{\pi}$ is a double covering of Σ so $T_x\hat{\pi} \colon T_x\hat{\Sigma} \to T_{\hat{\pi}(x)}\Sigma$ is a linear isomorphism. Therefore, any Riemannian metric g on Σ can be pulled pack to $\hat{\Sigma}$ along $\hat{\pi}$. The fiber $\hat{\Sigma}|_x := \hat{\pi}^{-1}(\{x\})$ consists precisely of the two possible orientations on $T_x\Sigma$. Thus, $\hat{\Sigma}$ is oriented in a canonical way: The orientation of $T_p\hat{\Sigma} \cong T_{\hat{\pi}(p)}\Sigma$ is p itself. Every $f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m)$ can be pulled back along $\hat{\pi}$ leading to the isometric embedding

$$\hat{\pi}^\# \colon (\mathrm{Imm}(\Sigma; \mathbb{R}^m), d_{\mathrm{Imm}}) \to (\mathrm{Imm}(\hat{\Sigma}; \mathbb{R}^m), d_{\mathrm{Imm}}), \qquad f \mapsto f \circ \hat{\pi}.$$

Let $\widehat{\mathrm{Gr}}_k(\mathbb{R}^m)$ be the oriented Grassmannian, i.e., the smooth manifold of oriented k-dimensional vector subspaces in $\mathbb{R}^m$. Every $\hat{f} \in \mathrm{Imm}(\hat{\Sigma}; \mathbb{R}^m)$ induces a *Gauss map* $\tau(\hat{f}) \in L^\infty(\hat{\Sigma}; \widehat{\mathrm{Gr}}_k(\mathbb{R}^m))$ via

$$\tau(\hat{f})|_p := \begin{cases} (\mathrm{im}(\mathrm{d}\hat{f}|_p), \hat{f}_\# p), & \mathrm{d}\hat{f}|_p \text{ exists and is injective,} \\ \text{undefined,} & \text{else,} \end{cases}$$

leading to the mapping $\tau \colon \mathrm{Imm}(\hat{\Sigma}; \mathbb{R}^m) \to L^\infty(\hat{\Sigma}; \widehat{\mathrm{Gr}}_k(\mathbb{R}^m))$. Here, $\hat{f}_\# p$ denotes the orientation on $\mathrm{im}(\mathrm{d}\hat{f}|_p)$ that makes $\mathrm{d}\hat{f}|_p \colon (T_p\hat{\Sigma}, p) \to (\mathrm{im}(\mathrm{d}\hat{f}|_p), \hat{f}_\# p)$ orientation-preserving. Any $\mathrm{GL}(\mathbb{R}^m)$-invariant Riemannian metric g_{Gr} on $\mathrm{Gr}_k(\mathbb{R}^m)$ can be lifted to a Riemannian metric $g_{\widehat{\mathrm{Gr}}}$ on $\widehat{\mathrm{Gr}}_k(\mathbb{R}^m)$, leading to a geodesic distance $d_{\widehat{\mathrm{Gr}}}$. Via the Gauss map, we may define the following augmented metric d_{Imm}^τ on $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$:

$$d_{\mathrm{Imm}}^\tau(f_1, f_2) := d_{\mathrm{Imm}}(f_1, f_2) + \underset{p \in \hat{\Sigma}}{\mathrm{ess\,sup}}\, d_{\widehat{\mathrm{Gr}}}(\tau(f_1 \circ \hat{\pi})|_p, \tau(f_2 \circ \hat{\pi})|_p),$$

for every $f_1, f_2 \in \mathrm{Imm}(\Sigma; \mathbb{R}^m)$.

5.2 Diffeomorphism Group

Let (Σ, g) be a compact, smooth Riemannian manifold. We define the group of *Lipschitz diffeomorphisms* by

$$\mathrm{Diff}(\Sigma) := \{\, \varphi \in W_g^{1,\infty}(\Sigma; \Sigma) \mid \varphi \text{ is a bi-Lipschitz homeomorphism}\,\},$$

where the group structure is given by

$$\mu\colon \mathrm{Diff}(\Sigma) \times \mathrm{Diff}(\Sigma) \to \mathrm{Diff}(\Sigma), \qquad (\varphi, \psi) \mapsto \varphi \circ \psi,$$
$$\iota\colon \mathrm{Diff}(\Sigma) \to \mathrm{Diff}(\Sigma), \qquad \varphi \mapsto \varphi^{-1}.$$

Note that the space $\mathrm{Diff}(\Sigma)$ does not depend on the choice of the Riemannian metric g. The existence of global Lipschitz constants for φ and φ^{-1} implies $\varphi^\# g$, $(\varphi^{-1})^\# g \in \mathcal{P}(\Sigma)$ for every Riemannian metric g on Σ. To give $\mathrm{Diff}(\Sigma)$ a topology, we define the distance

$$d_{\mathrm{Diff}}(\varphi_1, \varphi_2) := \operatorname*{ess\,sup}_{x \in \Sigma} d_g(\varphi_1(x), \varphi_2(x)) + \operatorname*{ess\,sup}_{x \in \Sigma} d_g(\varphi_1^{-1}(x), \varphi_2^{-1}(x))$$
$$+ d_P(\varphi_1^\# g, \varphi_2^\# g) + d_P((\varphi_1^{-1})^\# g, (\varphi_2^{-1})^\# g)$$

for all φ_1, $\varphi_2 \in \mathrm{Diff}(\Sigma)$. By construction, $\iota\colon (\mathrm{Diff}(\Sigma), d_{\mathrm{Diff}}) \to (\mathrm{Diff}(\Sigma), d_{\mathrm{Diff}})$ is an isometric isomorphism, thus Lipschitz continuous. That μ is locally Lipschitz continuous can be checked easily with the triangle inequality, the chain rule, and the fact $\|T\varphi\|_{L^\infty}, \|T(\varphi^{-1})\|_{L^\infty} \le d_{\mathrm{Diff}}(\varphi, \mathrm{id}_\Sigma) < \infty$. Thus, $(\mathrm{Diff}(\Sigma), \mu, \iota)$ is a topological group. Note that $\mathrm{Diff}_0(\Sigma) := \{\, \varphi \in \mathrm{Diff}(\Sigma) \mid \varphi|_{\partial\Sigma} = \mathrm{id}_{\partial\Sigma}\,\}$ is a closed normal subgroup of $\mathrm{Diff}(\Sigma)$. The topological group $\mathrm{Diff}(\Sigma)$ acts continuously from the right on $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$ via

$$L_\varphi\colon \mathrm{Imm}(\Sigma; \mathbb{R}^m) \to \mathrm{Imm}(\Sigma; \mathbb{R}^m), \qquad f \mapsto f \circ \varphi,$$

for all $\varphi \in \mathrm{Diff}(\Sigma)$.

For $\gamma \in \mathrm{Imm}(\partial\Sigma; \mathbb{R}^m)$, one may also define the space of immersions under boundary conditions:

$$\mathrm{Imm}_\gamma(\Sigma; \mathbb{R}^m) := \{\, f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m) \mid f|_{\partial\Sigma} = \gamma\,\}.$$

Note however, that for arbitrary $f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m)$, the restriction $f|_{\partial\Sigma}$ need not be Lipschitz continuous. The action of $\mathrm{Diff}_0(\Sigma)$ on $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$ restricts to an action on $\mathrm{Imm}_\gamma(\Sigma; \mathbb{R}^m)$.

Finally, we point out that $\mathrm{Diff}(\Sigma)$ acts on $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$ through isometries with respect to both d_{Imm} and d_{Imm}^τ—a fact that we utilize to analyze the quotient metric of d_{Imm} and d_{Imm}^τ.

5.3 Quotient Space

One may consider the following quotient spaces

$$\mathrm{Imm}(\Sigma; \mathbb{R}^m)/\,\mathrm{Diff}(\Sigma), \quad \mathrm{Imm}(\Sigma; \mathbb{R}^m)/\,\mathrm{Diff}_0(\Sigma), \quad \text{and} \quad \mathrm{Imm}_\gamma(\Sigma; \mathbb{R}^m)/\,\mathrm{Diff}_0(\Sigma),$$

where $\gamma \in \mathrm{Imm}(\partial\Sigma; \mathbb{R}^m)$. Let $(C, \mathcal{G})$ be one of the pairs $(\mathrm{Imm}(\Sigma; \mathbb{R}^m), \mathrm{Diff}(\Sigma))$, $(\mathrm{Imm}(\Sigma; \mathbb{R}^m), \mathrm{Diff}_0(\Sigma))$, or $(\mathrm{Imm}_\gamma(\Sigma; \mathbb{R}^m), \mathrm{Diff}_0(\Sigma))$. Denote by $\mathcal{X} := C/\mathcal{G}$ the quotient space and by $\Pi \colon C \to \mathcal{X}$ the canonical map. The metrics d_{Imm}, d_{Imm}^τ on C descend to quotient semi-metrics $d_{\mathcal{X}}$, $d_{\mathcal{X}}^\tau$ on the quotient $\mathcal{X}$ (see [8, p. 62] for a definition of the quotient semi-metric). In general, these quotient semi-metrics may be quite cumbersome to work with. In our case, the group $\mathcal{G}$ acts on C through isometries, hence the quotient semi-metrics satisfy

$$d_{\mathcal{X}}(\Pi(f_1), \Pi(f_2)) = \inf_{\varphi \in \mathcal{G}} d_{\mathrm{Imm}}(f_1, f_2 \circ \varphi) \text{ and}$$

$$d_{\mathcal{X}}^\tau(\Pi(f_1), \Pi(f_2)) = \inf_{\varphi \in \mathcal{G}} d_{\mathrm{Imm}}^\tau(f_1, f_2 \circ \varphi),$$

where $f_1, f_2 \in C$.

Lemma 5.3 *Let Σ be a compact smooth manifold with boundary. Then both $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{X}, d_{\mathcal{X}}^\tau)$ are metric spaces.*

PROOF. It suffices to show that $d_{\mathcal{X}}$ is a metric. Let $f, h \in C$ and let $\varphi_n \in \mathcal{G}$ be a sequence with $d_{\mathrm{Imm}}(f, h \circ \varphi_n) \to 0$. We have to show that there is a $\varphi \in \mathcal{G}$ with $f = \varphi^\# h$.

We start by choosing a smooth Riemannian metric g on Σ such that the boundary (if it exists) is totally geodesic. This way, for every point $x \in \Sigma$, every neighborhood U of x contains a geodesically convex neighborhood of x. Such a Riemannian metric can be constructed, for example, by choosing a cylinder metric on a smooth collar of Σ and extending it smoothly.[2]

Observe that $h_n := h \circ \varphi_n$ converges uniformly to f. Moreover, being convergent, $h_n^\# g_0$ is a bounded sequence in $\mathcal{P}_\Sigma$. Hence there is some $\Lambda_0 \geq 0$ with

$$\|\mathrm{d}h_n^{\dagger_g}\|_{L_g^\infty}, \ \|\mathrm{d}h_n\|_{L_g^\infty} \leq \Lambda_0.$$

The chain rule for weak derivatives of Lipschitz mappings yields

$$\mathrm{d}h_n|_x = \mathrm{d}h|_{\varphi_n(x)} \cdot T_x\varphi_n \quad \text{and}$$

$$\mathrm{d}h|_x = \mathrm{d}h_n|_{\varphi_n^{-1}(x)} \cdot T_x(\varphi_n^{-1}),$$

---

[2] A *smooth collar* of Σ is a smooth embedding $\Phi \colon [0, 1[\times \partial\Sigma \to \Sigma$ such that $\Phi(0, x) = x$ holds for all $x \in \partial\Sigma$. Every paracompact smooth manifold with boundary has a smooth collar.

hence one obtains $T_x\varphi_n = (\mathrm{d}h|_{\varphi_n(x)})^{\dagger_g} \cdot \mathrm{d}h_n|_x$ for almost all $x \in \Sigma$. Thus, there is a $\Lambda \geq 0$ with

$$\|T\varphi_n\|_{L_g^\infty} \leq \|(\mathrm{d}h)^{\dagger_g}\|_{L_g^\infty} \cdot \|\mathrm{d}h_n\|_{L_g^\infty} \leq \Lambda \quad \text{and}$$

$$\|(T\varphi_n)^{-1}\|_{L_g^\infty} \leq \|(\mathrm{d}h_n)^{\dagger_g}\|_{L_g^\infty} \cdot \|\mathrm{d}h\|_{L_g^\infty} \leq \Lambda,$$

showing that the families $(\varphi_n)_{n\in\mathbb{N}}$ and $(\varphi_n^{-1})_{n\in\mathbb{N}}$ are equicontinuous. Because Σ is a compact metric space, the families $(\varphi_n)_{n\in\mathbb{N}}$ and $(\varphi_n^{-1})_{n\in\mathbb{N}}$ are also pointwise relatively compact. Thus, the Arzelà-Ascoli theorem (see, e.g., [28, Theorem 47.1]) implies the existence of a subsequence (which we also denote by $(\varphi_n)_{n\in\mathbb{N}}$) such that both $\varphi_n \to \varphi$ and $\varphi_n^{-1} \to \varphi^{-1}$ converge in the compact-open topology on $C(\Sigma; \Sigma)$.

Up to now, we know that $\varphi \colon \Sigma \to \Sigma$ is a homeomorphism (probably fixing the boundary) and that $f = h \circ \varphi$. We are left to show that both φ are φ^{-1} are Lipschitz.

Let $V_1, \ldots, V_\alpha$ with some $\alpha \in \mathbb{N}$ be a covering of Σ by open, relatively compact, and geodesically convex sets. Choose a covering $U_1, \ldots, U_\beta$ with some $\beta \in \mathbb{N}$ of Σ by open, relatively compact and geodesically convex sets such that each $\varphi(U_i)$ is contained in some V_j. Then one has for all $x, y \in U_i$:

$$d_\Sigma(\varphi(x), \varphi(y)) \leq d_g(\varphi(x), \varphi_n(x)) + d_g(\varphi_n(x), \varphi_n(y)) + d_g(\varphi_n(y), \varphi(y))$$
$$\leq d_g(\varphi(x), \varphi_n(x)) + \|T\varphi_n\|_{L_g^\infty} d_g(x, y) + d_g(\varphi_n(y), \varphi(y)).$$

Applying $\limsup_{n\to\infty}$ yields $\|T\varphi\|_{L_g^\infty} \leq \Lambda$, hence φ is Lipschitz continuous. The same argument shows that φ^{-1} is Lipschitz continuous, too. $\qquad\square$

5.4 Embedding Theorems

Here we gather several technical lemmata for later use. Moreover, we analyze the relationship between $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$ and $W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$.

Lemma 5.4 *Fix b, $g \in P(V)$ and let $X \in T_b P(V) = \mathrm{Sym}(V)$ with $|X|_g < \mathrm{e}^{-d_P(b,g)}$. Then $b + X$ is also contained in $P(V)$ and one has*

$$d_P(b, b + X) \leq \mathrm{e}^{d_P(g,b)} |X|_g.$$

In particular, $P(V)$ is open in $\mathrm{Sym}(V)$.

Proof. Choose a g-orthonormal basis e of V and define $\mathbf{B} := \mathbf{G}_e(b)$, $\mathbf{X} := \mathbf{G}_e(X)$. Let $0 < \lambda_1 \leq \cdots \leq \lambda_k$ be the eigenvalues of $\mathbf{B}$. Observe that

$$\|\mathbf{B}^{-1}\| = \frac{1}{\lambda_1} = \exp\left(|\log(\tfrac{1}{\lambda_1})|\right) \leq \exp\sqrt{\sum_{i=1}^k |\log(\tfrac{1}{\lambda_1})|^2} = \mathrm{e}^{d_P(b,g)}.$$

The estimate

$$\|\mathbf{B}^{-\frac{1}{2}}\mathbf{X}\mathbf{B}^{-\frac{1}{2}}\| \le \|\mathbf{B}^{-1}\|\,\|\mathbf{X}\| \le \|\mathbf{B}^{-1}\|\,|\mathbf{X}| \le \mathrm{e}^{d_P(b,g)} \cdot |X|_g < 1$$

shows that $\mathbf{B} + t\,\mathbf{X} = \mathbf{B}^{\frac{1}{2}}(\mathbf{I}_n + t\,\mathbf{B}^{-\frac{1}{2}}\mathbf{X}\mathbf{B}^{-\frac{1}{2}})\mathbf{B}^{\frac{1}{2}}$ is invertible for all $t \in [0,1]$. This implies that $\mathbf{B} + \mathbf{X}$ and thus $b + X$ are positive definite. Now, we have

$$d_P(b, b + X) = |\log(\mathbf{B}^{-\frac{1}{2}}(\mathbf{B} + \mathbf{X})\mathbf{B}^{-\frac{1}{2}})| \le |\log(\mathbf{I}_k + \mathbf{B}^{-\frac{1}{2}}\mathbf{X}\mathbf{B}^{-\frac{1}{2}})|$$

$$\le |\mathbf{B}^{-\frac{1}{2}}\mathbf{X}\mathbf{B}^{-\frac{1}{2}}| = |\mathbf{B}^{-1}\mathbf{X}| \le |\mathbf{B}^{-1}|\,|\mathbf{X}| \le \mathrm{e}^{d_P(g,b)}\,|X|_g,$$

from which the stated estimate follows. □

Lemma 5.5 *Let (V_i, g_i) be finite-dimensional Euclidean spaces for $i = 1, 2$ and let $A, U \in \mathrm{Hom}(V_1, V_2)$ with A injective and U fulfilling*

$$|U|_{g_1, g_2} < \sqrt{\mathrm{e}^\ell + \mathrm{e}^{-\ell}} - \sqrt{\mathrm{e}^\ell},$$

where $\ell := d_P(g_1, A^\# g_2)$. Then $(A + U)^\# g_2$ is also contained in $P(V_1)$ and one has

$$d_P(A^\# g_2, (A + U)^\# g_2) \le (\sqrt{\mathrm{e}^{3\ell} + \mathrm{e}^\ell} + \sqrt{\mathrm{e}^{3\ell}}) \cdot |U|_{g_1, g_2}.$$

PROOF. We use the preceding lemma with $g = g_1$, $b := A^\# g_2$, and $X := (A + U)^\# g_2 - b$. Choose g_i-orthonormal bases e_i of V_i for $i = 1, 2$ and write $\mathbf{B} := \mathbf{G}_{e_1}(A^\# g_2)$ and $\mathbf{X} := \mathbf{G}_{e_1}(X)$. Let $\mathbf{A}$ and $\mathbf{U}$ be the matrix representations of A and U, respectively, with respect to these chosen bases. Since $\mathbf{X} = \mathbf{A}^\mathsf{T}\mathbf{U} + \mathbf{U}^\mathsf{T}\mathbf{A} + \mathbf{U}^\mathsf{T}\mathbf{U}$ one obtains

$$|X|_g = |\mathbf{X}| \le 2\|\mathbf{A}^\mathsf{T}\mathbf{A}\|^{\frac{1}{2}}|\mathbf{U}| + |\mathbf{U}|^2 \le 2\,\sqrt{\mathrm{e}^\ell} \cdot |U|_{g_1, g_2} + |U|^2_{g_1, g_2}$$

$$< (\sqrt{\mathrm{e}^\ell + \mathrm{e}^{-\ell}} + \sqrt{\mathrm{e}^\ell}) \cdot |U|_{g_1, g_2} < \mathrm{e}^{-\ell},$$

whenever $|U|_{g_1, g_2} < \sqrt{\mathrm{e}^\ell + \mathrm{e}^{-\ell}} - \sqrt{\mathrm{e}^\ell}$. Finally, one has

$$d_P(b, b + X) \le (\sqrt{\mathrm{e}^{3\ell} + \mathrm{e}^\ell} + \sqrt{\mathrm{e}^{3\ell}}) \cdot |U|_{g_1, g_2}$$

by Lemma 5.4. □

Corollary 5.6 *Let $f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m)$ and $g \in \mathcal{P}(\Sigma)$. Then the ball $B(f, r_g(f))$ with respect to the norm $\|\cdot\|_{W_g^{1,\infty}}$ of radius $r_g(f) := \sqrt{\mathrm{e}^\ell + \mathrm{e}^{-\ell}} - \sqrt{\mathrm{e}^\ell}$ is also contained in $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$, where $\ell := d_P(g, f^\# g_0)$. In particular $\mathrm{Imm}(\Sigma; \mathbb{R}^m) \subset W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$ is open.*

Corollary 5.7 *Let g be a Riemannian metric on Σ and let $d_{W_g^{1,\infty}}$ be the distance induced by $\|\cdot\|_{W_g^{1,\infty}}$. Then the identity mapping*

$$\mathrm{id}\colon (\mathrm{Imm}(\Sigma;\mathbb{R}^m), d_{W_g^{1,\infty}}) \to (\mathrm{Imm}(\Sigma;\mathbb{R}^m), d_{\mathrm{Imm}}^\tau)$$

is locally Lipschitz continuous.

PROOF. Fix $f \in \mathrm{Imm}(\Sigma;\mathbb{R}^m)$ and let $h \in \mathrm{Imm}(\Sigma;\mathbb{R}^m)$ with $\|f - h\|_{W_g^{1,\infty}} \le r_g(f)$. By Lemma 5.5, we have

$$d_{\mathcal{P}}(f^\# g_0, h^\# g_0) \le (\sqrt{\mathrm{e}^{3\ell} + \mathrm{e}^{\ell}} + \sqrt{\mathrm{e}^{3\ell}}) \cdot \|\mathrm{d}f - \mathrm{d}h\|_{L_g^\infty},$$

showing the local Lipschitz continuity of id with respect to d_{Imm}.

The local Lipschitz continuity of id with respect to d_{Imm}^τ can be deduced from the fact that $\tau(f)$ is a lift of $p(f) \circ \hat{\pi}$ along the canonical covering $\kappa\colon \widehat{\mathrm{Gr}}_k(\mathbb{R}^m) \to \mathrm{Gr}_k(\mathbb{R}^m)$

where $p(f)$ can be expressed as the orhtoprojector-valued mapping

$$p(f) = \mathrm{d}f(\mathrm{d}f^{*_g}\mathrm{d}f)^{-1}\mathrm{d}f^{*_g}.$$

The map $p\colon (\mathrm{Imm}(\Sigma;\mathbb{R}^m), d_{W_g^{1,\infty}}) \to L^\infty(\Sigma; \mathrm{Gr}_k(\mathbb{R}^m))$ is locally Lipschitz continuous and κ is a Riemannian submersion. Thus, the map $f \mapsto \tau(f)$ is also locally Lipschitz continuous. $\square$

Definition 5.8 Let $\mathcal{B} \subset L_g^\infty(\Sigma;\mathbb{R}^m)$ be a Banach space. We say, $\mathcal{B}$ is *compactly embedded* into $W_g^{1,\infty}(\Sigma;\mathbb{R}^m)$, if the canonical embedding $\mathcal{B} \hookrightarrow L_g^\infty(\Sigma;\mathbb{R}^m)$ has its image contained in $W_g^{1,\infty}(\Sigma;\mathbb{R}^m)$ and the induced linear embedding $i\colon \mathcal{B} \hookrightarrow W_g^{1,\infty}(\Sigma;\mathbb{R}^m)$ is compact.

