Gradient Methods for Membrane-Mediated Particle Interactions

Inauguraldissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

eingereicht am
Fachbereich Mathematik und Informatik
der Freien Universität Berlin

vorgelegt von
Tobias Kies

Berlin, Dezember 2018

Betreuer:	**Prof. Dr. Ralf Kornhuber**
Erstgutachter:	**Prof. Dr. Ralf Kornhuber**
Zweitgutachter:	**Prof. Dr. Charles M. Elliott**
Tag der Disputation:	**23. Mai 2019**

Contents

Introduction **7**

1 Hybrid models for membrane-mediated particle interactions **13**
1.1 Elastic Canham–Helfrich type membrane bending energy 13
1.1.1 Parameterization over a reference domain 15
1.1.2 Geometric linearization 19
1.2 Constraint-based membrane–particle coupling 23
1.2.1 Point value constraints 23
1.2.2 Curve constraints 25
1.2.3 Parametric particle configurations 26
1.3 Overall model and the interaction potential 27
1.3.1 Further remarks 30

2 Differentiation of the interaction potential **33**
2.1 Trace and constraint preserving maps 34
2.2 Differentiability 42
2.3 Volume integral representation of the gradient 46

3 Discretization for linear models **53**
3.1 Preliminaries 53
3.2 Fictitious domain stabilized Nitsche method for curve constraints 58
3.2.1 Discretization scheme 58
3.2.2 Existence and uniqueness of the discrete solution 61
3.2.3 Discretization error estimate 66
3.3 Conforming discretization for point value constraints 70
3.4 Discretization error in the gradient of the interaction potential 73

4 Implementation and numerical experiments **75**
4.1 Implementation remarks 75
4.1.1 Level set description of parametric interfaces and constraints . . 75
4.1.2 Approximation algorithm for stationary membranes 77
4.1.3 Evaluation of gradients of the interaction potential 79
4.2 Curve constraints: Optimal orders of convergence 80
4.2.1 Radially symmetric example 80
4.2.2 Non-symmetric example 82
4.3 Point value constraints: Non-optimal regularity 86
4.3.1 Radially symmetric example 86

4.3.2 Non-symmetric example 87
4.4 Gradient evaluation: Non-optimal error estimates 91
4.4.1 Curve constraints 91
4.4.2 Point value constraints 92

5 Applications **95**
5.1 Perturbed gradient method 95
5.2 Pattern formation of isotropic conical inclusions in the Monge-gauge . . 97
5.3 Aggregation of anisotropic ellipsoidal protein scaffolds in the Monge-gauge 101
5.4 Interaction of isotropic conical inclusions on a tube 102
5.5 Towards simulating superstructures induced by FCHo2 F-BAR domains 105

6 Outlook: Free energy computations **113**

Appendix **119**
A Matrix calculus: Differentiation of determinants and inverse matrices . . 119
B Differential geometry for perturbed surfaces 120
B.1 Material derivatives of common quantities 121
B.2 Derivatives of integrals over geometric quantities 127
C Ordinary differential equations: Existence and sensitivity 130
D Schemes related to the fictitious domain stabilized Nitsche method . . . 131
E A smooth repulsion potential for FCHo2 particle–particle interactions . 134

Bibliography **139**

Introduction

Background

A plethora of processes in cellular biology involve the deformation of biomembranes, which in various cases is known or suspected to be essentially driven by purely mechanical interactions between the membrane itself and the non-membrane particles in its vicinity. Those non-membrane particles are generally in some way connected to the membrane and are able to reshape the membrane locally. Typical deformation processes in this sense are, for example, several types and stages of endocytosis where even a multitude of different deformation mechanisms come into play (cf. [MG05]). While the overall structure of most processes is well-understood, many aspects of the spatial organization of curvature-generating particles and their induced membrane-shaping abilities remain still unknown or are up to debate.

Within the context of biophysics, a regularly employed approach for describing mechanical membrane–particle interactions are so-called continuum–discrete elasticity models for lipid bilayers, based on the famous works of Canham, Helfrich, and Goulian et al. [Can70, Hel73, GBP93]. In those models the biomembrane is assumed to consist solely out of a bilayer of lipid molecules with hydrophilic heads and hydrophobic tails. In a further modeling step, this bilayer is approximated by a two-dimensional hypersurface whose shape is determined by minimization of the so-called Canham–Helfrich elastic energy. This meso-scale description of the membrane generally constitutes the continuum part of a particle–membrane interaction model. The complementing discrete part is given by a model for the particles interacting with the lipid membrane. There the particles are considered to be rigid objects that, depending on their position and orientation, constrain the membrane's shape locally.

The combination of these two models yields a rich interplay between the membrane and the interacting particles: while the membrane shape is still essentially governed by the minimization of the globally acting Canham–Helfrich energy, there are now additional local position-dependent constraints on the membrane shape that originate from the non-membrane particles. In particular, the non-membrane particles indirectly act globally on the membrane through the mechanical elastic energy, and thus affect the membrane shape significantly. Conversely, each particle's position and orientation is subject to changes as the membrane's relaxation exerts a force on the particles. As a consequence, shaping of the membrane and organization of particles are strongly coupled. In summary, such a hybrid model describes a purely mechanical and membrane-mediated interaction between non-membrane particles.

The membrane-mediated interaction is expressed energetically via the so-called in-

teraction potential. Under the implicit assumption that membrane relaxation occurs on a faster time scale than particle relocation, it describes the overall membrane energy in its stationary state solely with respect to a given particle configuration. This makes the interaction potential the main object of interest in the investigation of spatial structures of particles in membranes, and it plays an eminent role both in the study of pairwise or multi-body particle interactions [GBP93, DF99, KNO00, BF03, MD10, RD11, SK15, FG15], and in the detection of optimal particle configurations [KNO98, DF99, DF02, Wei02, AS07, ISSS10, KKKH16a, KKKH16b].

Analytic expressions for the interaction potential are available under simplifying assumptions such as circular particle shapes, low particle counts, and almost flat membranes. However, this turns out to be insufficient for most applications as realistic processes involve large numbers of particles and because the impact of non-circular shapes is not negligible in general [KNO00]. As such, it is necessary to evaluate the interaction potential for more complicated systems by the use of numerical discretization schemes. When it comes to the computation of stationary particle configurations via minimization of the interaction potential, optimization algorithms applied in previous works generally rely on Monte–Carlo type methods [DF99, DF02, Wei02, AS07, ISSS10, KKKH16a]. From an optimization viewpoint this potentially is a critical source of inefficiency because no higher order information of the interaction potential is incorporated. Moreover, a rigorous error analysis has so far been largely omitted in literature, which implies the risk that convergence of the applied schemes to an optimal state is not necessarily guaranteed.

An important first step towards a numerically sound analysis is done by the work of Elliott, Gräser, Hobbs, Kornhuber and Wolf [EGH$^+$16] which presents a unifying variational framework for hybrid models of membrane-mediated particle interactions in lipid bilayers. The framework resulting from the variational ansatz makes models readily available to finite element discretizations and corresponding discretization error estimates. In particular, one may expect in many cases to derive provably convergent approximation schemes for the interaction potential by means of known finite element methods.

Research focus and value

This thesis builds on top of the variational framework from [EGH$^+$16] and investigates the differentiability of the interaction potential mathematically in a further step with the goal of a subsequent application in numerical schemes. The guiding questions are therefore:

> Is the interaction potential differentiable, and if so, how can differentiability be exploited in order to develop efficient, reliable and robust methods for the numerical simulation of membrane-mediated particle interactions?