Lemma 5.9 *Let $i\colon \mathcal{B} \hookrightarrow W_g^{1,\infty}(\Sigma;\mathbb{R}^m)$ be a compactly embedded Banach space. Then*

$$\mathcal{A}^s := \{ f \in \mathrm{Imm}(\Sigma;\mathbb{R}^m) \cap \mathcal{B} \mid d_{\mathcal{P}}(g, f^\# g_0) \le s, \ \|f\|_{\mathcal{B}} \le s \}$$

is a compact set in $(\mathrm{Imm}(\Sigma;\mathbb{R}^m), d_{\mathrm{Imm}}^\tau)$.

Proof. Let $f_n \in \mathcal{A}^s$ for all $n \in \mathbb{N}$. By compactness of i, there is an $f \in W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$ and a subsequence $(f_{n_l})_{l \in \mathbb{N}}$ with $f_{n_l} \to f$ in $W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$ as $l \to \infty$. Let $l \in \mathbb{N}$ be large enough such that

$$\|f_{n_l} - f\|_{W_g^{1,\infty}} < \sqrt{e^s + e^{-s}} - \sqrt{e^s}.$$

By Corollary 5.6 we have $f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m)$. Now, Corollary 5.7 implies $f_{n_l} \xrightarrow{l \to \infty} f$ with respect to d_{Imm}^τ. $\qquad\square$

Lemma 5.10 *Let $i \colon \mathcal{B} \hookrightarrow W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$ be a continuously embedded Banach space. Then the space $\mathrm{Imm}(\Sigma; \mathbb{R}^m) \cap \mathcal{B}$ with the metric*

$$d_{\mathcal{B}}(f_1, f_2) := d_{\mathrm{Imm}}^\tau(f_1, f_2) + \|f_1 - f_2\|_{\mathcal{B}} \quad \text{for all } f_1, f_2 \in \mathrm{Imm}(\Sigma; \mathbb{R}^m) \cap \mathcal{B}$$

is complete.

5.5 Volume Functionals

The volume functionals on the space of Lipschitz immersions and on shape space are essential for the treatment of least volume problems. In this section, we establish their local Lipschitz continuity.

Throughout, let Σ be a compact, k-dimensional smooth manifold with boundary.

Lemma 5.11 *The* volume functional

$$\mathcal{J} \colon (\mathcal{P}(\Sigma), d_{\mathcal{P}}) \to \mathbb{R}, \qquad g \mapsto \int_\Sigma \mathrm{vol}_g$$

has its modulus of continuity $\omega_{\mathcal{J}}(g, t)$ bounded by $\mathcal{J}(g) \, e^{\sqrt{k}\,t} \sqrt{k}\,t$, i.e.,

$$|\mathcal{J}(b) - \mathcal{J}(g)| \leq \mathcal{J}(g) \, e^{\sqrt{k}\,d_{\mathcal{P}}(g,b)} \sqrt{k}\,d_{\mathcal{P}}(g, b) \quad \text{for all } g, b \in \mathcal{P}(\Sigma).$$

Proof. We abbreviate $\Lambda := \frac{\mathrm{vol}_b}{\mathrm{vol}_g}$. From Lemma 4.15, we know that

$$|\log(\Lambda(x))| = d_{\mathrm{Vol}}(\mathrm{vol}_g|_x, \mathrm{vol}_b|_x) \leq \sqrt{k}\,d_P(g|_x, b|_x) \leq \sqrt{k}\,d_{\mathcal{P}}(g, b).$$

Together with the estimate $|t - 1| \leq |\log(t)|\, e^{|\log(t)|}$ for all $t > 0$, we obtain

$$|\mathcal{J}(b) - \mathcal{J}(g)| = \left| \int_\Sigma (\mathrm{vol}_b - \mathrm{vol}_g) \right|$$

$$\leq \int_\Sigma |\Lambda - 1| \, \mathrm{vol}_g \leq \mathcal{J}(g) \, e^{\sqrt{k}\,d_{\mathcal{P}}(g,b)} \sqrt{k}\,d_{\mathcal{P}}(g, b). \qquad\square$$

Corollary 5.12 *Let $\omega_{\mathcal{F}}$ be the modulus of continuity of the* volume functional

$$\mathcal{F}\colon (\mathrm{Imm}(\Sigma; \mathbb{R}^m), d_{\mathrm{Imm}}) \to \mathbb{R}, \qquad f \mapsto \int_\Sigma \mathrm{vol}_{f^\# g_0}\,.$$

Then one has $\omega_{\mathcal{F}}(f, t) \le \mathcal{F}(f)\, \mathrm{e}^{\sqrt{k}\, t}\, \sqrt{k}\, t$.

Corollary 5.13 *With the nomenclature of Section 5.3, let $\Pi_\#\mathcal{F}\colon (\mathcal{X}, d_{\mathcal{X}}) \to \mathbb{R}$ be the pushforward of $\mathcal{F}$ along Π. Its modulus of continuity $\omega_{\Pi_\#\mathcal{F}}$ satisfies*

$$\omega_{\Pi_\#\mathcal{F}}(x, t) \le (\Pi_\#\mathcal{F})(x)\, \mathrm{e}^{\sqrt{k}\, t}\, \sqrt{k}\, t.$$

PROOF. For $x, y \in \mathcal{X}$ and $f \in \Pi^{-1}(x)$, $h \in \Pi^{-1}(y)$ observe

$$\begin{aligned}
|(\Pi_\#\mathcal{F})(y) - (\Pi_\#\mathcal{F})(x)| &= \inf_{\varphi \in \mathcal{G}} |\mathcal{F}(h \circ \varphi) - \mathcal{F}(f)| \\
&\le \inf_{\varphi \in \mathcal{G}} \mathcal{F}(f)\, \mathrm{e}^{\sqrt{k}\, d_{\mathrm{Imm}}(f, h \circ \varphi)}\, \sqrt{k}\, d_{\mathrm{Imm}}(f, h \circ \varphi) \\
&= (\Pi_\#\mathcal{F})(x)\, \mathrm{e}^{\sqrt{k}\, d_{\mathcal{X}}(x, y)}\, \sqrt{k}\, d_{\mathcal{X}}(x, y). \qquad \square
\end{aligned}$$

Lemma 5.14 *Let $k \le m$, let V be a k-dimensional real vector space, and let $\mathrm{Hom}_k(V; \mathbb{R}^m)$ be the open subset of $\mathrm{Hom}(V; \mathbb{R}^m)$ consisting of the linear mappings with full rank. The mapping*

$$F\colon \mathrm{Hom}_k(V; \mathbb{R}^m) \to \Omega(V), \quad A \mapsto \mathrm{vol}_{A^\# g_0}$$

is differentiable with derivative given by

$$\langle \mathrm{d}F|_A, U \rangle = \langle A, U \rangle_{A^\# g_0}\, \mathrm{vol}_{A^\# g_0} \quad \text{for all } A \in \mathrm{Hom}_k(V; \mathbb{R}^m),\ U \in \mathrm{Hom}(V; \mathbb{R}^m).$$

PROOF. Choose an arbitrary $b \in P(V)$, a b-orthonormal basis e of V and a g_0-orthonormal basis f of $\mathbb{R}^m$. With $\mathbf{B} = \mathbf{G}_e(b)$, $\mathbf{A}_{ij} = g_0(f_i, A\, e_j)$, and $\mathbf{U}_{ij} = g_0(f_i, U\, e_j)$ we have $F(A) = \det(\mathbf{A}^\mathsf{T}\mathbf{A})^{\frac{1}{2}}\, \mathrm{vol}_b$ by (4.2), thus

$$\langle \mathrm{d}F|_A, U \rangle = \tfrac{1}{2} \det(\mathbf{A}^\mathsf{T}\mathbf{A})^{-\frac{1}{2}} \langle \mathrm{d}\det|_{\mathbf{A}^\mathsf{T}\mathbf{A}}, \mathbf{A}^\mathsf{T}\mathbf{U} + \mathbf{U}^\mathsf{T}\mathbf{A} \rangle\, \mathrm{vol}_b\,.$$

The derivative of the determinant $\det\colon \mathrm{GL}_k(\mathbb{R}^k) \to \mathbb{R}$ is given by

$$\langle \mathrm{d}\det|_\mathbf{B}, \mathbf{X} \rangle = \det(\mathbf{B})\, \mathrm{tr}(\mathbf{B}^{-\mathsf{T}}\mathbf{X}) \quad \text{for all } \mathbf{B} \in \mathrm{GL}_k(\mathbb{R}),\ \mathbf{X} \in \mathrm{Mat}_{k \times k}(\mathbb{R}).$$

Thus one may compute

$$\begin{aligned}
\langle \mathrm{d}F|_A, U \rangle &= \tfrac{1}{2} \det(\mathbf{A}^\mathsf{T}\mathbf{A})^{-\frac{1}{2}} \det(\mathbf{A}^\mathsf{T}\mathbf{A})\, \mathrm{tr}\left((\mathbf{A}^\mathsf{T}\mathbf{A})^{-\mathsf{T}}(\mathbf{A}^\mathsf{T}\mathbf{U} + \mathbf{U}^\mathsf{T}\mathbf{A})\right) \mathrm{vol}_b \\
&= \tfrac{1}{2} \det(\mathbf{A}^\mathsf{T}\mathbf{A})^{\frac{1}{2}} \left(\mathrm{tr}\left((\mathbf{A}^\mathsf{T}\mathbf{A})^{-\mathsf{T}}\mathbf{A}^\mathsf{T}\mathbf{U}\right) + \mathrm{tr}\left((\mathbf{A}^\mathsf{T}\mathbf{A})^{-\mathsf{T}}\mathbf{U}^\mathsf{T}\mathbf{A}\right) \right) \mathrm{vol}_b \\
&= \tfrac{1}{2} \left(\mathrm{tr}\left((\mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1})^\mathsf{T}\mathbf{U}\right) + \mathrm{tr}\left((\mathbf{U}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1})^\mathsf{T}\mathbf{A}\right) \right) \mathrm{vol}_{A^\# g_0} \\
&= \langle A, U \rangle_{A^\# g_0}\, \mathrm{vol}_{A^\# g_0}\,. \qquad \square
\end{aligned}$$

Lemma 5.15 *Let (Σ, g) be a compact, k-dimensional Riemannian manifold with boundary. The volume functional $\mathcal{F}: (\mathrm{Imm}(\Sigma; \mathbb{R}^m), \|\cdot\|_{W_g^{1,\infty}}) \to \mathbb{R}$ is differentiable and its derivative is given by*

$$\langle d\mathcal{F}|_f, u \rangle = \int_\Sigma \langle df, du \rangle_{f^\# g_0} \, \mathrm{vol}_{f^\# g_0} \quad \textit{for all } f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m), \, u \in W_g^{1,\infty}(\Sigma; \mathbb{R}^m).$$

PROOF. Let $f \in \mathrm{Imm}(\Sigma; \mathbb{R}^m)$. First, note that $\mathcal{F}(f) = \int_\Sigma F(df)$ with the bundle map $F: \mathrm{Hom}_k(T\Sigma; \mathbb{R}^m) \to \Omega(TM)$ given fiberwise as in the previous lemma. Second, one has $df|_x \in \mathrm{Hom}_k(T_xM; \mathbb{R}^m)$ for vol_g-almost all $x \in \Sigma$. Integration is linear and $\mathrm{Imm}(\Sigma; \mathbb{R}^m) \subset W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$ is an open set, hence the previous lemma yields the claim. $\qquad\square$

6 Approximation Theory for Lipschitz Immersions

In Chapters 7 and 8 we will focus on the discretization of variational problems on the space of Lipschitz immersions with fixed topological type under Dirichlet boundary conditions. The central tools for the discretization are simplicial submanifolds of Euclidean space and smooth triangulations of smooth manifolds with boundary.

In the present chapter, we couch these tools in the language of Chapter 2 by identifying:

- the configuration spaces C, C_n;

- meaningful a priori information sets $\mathcal{A}$, $\mathcal{A}_n$; and

- sampling and reconstruction operators $\mathcal{S}_n$, $\mathcal{R}_n$.

Moreover, we lay the foundation for proximity estimates.

6.1 Smooth Triangulations

Denote the *standard simplex* by

$$\Delta_k := \left\{ y = (y_0, \ldots, y_k) \in [0,1]^{k+1} \; \middle| \; \textstyle\sum_{i=0}^{k} y_i = 1 \right\}.$$

Denote by $e_0, \ldots, e_k$ the standard basis of $\mathbb{R}^{k+1}$ and define the set of *d-faces*

$$F_d := \{ \operatorname{conv}(e_{i_0}, \ldots, e_{i_d}) \mid 0 \le i_0 < \cdots < i_d \le k \},$$

where conv denotes the convex hull. For a smooth embedding $\sigma \colon \Delta_k \to \Sigma$, define the *vertex set*

$$V(\sigma) := \{ \sigma(e_0), \ldots, \sigma(e_k) \} = \sigma(F_0).$$

Definition 6.1 Let M be a smooth, k-dimensional manifold with corners. A *smooth triangulation* of Σ is a family

$$\mathcal{T} \subset \{\sigma\colon \Delta_k \to M \mid \sigma \text{ is a smooth embedding}\}$$

with the following properties:

1. $M = \bigcup_{\sigma \in \mathcal{T}} \sigma(\Delta_k)$.

2. For each pair σ, $\tau \in \mathcal{T}$ with $\sigma(\Delta_k) \cap \tau(\Delta_k) \neq \emptyset$, both $\sigma^{-1}(\sigma(\Delta_k) \cap \tau(\Delta_k))$ and $\tau^{-1}(\sigma(\Delta_k) \cap \tau(\Delta_k))$ are d-faces of Δ_k for some $0 \leq d \leq k-1$ and the mapping

$$\tau^{-1} \circ \sigma\colon \sigma^{-1}(\sigma(\Delta_k) \cap \tau(\Delta_k)) \to \tau^{-1}(\sigma(\Delta_k) \cap \tau(\Delta_k))$$

   is affine.

3. For each $\sigma \in \mathcal{T}$ with $\sigma(\Delta_k) \cap \partial M \neq \emptyset$, the set $\sigma^{-1}(\sigma(\Delta_k) \cap M)$ is a d-face of Δ_k for some $0 \leq d \leq k-1$.

We distinguish between *boundary vertices* and *interior vertices*:

$$V(\mathcal{T}) := \bigcup_{\sigma \in \mathcal{T}} V(\sigma), \quad V_{\mathrm{b}}(\mathcal{T}) := V(\mathcal{T}) \cap \partial M \quad \text{and} \quad V_{\mathrm{i}}(\mathcal{T}) := V(\mathcal{T}) \setminus V_{\mathrm{b}}(\mathcal{T}).$$

A smooth triangulation is called *finite* if its cardinality is finite.

Definition 6.2 Let $\mathcal{T}$ be a smooth triangulation of a k-dimensional smooth manifold M with boundary. Then $\mathcal{T}$ induces a smooth triangulation $\mathcal{T}|_{\partial M}$ of the boundary ∂M by:

$$\mathcal{T}|_{\partial M} := \{\sigma|_A \mid \sigma \in \mathcal{T}, A \in F_{k-1}\colon \sigma(A) \subset \partial M\}$$

Remark 6.3 Every smooth manifold with boundary admits a smooth triangulation (see [38]).

6.2 Configuration Spaces and Sampling Operator

Let k and m be positive integers with $k \leq m$ and let Σ be a compact, k-dimensional smooth manifold with boundary. We abbreviated $\tilde{C} := \mathrm{Imm}(\Sigma; \mathbb{R}^m)$. For a fixed boundary condition $\gamma \in \mathrm{Imm}(\partial\Sigma; \mathbb{R}^m)$ we define the *(smooth) configuration space* $C := \mathrm{Imm}_\gamma(\Sigma; \mathbb{R}^m)$. Note that $C \subset \tilde{C}$ is a closed set.

Let $\mathcal{T}$ be a smooth triangulation of Σ. We define the *discrete configuration space* or space of *discrete immersions*:

$$C_{\mathcal{T}} := \left\{ f \colon V(\mathcal{T}) \to \mathbb{R}^m \, \middle| \, \begin{array}{c} f|_{V_{\mathrm{b}}(\mathcal{T})} = \gamma|_{V_{\mathrm{b}}(\mathcal{T})}, \\ \forall \sigma \in \mathcal{T} \colon f(V(\sigma)) \text{ in general position} \end{array} \right\} \subset \mathbb{R}^{m \, \mathrm{card}(V_{\mathrm{i}}(\mathcal{T}))}.$$

While a *sampling operator*

$$\mathcal{S}_{\mathcal{T}} \colon C \subset W_g^{1,\infty}(\Sigma; \mathbb{R}^m) \to C_{\mathcal{T}}, \qquad f \mapsto f|_{V(\mathcal{T})},$$

is easily defined, we need some more effort to obtain a reasonable reconstruction operator. For each $p \in V(\mathcal{T})$ denote by $\lambda_p \colon \Sigma \to \mathbb{R}$ the continuous, piecewise smooth function defined by:

1. $\lambda_p(p) = 1$.

2. $\lambda_p(q) = 0$ for all $q \in V(\mathcal{T})$.

3. $\lambda_p \circ \sigma \colon \Delta_k \to \mathbb{R}$ is the restriction of an affine function for each $\sigma \in \mathcal{T}$.

This allows us to define a preliminary reconstruction operator

$$\tilde{\mathcal{R}}_{\mathcal{T}} \colon C_{\mathcal{T}} \to \tilde{C}, \qquad \tilde{\mathcal{R}}_{\mathcal{T}}(f)(x) := \sum_{p \in V(\mathcal{T})} \lambda_p(x) \, f(p).$$

Note that for every $f \in C_{\mathcal{T}}$, the image of $\tilde{\mathcal{R}}_{\mathcal{T}}(f)$ in $\mathbb{R}^m$ is a union of non-degenerate k-dimensional Euclidean simplices. We define the piecewise smooth mapping

$$\gamma_{\mathcal{T}} \colon \partial\Sigma \to \mathbb{R}^m, \qquad \gamma_{\mathcal{T}}(x) := \sum_{p \in V_{\mathrm{b}}(\mathcal{T})} \lambda_p(x) \, \gamma(p). \tag{6.1}$$

Observe that for each $f \in C_{\mathcal{T}}$, the preliminary reconstruction $\tilde{\mathcal{R}}_{\mathcal{T}}(f)$ restricted to $\partial\Sigma$ is identical to $\gamma_{\mathcal{T}}$. Moreover, the image of $\gamma_{\mathcal{T}}$ is a union of embedded $(k-1)$-dimensional simplices. In general, $\gamma_{\mathcal{T}}$ and γ need not to be equal which is why we have to modify $\tilde{\mathcal{R}}_{\mathcal{T}}$ later (see Section 6.4).

6.3 Relative Approximation Errors

Throughout, we let (Σ, g) be a k-dimensional Riemannian manifold with boundary and $\mathcal{T}$ be a smooth triangulation of Σ.

Definition 6.4 We define the *relative approximation errors* $\varrho^0(\mathcal{T})$, $\varrho^1(\mathcal{T})$ of the smooth triangulation $\mathcal{T}$ by

$$\varrho^0(\mathcal{T}) := \sup_{\substack{f \in W_g^{2,\infty}(\Sigma;\mathbb{R}^m) \\ \mathrm{d}f \neq 0}} \frac{\|f - f_{\mathcal{T}}\|_{L_g^\infty}}{\|\mathrm{d}f\|_{W_g^{1,\infty}}} \quad \text{and} \quad \varrho^1(\mathcal{T}) := \sup_{\substack{f \in W_g^{2,\infty}(\Sigma;\mathbb{R}^m) \\ \mathrm{d}f \neq 0}} \frac{\|\mathrm{d}f - \mathrm{d}f_{\mathcal{T}}\|_{L_g^\infty}}{\|\mathrm{d}f\|_{W_g^{1,\infty}}},$$

where $f_{\mathcal{T}} := (\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)$.

Definition 6.5 For $\sigma \in \mathcal{T}$ define its *approximation characteristics*

$$\chi^{(d)}(\sigma) := \|(T\sigma)^{-1}\|_{L_g^\infty}^d (\|T\sigma\|_{L_g^\infty}^2 + \|\mathrm{Hess}^g(\sigma)\|_{L_g^\infty}), \quad d = 0, 1.$$

We define the *approximation characteristics of* $\mathcal{T}$ as $\chi^{(d)}(\mathcal{T}) := \sup_{\sigma \in \mathcal{T}} \chi^{(d)}(\sigma)$.

Approximation characteristics provide upper bounds on relative approximation errors:

Lemma 6.6 *Let (Σ, g) be a k-dimensional smooth Riemannian manifold with boundary and $\mathcal{T}$ be a smooth triangulation of Σ. Then the relative approximation errors are bounded by:*

$$\varrho^{(0)}(\mathcal{T}) \leq (1 + R_k)\chi^{(0)}(\mathcal{T}) \quad \text{and} \quad \varrho^{(1)}(\mathcal{T}) \leq R_k \chi^{(1)}(\mathcal{T}) \quad \text{where } R_k := 2\,(k+1)^{\frac{1}{2}}.$$

PROOF. Let $p = (\frac{1}{k+1}, \ldots, \frac{1}{k+1}) \in \mathbb{R}^{k+1}$ be the barycenter of Δ_k. Let $f \in W_g^{2,\infty}(\Sigma; \mathbb{R}^m)$ and $\sigma \in \mathcal{T}$. We abbreviate $f_{\mathcal{T}} := (\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)$, $h := f \circ \sigma$, and $h_{\mathcal{T}} := f_{\mathcal{T}} \circ \sigma$. Since $h_{\mathcal{T}} \colon \Delta_k \to \mathbb{R}^m$ is an affine map, we may write it as

$$h_{\mathcal{T}}(y) = h(e_0) + A\,(y - e_0),$$

where $A \colon p^\perp \to \mathbb{R}^m$ is the linear map defined by $A\,(e_i - e_0) = h(e_i) - h(e_0)$ for $i = 1, \ldots, k$. Let $\beta(y) := h(y) - h(p) - \mathrm{d}h|_p(y - p)$. By Taylor's theorem, we have for each $y \in \Delta_k$:

$$|\beta(y)| \leq \frac{1}{2}\|\mathrm{Hess}(h)\|_{L^\infty}|y - p|^2 \leq \frac{1}{2}\frac{k}{k+1}\|\mathrm{Hess}(h)\|_{L^\infty}.$$

Note that

$$
\begin{aligned}
A(e_i - e_j) &- \mathrm{d}h|_p(e_i - e_j) \\
&= h(e_i) - h(e_j) - h(p) - \mathrm{d}h|_p(e_i - p) + h(p) + \mathrm{d}h|_p(e_j - p) \\
&= \left(h(e_i) - h(p) - \mathrm{d}h|_p(e_i - p) \right) - \left(h(e_j) - h(p) - \mathrm{d}h|_p(e_j - p) \right) \\
&= \beta(e_i) - \beta(e_j).
\end{aligned}
$$

Let $u \in p^\perp$. With $\alpha_i = \langle e_i - p, u \rangle$ for $i = 1, \ldots, k$ one may write $u = \sum_{i=1}^k \alpha_i(e_i - e_0)$ and, together with the above, obtain:

$$
|A u - \mathrm{d}h|_p u| \le \sum_{i=1}^k |\alpha_i|(|\beta_i| + |\beta_0|) \le \frac{k}{k+1} \|\mathrm{Hess}(h)\|_{L^\infty} \sum_{i=1}^k |\alpha_i|.
$$

The linear map

$$
L \colon p^\perp \to \mathbb{R}^k, \quad u \mapsto (\langle e_1 - p, u \rangle, \ldots, \langle e_k - p, u \rangle)
$$

has operator norm $\|L\|_{\ell^1 \to \ell^1} = \frac{2k-1}{k+1} \le 2$, and because of $\|u\|_1 \le (k+1)^{\frac{1}{2}} |u|$, we obtain

$$
\|\mathrm{d}h_{\mathcal{T}} - \mathrm{d}h|_p\| = \|A - \mathrm{d}h|_p\| \le R_k \|\mathrm{Hess}(h)\|_{L^\infty}. \tag{6.2}
$$

This estimate leads to

$$
\begin{aligned}
\|\mathrm{d}f|_{\sigma(y)} - \mathrm{d}(f_{\mathcal{T}} \circ \sigma)|_{\sigma(y)}\| &= \|(\mathrm{d}h|_y - \mathrm{d}h_{\mathcal{T}}|_y)(T_y\sigma)^{-1}\| \\
&\le R_k \|(T\sigma)^{-1}\|_{L_g^\infty} \|\mathrm{Hess}(h)\|_{L^\infty}
\end{aligned}
$$

and

$$
\begin{aligned}
|f(\sigma(y)) - f_{\mathcal{T}}(\sigma(y))| &= |h(y) - h_{\mathcal{T}}(y)| = |h(y) - h(e_0) - A(y - e_i)| \\
&= |\beta(y) - \beta(e_0) + (\mathrm{d}h|_p(y - e_i) - A(y - e_i))| \\
&\le |\beta(y)| + |\beta(e_0)| + |\mathrm{d}h|_p(y - e_i) - A(y - e_i)| \\
&\le (1 + R_k) \|\mathrm{Hess}(h)\|_{L^\infty}.
\end{aligned}
$$

For $y \in \Delta_k$ and $u, v \in T_y\Delta_k$, the chain and product rules imply

$$
\begin{aligned}
\mathrm{Hess}_y(h)(u, v) &= \mathrm{Hess}_y(f \circ \sigma)(u, v) \\
&= \mathrm{Hess}_{\sigma(y)}^g(f)(T_y\sigma\, u, T_y\sigma\, v) + \mathrm{d}_{\sigma(y)}f\ \mathrm{Hess}_y^g(\sigma)(u, v),
\end{aligned}
$$

which yields the estimate

$$
\|\mathrm{Hess}(h)\|_{L^\infty} \le \|\mathrm{d}f\|_{W_g^{1,\infty}} \sup_{\sigma \in \mathcal{T}} \left(\|T\sigma\|_{L_g^\infty}^2 + \|\mathrm{Hess}^g(\sigma)\|_{L_g^\infty} \right) \qquad \square
$$

It is crucial for the convergence analysis of finite elements that there are smooth triangulations of arbitrarily small relative approximation errors. Fortunately, affine subdivision can be used to improve upon the approximation characteristics (and thus upon the relative approximation errors):

Definition 6.7 Let Σ be a smooth manifold with boundary and let $\mathcal{T}$ be a smooth triangulation of Σ. A *subdivision* of $\mathcal{T}$ is a family $\Phi = (\Phi_\sigma)_{\sigma \in \mathcal{T}}$ of smooth triangulations of Δ_k such that $\Phi(\mathcal{T}) := \{ \sigma \circ \varphi \mid \sigma \in \mathcal{T}, \ \varphi \in \Phi_\sigma \}$ is a smooth triangulation of Σ.[12]

We call a subdivision Φ *affine* if it consists of affine mappings only, i.e., if each $\varphi \in \Phi_\sigma$ is affine for all $\sigma \in \mathcal{T}$.

Lemma 6.8 *Let (Σ, g) be a k-dimensional smooth Riemannian manifold with boundary, $\mathcal{T}$ be a smooth triangulation of Σ, and Φ be an affine subdivision of $\mathcal{T}$. Then the approximation characteristics of $\Phi(\mathcal{T})$ can be estimated by:*

$$\chi^{(0)}(\Phi(\mathcal{T})) \leq \chi^{(0)}(\mathcal{T}) \sup_{\sigma \in \mathcal{T}} \sup_{\varphi \in \Phi_\sigma} \|T\varphi\|_{L^\infty}^2,$$

$$\chi^{(1)}(\Phi(\mathcal{T})) \leq \chi^{(1)}(\mathcal{T}) \sup_{\sigma \in \mathcal{T}} \sup_{\varphi \in \Phi_\sigma} \|T\varphi\|_{L^\infty}^2 \|(T\varphi)^{-1}\|_{L^\infty}.$$

PROOF. Let $\sigma \in \mathcal{T}$ and $\varphi \in \Phi_\sigma$. Because φ is the restriction of an affine map to an affine subspace, we have $\text{Hess}(\varphi) = 0$. Thus one obtains

$$T(\sigma \circ \varphi) = T\sigma \, T\varphi \quad \text{and} \quad \text{Hess}^g(\sigma \circ \varphi) = \text{Hess}^g(\sigma)(T\varphi \cdot, T\varphi \cdot).$$

This supplies us with the estimates

$$\|T(\sigma \circ \varphi)\|_{L_g^\infty} \leq \|T\sigma\|_{L_g^\infty} \|T\varphi\|_{L^\infty},$$

$$\|(T(\sigma \circ \varphi))^{-1}\|_{L_g^\infty} \leq \|(T\sigma)^{-1}\|_{L_g^\infty} \|(T\varphi)^{-1}\|_{L^\infty},$$

$$\|\text{Hess}^g(\sigma \circ \varphi)\|_{L_g^\infty} \leq \|\text{Hess}^g(\sigma)\|_{L_g^\infty} \|T\varphi\|_{L^\infty}^2,$$

which imply the statement. □

The previous lemma shows that the approximation characteristics of $\mathcal{T}$ may be decreased by choosing an affine subdivision Φ with

$$\eta := \sup_{\sigma \in \mathcal{T}} \sup_{\varphi \in \Phi_\sigma} \|T\varphi\|$$

---

[1]Note that Δ_k itself is a smooth manifold with corners.

[2]This demand includes a certain compatibility between the triangulations Φ_σ, Φ_τ for neighboring σ, $\tau \in \mathcal{T}$, i.e., for those σ, $\tau \in \mathcal{T}$ with $\sigma(\Delta_k) \cap \tau(\Delta_k) \neq \emptyset$.

small and

$$\vartheta := \sup_{\sigma \in \mathcal{T}} \sup_{\varphi \in \Phi_\sigma} \|(T\varphi)^{-1}\| \, \|T\varphi\|$$

bounded. While η is related to the maximal *size* of the cells $\varphi(\Delta_k)$, the quantity ϑ is precisely the maximal *thickness* of the cells $\varphi(\Delta_k)$.

This also shows that many mesh refinement schemes that work for triangulations of domains in Euclidean space can also applied here.

Example 6.9 In the case $k = 2$, one may apply 4:1 subdivision: The simplex Δ_2 is split into four by inserting a vertex on the midpoint of each edge. In that case, one has $\|T\varphi\| = \frac{1}{2}$ and $\|T\varphi^{-1}\| = 2$ such that one obtains $\chi^{(d)}(\Phi(\mathcal{T})) \le 2^{d-2}\chi^{(d)}(\mathcal{T})$ for $d = 0, 1$.

Of course, there may be many other ways to construct smooth triangulations with arbitrarily small relative approximation error, e.g., Karcher coordinates (see [12]). The relevant facts for our convergence analysis are condensed in the following statement:

Corollary 6.10 *Let (Σ, g) be a compact, smooth Riemannian manifold with boundary. Then there are finite smooth triangulations with arbitrary small relative approximation errors, i.e., for every $\varepsilon > 0$ there is a finite smooth triangulation $\mathcal{T}$ of Σ with*

$$\varrho(\mathcal{T}) := \max\{\varrho^{(0)}(\mathcal{T}), \, \varrho^{(0)}(\mathcal{T}|_{\partial\Sigma}), \, \varrho^{(1)}(\mathcal{T}), \, \varrho^{(1)}(\mathcal{T}|_{\partial\Sigma})\} \le \varepsilon.$$

6.4 Reconstruction Operator

For $\gamma \in \operatorname{Imm}(\partial\Sigma; \mathbb{R}^m) \cap W^{2,\infty}(\partial\Sigma; \mathbb{R}^m)$, $f \in \operatorname{Imm}(\Sigma; \mathbb{R}^m) \cap W^{2,\infty}(\Sigma; \mathbb{R}^m)$, and $\gamma_\mathcal{T}$ as defined in (6.1) we obtain the relative approximation errors:

$$\|\gamma - \gamma_\mathcal{T}\|_{W^{1,\infty}_g} \le \|d\gamma\|_{W^{1,\infty}_g} \varrho(\mathcal{T}), \tag{6.3}$$

$$\|f - (\tilde{\mathcal{R}}_\mathcal{T} \circ \mathcal{S}_\mathcal{T})(f)\|_{W^{1,\infty}_g} \le \|df\|_{W^{1,\infty}_g} \varrho(\mathcal{T}). \tag{6.4}$$

Let $\mathrm{ext}\colon W^{1,\infty}_{g_{|\partial\Sigma}}(\partial\Sigma;\mathbb{R}^m) \to W^{1,\infty}_g(\Sigma;\mathbb{R}^m)$ be a continuous, linear extension operator[3] and let $u_{\mathcal{T}} := \mathrm{ext}(\gamma - \gamma_{\mathcal{T}})$. Now, (6.3) provides us with the estimate

$$\|u_{\mathcal{T}}\|_{W^{1,\infty}_g} \leq \|\mathrm{ext}\|\,\|\gamma - \gamma_{\mathcal{T}}\|_{W^{1,\infty}_g} \leq \|\mathrm{ext}\|\,\|\mathrm{d}\gamma\|_{W^{1,\infty}_g}\,\varrho(\mathcal{T}). \tag{6.5}$$

For $r > s > 0$, define the sets

$$\mathscr{A}^s := \{\, f \in C \cap W^{2,\infty}_g(\Sigma;\mathbb{R}^m) \mid d_{\mathcal{P}}(g, f^\# g_0) \leq s,\ \|\mathrm{d}f\|_{W^{1,\infty}_g} \leq s \,\}, \tag{6.6}$$

$$\mathscr{A}^r_{\mathcal{T}} := \{\, f \in C_{\mathcal{T}} \mid d_{\mathcal{P}(\Sigma)}(g, \tilde{\mathscr{R}}_{\mathcal{T}}(f)^\# g_0) \leq r \,\}, \tag{6.7}$$

and the operator

$$\mathscr{R}_{\mathcal{T}}\colon \mathscr{A}^r_{\mathcal{T}} \to W^{1,\infty}_g(\Sigma;\mathbb{R}^m), \qquad f \mapsto \mathscr{R}_{\mathcal{T}}(f) + u_{\mathcal{T}}.$$

As the final result of this chapter, we show that $\mathscr{R}_{\mathcal{T}}$ is a reconstruction operator, i.e., $\mathscr{R}_{\mathcal{T}}(\mathscr{A}^r_{\mathcal{T}}) \subset C$, at least for sufficiently "fine" triangulations. We also verify the condition $\mathcal{S}_{\mathcal{T}}(\mathscr{A}^s) \subset \mathscr{A}^r_{\mathcal{T}}$ of Theorem 2.28:

Lemma 6.11 *Let $r > s > 0$ and $c > 0$. Then there is $\varrho_0 > 0$ such that for every smooth triangulation $\mathcal{T}$ with $\varrho(\mathcal{T}) \leq \varrho_0$ the following hold:*

$$\mathcal{S}_{\mathcal{T}}(\mathscr{A}^s) \subset \mathscr{A}^r_{\mathcal{T}}, \quad \text{and} \quad \mathscr{R}_{\mathcal{T}}(\mathscr{A}^r_{\mathcal{T}}) \subset C.$$

Proof. Let $f \in \mathscr{A}^s$ and put $f_{\mathcal{T}} := (\tilde{\mathscr{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)$. By (6.4), we have $\|f - f_{\mathcal{T}}\|_{W^{1,\infty}_g} \leq s\,\varrho(\mathcal{T}) \leq s\,\varrho_0$. Corollary 5.6 tells us how small ϱ_0 has to be (depending on s only) so that $f_{\mathcal{T}} \in \mathrm{Imm}(\Sigma;\mathbb{R}^m)$ and thus $\mathcal{S}_{\mathcal{T}}(f) \in C_{\mathcal{T}}$. By Corollary 5.7, one has the inequality

$$d_{\mathcal{P}(\Sigma)}(g, f^\#_{\mathcal{T}} g_0) \leq d_{\mathcal{P}(\Sigma)}(g, f^\# g_0) + d_{\mathcal{P}(\Sigma)}(f^\# g_0, f^\#_{\mathcal{T}} g_0) \leq s + C(s)\,\varrho_0$$

which shows that $\mathcal{S}_{\mathcal{T}}(f) \in \mathscr{A}^r_{\mathcal{T}}$ if ϱ_0 is sufficiently small.

Now, let $f \in \mathscr{A}^r_{\mathcal{T}}$. We have $\tilde{\mathscr{R}}_{\mathcal{T}}(f) \in \mathrm{Imm}(\Sigma;\mathbb{R}^m)$ and $d_{\mathcal{P}(\Sigma)}(g, \tilde{\mathscr{R}}_{\mathcal{T}}(f)^\# g_0) \leq r$. By (6.5), we obtain

$$\|\mathscr{R}_{\mathcal{T}}(f) - \tilde{\mathscr{R}}_{\mathcal{T}}(f)\|_{W^{1,\infty}_g} = \|u_{\mathcal{T}}\|_{W^{1,\infty}_g} \leq C\,\|\mathrm{d}\gamma\|_{W^{1,\infty}_g}\,\varrho_0.$$

Again, Corollary 5.6 tells us how ϱ_0 has to be chosen depending on r such that $\mathscr{R}_{\mathcal{T}}(f) \in \tilde{C}$. Since $\mathscr{R}_{\mathcal{T}}(f)$ fulfills the boundary conditions by construction, we obtain $\mathscr{R}_{\mathcal{T}}(f) \in C$. $\qquad\square$

---

[3]Such an operator can be obtained, e.g., by choosing a smooth collar $\Phi\colon \partial\Sigma \times [0,1[\;\tilde{\to}\; U \subset \Sigma$ and by using the function $\chi\colon [0,1[\to \mathbb{R}, \chi(t) = \exp\left(\frac{t^2}{t^2-t}\right)$: Then

$$\mathrm{ext}(u)(x) := \begin{cases} (u \otimes \chi) \circ \Phi^{-1}(x), & x \in U, \\ 0, & x \in \Sigma \setminus U. \end{cases}$$

is the desired extension operator.

7 Minimal Surfaces

As a first extensive application of the theory developed in Chapter 2, in particular of Theorem 2.28, we discuss a variant of the *Douglas-Courant problem* or *least area/volume problem*: Among the immersed k-dimensional surfaces in $\mathbb{R}^m$ with prescribed topology and Dirichlet boundary conditions find those of minimal k-volume. For $k = 2$, C^2-minimizers are examples of minimal surfaces (see Lemma 7.6).[1]

We discretize this problem by searching for volume-minimizers among immersed k-dimensional simplicial meshes of fixed combinatorics bounded by a given, closed $(k-1)$-dimensional simplicial mesh. To some extent, this approach can be understood as a nonconforming Ritz-Galerkin method with first order Lagrange elements (piecewise linear finite elements).

We primarily aim at a convergence analysis for discrete minimizers, but we also discuss some numerical methods for obtaining them (Section 7.3 and Section 7.7). The point we would like to make is this: Given a sufficiently well-posed Plateau problem, i.e., the boundary conditions are such that volume minimizers within a certain topological class exist *and* have a certain uniform regularity, the set of solutions can be approximated by solutions of a discrete Plateau problem.

We start our exposition by giving a precise definition for minimal surfaces and by stating both the Douglas-Courant problem, and the least area problem (Section 7.1). After a brief overview of the classical theory of minimal surfaces (Section 7.2), we compare some of the pre-existing numerical methods for computing minimal surfaces (Section 7.3). Afterwards, we discretize the least area problem and identify the relevant entities occurring in Theorem 2.28, namely the smooth and discrete configuration spaces, functionals and test mappings, as well as the sampling and reconstruction operators (Section 7.5). Our convergence result then follows from an analysis of consistency and proximity errors (Section 7.6). Finally, we discuss a certain gradient-like flow that was introduced in [30] and which is very efficient for solving the discrete least area problem (Section 7.7).

---

[1]It is an unparalleled obscurity that also non-minimizers are called "minimal". Alas, this convention has grown historically.

7.1 General Theory

Definition 7.1 Let Σ be a 2-dimensional manifold with boundary and let (M_0, g_0) be a smooth Riemannian manifold of dimension $m \geq 3$. A mapping $f \in C^0(\Sigma; M_0) \cap C^2(\Sigma^\circ; M_0)$ is called a *minimal surface* if there is a Riemannian metric g of class C^1 in the interior $\Sigma^\circ := \Sigma \setminus \partial\Sigma$ and a function $\varrho \in C^1(\Sigma^\circ; [0, \infty[)$ with

$$\Delta^{g,g_0} f := \operatorname{tr}_g \operatorname{Hess}^{g,g_0}(f) = 0 \quad \text{and} \quad f^\# g_0 = \varrho\, g \quad \text{in } \Sigma^\circ. \tag{7.1}$$

Remark 7.2 The condition $f^\# g_0 = \varrho\, g$ for some $\varrho \in C^1(\Sigma^\circ; [0, \infty[)$ is often referred to as the "conformality condition" in the literature. Note however that a conformal map by definition fulfills $f^\# g_0 = \varrho\, g$, with a *nowhere* vanishing function ϱ, that is, a conformal map is necessarily *regular*.

The *Douglas-Courant problem*, also called the *Plateau-Douglas problem*, can be formulated as follows (see [17] or [10]):

Problem 7.3 (Douglas-Courant)
Let Σ be a 2-dimensional smooth manifold with boundary and let $\gamma \in C^0(\partial\Sigma; M_0)$ be an embedding. Find all minimal surfaces f with $f|_{\partial\Sigma} = \gamma \circ \varphi$ for some homeomorphism $\varphi \colon \partial\Sigma \to \partial\Sigma$.

In the case that $\Sigma = D$ is the closed unit disk and $M_0 = \mathbb{R}^3$, this is traditionally referred to as the *Plateau problem*.

The notion of minimal surfaces has its origin in the least area problem, the 2-dimensional instance of the least volume problem. We give a formulation of this problem in terms of Lipschitz Immersions:

Problem 7.4 (Least volume problem) *Let Σ be a compact, k-dimensional smooth manifold with boundary. Let (M_0, g_0) be a smooth, m-dimensional Riemannian manifold with $m > k$ and let $\gamma \in \operatorname{Imm}(\partial\Sigma; M_0)$ be a Lipschitz immersion. Given Σ and γ, minimize the volume functional*

$$\mathcal{F}(f) = \int_\Sigma \operatorname{vol}_{f^\# g_0}$$

on the space $C := \operatorname{Imm}_\gamma(\Sigma; M_0)$ of Lipschitz immersions that restrict on the boundary to γ (see Chapter 5).

Remark 7.5 Note that by using Lipschitz immersions as configuration space, we exclude "hairy" mappings (which is desired), but we also exclude continuously differentiable mappings with isolated branch points.

The Douglas-Courant problem and the least area problem are closely related:

Lemma 7.6 *Let Σ be a compact, 2-dimensional smooth manifold with boundary and (M_0, g_0) a smooth Riemannian manifold without boundary. Let $\gamma \in \mathrm{Imm}(\partial\Sigma; M_0)$ be a topological embedding and $f \in C \cap C^2(\Sigma^\circ; M_0)$ a Lipschitz immersion that is of class C^2 in the interior of Σ.*

Then f is a minimal surface if and only if it is a critical point of $\mathcal{F}|_C$.

PROOF. Note that an argumentation analogous to the proof of Lemma 5.15 shows that

$$\langle \mathrm{d}\mathcal{F}|_f, u \rangle = \int_\Sigma \langle Tf, \nabla^{g_0} u \rangle_{f^\# g_0} \, \mathrm{vol}_{f^\# g_0}$$

for all $f \in \mathrm{Imm}(\Sigma; M_0)$ and all vector fields $u \in \Gamma(\Sigma; f^\# TM_0)$ along f of class $W_g^{1,\infty}$.

Let $f \in C \cap C^2(\Sigma^\circ; M_0)$ be a critical point of $\mathcal{F}|_C$. By partial integration, one has for each vector field $u \in \Gamma_0(\Sigma^\circ; f^\# TM_0)$ along f of class C^1 with compact support:

$$0 = \langle \mathrm{d}\mathcal{F}|_f, u \rangle = \int_\Sigma \langle Tf, \nabla^{g_0} u \rangle_{f^\# g_0} \, \mathrm{vol}_{f^\# g_0} = - \int_\Sigma \langle \Delta^{f^\# g, g_0} f, u \rangle_{f^\# g_0, g_0} \, \mathrm{vol}_{f^\# g_0}.$$

Thus f is harmonic with respect to the Riemannian metric $f^\# g_0$. Moreover,

$$f \colon (\Sigma, f^\# g_0) \to (M_0, g_0)$$

is a Riemannian isometry, hence a conformal map. Thus, f is a minimal surface.

Let $f \in C \cap C^2(\Sigma^\circ; M_0)$ be a minimal surface, $\varrho \in C^1(\Sigma^\circ;]0, \infty[)$ a function, and g a Riemannian metric of class C^1 on Σ° with $\Delta^g f = 0$ and $f^\# g_0 = \varrho\, g$. Since f is a Lipschitz immersion and ϱ is continuous, one has $\varrho > 0$, hence f is conformal. By Lemma 7.7 below, one has for $u \in \Gamma_0(\Sigma^\circ; f^\# TM_0)$ of class C^1:

$$0 = - \int_\Sigma \langle \Delta^{g, g_0} f, u \rangle_{g, g_0} \, \mathrm{vol}_g = - \int_\Sigma \langle \Delta^{f^\# g_0, g_0} f, u \rangle_{f^\# g, g_0} \, \mathrm{vol}_{f^\# g_0} = \langle \mathrm{d}\mathcal{F}|_f, u \rangle. \qquad \square$$

Lemma 7.7 *In the case that $\dim(\Sigma) = 2$, the Laplacian Δ^{g, g_0} transforms under conformal changes $g_\lambda := \mathrm{e}^{2\lambda} g$ of the metric with $\lambda \in C^1(\Sigma^\circ; \mathbb{R})$ as follows:*

$$\int_{\Sigma^\circ} \langle \Delta^{g_\lambda, g_0} f, u \rangle_{g_0} \, \mathrm{vol}_{g_\lambda} = \int_{\Sigma^\circ} \langle \Delta^{g, g_0} f, u \rangle_{g_0} \, \mathrm{vol}_g$$

for all $f \in C^2(\Sigma^\circ; M_0)$ and all smooth vector fields $u \in \Gamma_0(\Sigma^\circ; f^\# TM_0)$ along f with compact support.