These questions ideally need to be answered for a reasonably large class of interaction models in order to open up as many applications as possible. Moreover, it is necessary

to ensure that the derived methods are indeed accessible for practical implementations and applications. As a compromise between scope and generality, this work considers model formulations within which membranes are parameterized over fixed reference domains and it uses only particle models that induce so-called curve and point-value constraints. Furthermore, the classical gradient-type method is chosen as the prototypical optimization method which is to be made accessible by the subsequent investigations on differentiability.

Under these framework conditions the central focus of attention for this thesis lies on the following three core objectives:

1. Prove differentiability of the interaction potential.
2. Derive a numerically feasible representation of the gradient.
3. Develop robust discretization schemes both for stationary membranes and the gradient and prove approximation error estimates.

The computation of stationary membranes naturally leads to the solution of partial differential equations (PDEs) whose domain and boundary is parametric due to their dependence on the particle states. Therefore, there is a tight link to free boundary problems and it appears natural to approach the issue of differentiability in the first objective by means of shape calculus. Indeed, it becomes clear in the subsequent chapters that differentiability, even for higher orders, can be proven based on the so-called velocity method (cf. [SZ92, DZ11]) and the well-known implicit function theorem (cf. [Zei85]).

This analysis may be carried out in such a way that an explicit and numerically accessible formula of the gradient is then obtainable via standard methods of matrix calculus (cf. [MN88]), which clears the second objective. A pleasant feature of both approaches is that they are in principle applicable to a large class of models, including almost arbitrary membrane energies over arbitrary parameterizable domains, as long as stationary membranes are locally unique and computable.

Concerning the discretization objective, it is in the nature of things that there is no general scheme for all types of models. Hence the focus is on a subset of linear models, which includes popular and numerically readily accessible choices such as the so-called Monge-gauge formulation with curve constraints, but also so-called point value constraints or tubular reference membrane shapes. In order to avoid frequent remeshing and therefore to allow an efficient treatment of moving particle interfaces, the so-called fictitious domain stabilized Nitsche method is applied, which is known to be capable of yielding optimal rates of convergence and to have the essential robustness properties (cf. [BH12]).

Altogether, this work creates a basis for the development and application of convergent and robust gradient-type methods in the context of membrane-mediated particle interactions. Typical use cases might be the exploration of stationary particle configurations and pattern formation. In addition, also the simulation of simple thermodynamic models based on the interaction potential and Langevin-type equations becomes feasible. Another benefit of the developed framework is that it only requires relatively few

simplifying assumptions and therefore the methods from the following chapters may additionally be valuable during the development of new interaction models and simulation methods which incorporate a greater level of physical detail than the currently established ones.

Many further aspects and questions that arise naturally from the results in this thesis exceed the defined scope but might be of interest for future research. Among these are the numerical approximation of higher order derivatives, possibly along the lines of [DZ91], the discretization and analysis of nonlinear models, potentially in the spirit of [DD06, DKS15, DGR17], as well as the implementation of more efficient and robust numerical algorithms in general, for example by including recently developed geometric multigrid methods as in [LGR18]. Even the extension to parametric formulations in the sense of [DE13] or to stochastic models as treated in [Bro08, Bro11] might be considered, although the problems originating from these topics are arguably more involved and therefore pose a greater challenge.

To the best of the author's knowledge, no previous research regarding the differentiability of the interaction potential for the sake of numerical applications has been conducted in a comparably general context. A method of differentiation by directly relating the surface geometry to its transmitted forces is considered in [MDG05a, MDG05b, Des15], but the numerical applicability of the approach presented therein remains to be investigated.

Outline

In Chapter 1 the general model framework is introduced within which the differentiability is considered later on. In particular, also the interaction potential and other related quantities are defined there. This part is mainly based on the variational framework from [EGH$^+$16, Hob16] but introduces a novel geometry-based formalism for the modeling of parametric particles, where the focus is on so-called curve and point value constraints. Concerning the elastic membrane energy, a Canham–Helfrich type functional is chosen as a root model, which subsequently is simplified by parameterization and linearization techniques. Most importantly, in the absence of particles the membrane is generally assumed to be parameterized over a fixed reference domain. While the resulting framework is compatible with a wide array of possible non-linearized particle–membrane interaction models, the main focus of attention lies on linearized models with either parametric curve or point value constraints during the remaining chapters.

The core analytical results are shown in Chapter 2 where differentiability of the interaction potential for general parameterized energies is proven. Velocity methods from shape calculus are used in order to construct feasible domain perturbations that maintain the particle-induced constraints in an appropriate sense. The actual differentiability result is due to a transformation of the relevant function spaces and energies onto a reference set and an elementary application of the implicit function theorem. Under mild assumptions this even yields smoothness of the interaction potential in the interior of its domain of definition. Application of standard methods of matrix calculus enable the

derivation of a volume-representation of the gradient in dependence of the corresponding stationary membrane solution. The resulting analytic formula only contains terms that are both computable and controllable within a finite element scheme, which makes it numerically feasible.

Afterwards, Chapter 3 restricts the attention completely to linearized elliptic energies subject to curve or point value constraints. For the discrete approximation of stationary membranes subject to curve constraints a so-called fictitious domain stabilized Nitsche method is introduced for shape-regular unfitted grids of quadrilaterals. For membranes subject to point value constraints a conforming finite element discretization is introduced based on a standard Galerkin method with the help of local QR decompositions. In both cases optimal a priori error estimates are proven. In addition, it is also shown for such linearized models that the error in the discrete approximation of the interaction potential's gradient inherits the rate of convergence that is obtained for the discretization of the stationary membrane.

Chapter 4 explains a possible implementation of these methods in conjunction with a level set ansatz for the description of parametric interfaces. It then proceeds to test previously obtained analytic error estimates for various different examples. In particular, rates of convergence for the computation of stationary membrane shapes and for the approximation of the interaction potential gradient are explored and discussed with respect to their optimality and efficiency.

In Chapter 5 possible applications of the developed framework are demonstrated. The clustering behavior of isotropic and certain anisotropic particles is examined in Monge-gauge, and also the interaction of isotropic particles in a tubular membrane is illustrated. Moreover, a model for the membrane-mediated aggregation of FCHo2 F-BAR domains is proposed based on physical data obtained from protein data base (PDB) files. The computation of stationary configurations in those examples is accomplished via a perturbed gradient method.

Finally, Chapter 6 gives a short outlook on a simple extension of the previously considered models to include thermal fluctuations of particles via Langevin dynamics. This again leads naturally to the task of computing quantities from statistical mechanics such as free energies, which can be tackled with the help of now newly accessible gradient-type methods, such as the so-called energy perturbation method.

Acknowledgments

During my work on this thesis I received most valuable support from a variety of people. In this spirit, I would like to thank my advisor Prof. Dr. Ralf Kornhuber for his guidance on mathematical topics and beyond, and for offering me the space to develop my ideas and skills freely. The same I would like to thank Prof. Dr. Carsten Gräser for his miraculous stamina during mathematical discussions which made the collaborative work highly enjoyable. Special thanks go to our collaborator Prof. Dr. Michael Kozlov for his contribution to the idea to model FCHo2 particles via point value constraints, to Dr. Leonardo Almeida-Souza for supplying and explaining the relevant information from the protein data base, and to Prof. Dr. Luigi Delle Site for his insights on statistical mechanics. Furthermore, I also would like to thank all my colleagues at the department of mathematics who helped to create such a wonderful and respectful work environment. Last in order but certainly not of least importance is the emotional backup granted by my family and friends. I am especially grateful for my endlessly loving and stunning wife Elizabeth who not only supports me in every conceivable and therefore in (almost-surely uncountably) infinite ways, but who also appreciates bad puns.