Proof. Note that $\mathrm{vol}_{g_\lambda} = \mathrm{e}^{2\lambda}\,\mathrm{vol}_g$ and that $\langle \xi, \eta \rangle_{g_\lambda} = \mathrm{e}^{-2\lambda}\langle \xi, \eta \rangle_g$ for all $\xi, \eta \in T'_x\Sigma$. Testing with $u \in \Gamma_0(\Sigma^\circ; f^\#TM_0)$, one obtains

$$\int_{\Sigma^\circ} \langle \Delta^{g_\lambda, g_0} f, u \rangle_{g_0}\,\mathrm{vol}_{g_\lambda} = -\int_{\Sigma^\circ} \langle Tf, \nabla^{g_0}u \rangle_{g_\lambda, g_0}\,\mathrm{vol}_{g_\lambda}$$

$$= -\int_{\Sigma^\circ} \mathrm{e}^{-2\lambda}\langle Tf, \nabla^{g_0}u \rangle_{g, g_0}\,\mathrm{e}^{2\lambda}\,\mathrm{vol}_g = \int_{\Sigma^\circ} \langle \Delta^{g, g_0} f, u \rangle_{g_0}\,\mathrm{vol}_g.\;\square$$

From now on, we exclusively discuss the case of $M_0 = \mathbb{R}^m$ being the Euclidean space and g_0 the Euclidean metric.

7.2 Existence Theorem for Disk-like Minimal Surfaces

Let $\Sigma = D$ be the closed unit disk in $\mathbb{R}^2$ and fix an embedding $\gamma \in C^0(D; \mathbb{R}^m)$. By the Riemann mapping theorem, there is exactly one conformal structure on D. Let g be the Euclidean metric on D. In light of Lemma 7.7, it is immediate that $f \in C^0(D; \mathbb{R}^m) \cap C^2(D^\circ; \mathbb{R}^m)$ is a minimal surface if and only if (7.1) is fulfilled for the Euclidean metric g in D.

Define the *Dirichlet functional* $\mathcal{D}$ by

$$\mathcal{D}(f) := \frac{1}{2} \int_D \langle \mathrm{d}f, \mathrm{d}f \rangle_g\,\mathrm{vol}_g \quad \text{for all } f \in W^{1,2}_g(D; \mathbb{R}^m).$$

Equip ∂D with some orientation and denote by $\mathfrak{M}$ the closure of the set of homeomorphisms $\partial D \to \partial D$ with mapping degree 1 in the topology of uniform convergence. Elements of $\mathfrak{M}$ are usually called *monotonic*. Fix three distinct points $p_1, p_2, p_3 \in \partial D$. Define

$$\mathfrak{M}^\times := \{ \varphi \in \mathfrak{M} \mid \varphi(p_i) = p_i,\; i = 1, 2, 3 \}$$

and the spaces of *admissible functions*

$$\mathfrak{C} := \{ f \in C^0(D; \mathbb{R}^m) \cap W^{1,2}_g(D; \mathbb{R}^m) \mid \exists \varphi \in \mathfrak{M}: f|_{\partial D} = \gamma \circ \varphi \},$$

$$\mathfrak{C}^\times := \{ f \in C^0(D; \mathbb{R}^m) \cap W^{1,2}_g(D; \mathbb{R}^m) \mid \exists \varphi \in \mathfrak{M}^\times: f|_{\partial D} = \gamma \circ \varphi \}.$$

The area functional $\mathcal{F}$ is continuously extendable to $\mathfrak{C}$ and one has

$$\inf_{f \in C} \mathcal{F}(f) = \inf_{f \in \mathfrak{C}} \mathcal{F}(f) \le \inf_{f \in \mathfrak{C}} \mathcal{D}(f) = \inf_{f \in \mathfrak{C}^\times} \mathcal{D}(f). \tag{7.2}$$

This first equality holds since $\tilde{C}$ is dense in $W^{1,2}(\Sigma; \mathbb{R}^m)$; since every continuous homeomorphism $\varphi\colon \partial D \to \partial D$ with mapping degree 1 can be a extended to a homeomorhism $D \to D$ which is of class C^1 in the interior[2]; and because of the invariance properties of $\mathcal{F}$. The inequality in (7.2) follows from $\det(\mathbf{A}) \leq \frac{1}{2}|\mathbf{A}|^2$ for all $\mathbf{A} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. The Dirichlet functional is invariant under conformal mappings (see the proof of Lemma 7.7); every conformal mapping $D \to D$ is uniquely and well-defined by prescribing its values on three distinct points. This implies the second equality.

The main theorem in the theory of disk-like minimal surfaces is the following existence theorem, proved by Douglas [16] and simplified later by Courant [10]. A modern account can be found in [13, Chapter 4].

Theorem 7.8 (Existence of disk-like minimal surfaces)
The infimal value of $\mathcal{D}$ on $\mathfrak{C}^\times$ is attained and every such minimizer f fulfills:

1. $f \in C^0(D; \mathbb{R}^m) \cap C^2(D^\circ; \mathbb{R}^m)$.

2. $f|_{\partial D} = \gamma \circ \varphi$ with a homeomorphism $\varphi\colon \partial D \to \partial D$.

3. $\Delta^g f = 0$ in D°.

4. $f^\# g_0 = \varrho\, g$ in D° with some $\varrho \in C^1(D^\circ; [0, \infty[)$.

5. $\mathcal{F}(f) = \inf_{h \in \mathfrak{C}} \mathcal{F}(h)$.

Thus, f is an area minimizing minimal surface.

Regularity theory for minimal surfaces can be found, e.g., in [14, Section 2.3, Theorem 1]. For our exposition, the essential result can be be summarized by:

Theorem 7.9 (Regularity of minimal surfaces)
Assume that $\gamma \in C^{l,\alpha}(\partial D; \mathbb{R}^m)$ with $l \in \mathbb{N}$ and $0 < \alpha < 1$. Then every minimal surface $f \in \mathfrak{C}$ is of class $C^{l,\alpha}$. In particular, the boundary trace $f|_{\partial D}$ is of class $C^{l,\alpha}$.

---

[2]Every homeomorphism of mapping degree 1 is homotopic to the identity along a homotopy $H\colon \partial D \times [0, 1] \to \partial D$ that can be chosen to be smooth on $\partial D \times [0, 1[$. Via polar coordinates, the homotopy can be interpreted as a homeomorphism from an annulus to itself. Gluing the identity of a small disk to this mapping yields a homeomorphism of a disk that has the original mapping as boundary conditions.

7.3 Numerical Methods Derived from Douglas' Existence Proof

Denote by $\mathcal{P} \colon C^0(\partial D; \mathbb{R}^m) \to C^0(D; \mathbb{R}^m) \cap C^2(D^\circ; \mathbb{R}^m)$ the Poisson integral operator:

$$\mathcal{P}(u)(x) := \frac{1}{2\pi} \int_{\partial D} \frac{1 - |x|^2}{|z - x|^2} \, u(z) \, \mathrm{d}z \quad \text{for all } u \in C^0(\partial D; \mathbb{R}^m) \text{ and all } x \in D^\circ.$$

The classical theory on the Possion kernel shows that for each $\varphi \in \mathfrak{M}^\times$, one has:

1. $\mathcal{P}(\gamma \circ \varphi)|_{\partial D} = \gamma \circ \varphi$, thus $\mathcal{P}(\gamma \circ \varphi) \in \mathfrak{C}^\times$.

2. $(\Delta^g \circ \mathcal{P})(\gamma \circ \varphi) = 0$ in D°.

3. $\mathcal{P}(\gamma \circ \varphi) = \arg\min \{ \mathcal{D}(f) \mid f \in C^0(D; \mathbb{R}^m) \cap C^2(D^\circ; \mathbb{R}^m) \text{ with } f|_{\partial D} = \gamma \circ \varphi \}$.

Define the *Douglas functional*[3]

$$\mathcal{J} \colon \mathfrak{M}^\times \to \mathbb{R}, \qquad \mathcal{J}(\varphi) := (\mathcal{D} \circ \mathcal{P})(\gamma \circ \varphi).$$

Thus, an area minimizer $f \in \mathfrak{C}^\times \cap C^2(D^\circ; \mathbb{R}^m)$ as in Theorem 7.8 can be obtained by finding a minimizer φ of $\mathcal{J}$ and putting $f = \mathcal{P}(\gamma \circ \varphi)$. Whenever γ has a certain regularity, say $\gamma \in C^{l,\alpha}(\partial D; \mathbb{R}^m)$ with $l \geq 1$, the regularity theorem above shows that each minimizer $\varphi = \gamma^{-1} \circ f|_{\partial D}$ has the same regularity. Thus, φ can be approximated by closed, piecewise geodesic curves in $S^1 = \partial D$ with mapping degree 1 with a certain convergence rate. This makes it possible to discretize the space $\mathfrak{M}^\times$ and the Douglas functional $\mathcal{J}$. Having found a discrete minimizer φ_n, one may obtain a triangle mesh $f_n = \mathcal{P}_n(\gamma \circ \varphi_n)$ by solving a discrete Poisson equation with boundary condition $\gamma \circ \varphi_n$ (or rather a polygonal approximation of it) in the finite element space of continuous piecewise-linear functions. These triangle meshes f_n are frequently called *discrete minimal surfaces*.

Several authors follow this approach in order to compute numerical approximations of minimal surfaces, e.g., Wilson [39], Tsuchyia [36], Hinze [24], Dzuik and Hutchinson [19], and Pozzi [31]. So far, it was the only approach for which convergence analysis was available (see [20] and [32]). Translated into our notation, the

---

[3]Douglas actually used the integral representation

$$\mathcal{J}(\varphi) = \frac{1}{16\pi^2} \int_{\partial D} \int_{\partial D} \frac{|(\gamma \circ \varphi)(x) - (\gamma \circ \varphi)(y)|^2}{\sin^2\left(\frac{x-y}{2}\right)} \, \mathrm{d}x \, \mathrm{d}y,$$

which can be deduced from Poisson's integral representation.

authors use the test mappings $\Psi(\varphi) := \mathcal{P}(\gamma \circ \varphi)$, $\Psi_n(\varphi_n) := \mathcal{P}_n(\gamma \circ \varphi_n)$ and state for the *critical points* $\mathcal{N} := \{ \varphi \in \mathfrak{M}^\times \mid d\mathcal{J} = 0 \}$:

$$\Psi(\mathfrak{A}^{s,\lambda} \cap \mathcal{N}) \subset \bar{B}(\Psi_n(\mathcal{N}_n), C(s)\,\lambda^{-1}\chi_n) \quad \text{with respect to the } W^{1,2}\text{-norm,}$$

where $\mathcal{N}_n$ denotes the set of critical points $\mathcal{J}_n$ and χ_n denotes the approximation quality of the used triangle mesh. For $s \geq 0$, $\lambda > 0$ the set

$$\mathfrak{A}^{s,\lambda} := \{ \varphi \in \mathfrak{M}^\times \mid |\mathrm{Hess}(\mathcal{J})(\varphi)(u,u)| \geq \lambda \|u\|^2_{W^{\frac{1}{2},2}} \}$$

can be interpreted as a priori information. For the proof, the authors seem to use a variant of the first statement in Theorem 3.30 together with the implicit function theorem (see also Example 3.31).

Albeit theoretically very elegant, this approach has some considerable drawbacks:

1. Because of the extensive use of conformal arguments, it is restricted to the case $\dim(\Sigma) = 2$.

2. For non-disk surfaces Σ, one also has to vary the conformal structure of Σ. This is cumbersome but not impossible as can be seen in the works of Pozzi [31, 32] who treats the case that Σ has the topology of a cylinder.

3. The method does not apply when surface area $\mathcal{F}$ is coupled to some other, conformally non-invariant functional. For examples, this is of interest in the physics of membranes, since surface area is proportional to surface tension energy. Another situation where coupling is desired is when one attempts to use $\mathcal{F}$ as a regularizer for another energy.

In contrary, the direct method of minimizing area among simplicial manifolds is capable of treating any genus and orientability with a single algorithm (see e.g. Figure 7.1) and coupling is available. It is even possible to treat non-manifold examples with the same method.
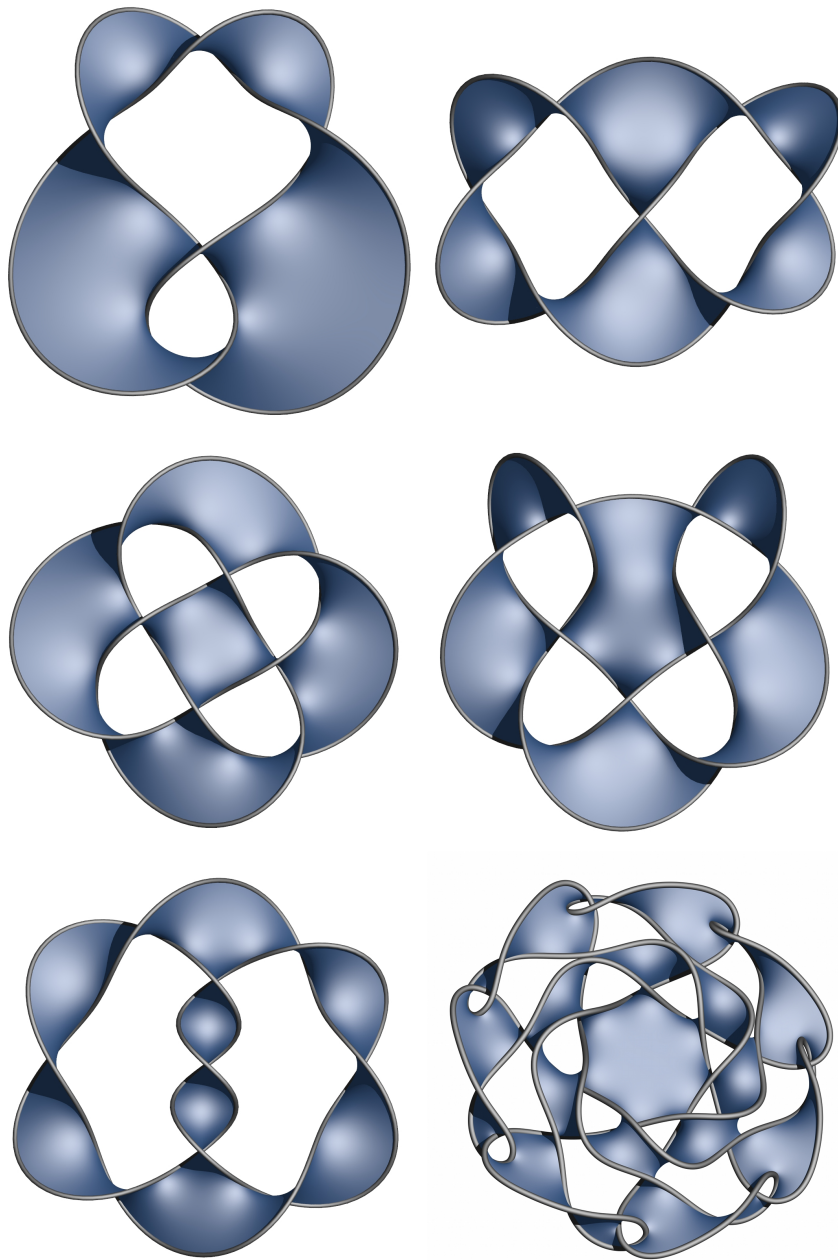
Figure 7.1: Orientable and nonorientable discrete minimal surfaces of nontrivial genus. (Thanks to Rob Scharein for providing the boundary curves in his knot zoo [35].)

7.4 Smooth Setting

In the following we use the abbreviations $\tilde{C} := \text{Imm}(\Sigma; \mathbb{R}^m)$ and $\mathcal{G} := \text{Diff}(\Sigma)$. Define $\mathcal{X} := \tilde{C}/\mathcal{G}$ and denote by $\Psi := \Pi \circ \iota \colon C \to \mathcal{X}$ the composition of the inclusion $\iota \colon C \hookrightarrow \tilde{C}$ and the quotient map $\Pi \colon \tilde{C} \to \mathcal{X} = \tilde{C}/\mathcal{G}$. By Lemma 5.3, $\mathcal{X}$ equipped with the quotient metric $d_{\mathcal{X}}$ induced by d_{Imm}^{τ} is a metric space.

Lemma 7.10 *The function $\mathcal{F}|_C$ is topologically stable along Ψ over $\mathcal{X}$.*

PROOF. *Claim 1.*: $\Psi_{\#}(\mathcal{F}|_C)$ is lower semi-continuous.
We abbreviate $\mathcal{H} := \Psi_{\#}(\mathcal{F}|_C)$ and observe $\mathcal{H} = \Pi_{\#}\iota_{\#}(\mathcal{F}|_C)$. Let $x, x_n \in \mathcal{X}$, $n \in \mathbb{N}$ with $x_n \to x$. We have to show that $\mathcal{H}(x) \leq \liminf_{n\to\infty} \mathcal{H}(x_n)$. The case $\liminf_{n\to\infty} \mathcal{H}(x_n) = \infty$ is trivial. Thus, we may assume that there is a subsequence $(n_l)_{l\in\mathbb{N}}$ such that $\mathcal{H}(x_{n_l}) < \infty$ and $\lim_{l\to\infty} \mathcal{H}(x_{n_l}) = \liminf_{n\to\infty} \mathcal{H}(x_n)$. We may choose $f, f_l \in \tilde{C}$ such that $x = \Pi(f)$, $x_{n_l} = \Pi(f_l)$, $\iota_{\#}(\mathcal{F}|_C)(f_l) \leq \mathcal{H}(x_{n_l}) + \frac{1}{l}$ and such that f_l converges to f in the distance d_{Imm}^{τ} as $l \to \infty$. Since $\iota_{\#}(\mathcal{F}|_C)(f_l)$ is finite and $\iota_{\#}(\mathcal{F}|_C)$ is an extension by infinity, we actually have $f_l \in C$. The set C is closed in $\tilde{C}$, thus $f \in C$ as well. By Corollary 5.12, $\mathcal{F}$ is continuous, thus

$$\mathcal{H}(x) \leq (\mathcal{F}|_C)(f) = \mathcal{F}(f) = \lim_{l\to\infty} \mathcal{F}(f_l) = \lim_{l\to\infty} \mathcal{H}(x_{n_l}) = \liminf_{n\to\infty} \mathcal{H}(x_n).$$

Claim 2.: $\Psi(\arg\min^{\varrho}(\mathcal{F}|_C)) = \arg\min^{\varrho}(\Psi_{\#}(\mathcal{F}|_C))$ for all $\varrho \in [0, \infty[$.
For $x \in \mathcal{X}$, one has $\mathcal{H}(x) = \inf_{y\in\Psi^{-1}(x)} \mathcal{F}(y)$. Observe that $\Psi^{-1}(x)$ is either empty or a whole $\text{Diff}_0(\Sigma)$-orbit in C. The function $\mathcal{F}|_C$ is invariant under $\text{Diff}_0(\Sigma)$. Thus, we have

$$\mathcal{H}(x) = \begin{cases} \mathcal{F}(f), & \text{if there exists } f \in C \text{ with } x = \Psi(f), \\ \infty, & \text{else,} \end{cases}$$

which shows the claim.

Finally, Lemma 2.31 finishes the proof. $\qquad\square$

In the convergence analysis, we will assume that $\gamma \in \text{Imm}(\partial\Sigma; \mathbb{R}^m) \cap W_g^{2,\infty}(\partial\Sigma; \mathbb{R}^m)$. As *a priori information*, we assume that there is a $s \geq 0$ with $\Psi(\mathcal{A}^s \cap \mathcal{M}) \supset \Psi(\mathcal{M})$, where

$$\mathcal{A}^s := \{ f \in C \cap W_g^{2,\infty}(\Sigma; \mathbb{R}^m) \mid d_{\mathcal{P}}(g, f^{\#}g_0) \leq s, \|f\|_{W_g^{2,\infty}} \leq s \} \quad \text{for all } s \geq 0.$$

That means, every minimizer $\Psi(f) \in \Psi(\mathcal{M}) = \arg\min(\Psi_{\#}\mathcal{F})$ allows for a "nice" parametrization $f \colon \Sigma \to \mathbb{R}^m$ with injective differentials, controlled distortion, and

controlled $W_g^{2,\infty}$-norm.[4] The assumption $\Psi(\mathcal{A}^s \cap \mathcal{M}) \supset \Psi(\mathcal{M})$ is satisfied in certain cases: We refer to the detailed regularity theory in [22], in particular to Sections 12.3 and 12.4, where the most relevant case $k = 2$, $n = 3$ is discussed. We point out that we do *not* state, that our a priori assumptions are always satisfied—not even in the case $k = 2$, $m = 3$—but at least for a variety of pairs (Σ, γ). For $k \geq 8$ or for codimensions $m - k$ other than 1, there are actually known counterexamples.

7.5 Discrete Setting

With a smooth triangulation $\mathcal{T}$ of Σ, we can formulate the discrete minimization problem:

Problem 7.11 (Discrete least area problem)
Minimize the discrete volume functional

$$\mathcal{F}_{\mathcal{T}}(f) := \sum_{\sigma \in \mathcal{T}} \int_{\mathrm{conv}(f(V(\sigma)))} \mathrm{vol}_{g_0},$$

on the discrete configuration space

$$C_{\mathcal{T}} := \left\{ f \colon V(\mathcal{T}) \to \mathbb{R}^m \; \middle| \; \begin{array}{c} f|_{V_b(\mathcal{T})} = \gamma|_{V_b(\mathcal{T})}, \\ \forall \sigma \in \mathcal{T} \colon f(V(\sigma)) \textit{ in general position} \end{array} \right\} \subset \mathbb{R}^{m \, \mathrm{card}(V_i(\mathcal{T}))}.$$

Note that the manifold Σ and its triangulation $\mathcal{T}$ do not occur explicitly in the formulation of the problem, only the combinatorics of $\mathcal{T}$. The same is true for many algorithms that were developed for finding critical points of $\mathcal{F}_{\mathcal{T}}$. Wagner [37] applied Newton-like methods for finding zeroes of $d\mathcal{F}_{\mathcal{T}}$. Dzuik [18] and Brakke [7] applied the discrete mean curvature flow in order to produce discrete minimizers. In [30], Pinkall and Polthier presented a further iterative algorithm for minimization of $\mathcal{F}_{\mathcal{T}}$ (see Section 7.7).

In the following, we use the operators $\mathcal{S}_{\mathcal{T}}$, $\tilde{\mathcal{R}}_{\mathcal{T}}$, and $\mathcal{R}_{\mathcal{T}}$ from Section 6.2 and Section 6.4. Recall the definition of the set of discrete a priori information

$$\mathcal{A}_{\mathcal{T}}^r := \{ \, f \in C_{\mathcal{T}} \mid d_{\mathcal{P}(\Sigma)}(g, \tilde{\mathcal{R}}_{\mathcal{T}}(f)^{\#} g_0) \leq r \, \}$$

---

[4]Most of the results of this chapter remain true if one uses—instead of $W_g^{2,\infty}(\Sigma; \mathbb{R}^m)$—any other Banach space $\mathcal{B} \supset C^\infty(\Sigma; \mathbb{R}^m)$ that embeds compactly into $W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$. Other natural choices are $C_g^{\ell,\alpha}(\Sigma; \mathbb{R}^m)$ for $\ell \geq 1$, $\alpha \in \,]0, 1[$ or $W_g^{\ell,p}(\Sigma; \mathbb{R}^m)$ for $\ell \geq 2$, $p > k$. The same applies to the Banach space that describes the regularity of γ.
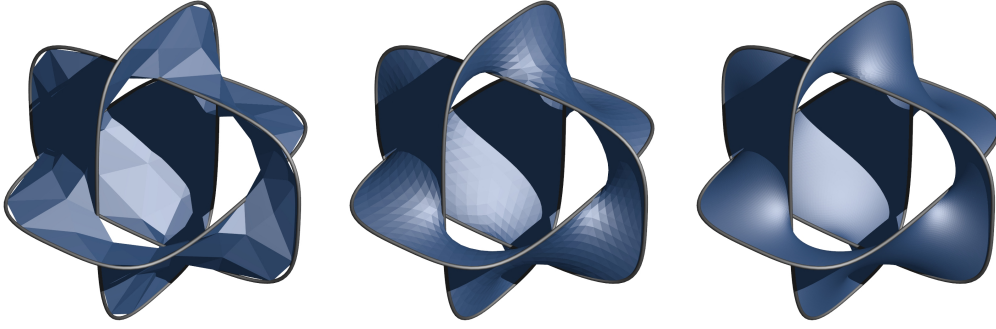
Figure 7.2: Some minimizers of the discrete Plateau problem with Borromean rings as boundary conditions at increasing mesh resolutions (144, 2304, and 36 864 faces respectively).

form (6.7). Additionally, we use

$$\Psi_{\mathcal{T}} := \Pi \circ \tilde{\mathcal{R}}_{\mathcal{T}} : \mathcal{A}_{\mathcal{T}}^r \to \mathcal{X}$$

as discrete test functional. By Lemma 6.11, we may assume that $\mathcal{R}_{\mathcal{T}} : \mathcal{A}_{\mathcal{T}}^r \to C$ is well defined by allowing only triangulations with $\varrho(\mathcal{T})$ sufficiently small, where

$$\varrho(\mathcal{T}) := \max\{\varrho^{(0)}(\mathcal{T}),\ \varrho^{(0)}(\mathcal{T}|_{\partial\Sigma}),\ \varrho^{(1)}(\mathcal{T}),\ \varrho^{(1)}(\mathcal{T}|_{\partial\Sigma})\}.$$

7.6 Convergence Analysis

Lemma 7.12 (Proximity errors) *Let $r > s > 0$. Then there is $\varrho_0 > 0$ and a constant $C \geq 0$ such that for all smooth triangulations with $\varrho(\mathcal{T}) \leq \varrho_0$, one has the following estimates for sampling and reconstruction proximity errors $\varepsilon_{\mathcal{T}}^{\mathcal{S}}$, $\varepsilon_{\mathcal{T}}^{\mathcal{R}}$ of $(\Psi, \Psi_{\mathcal{T}}, \mathcal{S}_{\mathcal{T}}, \mathcal{R}_{\mathcal{T}})$ on $(\mathcal{A}^s, \mathcal{A}_{\mathcal{T}}^r)$:*

$$\varepsilon_{\mathcal{T}}^{\mathcal{S}} \leq C\,\varrho(\mathcal{T}) \quad \text{and} \quad \varepsilon_{\mathcal{T}}^{\mathcal{R}} \leq C\,\varrho(\mathcal{T}).$$

Proof. From Corollary 5.7 and (6.4) one obtains for the sampling proximity error:

$$\varepsilon_{\mathcal{T}}^{\mathcal{S}} = \sup_{f \in \mathcal{A}^s} d_X^\tau(\Psi(f), (\Psi_n \circ \mathcal{S}_{\mathcal{T}})(f))$$

$$= \sup_{f \in \mathcal{A}^s} d_X^\tau(\Pi(f), (\Pi \circ \tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f))$$

$$\leq \sup_{f \in \mathcal{A}^s} d_{\mathrm{Imm}}^\tau(f, (\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f))$$

$$\leq C(s) \sup_{f \in \mathcal{A}^s} \|f - (\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)\|_{W_g^{1,\infty}} \leq C(s)\varrho(\mathcal{T}).$$

Corollary 5.7 and (6.5) imply for the reconstruction proximity error:

$$\varepsilon_{\mathcal{T}}^{\mathcal{R}} = \sup_{f \in \mathcal{A}_{\mathcal{T}}^s} d_X^\tau(\Psi_{\mathcal{T}}(f), (\Psi \circ \mathcal{R}_{\mathcal{T}})(f))$$

$$= \sup_{f \in \mathcal{A}_{\mathcal{T}}^s} d_X^\tau((\Pi \circ \tilde{\mathcal{R}}_{\mathcal{T}})(f), (\Pi \circ \mathcal{R}_{\mathcal{T}})(f))$$

$$\leq \sup_{f \in \mathcal{A}_{\mathcal{T}}^s} d_{\mathrm{Imm}}^\tau(\tilde{\mathcal{R}}_{\mathcal{T}}(f), \mathcal{R}_{\mathcal{T}}(f))$$

$$\leq C(s) \sup_{f \in \mathcal{A}_{\mathcal{T}}^s} \|\tilde{\mathcal{R}}_{\mathcal{T}}(f) - \mathcal{R}_{\mathcal{T}}(f)\|_{W_g^{1,\infty}} = C(s) \|u_{\mathcal{T}}\|_{W_g^{1,\infty}} \leq C(s)\varrho(\mathcal{T}). \qquad \square$$

Lemma 7.13 (Consistency errors) *Let $r > s > 0$. Then there is $\varrho_0 > 0$ and a constant $C \geq 0$ such that for all smooth triangulations with $\varrho(\mathcal{T}) \leq \varrho_0$, one has the following estimates for sampling and reconstruction consistency errors $\delta_{\mathcal{T}}^{\mathcal{S}}$, $\delta_{\mathcal{T}}^{\mathcal{R}}$ of $(\mathcal{F}_{\mathcal{T}}, \mathcal{S}_{\mathcal{T}}, \mathcal{R}_{\mathcal{T}})$ with respect to $\mathcal{F}$ on $(\mathcal{A}^s, \mathcal{A}_{\mathcal{T}}^r)$:*

$$\delta_{\mathcal{T}}^{\mathcal{S}} \leq C\varrho(\mathcal{T}) \quad and \quad \delta_{\mathcal{T}}^{\mathcal{R}} \leq C\varrho(\mathcal{T}).$$

Proof. Fix $f \in \mathcal{A}^s$ and abbreviate $f_{\mathcal{T}} := (\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)$. Observe $(\mathcal{F}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f) = \mathcal{F}(f_{\mathcal{T}})$. By Corollary 5.12, Lemma 5.5, and (6.4), we obtain

$$(\mathcal{F}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f) - \mathcal{F}(f) = \mathcal{F}(f_{\mathcal{T}}) - \mathcal{F}(f)$$

$$\leq \mathcal{F}(f) \, e^{\sqrt{k}\, d_C(f_{\mathcal{T}}, f)} \, \sqrt{k}\, d_C(f_{\mathcal{T}}, f)$$

$$\leq C(s) \|f_{\mathcal{T}} - f\|_{W_g^{1,\infty}} \leq C(s)\varrho(\mathcal{T}),$$

thus the sampling consistency error $\delta_{\mathcal{T}}^{\mathcal{S}}$ is bounded by $\delta_{\mathcal{T}}^{\mathcal{S}} \leq C\varrho(\mathcal{T})$.

For $f \in \mathcal{A}_{\mathcal{T}}^r$, put $f_{\mathcal{T}} = \tilde{\mathcal{R}}_{\mathcal{T}}(f)$. From local Lipschitz continuity of $\mathcal{F}$ (see Corollary 5.12) and (6.5), we deduce

$$(\mathcal{F} \circ \mathcal{R}_{\mathcal{T}})(f) - \mathcal{F}_{\mathcal{T}}(f) = \mathcal{F}(f_{\mathcal{T}} + u_{\mathcal{T}}) - \mathcal{F}(f_{\mathcal{T}})$$

$$\leq \mathcal{F}(f_{\mathcal{T}}) \, C(r) \|u_{\mathcal{T}}\|_{W_g^{1,\infty}} \leq C(r)\varrho(\mathcal{T}),$$

obtaining $\delta_{\mathcal{T}}^{\mathcal{R}} \leq C\varrho(\mathcal{T})$. $\qquad \square$

Let Σ be a compact, k-dimensional smooth manifold with boundary, let g be a smooth Riemannian metric on Σ and let $\gamma \in \mathrm{Imm}(\partial\Sigma; \mathbb{R}^m) \cap W^{2,\infty}(\partial\Sigma; \mathbb{R}^m)$ be an immersion of class $W^{2,\infty}$. Let $(\mathcal{T}_n)_{n \in \mathbb{N}}$ be a sequence of smooth triangulations of (Σ, g) with $\varrho_n := \varrho(\mathcal{T}_n) \overset{n \to \infty}{\longrightarrow} 0$. Instead of $C_{\mathcal{T}_n}, \mathcal{F}_{\mathcal{T}_n}, \mathcal{A}^r_{\mathcal{T}_n}, \ldots$ we shall write $C_n, \mathcal{F}_n, \mathcal{A}^r_n, \ldots$ As in Chapter 2, we denote the sets of δ-minimizers by $\mathcal{M}^\delta := \arg\min^\delta(\mathcal{F}|_C)$ and $\mathcal{M}^\delta_n := \arg\min^\delta(\mathcal{F}_n)$.

Theorem 7.14 (Kuratowksi convergence of minimizers)
Assume $\emptyset \neq \Psi(\mathcal{M}) \subset \Psi(\mathcal{A}^s \cap \mathcal{M})$ for some $s \in \,]0, \infty[$ and that the $\mathcal{A}^r_n$ are valid[5] for some $r \in \,]s, \infty[$ and all $n \in \mathbb{N}$. Then there is a constant $C \geq 0$ depending on Σ, g, γ, s, r only, such that

$$\underset{n \to \infty}{\mathrm{Ls}}\ \Psi_n(\mathcal{A}^r_n \cap \mathcal{M}_n) \subset \Psi(\mathcal{M}) = \underset{n \to \infty}{\mathrm{Lt}}\ \Psi_n(\mathcal{A}^r_n \cap \mathcal{M}^{\delta_n}_n)$$

holds with $\delta_n \leq C \varrho_n$ for sufficiently large $n \in \mathbb{N}$. The convergence is with respect to the topology generated by $d^\tau_\mathcal{X}$.

PROOF. We would like to apply Theorem 2.28 with $\mathcal{K} := \mathcal{X}$. Hence, all we have to do is to check its conditions:

Proximity and consistency follow from Lemma 7.12 and Lemma 7.13. The inclusion $\mathcal{S}_n(\mathcal{A}^s) \subset \mathcal{A}^r_n$ was shown in Lemma 6.11. According to Lemma 7.10, $\mathcal{F}$ is topologically stable along Ψ over $\mathcal{X}$. Validity of $\mathcal{A}^s$ with respect to $(\mathcal{F}|_C, \mathcal{S}_n)$ follows from $\emptyset \neq \Psi(\mathcal{M}) \subset \Psi(\mathcal{A}^s \cap \mathcal{M})$ and the validity of $\mathcal{A}^r_n$ was imposed as a condition. $\square$

Remark 7.15 If each connected component of Σ has nontrivial boundary, the existence of some $s \in \,]0, \infty[$ such that $\mathcal{A}^s$ is valid also implies the existence of minimizers as we show now. Recall the definition of $\mathcal{A}^s$ from (6.6):

$$\mathcal{A}^s := \{ f \in C \cap W^{2,\infty}_g(\Sigma; \mathbb{R}^m) \mid d_\mathcal{P}(g, f^\# g_0) \leq s, \|\mathrm{d}f\|_{W^{1,\infty}_g} \leq s \}.$$

The Arzelà-Ascoli theorem states that $\mathcal{B} := W^{2,\infty}_g(\Sigma; \mathbb{R}^m)$ is compactly embedded into $W^{1,\infty}_g(\Sigma; \mathbb{R}^m)$. If the boundary of each connected component of Σ is nontrivial, bounds on $\|\mathrm{d}f\|_{W^{1,\infty}_g}$ and the imposed boundary conditions imply also bounds on $\|f\|_{W^{2,\infty}_g}$. Thus, $\mathcal{A}^s$ is bounded in $\mathcal{B}$ and by Lemma 5.9, it is a relatively compact subset in $(C, d^\tau_{\mathrm{Imm}})$. Since $\mathcal{F}|_C$ is lower semi-continuous, the direct method of the calculus of variations and $\inf(\mathcal{F}|_{\mathcal{A}^s}) = \inf(\mathcal{F}|_C)$ imply $\emptyset \neq \arg\min(\mathcal{F}|_{\mathcal{A}^s}) = \mathcal{A}^s \cap \mathcal{M} \subset \mathcal{M}$.

---

[5]see Definition 2.4.
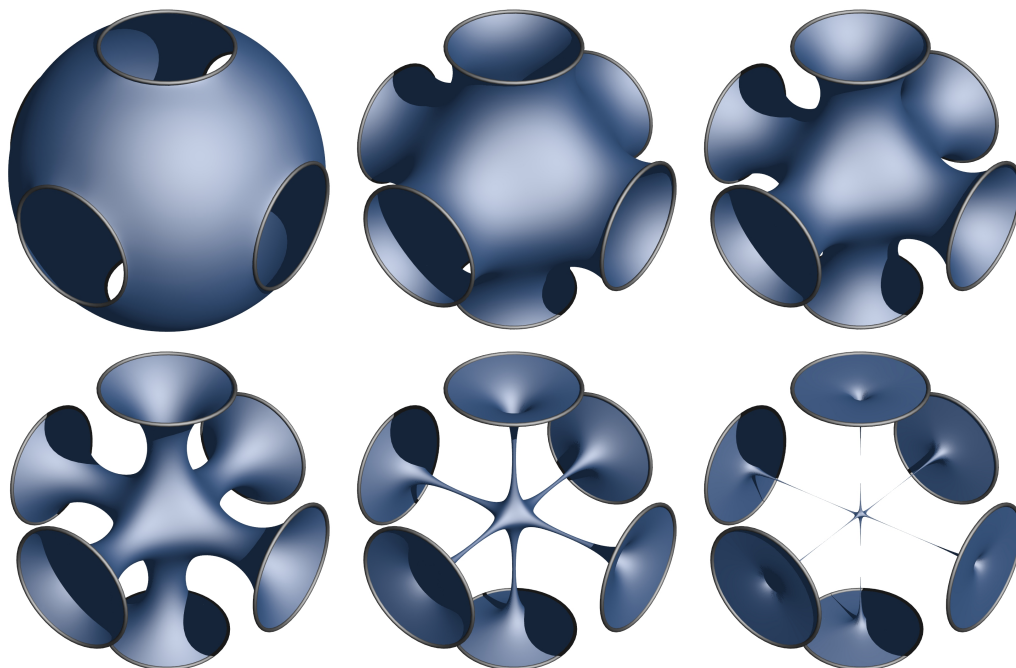
Figure 7.3: Sequence of surfaces obtained by a descending flow starting from a sixfold perforated sphere and degenerating into six disks.

Remark 7.16 The validity of the discrete a priori assumptions $\mathcal{A}_{\mathcal{T}}^r$ amounts to the existence of a non-degenerating minimizing sequence of simplicial meshes in $\mathbb{R}^m$. More precisely, $\mathcal{A}_{\mathcal{T}}^r$ is valid if and only if for each $l \in \mathbb{N}$ there is $f_l \in C_{\mathcal{T}}$ such that $(\tilde{\mathcal{R}}_{\mathcal{T}}(f_l))_{l \in \mathbb{N}}$ is bounded in d_{Imm} and $\lim_{l \to \infty} \mathcal{F}_{\mathcal{T}}(f_l) = \inf(\mathcal{F}_{\mathcal{T}})$. In particular, for each $\sigma \in \mathcal{T}$, the affine mappings $h_l := \tilde{\mathcal{R}}_{\mathcal{T}}(f_l) \circ \sigma \colon \Delta_k \to \mathbb{R}^m$ have to be bounded in $(\mathrm{Imm}(\Delta_k; \mathbb{R}^m), d_{\mathrm{Imm}})$. This implies uniform bounds on $\|\mathrm{d}h_l\|$ and $\|\mathrm{d}h_l^\dagger\|$ for all $l \in \mathbb{N}$. Note that $\|\mathrm{d}h_l\|$ and $\|\mathrm{d}h_l^\dagger\|$ are descriptors for the quality of the simplex $h_l(\Delta_k)$ since h_l is affine.

In general, $\mathcal{A}_{\mathcal{T}}^r$ needs not to be valid as Figure 7.3 illustrates. Note that in the shown case, the degenerating minimizing sequence indicates the non-existence of minimizers in the given topological class.[6]

Maybe, the best one may hope to find in the course of a discrete regularity theory is that $\mathcal{A}_{\mathcal{T}}^r$ is valid if $\mathcal{A}^s$ is. Another rewarding task might be the further modification

---

[6]Still, the sequence seems to converge in a weaker topology to a "minimizer", e.g., in the sense of integral currents.

of $\mathcal{A}_{\mathcal{T}}^r$ and $\mathcal{R}_{\mathcal{T}}$ in order to establish $\mathcal{R}_{\mathcal{T}}(\mathcal{A}_{\mathcal{T}}^r) \subset \mathcal{A}^t$ for some $t > r$ and all sufficiently fine $\mathcal{T}$, since this would show Hausdorff convergence of the solution sets.[7]

Remark 7.17 Adding quantitative stability, our results would yield certain convergence rates (see Theorem 2.41). Although it seems reasonable to analyze the Hessian of the volume functional around minimizers in order to obtain such estimates, we have currently no idea how to do so in an a priori fashion.

7.7 Descending Flows

For the practical application of the presented theory, one needs numerical methods for finding minimizers of $\mathcal{F}_{\mathcal{T}}$ in $\mathcal{C}_{\mathcal{T}}$. This can be accomplished, e.g., by following volume decreasing flows on the configuration space. We consider two such flows on the smooth configuration space C before we discuss their discretizations on $\mathcal{C}_{\mathcal{T}}$.

We recall that the Lipschitz immersions $C \subset W_g^{1,\infty}(\Sigma; \mathbb{R}^m)$ form an open set with respect to $\|\cdot\|_{W_g^{1,\infty}}$. Note that one has $T_f C = W_{0,g}^{1,\infty}(\Sigma; \mathbb{R}^m)$ and that the differential $d\mathcal{F}$ of the volume functional is given by (see Lemma 5.15)

$$\langle d\mathcal{F}|_f, u \rangle = \int_{\Sigma} \langle df, du \rangle_{f^\# g_0} \, \text{vol}_{f^\# g_0} \quad \text{for all } u \in W_{0,g}^{1,\infty}(\Sigma; \mathbb{R}^m).$$

7.7.1 Mean curvature flow

The mean curvature flow, used for example by Brakke [7] and Dzuik [18] is defined by the partial differential equation $\frac{\partial}{\partial t} F(t) = \vec{H}_{F(t)}$, $F(0) = f \in C$. Here the *mean curvature vector* $\vec{H}_f$ for $f \in C$ is defined by the weakly formulated equation

$$\int_{\Sigma} \langle \vec{H}_f, v \rangle_{f^\# g_0} \, \text{vol}_{f^\# g_0} = -\langle d\mathcal{F}|_f, v \rangle \quad \text{for all } v \in C_0^\infty(\Sigma; \mathbb{R}^m). \tag{7.3}$$

Often, $-\vec{H}_f$ is incorrectly referred to as *the* gradient of the volume functional $\mathcal{F}$. Although it may seem instructive to do so, there are two substantial issues with this point of view:

1. A gradient is by definition a vector field. The mapping $f \mapsto \vec{H}_f$ does not constitute a vector field: If $f \in C \cap W_g^{l,\infty}(\Sigma; \mathbb{R}^m)$ and $l \geq 2$, one may apply partial integration on the right hand side in order to interpret $\vec{H}_f = \Delta^{f^\# g_0} f$ as

---

[7]Showing that $\Psi_{\mathcal{T}}(\mathcal{A}_{\mathcal{T}}^r)$ is contained in *some* relatively compact set $\mathcal{K} \subset \mathcal{X}$ would suffice.

an element of $W_g^{l-2,\infty}(\Sigma; \mathbb{R}^m)$, where $\Delta_f = -\operatorname{div}^{f^\# g_0} \operatorname{grad}^{f^\# g_0}$. But in general, $\vec{H}_f$ can only be understood as a distribution and not as a measurable function $\Sigma \to \mathbb{R}^m$, let alone as an element of $W_{0,g}^{1,\infty}(\Sigma; \mathbb{R}^m) = T_f C$.

2. One has to specify a Riemannian metric with respect to which a gradient should be defined. In view of (7.3), one could call $\vec{H}_f$ the gradient with respect to the L^2-inner product—if the L^2-inner product were an inner product on the tangent bundle $T_f C$.

This is also why the equation $\frac{\partial}{\partial t} F(t) = \vec{H}_{F(t)}$ cannot be treated as an ordinary differential equation. It has to be analyzed as a nonlinear parabolic partial differential equation with all its hardships. For example, even proving the existence of short-term solutions is non-trivial.

7.7.2 Another gradient-like flow

We may extend $d\mathcal{F}|_f$ to the continuous linear form $\alpha|_f \in (W_{0,g}^{1,1}(\Sigma; \mathbb{R}^m))'$, leading to the operator $\alpha \colon C \to (W_{0,g}^{1,1}(\Sigma; \mathbb{R}^m))'$. The Laplacian

$$\bar{A}_f \colon W_{0,g}^{1,\infty}(\Sigma; \mathbb{R}^m) \to (W_{0,g}^{1,1}(\Sigma; \mathbb{R}^m))', \quad u \mapsto \left(v \mapsto \int_\Sigma \langle du, dv \rangle_{f^\# g_0} \operatorname{vol}_{f^\# g_0} \right)$$

is an isomorphism of Banach spaces according to Lemma 1.2 (see also Remark 3.14). By Lemma 4.15 and Lemma 5.5, both

$$\alpha \colon C \to (W_{0,g}^{1,1}(\Sigma; \mathbb{R}^m))', \quad \text{and}$$
$$\bar{A} \colon C \to L(W_{0,g}^{1,\infty}(\Sigma; \mathbb{R}^m); (W_{0,g}^{1,1}(\Sigma; \mathbb{R}^m))'), \quad f \mapsto \bar{A}_f$$

are locally Lipschitz continuous with respect to $\|\cdot\|_{W_g^{1,\infty}}$. This allows us to define a vector field $X \in \Gamma(C; TC)$ on C which is locally Lipschitz continuous with respect to $\|\cdot\|_{W_g^{1,\infty}}$ by

$$X|_f := \bar{A}_f^{-1} \alpha|_f \quad \text{for all } f \in C.$$

More explicitly, $X|_f \in W_{0,g}^{1,\infty}(\Sigma; \mathbb{R}^m)$ is the unique solution of the weakly formulated equation

$$\int_\Sigma \langle d(X|_f), dv \rangle_{f^\# g_0} \operatorname{vol}_{f^\# g_0} = \int_\Sigma \langle df, dv \rangle_{f^\# g_0} \operatorname{vol}_{f^\# g_0} \quad \text{for all } v \in W_{0,g}^{1,1}(\Sigma; \mathbb{R}^m). \quad (7.4)$$

Substituting $v = X|_f$, we obtain

$$\langle d\mathcal{F}|_f, X|_f \rangle = \int_\Sigma \langle df, d(X|_f) \rangle_{f^\# g_0} \, \mathrm{vol}_{f^\# g_0} \tag{7.5}$$

$$= \int_\Sigma \langle d(X|_f), d(X|_f) \rangle_{f^\# g_0} \, \mathrm{vol}_{f^\# g_0} \geq \tfrac{1}{(1+C_2^2)} \|X|_f\|^2_{W^{1,2}_{0,g}}. \tag{7.6}$$

Now, consider the ordinary differential equation

$$\dot{F}(t) = -X|_{F(t)}, \qquad F(0) = f$$

for a curve $F \colon]-\varepsilon, \varepsilon[\to C$, where $\varepsilon > 0$. The vector field X is locally Lipschitz continuous. Thus the Picard-Lindelöff theorem shows that the initial value problem has short-term solutions for all initial surfaces $f \in C$ and the trajectories are of class $C^{1,1}(]-\varepsilon, \varepsilon[; C)$ with respect to $\|\cdot\|_{W^{1,\infty}_g}$. Let F be such a trajectory. Then one has by (7.6):

$$\tfrac{d}{dt} \mathcal{F}(F(t)) = -\langle d\mathcal{F}|_{F(t)}, X|_{F(t)} \rangle \leq -\tfrac{1}{(1+C_2^2)} \|X|_{F(t)}\|^2_{W^{1,2}_{0,g}}.$$

This shows that stationary points of X are precisely the critical points of $\mathcal{F}$ in C and that the flow generated by X is descending with respect to $\mathcal{F}$, i.e., $\mathcal{F}(F(t)) \leq \mathcal{F}(F(s))$ for $t \geq s$.