This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project A07 "Langevin dynamics of particles in membranes."

1 Hybrid models for membrane-mediated particle interactions

This chapter introduces the general modeling framework within which membrane-mediated particle interactions are approached in the subsequent chapters. The first part introduces the Canham–Helfrich elastic model for lipid membranes and proceeds with simplification techniques for the resulting energy by parameterization over specific reference domains and by geometric linearization. The second part discusses the concept of constraint-based membrane–particle couplings on the basis of so-called curve and point value constraints. The concept of parameter-dependent particle configurations and constraints is also established. The chapter ends with the definition of the interaction potential and the statement of the stationary particle configuration problem.

1.1 Elastic Canham–Helfrich type membrane bending energy

At the core of the model lies the description of the lipid membrane as a sufficiently smooth surface $\mathcal{M} \subseteq \mathbb{R}^3$, which is depicted in Fig. 1.1.1. The energy associated to this surface $\mathcal{M}$ is given in terms of its *principal curvatures* κ_1, κ_2 by the *Canham–Helfrich energy with surface tension*

$$J_{\mathrm{CH}}(\mathcal{M}) := \int_{\mathcal{M}} \frac{1}{2}\kappa(H - c_0)^2 + \kappa_G K + \sigma \,\mathrm{d}\mathbb{H}(\mathcal{M}).$$

Here, $H := \kappa_1 + \kappa_2$ is the *mean curvature*, $K := \kappa_1\kappa_2$ is the *Gaussian curvature*, $c_0 \colon \mathcal{M} \to \mathbb{R}$ is the *spontaneous curvature*, $\kappa \in \mathbb{R}_{>0}$ and $\kappa_G \in \mathbb{R}$ are *bending rigidities*,

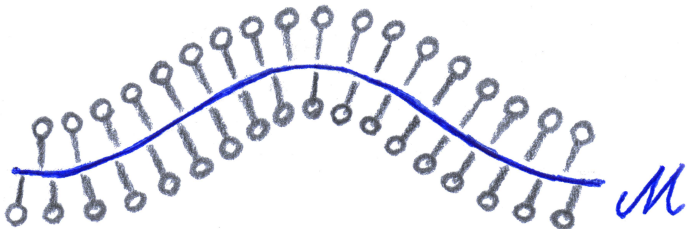

Figure 1.1.1: The biological membrane is a bilayer formed by lipids with hydrophobic tails and hydrophilic heads (rods with spheres). It is approximated by an infinitely thin single sheet along its center line (continuous blue line $\mathcal{M}$).

and $\sigma \in \mathbb{R}_{\geq 0}$ is the *surface tension*. The integration is carried out with respect to the *Hausdorff measure* $\mathbb{H}(\mathcal{M})$.

Stationary membranes are determined by minimization of this energy. Therefore, the direct use of the Canham–Helfrich energy in the interaction model implies that its associated Lagrange equation needs to be solved over a not yet defined set of admissible membranes. In the special case $c_0 = \kappa_G = \sigma = 0$ the above energy is reduced to the *Willmore energy* and the corresponding Lagrange equation reads

$$\Delta_{\mathcal{M}} H + 2H(H^2 - K) = 0.$$

Analogous formulations are known for the general case (cf. [DN12]). This defines a geometric fourth order partial differential equation, which is highly nonlinear in nature. The analytical and numerical analysis of such a problem pose a considerable challenge and the repeated numerical solution of this PDE, which is required during the numerical optimization of particle configurations, demands a substantial amount of computational power (cf. [Dzi08, BGN08]). In order to reduce complexity, it is desirable to incorporate further assumptions and approximations into the membrane model as far as reasonably possible.

Arguably, the most commonly used simplification relies on the famous *Gauss–Bonnet theorem*, which states

$$\int_{\mathcal{M}} K \,\mathrm{d}\mathbb{H}(\mathcal{M}) + \int_{\partial\mathcal{M}} k_g \,\mathrm{d}\mathbb{H}(\partial\mathcal{M}) = 2\pi_\chi(\mathcal{M}). \tag{1.1.1}$$

Here k_g denotes the *geodesic curvature* of $\partial\mathcal{M}$ and $\pi_\chi(\mathcal{M})$ is the *Euler characteristic* of the hypersurface $\mathcal{M}$. The equality reveals the negligibility of the Gaussian curvature term during minimization of the elastic energy J_{CH} as long as the integral over the geodesic curvature k_g on the boundary $\partial\mathcal{M}$ remains constant within a given set of admissible membranes. This property is true especially if the shape of $\partial\mathcal{M}$ is fixed and if also the unit normals along $\partial\mathcal{M}$ remain constant. In the majority of particle interaction models considered in the literature, this assumption is generally fulfilled, and henceforth the Gaussian curvature usually is rightfully neglected. However, in this work this simplification is *not* to be applied just yet because the negligibility depends on the precise particle–membrane interaction model and other model decisions concerning every particle's degrees of freedom which are still to be introduced. Therefore the further discussion of this issue is postponed until the end of this chapter in Section 1.3.1.

Two further simplifications are also possible under the assumption that stationary membranes are perturbations of an *a priori* known reference surface $\mathcal{M}_0$. The first simplification technique makes use of this situation by transforming the minimization problem via a parameterization of the reference surface $\mathcal{M}_0$ onto some parameterization domain Ω. This reduces the problem of computing stationary membranes, which usually corresponds to an energy minimization problem over a class of hypersurfaces, to a scalar non-linear PDE over the single fixed set Ω. The second technique imposes an additional smallness assumption on the perturbations of the reference surface $\mathcal{M}_0$, which vindicates the approximation of the elastic energy J_{CH} in terms of a second order Taylor expansion.

Hence, a stationary membrane shape is approximated by minimizers of a quadratic energy, which in turn corresponds to solving a linear PDE over the fixed domain Ω.

1.1.1 Parameterization over a reference domain

This section considers a parameterized bounded reference surface $\mathcal{M}_0$. For the sake of simplicity, suppose existence of an almost-global parameterization, i. e. a single bounded parameterization domain $\Omega \subseteq \mathbb{R}^2$ and a map $\varphi\colon \Omega \to \mathcal{M}_0$ which is a bijection up to a relative null set of $\mathcal{M}_0$. It is furthermore assumed that φ defines a 2-diffeomorphism. Given a unit normal field ν_0 on $\mathcal{M}_0$ and a perturbation function $u \in C^2(\Omega)$, the *perturbation of $\mathcal{M}_0$ by u (in normal direction ν_0)* is defined as

$$\mathcal{M} := \mathcal{M}(u) := \{\varphi(x) + u(x)\nu_0(\varphi(x)) \mid x \in \Omega\}\,.$$

The following paragraphs implicitly assume that $\|u\|_\infty$ is sufficiently small and, if applicable, fulfills suitable boundary conditions on Ω such that the perturbation $\mathcal{M}$ is again a surface without self-intersections and which is smooth enough for the energy $J_{\mathrm{CH}}(\mathcal{M})$ to be well-defined.

By standard means of differential geometry (as introduced e. g. in [Wil13]) the elastic energy $J_{\mathrm{CH}}(\mathcal{M})$ is expressable exclusively in terms of the perturbation function u and derivatives thereof as well as geometric quantities of the reference surface $\mathcal{M}_0$, but without relying explicitly on quantities of the perturbed surface $\mathcal{M}$.