Long-term trajectories need not exist. For example, Figure 7.3 shows how the flow (or rather its discretization, see below) arrives at the boundary of C within a finite amount of time. In terms of the metrics d_C and d_C^τ, the flow actually converges to infinity in the sense that it leaves every bounded set within a finite amount of time.

By a slight abuse of notation, the vector field $X|_f$ can be related to the mean curvature vector $\vec{H}_f$ through

$$X|_f = -\bar{A}_f^{-1} \vec{H}_f,$$

because $\vec{H}_f$ may be considered to be equal to the restriction of $\alpha|_f$ to $C_0^\infty(\Sigma; \mathbb{R}^m)$. In the same (non-)sense as $-\vec{H}_f$ can be interpreted as the L^2-gradient of $\mathcal{F}$, $X|_f$ can be seen as the $W^{1,2}$-gradient of $\mathcal{F}$.

7.7.3 Discretization

The discretization of the vector field X is straightforward: For a triangulation $\mathcal{T}$ with $N := \operatorname{card}(V(\mathcal{T}))$ and $N_{\mathrm{i}} := \operatorname{card}(V_{\mathrm{i}}(\mathcal{T}))$ consider $\mathcal{F}_{\mathcal{T}} \colon C_{\mathcal{T}} \subset \mathbb{R}^{mN_{\mathrm{i}}} \to \mathbb{R}$ as a function on $\mathbb{R}^{mN_{\mathrm{i}}}$. Let $f \in C_{\mathcal{T}}$ and consider the differential $\mathrm{d}_f \mathcal{F}_{\mathcal{T}}$ as an element in $(\mathbb{R}^{mN_{\mathrm{i}}})'$. Define the *discrete Laplacian* $A_{\mathcal{T},f}$ as the finite element discretization of $\bar{A}_f$, i.e., by the discrete weak formulation

$$\langle A_{\mathcal{T},f}\, u, v \rangle = \langle \bar{A}_{\tilde{\mathcal{R}}_{\mathcal{T}}(f)}\tilde{\mathcal{R}}_{\mathcal{T}}\, u, \tilde{\mathcal{R}}_{\mathcal{T}}\, v \rangle \quad \text{for all } u, v \in \mathbb{R}^{mN}.$$

This way, the discrete Laplacian A_f is a linear operator

$$A_{\mathcal{T},f} \colon \mathbb{R}^{mN} \to (\mathbb{R}^{mN})'.$$

When $k = 2$, $A_{\mathcal{T},f}$ is the cotan-Laplacian applied component-wise. As a bilinear form, it is merely the pullback of $\bar{A}$ along $\tilde{\mathcal{R}}_{\mathcal{T}}$. Now, we may define $X_{\mathcal{T}} \in \mathfrak{X}(C_{\mathcal{T}})$ by the discrete weak equation

$$\langle A_{\mathcal{T},f}\, (X_{\mathcal{T}}|_f), v \rangle = \langle \mathrm{d}_f \mathcal{F}_{\mathcal{T}}, v \rangle \quad \text{for all } v \in T_f C_{\mathcal{T}} \cong \mathbb{R}^{mN_{\mathrm{i}}}.$$

Moreover, one has $\langle \mathrm{d}_f \mathcal{F}_{\mathcal{T}}, v \rangle = \langle A_{\mathcal{T},f} f, v \rangle$ for all $f \in C_{\mathcal{T}}$ and $v \in T_f C_{\mathcal{T}}$ so that one only needs the discrete Laplacian in order to calculate $X_{\mathcal{T}}$:[8]

$$\langle A_{\mathcal{T},f}\, (X_{\mathcal{T}}|_f), v \rangle = \langle A_{\mathcal{T},f}\, f, v \rangle \quad \text{for all } v \in T_f C_{\mathcal{T}} \cong \mathbb{R}^{mN_{\mathrm{i}}}. \tag{7.7}$$

The flow generated by X may be discretized, e.g., by explicit-Euler integration of $X_{\mathcal{T}}$:

$$f_{\nu+1} := f_\nu - \eta\, X_{\mathcal{T}}|_{f_\nu}, \quad \nu = 0, 1, 2, \ldots$$

with some $\eta > 0$. We used this discrete flow to compute the discrete minimal surfaces depicted in this work (see Figures 0.1, 0.2, 7.1, and 8.1).

The discrete flow $X_{\mathcal{T}}$ is by no means new: For $\eta = 1$, it has already been proposed by Pinkall and Polthier in [30], even though it may be hard to recognize. Translated into our language, the authors define the element $f_{\nu+1} \in C_{\mathcal{T}}$ as the minimizer of the Dirichlet energy $\mathcal{D}_{\mathcal{T},f_\nu}$ in the affine space

$$\{\, h \colon V(\mathcal{T}) \to \mathbb{R}^m \,\big|\, h|_{V_{\mathrm{b}}(\mathcal{T})} = \gamma|_{V_{\mathrm{b}}(\mathcal{T})} \,\} = f_\nu + T_{f_\nu} C_{\mathcal{T}},$$

---

[8]We point out that mass matrices are *not* needed here because we do not use the L^2-Riesz isomorphism.
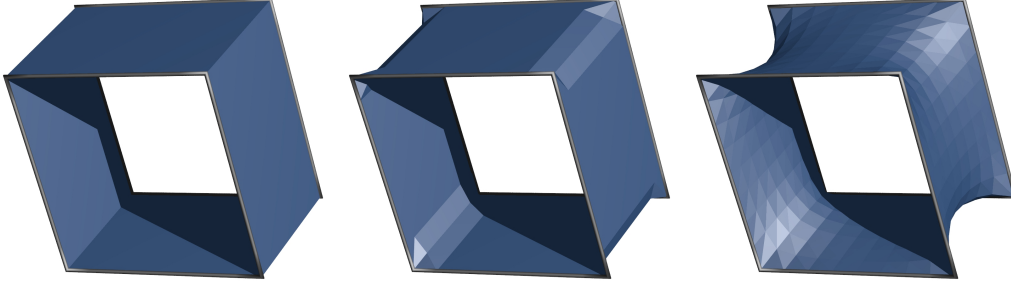
Figure 7.4: Left to right: Initial surface (512 faces); surface after one explicit Euler step in discrete mean curvature direction; surface after one explicit Euler step in direction $X_{\mathcal{T}}$.
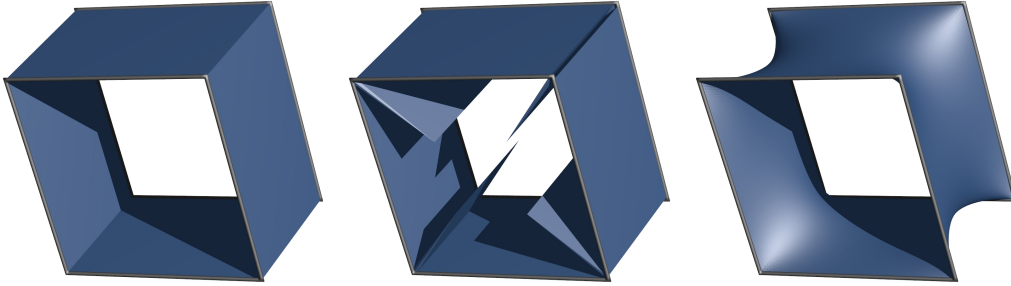
Figure 7.5: The same as in Figure 7.4 with the same step sizes, but with refined mesh (32 768 faces).

where $\mathcal{D}_{\mathcal{T},f}(h) := \frac{1}{2} \int_{\Sigma} |\mathrm{d}(\tilde{\mathcal{R}}_{\mathcal{T}}(h))|^2_{f^\# g_0} \, \mathrm{vol}_{f^\# g_0}$ for all $f \in C_{\mathcal{T}}$. Thus, the necessary and sufficient condition for $f_{\nu+1}$ to be a minimizer of $\mathcal{D}_{\mathcal{T},f_\nu}$ is:

$$0 = \langle \mathrm{d}\mathcal{D}_{\mathcal{T},f_\nu}|_{f_{\nu+1}}, v \rangle \quad \text{for all } v \in T_{f_\nu} C_{\mathcal{T}}.$$

Since

$$\langle \mathrm{d}\mathcal{D}_{\mathcal{T},f_\nu}|_{f_{\nu+1}}, v \rangle = \int_{\Sigma} \langle \mathrm{d}(\tilde{\mathcal{R}}_{\mathcal{T}}(f_{\nu+1})), \mathrm{d}(\tilde{\mathcal{R}}_{\mathcal{T}}(v)) \rangle_{\tilde{\mathcal{R}}_{\mathcal{T}}(f_\nu)^\# g_0} \, \mathrm{vol}_{\tilde{\mathcal{R}}_{\mathcal{T}}(f_\nu)^\# g_0}$$
$$= \langle \bar{A}_{\tilde{\mathcal{R}}_{\mathcal{T}}(f_\nu)} \circ \tilde{\mathcal{R}}_{\mathcal{T}}(f_{\nu+1}), \tilde{\mathcal{R}}_{\mathcal{T}}(v) \rangle = \langle A_{\mathcal{T},f_\nu}(f_{\nu+1}), v \rangle,$$

we obtain by writing $f_{n+1} = f_\nu + u$ with $u \in T_{f_\nu} C_{\mathcal{T}}$:

$$\langle A_{\mathcal{T},f_\nu} u, v \rangle = -\langle A_{\mathcal{T},f_\nu} f_\nu, v \rangle \quad \text{for all } v \in T_{f_\nu} C_{\mathcal{T}}.$$

Thus, by (7.7), we have $f_{\nu+1} = f_\nu + u = f_\nu - X_{\mathcal{T}}|_{f_\nu}$.

Although this discrete flow is already known, we contribute an explanation as to why this flow performs so much better in area minimization than the discrete mean curvature flow—a fact that was already observed by Pinkall and Polthier:

The discrete mean curvature flow involves a discretization of a nonlinear, parabolic *partial* differential equation. A striking consequence in practice is the need for decreasing the step size of time integrators along with mesh refinement. Otherwise, one risks a lot of overshoot: Parabolic partial differential equations have infinite propagation speed, but the discrete mean curvature flow propagates information only from one vertex to its neighbors. This has to be compensated by tiny step sizes.

In contrast, the flow of $X_{\mathcal{T}}$ is the discretization of an *ordinary* differential equation; one may use the same step size for the explicit-Euler integrator—independent of the mesh resolution—without disrupting the mesh (compare Figure 7.4 with Figure 7.5). Because the vector field $X_{\mathcal{T}}$ contains an inverse Laplacian, information at one vertex is immediately distributed along the whole mesh within a single time integration step.

8 Hencky Elasticity

Our second detailed example of an application of Chapter 2 is a discretization of a nonlinear elasticity model together with a qualitative convergence analysis. This elasticity model is usually attributed to Heinrich Hencky (see the English translation of some of Hencky's original work [29] and, in particular, the comments of the translators). The Hencky strain tensor, which we obtain as the gradient field of a squared distance function on $P(V)$ (see Section 4.4) is often referred to as logarithmic or "true" strain.

This chapter is actually an unintentional by-product of our treatment of discrete minimal surfaces (Chapter 7). More precisely, the need for a $\mathrm{Diff}(\Sigma)$-invariant distance on shape space $\mathrm{Imm}(\Sigma; \mathbb{R}^m)$ caused us to develop the distance d_{Imm} which happens to contain the Hencky strain tensor. It is essentially the (lower semi-)continuity of the elastic Hencky energy with respect to d_{Imm} that enables us to use Theorem 2.28.

As it turns out, Hencky elasticity is somewhat easier to treat since it involves *parameterized* immersions, so we do not have to descend to the quotient $\mathrm{Imm}(\Sigma; \mathbb{R}^m)/\mathrm{Diff}(\Sigma)$.

Exactly as in Chapter 7, the conducted analysis depends on existence and regularity assumptions in both the smooth *and* the discrete setting.

8.1 Hencky's Constitutive Function

Actually, the constitutive function occurred already in Lemma 4.23 as a squared distance function:

Definition 8.1 Let V be a finite-dimensional real vector space. For fixed $b_{\mathrm{ref}} \in P(V)$ define *Hencky's constitutive function* with *Lamé coefficients* $\mu > 0$ and $\lambda > -\frac{2\mu}{\dim(V)}$ by

$$W_{\mu,\lambda} \colon P(V) \to \mathbb{R}, \quad W_{\mu,\lambda}(b) := \tfrac{1}{2} \operatorname{dist}^2_{\mu,\lambda}(b, b_{\mathrm{ref}}) \quad \text{for all } b \in P(V).$$

Remark 8.2 In order to interpret Hencky's constitutive function in terms of engineering notation, choose $\mathbf{L}_{\mathrm{ref}}$, $\mathbf{L} \in \mathrm{Mat}_{m \times m}(\mathbb{R})$ with $\mathbf{L}^{\mathsf{T}}\mathbf{L} = \mathbf{G}_e(b)$ and $\mathbf{L}_{\mathrm{ref}}^{\mathsf{T}}\mathbf{L}_{\mathrm{ref}} = \mathbf{G}_e(b_{\mathrm{ref}})$. Interpreting $\mathbf{F} := \mathbf{L}\mathbf{L}_{\mathrm{ref}}^{-1}$ as the "deformation gradient", one obtains

$$
\begin{aligned}
W_{\mu,\lambda}(b) &:= \tfrac{1}{2}\,\mathrm{dist}_{\mu,\lambda}^2(b, b_{\mathrm{ref}}) \\
&= \mu\,\mathrm{tr}\log^2\left(\mathbf{L}^{-\mathsf{T}}\mathbf{L}_{\mathrm{ref}}^{\mathsf{T}}\mathbf{L}_{\mathrm{ref}}\mathbf{L}^{-1}\right) + \tfrac{\lambda}{2}\log^2\left(\tfrac{\det \mathbf{L}}{\det \mathbf{L}_{\mathrm{ref}}}\right) \\
&= \mu\,\mathrm{tr}\log^2\left(\mathbf{F}^{-\mathsf{T}}\mathbf{F}^{-1}\right) + \tfrac{\lambda}{2}\log^2\left(\det \mathbf{F}\right) \\
&= \mu\,|\log(\mathbf{F}^{\mathsf{T}}\mathbf{F})|^2 + \tfrac{\lambda}{2}\log^2\left(\det \mathbf{F}\right).
\end{aligned}
$$

Lemma 8.3 *The gradient of $W_{\mu,\lambda}$ with respect to $g_{\mu,\lambda}$ at $b \in P(V)$ is given by the Hencky field:*

$$
\mathrm{grad}^{\mu,\lambda}\,W_{\mu,\lambda}|_b = -H_{b_{\mathrm{ref}}}|_b.
$$

PROOF. Let $\gamma\colon [0,1] \to P(V)$ be the unique geodesic from b to b_{ref} with respect to both g_P and $g_{\mu,\lambda}$. The function $W_{\mu,\lambda}$ is a squared distance function, thus we have

$$
\mathrm{grad}^{\mu,\lambda}\,W_{\mu,\lambda}|_b = \dot\gamma(0) = \mathrm{grad}^{g_P}\,W_{\mu,\lambda}|_b = -H_{b_{\mathrm{ref}}}|_b. \qquad \square
$$

Lemma 8.4 *One has*

$$
W_{\mu,\lambda}(b_{\mathrm{ref}}) = 0, \quad \mathrm{d}W_{\mu,\lambda}|_{b_{\mathrm{ref}}} = 0, \quad \text{and} \quad \mathrm{Hess}^{\mu,\lambda}(W_{\mu,\lambda})|_{b_{\mathrm{ref}}} = \mathrm{Hess}^{g_P}(W_{\mu,\lambda})|_{b_{\mathrm{ref}}} = g_{\mu,\lambda}\,|_{b_{\mathrm{ref}}}.
$$

PROOF. The Levi-Civita connections of g_P and $g_{\mu,\lambda}$ coincide by Lemma 4.21. Thus, the same is true for their Hessian operators:

$$
\mathrm{Hess}^{g_P} = \nabla\mathrm{d} = \mathrm{Hess}^{\mu,\lambda}.
$$

Because $W_{\mu,\lambda}(b)$ is a squared distance function with center b_{ref}, the statements follow immediately. $\qquad \square$

8.2 Hencky's Elastic Energy

Definition 8.5 Let (M, g_{ref}), (M_0, g_0) be m-dimensional smooth Riemannian manifolds with boundary and let μ, $\lambda \in L_{g_{\mathrm{ref}}}^\infty(M; \mathbb{R})$ be measurable bounded functions with $\mu > 0$ and $\lambda > -\frac{2\mu}{m}$ almost everywhere. We define *Hencky's elastic energy* on $\mathrm{Imm}(M; M_0)$ by

$$
\mathcal{H}(f) := \int_M W_{\mu,\lambda}(f^{\#}g_0)\,\mathrm{vol}_{g_{\mathrm{ref}}} \quad \text{for all } f \in \mathrm{Imm}(M; M_0),
$$

with the constitutive bundle mapping

$$
\begin{array}{ccc}
P(TM) & \xrightarrow{\ W_{\mu,\lambda}\ } & M \times \mathbb{R} \\
& \searrow{\scriptstyle \pi} \qquad {\scriptstyle \mathrm{pr}_M}\swarrow & \\
& M &
\end{array}
$$

given by $W_{\mu,\lambda}|_x = W_{\mu(x),\lambda(x)} \colon P(T_xM) \to \mathbb{R}$ with $W_{\mu(x),\lambda(x)}(b) := \frac{1}{2}d^2_{\mu(x),\lambda(x)}(b|_x, g_{\mathrm{ref}}|_x)$.

We refer to M as the *material manifold*, to g_{ref} as the *reference metric* and to the functions μ and λ as the *Lamé coefficients*.

Remark 8.6 We are going to treat the case $M_0 = \mathbb{R}^m$, where g_0 is the Euclidean inner product on $\mathbb{R}^m$. The reference metric may be defined, for example, by $g_{\mathrm{ref}} := f_{\mathrm{ref}}^{\#}g_0$ with some *reference embedding* $f_{\mathrm{ref}} \in \mathrm{Imm}(M; \mathbb{R}^m)$. A frequent setting in elasticity theory is that the material manifold $M \subset \mathbb{R}^m$ is a compact set with smooth boundary and that the reference embedding $f_{\mathrm{ref}} \colon M \hookrightarrow \mathbb{R}^m$ is the canonical embedding so that $g_{\mathrm{ref}} = g_0$ is the Euclidean inner product. The Lamé coefficients μ and λ are often assumed to be *constants* describing the elastic properties of a homogeneous and isotropic elastic material.

Let e be an orthonormal basis of $\mathbb{R}^m$ and define the "deformation gradient" by

$$
\mathbf{F} \colon M \to \mathrm{Mat}_{m \times m}(\mathbb{R}), \qquad \mathbf{F}_{ij}|_x := g_0(e_i, \mathrm{d}f|_x\, e_j).
$$

Then $\mathbf{G}_e(f^{\#}g_0) = \mathbf{F}^{\mathsf{T}}\mathbf{F}$ is the *Cauchy-Green tensor*. With Remark 8.2, Hencky's energy can be written as

$$
\mathcal{H}(f) = \int_M \left(\mu\, |\log(\mathbf{F}^{\mathsf{T}}\mathbf{F})|^2 + \tfrac{\lambda}{2} \log^2(\det \mathbf{F}) \right) \mathrm{d}x.
$$

Note that by allowing g_{ref} to be arbitrary, one may also treat materials under internal prestress, e.g., materials with reference metric g_{ref} that—due to non-vanishing Riemannian curvature—cannot be isometrically embedded into $\mathbb{R}^m$. Such reference metrics may occur, e.g., when heat, moisture or chemical reactions change the material properties locally.