In order to carry out this transformation explicitly one defines a parameterization of the perturbed surface $\mathcal{M}$ by

$$\phi(x) := \varphi(x) + u(x)\nu_0(\varphi(x)).$$

In a slight abuse of notation the partial derivatives of ϕ are abbreviated by

$$\phi_i := \partial_i \phi, \qquad \phi_{ij} := \partial_{ij}\phi,$$

where it is noted that the following steps entirely avoid any explicit indexing of components of ϕ in order to rule out any possible ambiguity with this notation. The tangent space of the perturbed surface $\mathcal{M}$ in a point $\phi(x)$ is given by the image of the Jacobian

$$D\phi(x) = \big(\phi_1(x), \quad \phi_2(x)\big)$$

and hence a normal field ν on $\mathcal{M}$ is defined by

$$\widetilde{\nu} := \phi_1 \times \phi_2, \qquad Q := \|\widetilde{\nu}\|\,, \qquad \nu := \frac{\widetilde{\nu}}{Q}.$$

The *first fundamental* form, also known as *metric tensor*, is given by the symmetric matrix

$$\mathrm{I} := D\phi^T D\phi = \begin{pmatrix} \|\phi_1\|^2 & \phi_1 \cdot \phi_2 \\ \phi_1 \cdot \phi_2 & \|\phi_2\|^2 \end{pmatrix}.$$

It holds

$$\det(\mathrm{I}) = \|\phi_1\|^2 \|\phi_2\|^2 - (\phi_1 \cdot \phi_2)^2 = Q^2$$

and therefore

$$\mathrm{I}^{-1} = \frac{1}{Q^2} \begin{pmatrix} \mathrm{I}_{22} & -\mathrm{I}_{12} \\ -\mathrm{I}_{21} & \mathrm{I}_{11} \end{pmatrix} = \frac{1}{Q^2} \begin{pmatrix} \|\phi_2\|^2 & -\phi_1 \cdot \phi_2 \\ -\phi_1 \cdot \phi_2 & \|\phi_1\|^2 \end{pmatrix}.$$

The *second fundamental form* is defined as

$$\mathrm{II} := \begin{pmatrix} \partial_{11}\phi \cdot \nu & \partial_{12}\phi \cdot \nu \\ \partial_{21}\phi \cdot \nu & \partial_{22}\phi \cdot \nu \end{pmatrix} = \frac{1}{Q} \begin{pmatrix} \partial_{11}\phi \cdot \widetilde{\nu} & \partial_{12}\phi \cdot \widetilde{\nu} \\ \partial_{21}\phi \cdot \widetilde{\nu} & \partial_{22}\phi \cdot \widetilde{\nu} \end{pmatrix}$$

and the *shape operator* $S := \mathrm{II} \cdot \mathrm{I}^{-1}$ takes the form

$$S = \frac{1}{Q^3} \begin{pmatrix} \|\phi_2\|^2 (\phi_{11} \cdot \widetilde{\nu}) - (\phi_1 \cdot \phi_2)(\phi_{12} \cdot \widetilde{\nu}), & \|\phi_1\|^2 (\phi_{12} \cdot \widetilde{\nu}) - (\phi_1 \cdot \phi_2)(\phi_{11} \cdot \widetilde{\nu}) \\ \|\phi_2\|^2 (\phi_{21} \cdot \widetilde{\nu}) - (\phi_1 \cdot \phi_2)(\phi_{22} \cdot \widetilde{\nu}), & \|\phi_1\|^2 (\phi_{22} \cdot \widetilde{\nu}) - (\phi_1 \cdot \phi_2)(\phi_{21} \cdot \widetilde{\nu}) \end{pmatrix}.$$

The eigenvalues of the shape operator define the *principal curvatures* κ_1 and κ_2. Consequently, the *mean curvature* is

$$H = \mathrm{Tr}(S) = \frac{1}{Q^3} \left(\|\phi_1\|^2 (\phi_{22} \cdot \widetilde{\nu}) + \|\phi_2\|^2 (\phi_{11} \cdot \widetilde{\nu}) - 2(\phi_1 \cdot \phi_2)(\phi_{12} \cdot \widetilde{\nu}) \right)$$

and the *Gaussian curvature* is

$$K = \det(S) = \frac{1}{Q^4} \left((\phi_{11} \cdot \widetilde{\nu})(\phi_{22} \cdot \widetilde{\nu}) - (\phi_{12} \cdot \widetilde{\nu})^2 \right).$$

Finally, application of the transformation formula to J_{CH} with the diffeomorphism ϕ yields the *parameterized elastic energy*

$$J_{\mathrm{CH}}(\mathcal{M}) = \int_\Omega Q \left(\frac{\kappa}{2} (H - c_0)^2 + \kappa_G K + \sigma \right) \mathrm{d}x.$$

Similar to the original formulation, the minimization of this energy is in general equivalent to solving a highly non-linear fourth order PDE. However, due to the restriction to perturbations of $\mathcal{M}_0$, this PDE is now stated for a scalar function u over a single fixed reference domain Ω instead of for a class of hypersurfaces.

In the context of particle–membrane interaction models flat, tubular and spherical geometries are especially of interest. The examples below state the parameterized mean and Gaussian curvature as well as the volume element Q for the corresponding reference surfaces. This immediately yields the respective parameterized elastic energy by substitution of those quantities into the formula above. Although the main focus of attention in this thesis is on the linearized models presented in the next section, such non-linear parameterized formulations have a potential value in future applications for the modeling of particle interactions with locally strong deformations.

While generally relatively little is known about the regularity of solutions of the associated non-linear PDEs and about the approximation properties of related discretization schemes, there are, at least to an extent, such results for the closely related minimization problem of the Willmore functional in the graph case, see [DKS15, DGR17]. Those works give rise to the expectation that similar properties might be derivable for other related formulations as well, thus potentially making these problems numerically more tractable.

Example 1.1.1 (Euclidean plane)**.** *Suppose $\Omega \subseteq \mathbb{R}^2$ and define a flat reference surface by $\mathcal{M}_0 := \Omega \times \{0\}$, admitting the canonical parameterization $\varphi\colon \Omega \to \mathcal{M}_0$, $x \mapsto (x, 0)$. For notational convenience, denote the partial derivatives of u by $u_i := \partial_i u$ and $u_{ij} := \partial_{ij} u$. It holds*

$$\phi(x) = \begin{pmatrix} x_1 \\ x_2 \\ u(x) \end{pmatrix}, \quad \phi_1(x) = \begin{pmatrix} 1 \\ 0 \\ u_1(x) \end{pmatrix}, \quad \phi_2(x) = \begin{pmatrix} 0 \\ 1 \\ u_2(x) \end{pmatrix}, \quad \widetilde{\nu}(x) = \begin{pmatrix} -u_1(x) \\ -u_2(x) \\ 1 \end{pmatrix}$$

and

$$Q(x) = \sqrt{1 + u_1(x)^2 + u_2(x)^2}, \qquad \phi_{ij}(x) \cdot \widetilde{\nu}(x) = u_{ij}(x).$$

Therefore,

$$H = \frac{u_{22}(1+u_1^2) + u_{11}(1+u_2^2) - 2u_1u_2u_{12}}{(1+u_1^2+u_2^2)^{3/2}}, \qquad K = \frac{u_{11}u_{22} - u_{12}^2}{(1+u_1^2+u_2^2)^2}\ .$$

Example 1.1.2 (Tube)**.** *Let $L \in \mathbb{R}_{>0}$, $r \in \mathbb{R}_{>0}$ and suppose a tube of length L and radius r, given by*

$$\mathcal{M}_0 := \{x \in \mathbb{R}^3 \mid x_1 \in [0, L],\ x_2^2 + x_3^2 = r\}.$$

A global parameterization of the cylinder $\mathcal{M}_0$ up to a null set over the reference domain $\Omega = [0, L] \times [0, 2\pi)$ is given by the cylindrical coordinates

$$\varphi(x, y) = \begin{pmatrix} x \\ r\cos(y) \\ r\sin(y) \end{pmatrix}, \qquad \nu_0(x, y) = \begin{pmatrix} 0 \\ \cos(y) \\ \sin(y) \end{pmatrix}.$$

With the shorthands $c := \cos(y)$, $s := \sin(y)$ one has

$$\phi = \begin{pmatrix} x \\ (u+r)c \\ (u+r)s \end{pmatrix}, \quad \phi_1 = \begin{pmatrix} 1 \\ u_1 c \\ u_1 s \end{pmatrix}, \quad \phi_2 = \begin{pmatrix} 0 \\ u_2 c - (u+r)s \\ u_2 s + (u+r)c \end{pmatrix}, \quad \widetilde{\nu} = \begin{pmatrix} u_1(u+r) \\ -u_2 s - (u+r)c \\ u_2 c - (u+r)s \end{pmatrix}$$

with

$$\|\phi_1\|^2 = 1 + u_1^2, \quad \|\phi_2\|^2 = (u+r)^2 + u_2^2, \quad \phi_1 \cdot \phi_2 = u_1 u_2$$

as well as

$$\begin{aligned} Q &= \sqrt{(1+u_1^2)(u+r)^2+u_2^2} \\ \phi_{11}\cdot\widetilde{\nu} &= -u_{11}(u+r) \\ \phi_{12}\cdot\widetilde{\nu} &= u_1u_2-(u+r)u_{12} \\ \phi_{22}\cdot\widetilde{\nu} &= (u+r)^2-(u+r)u_{22}+2u_2^2\,. \end{aligned}$$

This yields

$$\begin{aligned} H = \frac{1}{Q^3}\big(&-(u+r)^3u_{11}+(u+r)^2(1+u_1^2) \\ &+(u+r)\left(2u_1u_2u_{12}-u_{22}(1+u_1^2)-u_2^2u_{11}\right)+2u_2^2\big) \end{aligned}$$

and

$$\begin{aligned} K = \frac{1}{Q^4}\big(&-(u+r)^3u_{11}+(u+r)^2(u_{11}u_{22}-u_{12}^2) \\ &+2(u+r)(u_1u_2u_{12}-u_2^2u_{11})-u_1^2u_2^2\big)\,. \end{aligned}$$

Note that here the perturbation function u is required to fulfill suitable periodic boundary conditions along the boundary $[0,L]\times\{0,2\pi\}$ in order to compensate for the fact that the function φ is not a global parameterization in the classical sense. This ensures $u\circ\varphi^{-1}\in C^2(\mathcal{M}_0)$.

Example 1.1.3 (Sphere)**.** *Suppose $r\in\mathbb{R}_{>0}$ and an r-sphere*

$$\mathcal{M}_0 := \left\{x\in\mathbb{R}^3 \mid \|x\| = r\right\}$$

and parameterize this surface up to a null set over the reference domain $\Omega=[0,2\pi)\times[0,\pi]$ by the spherical coordinates

$$\varphi(x,y) = r\begin{pmatrix}\cos(x)\sin(y)\\ \sin(x)\sin(y)\\ \cos(y)\end{pmatrix}, \qquad \nu_0(x,y) = \begin{pmatrix}\cos(x)\sin(y)\\ \sin(x)\sin(y)\\ \cos(y)\end{pmatrix}.$$

Using the shorthands $c_x=\cos(x), s_x=\sin(x), c_y=\cos(y), s_y=\sin(y)$, this leads to

$$\begin{aligned} \phi &= (r+u)\begin{pmatrix}\cos(x)\sin(y)\\ \sin(x)\sin(y)\\ \cos(y)\end{pmatrix}, \quad \phi_1 = \begin{pmatrix}s_y\left(c_xu_1-(u+r)s_x\right)\\ s_y\left(s_xu_1+(u+r)c_x\right)\\ c_yu_1\end{pmatrix}, \\ \phi_2 &= \begin{pmatrix}c_x\left(s_yu_2+(u+r)c_y\right)\\ s_x\left(s_yu_2+(u+r)c_y\right)\\ c_yu_2-(u+r)s_y\end{pmatrix}, \quad \widetilde{\nu} = (u+r)\begin{pmatrix}-(u+r)c_xs_y^2-u_1s_x+u_2c_xc_ys_y\\ -(u+r)s_xs_y^2+u_1c_x+u_2s_xc_ys_y\\ -s_y\left((u+r)c_y+u_2s_y\right)\end{pmatrix} \end{aligned}$$

and

$$\|\phi_1\|^2 = (u+r)^2s_y^2+u_1^2, \qquad \|\phi_2\|^2=(u+r)^2+u_2^2, \qquad \phi_1\cdot\phi_2 = u_1u_2$$

as well as

$$\begin{aligned}
Q &= |u+r|\sqrt{\left((u+r)^2+u_2^2\right)s_y^2+u_1^2}\\
\phi_{11}\cdot\widetilde{\nu} &= (u+r)\left(s_y^3(u+r)^2+s_y\left(2u_1^2-(u+r)u_{11}\right)-u_2c_ys_y^2(u+r)\right)\\
\phi_{12}\cdot\widetilde{\nu} &= (u+r)\left((u+r)\left(u_1c_y-u_{12}s_y\right)+2u_1u_2s_y\right)\\
\phi_{22}\cdot\widetilde{\nu} &= s_y(u+r)\left((u+r)^2-u_{22}(u+r)+2u_2^2\right).
\end{aligned}$$

Altogether,

$$\begin{aligned}
H = \frac{(u+r)^2}{Q^3}\Big(&2s_y^3(u+r)^3-s_y(u+r)^2\left(u_{11}+u_{22}s_y^2+c_ys_yu_2\right)+3s_y(u+r)\left(u_1^2+s_y^2u_2^2\right)\\
&+s_y\left(u_{11}u_2^2+u_{22}u_1^2-2u_1u_2u_{12}\right)+c_yu_2\left(2u_1^2+s_y^2u_2^2\right)\Big)
\end{aligned}$$

and

$$\begin{aligned}
K = \frac{(u+r)^3}{Q^4}\Big(&(u+r)^3s_y^4-(u+r)^2s_y^2\left(u_{11}+s_yc_yu_2+u_{22}s_y^2\right)\\
&+(u+r)s_y^2\left(u_{11}u_{22}-u_{12}^2+2u_1^2+2u_2^2s_y^2+s_yc_yu_2u_{22}\right)-(u+r)c_y^2u_1^2\\
&-2s_y^2\left(u_{11}u_2^2-2u_1u_2u_{12}+u_{22}u_1^2\right)+2s_yc_y\left((u+r)u_1u_{12}-2u_1^2u_2-u_2^3s_y^2\right)\Big).
\end{aligned}$$

As for the cylinder, the perturbation function is required to fulfill suitable periodic boundary conditions on the boundary $\partial\Omega$ in order to ensure smoothness of the perturbation surface $\mathcal{M}(u)$.