Remark 8.7 (Linear Elasticity) By Lemma 8.4, the second derivative of $\mathcal{H}(f)$ with respect to f at the reference embedding f_{ref} can be written as

$$
\begin{aligned}
D^2\mathcal{H}(f_{\mathrm{ref}})(u, u) &= \int_M \mathrm{Hess}^{\mu,\lambda}\, W_{\mu,\lambda}(g_{\mathrm{ref}})(X, X)\, \mathrm{vol}_{g_{\mathrm{ref}}} \\
&= \int_M \left(\mu\, |X|^2_{g_{\mathrm{ref}}} + \tfrac{\lambda}{2} g_{\mathrm{ref}}(g_{\mathrm{ref}}, X)^2 \right) \mathrm{vol}_{g_{\mathrm{ref}}},
\end{aligned}
$$

where $u \in T_{f_{\mathrm{ref}}} \operatorname{Imm}(M; \mathbb{R}^m)$ is an *infinitesimal deformation* and $X \in T_{g_{\mathrm{ref}}} \mathcal{P}(M)$ is the *infinitesimal distortion tensor induced by u*, given by

$$X(v, w)|_x = g_0(\mathrm{d}f_{\mathrm{ref}} v, \mathrm{d}u\, w) + g_0(\mathrm{d}u\, v, \mathrm{d}f_{\mathrm{ref}} w), \quad \text{for } v, w \in T_x M.$$

Let $M \subset \mathbb{R}^m$ be an m-dimensional manifold with boundary and $f_{\mathrm{ref}} = \operatorname{id}_{\mathbb{R}^m}$. By defining $\mathbf{X}$ and $\mathbf{u} \in \operatorname{Mat}_{m \times m}(\mathbb{R})$ by $\mathbf{X} := \mathbf{G}_e(X)$ and $\mathbf{u}_{ij} := \langle e_i, \mathrm{d}u\, e_j \rangle$, we obtain $\mathbf{X} = \mathrm{d}\mathbf{u}^\mathsf{T} + \mathrm{d}\mathbf{u}$. This leads to the well-known deformation energy $\mathcal{E}_{\mathrm{lin}}$ of linear elasticity (see [25, Equation (4.1)]):

$$\mathcal{E}_{\mathrm{lin}}(u) := D^2 \mathcal{H}(f_{\mathrm{ref}})(u, u) = \int_M \left(\mu\, |\mathbf{X}|^2 + \tfrac{\lambda}{2} \operatorname{tr}(\mathbf{X})^2 \right) \operatorname{vol}_{g_{\mathrm{ref}}}.$$

8.3 Smooth Setting

Let (M, g) be a connected, compact, m-dimensional, smooth Riemannian manifold with boundary and let $\Sigma \subset \partial M$ be a set with non-empty interior and smooth boundary. We are going to impose Dirichlet boundary conditions on the set Σ: For a given $\gamma \in \operatorname{Imm}(\Sigma; \mathbb{R}^m) \cap W_g^{2,\infty}(\Sigma; \mathbb{R}^m)$, we define the configuration space

$$C := \{ f \in \operatorname{Imm}(M; \mathbb{R}^m) \mid f|_\Sigma = \gamma \}$$

and the distance

$$d_C(f_1, f_2) := d_{\operatorname{Imm}}(f_1, f_2) + \|f_1 - f_2\|_{W_g^{1,\infty}} \quad \text{for all } f_1, f_2 \in C.$$

Let $g_{\mathrm{ref}} \in \mathcal{P}(M) \cap W_g^{1,\infty}(M; T'M \otimes T'M)$ be a reference metric, $V \in W_{\mathrm{loc}}^{1,\infty}(M \times \mathbb{R}^m, \mathbb{R})$ be a locally Lipschitz continuous potential, and $\mu, \lambda \in W^{1,\infty}(M, \mathbb{R})$ Lamé coefficients. For simplicity, we assume that μ and λ are constant. We define the energy

$$\mathcal{F} : C \to \mathbb{R}, \qquad \mathcal{F}(f) := \mathcal{H}(f) + \mathcal{V}(f),$$

where $\mathcal{V}(f) := \int_M (x \mapsto V(x, f(x)))\, \operatorname{vol}_{g_{\mathrm{ref}}}$. Here, $\mathcal{V}$ models certain types of potential energy, for example a gravitational potential:

Example 8.8 Choose $m = 3$ and $V(x, z) = r(x) \langle \vec{g}, z \rangle$ with some $\vec{g} \in (\mathbb{R}^3)'$. Then $\mathcal{V}(f)$ is the gravitational potential of $f(M)$ with mass distribution $r \colon M \to \mathbb{R}$.

Now, we define the (partially free) boundary value problem:

Problem 8.9 *Find the minimizers $\mathcal{M}$ of $\mathcal{F}$ on C.*

A solution f of this problem models an elastic material M immersed by f into $\mathbb{R}^m$ in static equilibrium under the influence of potential $\mathcal{V}$ and under partially free boundary conditions $f|_\Sigma = \gamma$.

In the upcoming convergence analysis, we use essentially the same a priori assumptions $\mathcal{A}^s$ as in Chapter 6; all we do is add an explicit bound on $\|f\|_{L_g^\infty}$:

$$\mathcal{A}^s := \{\, f \in C \cap W_g^{2,\infty}(\Sigma; \mathbb{R}^m) \mid d_{\mathcal{P}}(g, f^\# g_0) \leq s,\ \|f\|_{W_g^{2,\infty}} \leq s \,\}.$$

8.4 Discrete Setting

Let $\mathcal{T}$ be a finite smooth triangulation $\mathcal{T}$ of M such that a smooth triangulation $\mathcal{T}|_\Sigma$ of Σ is induced. Define the discrete configuration space

$$C_\mathcal{T} := \left\{\, f \colon V(\mathcal{T}) \to \mathbb{R}^m \;\middle|\; \begin{array}{l} f|_{V(\mathcal{T}|_\Sigma)} = \gamma|_{V(\mathcal{T}|_\Sigma)}, \\ \forall \sigma \in \mathcal{T} \colon f(V(\sigma)) \text{ in general position} \end{array} \,\right\}.$$

We use the same definitions for $\mathcal{S}_\mathcal{T}$, $\tilde{\mathcal{R}}_\mathcal{T}$ and $\gamma_\mathcal{T}$ as in Section 6.2:

$$\mathcal{S}_\mathcal{T} \colon C \to C_\mathcal{T}, \qquad f \mapsto f|_{V(\mathcal{T})},$$

$$\tilde{\mathcal{R}}_\mathcal{T} \colon C_\mathcal{T} \to \mathrm{Imm}(M; \mathbb{R}^m), \quad f \mapsto \sum_{p \in V(\mathcal{T})} \lambda_p\, f(p),$$

$$\gamma_\mathcal{T} := \sum_{p \in V(\mathcal{T}|_\Sigma)} \lambda_p\, \gamma(p).$$

As a priori information, we use

$$\mathcal{A}_\mathcal{T}^r := \{\, f \in C_\mathcal{T} \mid d_{\mathcal{P}(M)}(g, \tilde{\mathcal{R}}_\mathcal{T}(f)^\# g_0) \leq r \text{ and } \|\tilde{\mathcal{R}}_\mathcal{F}(g)\|_{L_g^\infty} \leq r \,\}.$$

In order to define a reconstruction operator, let

$$\mathrm{ext} \colon W_g^{1,\infty}(\Sigma; \mathbb{R}^m) \to W_g^{1,\infty}(M; \mathbb{R}^m)$$

be a continuous linear extension operator[1], let $u_\mathcal{T} := \mathrm{ext}(\gamma - \gamma_\mathcal{T})$ and define

$$\mathcal{R}_\mathcal{T} \colon C_\mathcal{T} \to W_g^{1,\infty}(M; \mathbb{R}^m), \qquad f \mapsto \tilde{\mathcal{R}}_\mathcal{T}(f) + u_\mathcal{T}.$$

---

[1]For example, one may use the technique described in Section 6.4 twice: Let $\mathrm{ext}_1 \colon W_g^{1,\infty}(\Sigma; \mathbb{R}^m) \to W_g^{1,\infty}(\partial M; \mathbb{R}^m)$ and $\mathrm{ext}_2 \colon W_g^{1,\infty}(\partial M; \mathbb{R}^m) \to W_g^{1,\infty}(M; \mathbb{R}^m)$ be continuous extension operators and define $\mathrm{ext} = \mathrm{ext}_2 \circ \mathrm{ext}_1$.

115

As the test space, we use $\mathcal{X} = \mathrm{Imm}(M; \mathbb{R}^m)$ with metric

$$d_{\mathcal{X}}(f_1, f_2) := d_{\mathrm{Imm}}(f_1, f_2) + \|f_1 - f_2\|_{W_g^{1,\infty}} \quad \text{for all } f_1, f_2 \in \mathcal{X}.$$

As test functionals, we use the canonical inclusion $\Psi : C \to \mathcal{X}$ and the mapping $\Psi_n = \tilde{\mathcal{R}}_{\mathcal{T}}$.

It turns out that we have to discretize g_{ref} as well. We introduce the operator

$$Q_{\mathcal{T}} : \mathcal{P}(M) \cap W_g^{1,\infty}(M; T'M \otimes T'M) \to \mathcal{P}(M)$$

by the following condition: For each $\sigma \in \mathcal{T}$ one has pointwise almost everywhere in Δ_m:

$$\sigma^\#(Q_{\mathcal{T}}(b)) := (\sigma^\# b)|_p \quad \text{for all } b \in \mathcal{P}(M) \cap W_g^{1,\infty}(M; T'M \otimes T'M),$$

where $p = (\frac{1}{m+1}, \ldots, \frac{1}{m+1}) \in \Delta_m$ is the barycenter of Δ_m.[2] Thus, $Q_{\mathcal{T}}$ can be seen as a generalization of a piecewise constant approximation operators in the same way as $\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}}$ can be seen as a piecewise affine approximation operator.

Now, we abbreviate $g_{\mathcal{T}} := Q_{\mathcal{T}}(g_{\mathrm{ref}})$ and discretize the Hencky energy as follows:

$$\mathcal{H}_{\mathcal{T}}(f) := \frac{1}{2} \int_M d_{\mu,\lambda}^2(\tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0, g_{\mathcal{T}}) \, \mathrm{vol}_{g_{\mathcal{T}}} \quad \text{for all } f \in C_{\mathcal{T}}.$$

Remark 8.10 It requires further analysis to treat the discrete Hencky energy numerically. With the help of the transformation law of integration, we may decompose the integral into

$$\mathcal{H}_{\mathcal{T}}(f) = \sum_{\sigma \in \mathcal{T}} \mathcal{H}_\sigma(f),$$

where

$$\mathcal{H}_\sigma(f) := \frac{1}{2} \int_{\Delta_m} \sigma^\# \Big(d_{\mu,\lambda}^2(\tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0, g_{\mathcal{T}}) \, \mathrm{vol}_{g_{\mathcal{T}}} \Big).$$

For $\sigma \in \mathcal{T}$, we abbreviate $h_\sigma := \tilde{\mathcal{R}}(f) \circ \sigma$. Observe that $h_\sigma^\# g_0$ and $\sigma^\# g_{\mathcal{T}} = (\sigma^\# g_{\mathrm{ref}})|_p$ are constant on Δ_m: This follows from the fact that h_σ is affine and from the definition of $g_{\mathcal{T}}$, respectivly. By the isometric properties of the distances on $P(V)$ (see Remark 4.16 and Lemma 4.17), we have

$$\begin{aligned}
\sigma^\#(d_{\mu,\lambda}^2(\tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0, g_{\mathcal{T}}) \, \mathrm{vol}_{g_{\mathcal{T}}}) &= d_{\mu,\lambda}^2(\sigma^\# \tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0, \sigma^\# g_{\mathcal{T}}) \, \sigma^\# \mathrm{vol}_{g_{\mathcal{T}}} \\
&= d_{\mu,\lambda}^2(h_\sigma^\# g_0, (\sigma^\# g_{\mathrm{ref}})|_p) \, (\sigma^\# \mathrm{vol}_{g_{\mathrm{ref}}})|_p \\
&= d_{\mu,\lambda}^2(h_\sigma^\# g_0|_p, (\sigma^\# g_{\mathrm{ref}})|_p) \, (\sigma^\# \mathrm{vol}_{g_{\mathrm{ref}}})|_p.
\end{aligned}$$

---

[2]Every other choice of points $p_\sigma \in \Delta_m$ for $\sigma \in \mathcal{T}$ would also do.

Thus, $\mathcal{H}_\sigma(f)$ can be easily computed since it is the sum of integrals of constant densities.

More precisely: Let $v = v_1, \ldots, v_m$ be some basis of $p^\perp = T_p\Delta_m$ and let $w = w_1, \ldots, w_m$ be some basis of $\mathbb{R}^m$. We write $(\mathbf{L}_\sigma)_{ij} = \langle w_i, d_p h_\sigma v_j \rangle$ and choose $(\mathbf{L}_{\text{ref},\sigma}) \in \text{Mat}_{m\times m}(\mathbb{R})$ with $\mathbf{L}_{\text{ref},\sigma}^\top \mathbf{L}_{\text{ref},\sigma} = \mathbf{G}_v(\sigma^\# g_\mathcal{T})$. With $\mathbf{F}_\sigma := \mathbf{L}_\sigma \mathbf{L}_{\text{ref},\sigma}^{-1}$ and Remark 8.2, we obtain

$$\mathcal{H}_\mathcal{T}(f) = \frac{1}{2m!} \sum_{\sigma \in \mathcal{T}} \left(\mu \, |\log(\mathbf{F}_\sigma^\top \mathbf{F}_\sigma)|^2 + \tfrac{\lambda}{2} \log^2(\det \mathbf{F}_\sigma) \right) |\det(\mathbf{L}_{\text{ref},\sigma})|.$$

A rather coarse but completely sufficient discretization of $\mathcal{V}$ can be defined by

$$\mathcal{V}_\mathcal{T}(f) := \sum_{\sigma \in \mathcal{T}} \int_{\Delta_m} V(\sigma(p), \tilde{\mathcal{R}}_\mathcal{T}(f)(p)) \, (\sigma^\# \text{vol}_{g_\mathcal{T}})$$

$$= \frac{1}{m!} \sum_{\sigma \in \mathcal{T}} V(\sigma(p), \tilde{\mathcal{R}}_\mathcal{T}(f)(p)) \, |\det(\mathbf{L}_{\text{ref},\sigma})|,$$

for $f \in C_\mathcal{T}$. Hence, we obtain the discrete problem

Problem 8.11 *Determine the minimizers $\mathcal{M}_\mathcal{T}$ of $\mathcal{F}_\mathcal{T} := \mathcal{H}_\mathcal{T} + \mathcal{V}_\mathcal{T}$ on $C_\mathcal{T}$.*

8.5 Convergence Analysis

Define the approximation characteristic

$$\chi^{(2)}(\mathcal{T}) := \sup_{\sigma \in \mathcal{T}} \|(T\sigma)^{-1}\|_{L_g^\infty}^2 \left(\|T\sigma\|_{L_g^\infty}^2 + 2\|\text{Hess}^g(\sigma)\|_{L^\infty} \right) \|T\sigma\|_{L_g^\infty}.$$

Lemma 8.12 *For $b \in \mathcal{P}(M) \cap W_g^{1,\infty}$ and a smooth triangulation $\mathcal{T}$ of M with $\chi^{(2)}(\mathcal{T}) < \|b\|_{W_g^{1,\infty}}^{-1} e^{-d_\mathcal{P}(b,g)}$, one has*

$$d_\mathcal{P}(Q_\mathcal{T}(b), b) \le e^{d_\mathcal{P}(b,g)} \|b\|_{W_g^{1,\infty}} \chi^{(2)}(\mathcal{T}).$$

Proof. Let $\sigma \in \mathcal{T}$. One has pointwise on $\sigma(\Delta_m)$:

$$\|b - Q_\mathcal{T}(b)\|_{L_g^\infty} \le \|(T\sigma)^{-1}\|_{L_g^\infty}^2 \|\sigma^\# b - \sigma^\# Q_\mathcal{T}(b)\|_{L^\infty}$$

$$= \|(T\sigma)^{-1}\|_{L_g^\infty}^2 \|\sigma^\# b - \sigma^\# b|_p\|_{L^\infty}$$

$$\le \|(T\sigma)^{-1}\|_{L_g^\infty}^2 \|\nabla(\sigma^\# b)\|_{L^\infty}.$$

In order to estimate $\|\nabla(\sigma^{\#}b)\|_{L^{\infty}}$, let $u, v, w \in \mathfrak{X}(\Delta_m)$ be (covariantly) constant vector fields. We compute

$$(\nabla(\sigma^{\#}b))(u, v, w) = \nabla_u(b(T\sigma v, T\sigma w))$$
$$= (\nabla^g_{T\sigma u}b)(T\sigma v, T\sigma w) + b(\mathrm{Hess}^g(\sigma)(u, v), T\sigma w) + b(T\sigma v, \mathrm{Hess}^g(\sigma)(u, w))$$

and deduce

$$\|\nabla(\sigma^{\#}b)\|_{L^{\infty}} \leq \|\nabla^g b\|_{L^{\infty}_g}\|T\sigma\|^3_{L^{\infty}_g} + 2\|b\|_{L^{\infty}_g}\|\mathrm{Hess}^g(\sigma)\|_{L^{\infty}}\|T\sigma\|_{L^{\infty}_g}$$
$$\leq \|b\|_{W^{1,\infty}_g}\Big(\|T\sigma\|^2_{L^{\infty}_g} + 2\|\mathrm{Hess}^g(\sigma)\|_{L^{\infty}}\Big)\|T\sigma\|_{L^{\infty}_g}.$$

Combining these estimates and using $\chi^{(2)}(\mathcal{T}) < \|b\|^{-1}_{W^{1,\infty}_g}\mathrm{e}^{-d_{\mathcal{P}}(b,g)}$, we obtain

$$\|b - Q_{\mathcal{T}}(b)\|_{L^{\infty}_g} \leq \|b\|_{W^{1,\infty}_g}\chi^{(2)}(\mathcal{T}) < \mathrm{e}^{-d_{\mathcal{P}}(b,g)}. \qquad \square$$

Remark 8.13 It follows immediately from the proof of Lemma 6.8 that $\chi^{(2)}$ behaves as follows under an affine subdivision Φ of $\mathcal{T}$:

$$\chi^{(2)}(\Phi(\mathcal{T})) \leq \chi^{(2)}(\mathcal{T}) \sup_{\sigma\in\mathcal{T}} \sup_{\varphi\in\Phi_{\sigma}} \|(T\varphi)^{-1}\|^2_{L^{\infty}}\|T\varphi\|^3_{L^{\infty}}.$$

In particular, $\chi^{(2)}(\mathcal{T})$ can be made arbitrarily small by affine subdivision (see also the discussion in Section 6.3).

In the following, we provide estimates with respect to the quantity

$$\varrho(\mathcal{T}) := \max\{\varrho^{(0)}(\mathcal{T}), \ \varrho^{(0)}(\mathcal{T}|_{\Sigma}), \ \varrho^{(1)}(\mathcal{T}), \ \varrho^{(1)}(\mathcal{T}|_{\Sigma}), \ \chi^{(2)}(\mathcal{T})\}.$$

Note that equations (6.3), (6.4), (6.5) as well as Lemma 6.11 remain true with this definition. Moreover, we point out that there are smooth triangulations with arbitrary low $\varrho(\mathcal{T})$.

Lemma 8.14 (Proximity errors) *For the sampling proximity error $\varepsilon^{\mathcal{S}}_{\mathcal{T}}$ and the reconstruction proximity error $\varepsilon^{\mathcal{R}}_{\mathcal{T}}$ on $\mathcal{A}^s$ and $\mathcal{A}^r_{\mathcal{T}}$, respectively, the following hold:*

$$\varepsilon^{\mathcal{S}}_{\mathcal{T}} \leq C(s)\,\varrho(\mathcal{T}) \quad \text{and} \quad \varepsilon^{\mathcal{R}}_{\mathcal{T}} \leq C(r)\,\|\mathrm{ext}\|\,\|\mathrm{d}\gamma\|_{W^{1,\infty}_g}\,\varrho(\mathcal{T}).$$

PROOF. Let $f \in \mathcal{A}^s$ and $h \in \mathcal{A}^r_n$. By (6.4) and (6.5), one has:

$$\|(\Psi_n \circ \mathcal{S}_{\mathcal{T}})(f) - \Psi(f)\|_{W^{1,\infty}_g} = \|(\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f) - f\|_{W^{1,\infty}_g} \leq C(s)\,\varrho(\mathcal{T}),$$
$$\|(\Psi \circ \mathcal{R}_{\mathcal{T}})(h) - \Psi_n(h)\|_{W^{1,\infty}_g} = \|\mathcal{R}_{\mathcal{T}}(h) - \tilde{\mathcal{R}}_{\mathcal{T}}(h)\|_{W^{1,\infty}_g}$$
$$= \|u_{\mathcal{T}}\|_{W^{1,\infty}_g} \leq \|\mathrm{ext}\|\,\|\mathrm{d}\gamma\|_{W^{1,\infty}_g}\,\varrho(\mathcal{T}),$$

respectively. Now, the statements of this lemma follow from Corollary 5.7. $\qquad \square$

Lemma 8.15 (Consistency errors of Hencky energy)
Let $r > s > 0$. Then there is a $\varrho_0 > 0$ and a constant $C \geq 0$ such that for all smooth triangulations with $\varrho(\mathcal{T}) \leq \varrho_0$, one has the following estimates for sampling and reconstruction consistency errors $\delta_{\mathcal{T}}^{\mathcal{S}}$, $\delta_{\mathcal{T}}^{\mathcal{R}}$ of $(\mathcal{H}_{\mathcal{T}}, \mathcal{S}_{\mathcal{T}}, \mathcal{R}_{\mathcal{T}})$ with respect to $\mathcal{H}$ on $(\mathcal{A}^s, \mathcal{A}_{\mathcal{T}}^r)$:

$$\delta_{\mathcal{T}}^{\mathcal{S}} \leq C \varrho(\mathcal{T}) \quad and \quad \delta_{\mathcal{T}}^{\mathcal{R}} \leq C \varrho(\mathcal{T}).$$

Proof. *Sampling consistency error:* Fix $f \in \mathcal{A}^s$ and abbreviate $f_{\mathcal{T}} := (\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)$. Note that we have

$$\mathcal{H}_{\mathcal{T}}(f) = \frac{1}{2} \int_{\Sigma} d_{\mu,\lambda}^2 (f_{\mathcal{T}}^{\#} g_0, g_{\mathcal{T}}) \, \text{vol}_{g_{\mathcal{T}}} \, .$$

Put $\ell := d_{\mathcal{P}}(g, g_{\text{ref}})$. From (6.4), we obtain $\|f - f_{\mathcal{T}}\|_{W_g^{1,\infty}} \leq s \varrho(\mathcal{T})$. Under the assumption that $\varrho(\mathcal{T})$ is sufficiently small, we may apply Corollary 5.6, Corollary 5.7, and Lemma 4.15 to obtain pointwise almost everywhere:

$$d_P(f^{\#} g_0, f_{\mathcal{T}}^{\#} g_0) \leq C(s, \ell) \varrho(\mathcal{T}) \quad and \quad d_{\text{Vol}}(\text{vol}_{f^{\#} g_0}, \text{vol}_{f_{\mathcal{T}}^{\#} g_0}) \leq C(s, \ell) \varrho(\mathcal{T}).$$

This immediately leads to

$$d_{\mu,\lambda}(f^{\#} g_0, f_{\mathcal{T}}^{\#} g_0) \leq (\mu + |\lambda|) C(s, \ell) \varrho(\mathcal{T}).$$

By Lemma 8.12, we may assume

$$d_{\mu,\lambda}(g_{\text{ref}}, g_{\mathcal{T}}) \leq (\mu + |\lambda|) C(\ell) \varrho(\mathcal{T}) \quad and \quad \text{vol}_{g_{\mathcal{T}}} \leq (1 + C(\ell) \varrho(\mathcal{T})) \text{vol}_{g_{\text{ref}}} . \quad (8.1)$$

The triangle inequality yields

$$d_{\mu,\lambda}(f_{\mathcal{T}}^{\#} g_0, g_{\mathcal{T}}) \leq d_{\mu,\lambda}(f_{\mathcal{T}}^{\#} g_0, f^{\#} g_0) + d_{\mu,\lambda}(f^{\#} g_0, g_{\text{ref}}) + d_{\mu,\lambda}(g_{\text{ref}}, g_{\mathcal{T}})$$
$$\leq d_{\mu,\lambda}(f^{\#} g_0, g_{\text{ref}}) + (\mu + |\lambda|) C(\ell, s) \varrho(\mathcal{T}).$$

Now standard estimates lead to

$$d_{\mu,\lambda}^2(f_{\mathcal{T}}^{\#} g_0, g_{\mathcal{T}}) \, \text{vol}_{g_{\mathcal{T}}} \leq d_{\mu,\lambda}^2(f^{\#} g_0, g_{\text{ref}}) \, \text{vol}_{g_{\text{ref}}} + C(\ell, s, \mu, |\lambda|) \varrho(\mathcal{T}) \, \text{vol}_{g_{\text{ref}}}$$

for sufficiently small $\varrho(\mathcal{T})$. After integration, this becomes

$$\mathcal{H}_{\mathcal{T}}(\mathcal{S}_{\mathcal{T}}(f)) \leq \mathcal{H}(f) + C(\ell, s, \mu, |\lambda|) \varrho(\mathcal{T}),$$

thus one has $\delta_{\mathcal{T}}^{\mathcal{S}} \leq C(\ell, s, \mu, |\lambda|) \varrho(\mathcal{T})$.