1.1.2 Geometric linearization

Next the linearization of the elastic bending energy J_{CH} via a second order Taylor expansion is considered. Such an approximation simplifies the computation of a stationary membrane shape to the solution of a linear fourth order partial differential equation.

Let $\mathcal{M}_0$ again be a sufficiently smooth reference surface and ν_0 a unit normal field on $\mathcal{M}_0$. Given a perturbation function $u \in C^2(\mathcal{M}_0)$, a family of perturbed surfaces is defined by

$$\mathcal{M}(t) := \{x + tu(x)\nu_0(x) \mid x \in \mathcal{M}_0\}.$$

Again, assuming that $\|u\|_{L^\infty(\mathcal{M}_0)}$ is small enough, $\mathcal{M}(t)$ is well-defined for $t \in [-1,1]$. With

$$J(t) := J_{\mathrm{CH}}(\mathcal{M}(t))$$

the linearization of J_{CH} in u and in normal direction ν_0 is given by

$$J_{\mathrm{lin}}(u) := J(0) + J'(0) + \frac{1}{2}J''(0) \approx J(1) = J_{\mathrm{CH}}(\mathcal{M}(u)).$$

Similar to the case of parameterized energies, it is possible to state the linearized energy exclusively in terms of the perturbation function u and its derivatives and in terms of geometric quantities of the reference surface $\mathcal{M}_0$ without relying on the $\mathcal{M}(t)$ explicitly.

An extensive computation of the linearization terms is carried out in Appendix B. For the reader's convenience the most important results are restated below. There, H_0 and K_0 denote the mean and Gaussian curvature on $\mathcal{M}_0$, respectively, and $\mathcal{H}_0$ and $\hat{\mathcal{H}}_0 := \mathcal{H}_0 + \nu_0 \otimes \nu_0$ are the extended Weingarten map and the regularized extended Weingarten map, respectively. Furthermore, $\nabla_{\mathcal{M}_0}$ and $D^2_{\mathcal{M}_0}$ are the surface gradient and surface Hessian, respectively. Finally, $\dot{\partial}$ denotes the so-called material derivative on the space–time manifold $\mathcal{M}_T := \bigcup_{t\in[-1,1]}\{t\} \times \mathcal{M}(t)$. See Appendix B for precise definitions of the involved quantities.

Proposition B.9. *It holds*

$$\begin{aligned}\dot{\partial} H|_{t=0} &= -u \left\|\mathcal{H}_0\right\|_F^2 - \Delta_{\mathcal{M}_0} u \\ \dot{\partial}^2 H|_{t=0} &= 2u^2 \operatorname{Tr}(\mathcal{H}_0^3) - \nabla_{\mathcal{M}_0} u^T \mathcal{H}_0 \nabla_{\mathcal{M}_0} u + 4u\mathcal{H}_0 : D^2_{\mathcal{M}_0} u \\ &\quad - \left\|\nabla_{\mathcal{M}_0} u\right\|^2 H_0 + 2u\Delta_{\mathcal{M}_0}\nu_0 \cdot \nabla_{\mathcal{M}_0} u\end{aligned}$$

where $\Delta_{\mathcal{M}_0}$ is the Laplace-Beltrami operator on $\mathcal{M}_0$ and $\|\cdot\|_F$ denotes the Frobenius norm with $:$ as the inducing matrix scalar product.

Proposition B.10. *Let* $\operatorname{adj}(\hat{\mathcal{H}})$ *denote the adjugate matrix of* $\hat{\mathcal{H}}$*. It holds*

$$\begin{aligned}\dot{\partial} K|_{t=0} &= \operatorname{Tr}(\operatorname{adj}(\hat{\mathcal{H}})\dot{\partial}\hat{\mathcal{H}})|_{t=0} \\ \dot{\partial}^2 K|_{t=0} &= \operatorname{Tr}(\dot{\partial}\operatorname{adj}(\hat{\mathcal{H}})\,\dot{\partial}\hat{\mathcal{H}})|_{t=0} + \operatorname{Tr}(\operatorname{adj}(\hat{\mathcal{H}})\dot{\partial}^2\hat{\mathcal{H}})|_{t=0}\end{aligned}$$

where expressions for $\dot{\partial}\hat{\mathcal{H}}|_{t=0}$ *and* $\dot{\partial}^2\hat{\mathcal{H}}|_{t=0}$ *are given in Lemma B.8. If furthermore* $K(0) \neq 0$*, let* $A := \hat{\mathcal{H}}^{-1}\dot{\partial}\mathcal{H}^{-1}|_{t=0}$*. Then*

$$\begin{aligned}\dot{\partial} K|_{t=0} &= K \operatorname{Tr}(A) \\ \dot{\partial}^2 K|_{t=0} &= K\left(\operatorname{Tr}(A)^2 - \|A\|_F^2 + \operatorname{Tr}\left(\hat{\mathcal{H}}^{-1}\dot{\partial}^2\hat{\mathcal{H}}\right)|_{t=0}\right).\end{aligned}$$

Theorem B.13. *It holds*

$$\frac{\partial}{\partial t}\int_{\mathcal{M}(t)} 1 \,\mathrm{d}\mathbb{H}\bigg|_{t=0} = \int_{\mathcal{M}_0} uH_0 \,\mathrm{d}\mathbb{H}$$

and

$$\frac{\partial^2}{\partial t^2}\int_{\mathcal{M}(t)} 1 \,\mathrm{d}\mathbb{H}\bigg|_{t=0} = \int_{\mathcal{M}_0} \left\|\nabla_{\mathcal{M}_0} u\right\|^2 + 2u^2 K_0 \,\mathrm{d}\mathbb{H}.$$

Theorem B.14. *It holds*

$$\begin{aligned}\frac{\partial}{\partial t}\frac{1}{2}\int_{\mathcal{M}(t)} (H - c_0)^2 \,\mathrm{d}\mathbb{H}\bigg|_{t=0} = \int_{\mathcal{M}_0} &-(H_0 - c_0)\left(u\left\|\mathcal{H}_0\right\|_F^2 + \Delta_{\mathcal{M}_0} u + \dot{\partial} c_0\right) \\ &+ \frac{1}{2} u H_0 (H_0 - c_0)^2 \,\mathrm{d}\mathbb{H}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial t^2}\frac{1}{2}\int_{\mathcal{M}(t)}(H-c_0)^2\,\mathrm{d}\mathbb{H}\bigg|_{t=0} &= \int_{\mathcal{M}_0}\left(u\,\|\mathcal{H}_0\|_F^2+\Delta_{\mathcal{M}_0}u+\dot\partial c_0|_{t=0}\right)^2\mathrm{d}\mathbb{H}\\ &\quad+\int_{\mathcal{M}_0}(H_0-c_0)(\dot\partial^2 H|_{t=0}-\dot\partial^2 c_0|_{t=0})\,\mathrm{d}\mathbb{H}\\ &\quad-\int_{\mathcal{M}_0}2uH_0(H_0-c_0)\left(u\,\|\mathcal{H}_0\|_F^2+\Delta_{\mathcal{M}_0}u+\dot\partial c_0|_{t=0}\right)\mathrm{d}\mathbb{H}\\ &\quad+\int_{\mathcal{M}_0}\frac{1}{2}(H_0-c_0)^2\left(\|\nabla_{\mathcal{M}_0}u\|^2+u^2K_0\right)\mathrm{d}\mathbb{H}\end{aligned}$$

where $\dot\partial^2 H|_{t=0}$ is as in Proposition B.9.

Theorem B.15. *It holds*

$$\frac{\partial}{\partial t}\int_{\mathcal{M}(t)}K\,\mathrm{d}\mathbb{H}\bigg|_{t=0}=\int_{\mathcal{M}_0}\dot\partial K|_{t=0}+uK_0H_0\,\mathrm{d}\mathbb{H}$$

and

$$\frac{\partial^2}{\partial t^2}\int_{\mathcal{M}(t)}K\,\mathrm{d}\mathbb{H}\bigg|_{t=0}=\int_{\mathcal{M}_0}\dot\partial^2 K|_{t=0}+2uH_0\dot\partial K|_{t=0}+K_0\left(\|\nabla_{\mathcal{M}_0}u\|^2+u^2K_0\right)\mathrm{d}\mathbb{H}$$

where $\dot\partial K|_{t=0}$ and $\dot\partial^2 K|_{t=0}$ are given in Proposition B.10.

Examples of linearized bending energies

Combining the results from Theorems B.13 to B.15 yields the linearized bending energy J_{lin}. For the special cases of flat and tubular reference surfaces the explicit formulas of the linearized energy are shown below where the energy is both linearized and parameterized over a suitable reference domain. In those examples the spontaneous curvature $c_0(t)$ is assumed to have a negligible dependence on the perturbation function u, which implies a vanishing material derivative $\dot\partial c_0=0$. Furthermore, u is directly expressed in local coordinates, i. e. as a function $u\colon\Omega\to\mathbb{R}$.

Example 1.1.4 (Euclidean plane)**.** *Suppose a flat surface $\mathcal{M}_0=\Omega\times\{0\}$. In that case ν_0 and $\hat{\mathcal{H}}_0|_{t=0}$ are constant and $\mathcal{H}_0$, K_0 and H_0 vanish. As a consequence, also* $\operatorname{adj}\hat{\mathcal{H}}|_{t=0}=0$ *holds and the derivative $\dot\partial K|_{t=0}$ vanishes as well. Explicit computation reveals*

$$\begin{aligned}\dot\partial^2 K|_{t=0} &= \dot\partial\operatorname{adj}(\hat{\mathcal{H}}):\dot\partial\hat{\mathcal{H}}\Big|_{t=0}\\ &= \begin{pmatrix}-u_{22}&u_{12}&0\\u_{12}&-u_{11}&0\\0&0&0\end{pmatrix}:\begin{pmatrix}u_{11}&u_{12}&-u_1\\u_{12}&u_{22}&-u_2\\-u_1&-u_2&0\end{pmatrix}\\ &= 2u_{12}^2-2u_{11}u_{22}.\end{aligned}$$

Moreover, the surface derivatives simplify to the standard derivatives in two dimensions, i. e. $\nabla_{\mathcal{M}_0}$ and $\Delta_{\mathcal{M}_0}$ are identified with ∇ and Δ on $\mathbb{R}^2$. Combining the above results yields, up to a u-independent constant, the linearized energy

$$J_{lin}(u) = \frac{1}{2}\int_\Omega \kappa(\Delta u)^2 + 2\kappa_G\left(u_{12}^2 - u_{11}u_{22}\right) + \left(\sigma + \frac{\kappa c_0^2}{2}\right)\|\nabla u\|^2 \,\mathrm{d}x + \int_\Omega \kappa\, c_0\, \Delta u \,\mathrm{d}x.$$

In case of vanishing spontaneous curvature $c_0 = 0$ and by neglecting the Gaussian curvature term, the well-known Canham–Helfrich energy in Monge-gauge *is recovered:*

$$J_{lin}(u) = \frac{1}{2}\int_\Omega \kappa(\Delta u)^2 + \sigma\|\nabla u\|^2 \,\mathrm{d}x.$$

Example 1.1.5 (Tube)**.** *Let $L, r \in \mathbb{R}_{>0}$ and suppose a tube of length L and radius r*

$$\mathcal{M}_0 = \left\{x \in \mathbb{R}^3 \mid x_1 \in [0, L],\ x_2^2 + x_3^2 = r^2\right\}$$

equipped with the cylindrical coordinates $\varphi(x, y) = (x, r\cos(y), r\sin(y))^T$ over the parameterization domain $\Omega := [0, L]\times[0, 2\pi]$. Moreover, for the sake of simplicity it is assumed that Gaussian curvature is negligible, i. e. $\kappa_G = 0$. Using the shorthands $c := \cos(y)$ and $s := \sin(y)$ the quantities of interest are given in local coordinates by

$$\widetilde{\nu}_0 = \begin{pmatrix}0\\ c\\ s\end{pmatrix}, \qquad \widetilde{\mathcal{H}}_0 = \frac{1}{r}\begin{pmatrix}0 & 0 & 0\\ 0 & s^2 & -cs\\ 0 & -cs & c^2\end{pmatrix},$$

which further implies $H_0 = \frac{1}{r}$, $\|\mathcal{H}_0\|_F^2 = \frac{1}{r^2}$ and $K_0 = 0$. The surface gradient and Hessian in local coordinates are

$$\nabla_{\mathcal{M}_0} u = \frac{1}{r}\begin{pmatrix}ru_1\\ -su_2\\ cu_2\end{pmatrix}, \qquad D^2_{\mathcal{M}_0} u = \frac{1}{r^2}\begin{pmatrix}r^2u_{11} & -ru_{12}s & ru_{12}c\\ -ru_{12}s & u_{22}s^2 + u_2cs & u_2s^2 - u_{22}cs\\ ru_{12}c & -u_2c^2 - u_{22}cs & u_{22}c^2 - u_2cs\end{pmatrix}.$$

Define $c := \frac{1}{r} - c_0$. Using the constant area element $Q \equiv r$ the parameterized linearized energy then reads, again up to a u-independent constant,

$$\begin{aligned}
J_{lin}(u) = {} & \frac{r\kappa}{2}\int_\Omega \left(u_{11} + \frac{u + u_{22}}{r^2}\right)^2 + c\left(\frac{2u^2}{r^3} - \frac{u_2^2}{r^3} + \frac{4uu_{22}}{r^3} - \frac{1}{r}\left(u_1^2 + \frac{u_2^2}{2}\right)\right)\mathrm{d}x\\
& + \frac{r\kappa}{2}\int_\Omega -\frac{2uc}{r}\left(u_{11} + \frac{u + u_{22}}{r^2}\right) + \frac{c^2}{2}\left(u_1^2 + \frac{u_2^2}{r^2}\right)\mathrm{d}x\\
& + \frac{r\sigma}{2}\int_\Omega \left(u_1^2 + \frac{u_2^2}{r^2}\right)\mathrm{d}x\\
& + r\kappa\int_\Omega -c\left(u_{11} + \frac{u + u_{22}}{r^2}\right) + \frac{uc^2}{2r}\,\mathrm{d}x\\
& + r\sigma\int_\Omega \frac{u}{r}\,\mathrm{d}x.
\end{aligned}$$

As in Example 1.1.2 for the parameterization over tubular surfaces it is again assumed that u fulfills appropriate periodic boundary conditions along the boundary $[0, L]\times\{0, 2\pi\}$.

1.2 Constraint-based membrane–particle coupling

Membrane–particle couplings are expressed via geometric constraints for the membrane surface which hold along an interface between membrane and particle. This concept leads to a wealth of possibilities for the modeling of such couplings and reflects on the diversity of membrane–particle interactions in biology. This view is in particular appropriate if the particles can be considered to be rigid and if such local deformations are enforced strongly. Within the variational framework, the works [EGH$^+$16] and [Hob16] give an overview of established and novel types of couplings for flat and spherical reference surfaces.