Reconstruction consistency error: Fix $f \in \mathcal{A}^r_{\mathcal{T}}$ and observe that

$$d_{\mu,\lambda}(\mathcal{R}_{\mathcal{T}}(f)^\# g_0, g_{\mathrm{ref}})$$
$$\leq d_{\mu,\lambda}(\mathcal{R}_{\mathcal{T}}(f)^\# g_0, \tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0) + d_{\mu,\lambda}(\tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0, g_{\mathcal{T}}) + d_{\mu,\lambda}(g_{\mathcal{T}}, g_{\mathrm{ref}}).$$

While the center term in the second line contributes to the discrete energy, the other two are small: We treat the last term with (8.1) and use (6.5) to obtain

$$\|\mathcal{R}_{\mathcal{T}}(f) - \tilde{\mathcal{R}}_{\mathcal{T}}(f)\|_{W^{1,\infty}_g} = \|u_{\mathcal{T}}\|_{W^{1,\infty}_g} \leq C \varrho(\mathcal{T})$$

and then apply Corollary 5.6, Corollary 5.7, and Lemma 4.15 again to find:

$$d_{\mu,\lambda}(\mathcal{R}_{\mathcal{T}}(f)^\# g_0, \tilde{\mathcal{R}}_{\mathcal{T}}(f)^\# g_0) \leq C(\ell, r, \mu, |\lambda|) \varrho(\mathcal{T}).$$

Now, an argument similar to that one given for the sampling consistency error yields

$$\mathcal{H}(f) \leq \mathcal{H}_{\mathcal{T}}(\mathcal{R}_{\mathcal{T}}(f)) + C(\ell, r, \mu, |\lambda|) \varrho(\mathcal{T}),$$

thus $\delta^{\mathcal{R}}_{\mathcal{T}} \leq C(\ell, r, \mu, |\lambda|) \varrho(\mathcal{T})$. $\qquad\square$

Lemma 8.16 (Consistency errors of potential)
Let $r > s > 0$. Then there is a $\varrho_0 > 0$ and a constant $C \geq 0$ such that for all smooth triangulations with $\varrho(\mathcal{T}) \leq \varrho_0$, one has the following estimates for sampling and reconstruction consistency errors $\delta^{\mathcal{S}}_{\mathcal{T}}, \delta^{\mathcal{R}}_{\mathcal{T}}$ of $(\mathcal{V}_{\mathcal{T}}, \mathcal{S}_{\mathcal{T}}, \mathcal{R}_{\mathcal{T}})$ with respect to $\mathcal{V}$ on $(\mathcal{A}^s, \mathcal{A}^r_{\mathcal{T}})$:

$$\delta^{\mathcal{S}}_{\mathcal{T}} \leq C \varrho(\mathcal{T}) \quad and \quad \delta^{\mathcal{R}}_{\mathcal{T}} \leq C \varrho(\mathcal{T}).$$

Proof. For $y \in \Delta_m$ and $\sigma \in \mathcal{T}$ we have $d_g(\sigma(y), \sigma(p)) \leq \|T\sigma\|_{L^\infty_g} |y - p| \leq \|T\sigma\|_{L^\infty_g}$. On the one hand, we have for all $h \in \mathcal{A}^r_{\mathcal{T}}$:

$$|\mathcal{R}_{\mathcal{T}}(h)(\sigma(y)) - \tilde{\mathcal{R}}_{\mathcal{T}}(h)(\sigma(p))|$$
$$\leq |\mathcal{R}_{\mathcal{T}}(h)(\sigma(y)) - \tilde{\mathcal{R}}_{\mathcal{T}}(h)(\sigma(y))| + |\tilde{\mathcal{R}}_{\mathcal{T}}(h)(\sigma(y)) - \tilde{\mathcal{R}}_{\mathcal{T}}(h)(\sigma(p))|$$
$$\leq C \varrho(\mathcal{T}) + \|\mathrm{d}(\tilde{\mathcal{R}}_{\mathcal{T}}(h))\|_{L^\infty} d_g(\sigma(y), \sigma(p))$$
$$\leq C \varrho(\mathcal{T}).$$

On the other, we have for all $f \in \mathcal{A}^s$:

$$|(\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)(\sigma(p)) - f(\sigma(y))|$$
$$\leq |(\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)(\sigma(p)) - f(\sigma(p))| + |f(\sigma(p)) - f(\sigma(y))|$$
$$\leq C \varrho(\mathcal{T}) + \|\mathrm{d}f\|_{L^\infty} d_g(\sigma(y), \sigma(p))$$
$$\leq C \varrho(\mathcal{T}) + C(s) \|T\sigma\|_{L^\infty_g} \leq C \varrho(\mathcal{T}).$$

Let $R > r$ and let $L \geq 0$ be the Lipschitz constant of V on $M \times \bar{B}(0, R)$. For sufficiently small $\varrho(\mathcal{T})$, we may assume that the images of $\mathcal{R}_{\mathcal{T}}(h)$ and $(\tilde{\mathcal{R}}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{T}})(f)$ are contained in $\bar{B}(0, R)$. Now, Lipschitz continuity of $V|_{M \times \bar{B}(0, R)}$ and (8.1) imply

$$\mathcal{V}_{\mathcal{T}}(\mathcal{S}_{\mathcal{T}}(f)) \leq \mathcal{V}(f) + L\,C\,\varrho(\mathcal{T}) \quad \text{and} \quad \mathcal{V}(\mathcal{R}_{\mathcal{T}}(h)) \leq \mathcal{V}_{\mathcal{T}}(h) + L\,C\,\varrho(\mathcal{T}). \qquad \square$$

Analogously to Section 7.6, let $(\mathcal{T}_n)_{n \in \mathbb{N}}$ be a sequence of smooth triangulations of (Σ, g) with $\varrho_n := \varrho(\mathcal{T}_n) \overset{n \to \infty}{\longrightarrow} 0$. Again, we write $C_n, \mathcal{F}_n, \mathcal{A}_n^r, \dots$ instead of $C_{\mathcal{T}_n}, \mathcal{F}_{\mathcal{T}_n}$, $\mathcal{A}_{\mathcal{T}_n}^r, \dots$. Moreover, we denote the sets of δ-minimizers by $\mathcal{M}^\delta := \arg\min^\delta(\mathcal{F})$ and $\mathcal{M}_n^\delta := \arg\min^\delta(\mathcal{F}_n)$.

Theorem 8.17 (Kuratowksi convergence of minimizers)
Assume $\emptyset \neq \Psi(\mathcal{M}) \subset \Psi(\mathcal{A}^s \cap \mathcal{M})$ for some $s > 0$ and that the $\mathcal{A}_n^r$ are valid[3] for some $r > s$ and all $n \in \mathbb{N}$. Then there is a constant $C \geq 0$ depending on Σ, g, γ, s, r only, such that one has $\delta_n \leq C\,\varrho_n$ for sufficiently large $n \in \mathbb{N}$ and such that

$$\underset{n \to \infty}{\mathrm{Ls}} \; \Psi_n(\mathcal{A}_n^r \cap \mathcal{M}_n) \subset \Psi(\mathcal{M}) = \underset{n \to \infty}{\mathrm{Lt}} \; \Psi_n(\mathcal{A}_n^r \cap \mathcal{M}_n^{\delta_n})$$

holds with respect to the metric d_C defined in Section 8.3. In particular, one has Kuratowski convergence with respect to the topology generated by $W_g^{1,\infty}$.

Proof. Once again, we may apply Theorem 2.28 with $\mathcal{K} := \mathcal{X}$ to prove the result:
Proximity follows from Lemma 8.14. Together, Lemma 8.15, Lemma 8.16, and Corollary 2.12 show consistency. The inclusion $\mathcal{S}_n(\mathcal{A}^s) \subset \mathcal{A}_n^r$ has already been shown in Lemma 6.11. The function $\mathcal{F} \colon (C, d_C) \to \mathbb{R}$ is continuous and Ψ is the inclusion of the closed set $C \subset \mathrm{Imm}(M; \mathbb{R}^m)$. Thus, $\Psi_\# \mathcal{F}$ is an extension by infinity so that Example 2.27 implies topological stability of $\mathcal{F}$ along Ψ over $\mathcal{X}$. Validity of $\mathcal{A}^s$ follows from $\emptyset \neq \Psi(\mathcal{M}) \subset \Psi(\mathcal{A}^s \cap \mathcal{M})$ and the validity of $\mathcal{A}_n^r$ was imposed as a condition. $\qquad \square$

Finally, we point out that this analysis can be conducted analogously for functionals $\mathcal{J}_i \colon C \to \mathbb{R}$ of the form

$$\mathcal{J}_1(f) = \int_M W_1(\mathrm{d}f) \, \mathrm{vol}_g \quad \text{and} \quad \mathcal{J}_2(f) = \int_M W_2(f^\# g_0) \, \mathrm{vol}_{f^\# g_0}$$

as long as the constitutive function W_1 is locally Lipschitz continuous with respect to $\|\cdot\|_{W^{1,g}}$ and the function W_2 is locally Lipschitz continuous with respect to $\mathrm{d}_{\mathrm{Imm}}$ or $\mathrm{d}_{\mathrm{Imm}}^r$, respectively.

---

[3] see Definition 2.4.
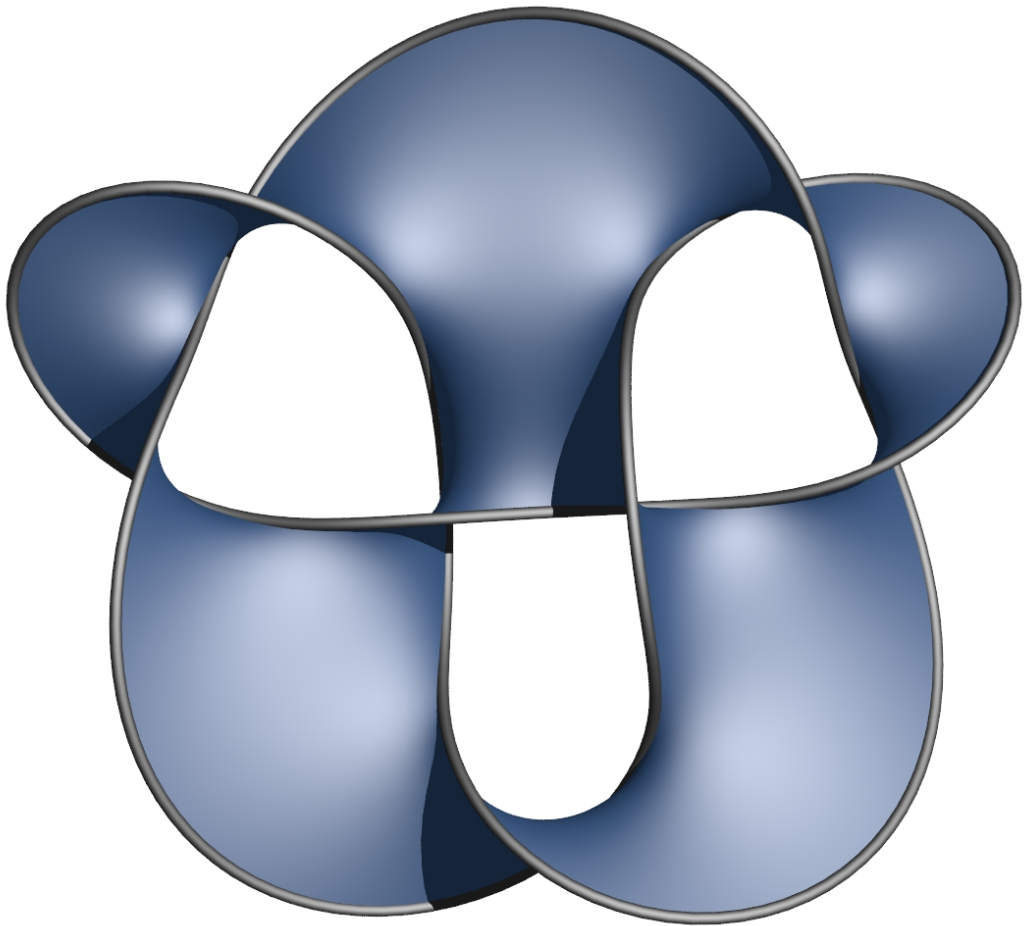
Figure 8.1: The "panda surface", a non-orientable minimal surface spanning a knot of class 7_7. (A 3D-print of the boundary curve is part of the *Göttingen Collection of Mathematical Models and Instruments*.)

Conclusion and Future Work

In the hope that it can be broadly applied, we developed a framework for the convergence analysis of parametric variational problems. We discussed two examples of nonconforming Ritz-Galerkin schemes in detail. These are nonconforming because of boundary conditions only. However, there are other, more severe, variational crimes to commit: For future work, we think of elastic energies of thin shells and rods. These depend nonlinearly on curvature, thus on second derivatives, as well. Hence discretizing them by piecewise linear structures such as simplicial meshes and polygonal lines leads inevitably to nonconforming schemes. There are already several discretizations that are geometrically meaningful *and* that also provide efficient and easily implementable algorithms. While the mean curvature vector and Gauss curvature of surfaces are known to be discretized well on simplicial surfaces (at least in a distributional sense), variational convergence is yet to be shown for elastic energies that depend nonlinearly on curvature.

We point out that, up to the current thesis, this was precisely the situation in which we found the theory of discrete minimal surfaces (in the sense of Pinkall and Polthier [30]): There were well working algorithms and they were quite *plausible*, but a rigorous proof for the variational convergence of these methods was missing.

Our given proof is still somewhat incomplete in the sense that a discrete regularity theory is yet to be developed. Moreover, there is still potential for improvement: First, as we pointed out at the end of Chapter 7, a careful manipulation of the reconstruction operators and narrower discrete a priori information could show *uniform* convergence of solution sets. Second, we think it is possible to derive a quantitative stability result at least in the case that each smooth solution f is a *stable minimal surface*, i.e., the Hessian $\mathrm{Hess}(\mathcal{F})|_f$, restricted to the sections in the normal bundle of f, is coercive (or has eigenvalues bounded away from 0).

List of Figures

Bibliography

[1]   Hédy Attouch and Roger J.-B. Wets.
      "Quantitative stability of variational systems. I. The epigraphical distance".
      In: *Trans. Amer. Math. Soc.* 328.2 (1991), pp. 695–729 (cit. on p. 14).

[2]   Hédy Attouch and Roger J.-B. Wets. "Quantitative stability of variational
      systems. II. A framework for nonlinear conditioning".
      In: *SIAM J. Optim.* 3.2 (1993), pp. 359–381 (cit. on pp. 14, 33).

[3]   Wilhelm Blaschke. *Kreis und Kugel*. Veit, Leipzig, 1916 (cit. on p. 8).

[4]   J. Frédéric Bonnans and Alexander Shapiro.
      "Nondegeneracy and quantitative stability of parameterized optimization
      problems with multiple solutions".
      In: *SIAM J. Optim.* 8.4 (1998), pp. 940–946 (cit. on p. 33).

[5]   J. Frédéric Bonnans and Alexander Shapiro.
      "Optimization problems with perturbations: a guided tour".
      In: *SIAM Rev.* 40.2 (1998), pp. 228–264 (cit. on p. 33).

[6]   Dietrich Braess. *Finite elements*. Third. Theory, fast solvers, and applications
      in elasticity theory, Translated from the German by Larry L. Schumaker.
      Cambridge University Press, Cambridge, 2007, pp. xviii+365
      (cit. on pp. 54, 56, 57).

[7]   Kenneth A. Brakke. "The surface evolver".
      In: *Experiment. Math.* 1.2 (1992), pp. 141–165.
      URL: http://projecteuclid.org/euclid.em/1048709050
      (cit. on pp. 100, 105).

[8]   Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*.
      Vol. 33. Graduate Studies in Mathematics.
      American Mathematical Society, Providence, RI, 2001, pp. xiv+415
      (cit. on p. 76).

[9] Manfredo Perdigão do Carmo. *Riemannian geometry*.
Mathematics: Theory & Applications.
Translated from the second Portuguese edition by Francis Flaherty.
Birkhäuser Boston, Inc., Boston, MA, 1992, pp. xiv+300 (cit. on p. 63).

[10] R. Courant.
Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces.
Appendix by M. Schiffer. Interscience Publishers, Inc., New York, N.Y., 1950,
pp. xiii+330 (cit. on pp. 92, 95).

[11] Gianni Dal Maso. *An introduction to Γ-convergence*.
Progress in Nonlinear Differential Equations and their Applications, 8.
Boston, MA: Birkhäuser Boston Inc., 1993, pp. xiv+340 (cit. on p. 13).

[12] Stefan von Deylen.
"Numerical Approximation in Riemannian Manifolds by Karcher Means".
PhD thesis. Freie Universität Berlin, 2013 (cit. on p. 89).

[13] Ulrich Dierkes, Stefan Hildebrandt, and Friedrich Sauvigny.
Minimal surfaces. second. Vol. 339. Grundlehren der Mathematischen
Wissenschaften [Fundamental Principles of Mathematical Sciences].
With assistance and contributions by A. Küster and R. Jakob.
Springer, Heidelberg, 2010, pp. xvi+688 (cit. on p. 95).

[14] Ulrich Dierkes, Stefan Hildebrandt, and Anthony J. Tromba.
Regularity of minimal surfaces. second. Vol. 340.
Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of
Mathematical Sciences]. With assistance and contributions by A. Küster.
Springer, Heidelberg, 2010, pp. xviii+623 (cit. on p. 95).

[15] A. L. Dontchev. *Perturbations, approximations and sensitivity analysis of
optimal control systems*. Vol. 52.
Lecture Notes in Control and Information Sciences.
Berlin: Springer-Verlag, 1983, pp. iv+158 (cit. on p. 37).

[16] Jesse Douglas. "Solution of the problem of Plateau".
In: *Trans. Amer. Math. Soc.* 33.1 (1931), pp. 263–321 (cit. on p. 95).

[17] Jesse Douglas. "The most general form of the problem of Plateau".
In: *Amer. J. Math.* 61 (1939), pp. 590–608 (cit. on p. 92).

[18] G. Dziuk. "An algorithm for evolutionary surfaces".
In: *Numer. Math.* 58.6 (1991), pp. 603–611 (cit. on pp. 100, 105).

[19] Gerhard Dziuk and John E. Hutchinson.
"The discrete Plateau problem: algorithm and numerics".
In: *Math. Comp.* 68.225 (1999), pp. 1–23 (cit. on p. 96).

[20] Gerhard Dziuk and John E. Hutchinson.
"The discrete Plateau problem: convergence results".
In: *Math. Comp.* 68.226 (1999), pp. 519–546 (cit. on pp. ix, 96).

[21] Lawrence C. Evans. *Partial differential equations*. Vol. 19.
Graduate Studies in Mathematics.
American Mathematical Society, Providence, RI, 1998, pp. xviii+662
(cit. on p. 3).

[22] Robert Hardt and Leon Simon. "Boundary regularity and embedded solutions
for the oriented Plateau problem".
In: *Ann. of Math. (2)* 110.3 (1979), pp. 439–486 (cit. on p. 100).

[23] Jeff Henrikson.
"Completeness and total boundedness of the Hausdorff metric".
In: *MIT Undergraduate Journal of Mathematics* (1999) (cit. on p. 8).

[24] Michael Hinze. "On the numerical approximation of unstable minimal
surfaces with polygonal boundaries".
In: *Numer. Math.* 73.1 (1996), pp. 95–118 (cit. on p. 96).

[25] L. D. Landau and E. M. Lifshitz. *Course of theoretical physics. Vol. 7*. Third.
Theory of elasticity, Translated from the Russian by J. B. Sykes and W. H.
Reid. Pergamon Press, Oxford, 1986, pp. viii+187 (cit. on p. 114).

[26] John M. Lee. *Introduction to smooth manifolds*. Vol. 218.
Graduate Texts in Mathematics. Springer-Verlag, New York, 2003,
pp. xviii+628 (cit. on pp. 67, 74).

[27] Maher Moakher and Mourad Zéraï.
"The Riemannian geometry of the space of positive-definite matrices and its
application to the regularization of positive-definite matrix-valued data".
In: *J. Math. Imaging Vision* 40.2 (2011), pp. 171–187 (cit. on p. 61).

[28] James R. Munkres. *Topology*. 2nd.
Prentice Hall, Inc., Upper Saddle River, N.J., 2000 (cit. on pp. 8, 77).

[29] Patrizio Neff, Bernhard Eidel, and Robert Martin. *The axiomatic deduction of
the quadratic Hencky strain energy by Heinrich Hencky*. 2014.
arXiv: 1402.4027 [math.HO] (cit. on p. 111).

[30] Ulrich Pinkall and Konrad Polthier.
      "Computing discrete minimal surfaces and their conjugates".
      In: *Experiment. Math.* 2.1 (1993), pp. 15–36.
      URL: http://projecteuclid.org/euclid.em/1062620735
      (cit. on pp. 91, 100, 108, 123).

[31] Paola Pozzi. "The discrete Douglas problem: convergence results".
      In: *IMA J. Numer. Anal.* 25.2 (2005), pp. 337–378 (cit. on pp. 96, 97).

[32] Paola Pozzi. "The discrete Douglas problem: theory and numerics".
      In: *Interfaces Free Bound.* 6.2 (2004), pp. 219–252 (cit. on pp. 96, 97).

[33] Tristan Rivière. "Lipschitz conformal immersions from degenerating
      Riemann surfaces with L^2-bounded second fundamental forms".
      In: *Adv. Calc. Var.* 6.1 (2013), pp. 1–31 (cit. on p. 74).

[34] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*. Vol. 317.
      Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of
      Mathematical Sciences]. Berlin: Springer-Verlag, 1998, pp. xiv+733
      (cit. on pp. 5, 12–14, 33, 38).

[35] Robert G. Scharein. URL: http://knotplot.com/zoo/ (cit. on p. 98).

[36] Takuya Tsuchiya.
      "Discrete solution of the Plateau problem and its convergence".
      In: *Math. Comp.* 49.179 (1987), pp. 157–165 (cit. on p. 96).

[37] H.-J. Wagner.
      "A contribution to the numerical approximation of minimal surfaces".
      In: *Computing* 19.1 (1977/78), pp. 35–58 (cit. on pp. ix, 100).

[38] J. H. C. Whitehead. "On C^1-complexes".
      In: *Ann. of Math. (2)* 41 (1940), pp. 809–824 (cit. on p. 84).

[39] Walter L. Wilson Jr. "On discrete Dirichlet and Plateau problems".
      In: *Numer. Math.* 3 (1961), pp. 359–373 (cit. on p. 96).

[40] Eberhard Zeidler. *Nonlinear functional analysis and its applications. I*.
      Fixed-point theorems, Translated from the German by Peter R. Wadsack.
      Springer-Verlag, New York, 1986, pp. xxi+897 (cit. on p. 40).

Index