This section introduces two of those constraint types, namely *point value constraints* and *curve constraints*. In comparison to the above-mentioned works an additional effort is made to put those constraints into a fully geometric perspective, which de-couples them from the specific choice of reference surface. In order to relate the geometric particle objects and their constraints to a membrane surface, a formalism is introduced based on local projections. Within this setting, particle degrees of freedom are then modeled through the direct application of a *transformation function* to the geometric particle object. For ease of presentation only one single particle is considered at first.

1.2.1 Point value constraints

As the name suggests, point value constraints prescribe the membrane shape in a finite number of points. An example application of such constraints is the modeling of filament-induced deformations of membranes. There the filament is essentially considered to be a rigid rod whose end is attached to the membrane. Thus, given a fixed position of the rod and assuming that its end is sufficiently small, it is justified to model such an interaction by simply prescribing the membrane position at a single point. Another slightly different example application is the modeling of BAR-domain proteins binding to biomembranes, where the protein is effectively rigid and where there is a small number of atoms which are likely to connect to the membrane. See especially Section 5.5 later for further details on this in the context of a concrete application to the aggregation of FCHo2 F-BAR domains. Both concepts are illustrated graphically in Fig. 1.2.1.

Mathematically, these situations are each described by a finite set $\mathcal{G} \subseteq \mathbb{R}^3$ of discrete points that prescribe those areas of the particle to which the biomembrane binds. Point value constraints then require the surface $\mathcal{M}$ to contain all the points in $\mathcal{G}$, which simply is stated as the geometric inclusion condition

$$\mathcal{G} \subseteq \mathcal{M}.$$

Note that point value constraints are not to be confused with *point curvature constraints*, which are treated to a greater extent for example in [Hob16].

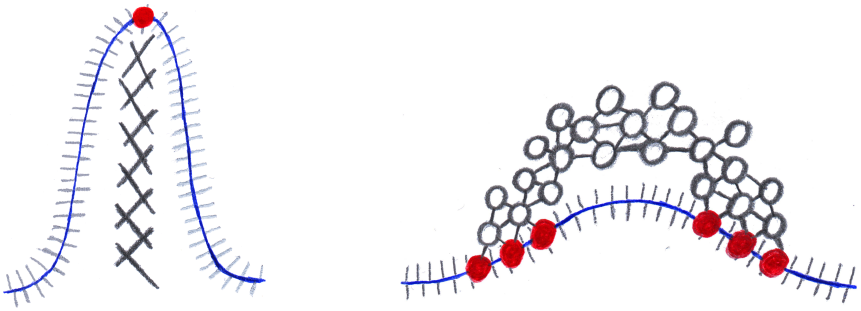

Figure 1.2.1: Schematics of particle–membrane interactions modeled by point value constraints. Blue curves indicate membrane shape and red dots are constraint points. Left: Filament pushing against a membrane. Right: Residues of a BAR-protein binding to a membrane.

Parameterization over reference set

Suppose a reference surface $\mathcal{M}_0$ with a unit normal field ν_0 and a parameterization function $\varphi\colon \Omega \to \mathcal{M}_0$. Given a perturbation functions $u\colon \Omega \to \mathbb{R}$, which in turn induces a perturbed surface

$$\mathcal{M}(u) := \{\varphi(x) + u(x)\nu_0(\varphi(x)) \mid x \in \Omega\},$$

the condition $\mathcal{G} \subseteq \mathcal{M}(u)$ may be expressed in terms of u alone. To this end, the *local projection function* $\pi\colon \mathbb{R}^3 \to \mathcal{M}_0$ and the *local distance function* $d\colon \mathbb{R}^3 \to \mathbb{R}$ are introduced on $\mathcal{M}_0$. For all x in a sufficiently small neighborhood of $\mathcal{M}_0$ they are uniquely defined by the property

$$x = \pi(x) + d(x)\,\nu_0(\pi(x)).$$

One then further arranges the points in $\mathcal{G}$ as $\mathcal{G} = \{x_1, \dots, x_N\}$ and defines

$$\widetilde{x}_i := \varphi^{-1}(\pi(x_i)), \qquad g_i := d(x_i).$$

The inclusion $\mathcal{G} \subseteq \mathcal{M}(u)$ is equivalently stated as

$$\forall i \in \{1, \dots, N\}\colon \; u(\widetilde{x}_i) = g_i,$$

or even shorter with $\Gamma := \{\widetilde{x}_1, \dots, \widetilde{x}_N\}$ and $g := (g_1, \dots, g_N) \in \mathbb{R}^N$ as

$$u|_\Gamma = g.$$

It quickly becomes evident that this representation need not be well-defined for general point sets $\mathcal{G}$. Problems occur in particular when $\pi(x_i) = \pi(x_j)$ holds for $i \neq j$ or when an x_i is outside the domain of definition of π or d. Therefore, well-definedness is imposed as an assumption, which is reasonable if the particle described by $\mathcal{G}$ has an appropriate orientation relative to the reference membrane and induces only a small deformation.

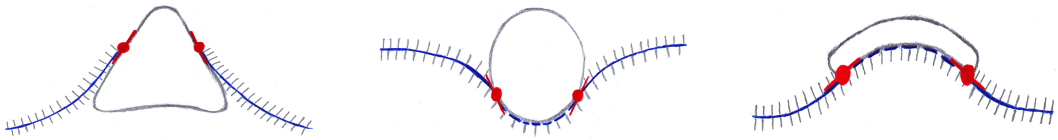

Figure 1.2.2: Schematics of objects interacting with a lipid bilayer leading to curve constraints. Blue curves indicate membrane shape, red dots are constraint points, and red lines show slope constraints. Left: Membrane inclusion. Middle: Partially wrapped nanoparticle. Right: Membrane scaffold.

1.2.2 Curve constraints

Curve constraints are typically used when the interface between particle and membrane forms a simple closed curve $\mathcal{G} \subseteq \mathbb{R}^3$. An example of this are transmembrane proteins with a so-called hydrophobic belt that forms the interface to which the membrane's lipids are likely to attach. Usually it is assumed that the structure of the belt furthermore prescribes a fixed slope for the membrane along the interface curve $\mathcal{G}$. Mathematically, the inclusion condition is directly stated as

$$\mathcal{G} \subseteq \mathcal{M}.$$

The slope-condition is interpreted such that for every interface point $x \in \mathcal{G}$ a two-dimensional vector space $\widetilde{T}_x\mathcal{G} \subseteq \mathbb{R}^3$ must be given which also contains the tangent space of $\mathcal{G}$, i. e. $T_x\mathcal{G} \subseteq \widetilde{T}_x\mathcal{G}$ holds. The slope-constraint is then stated by prescribing the tangent spaces of $\mathcal{M}$ along the interface $\mathcal{G}$ as

$$\forall x \in \mathcal{G}\colon\ T_x\mathcal{M} = \widetilde{T}_x\mathcal{G}.$$

Further examples for the application of curve constraints are the modeling of scaffolding particles or wrapped nanoparticles. While for the case of transmembrane proteins the membrane area "inside" the curve $\mathcal{G}$ is not present at all, these two other models would at first prescribe the membrane shape to be fixed on the inside. However, due to the fixed shape the total bending energy in that area remains constant over all feasible membrane shapes and is therefore negligible during energy minimization. Hence, scaffolding and wrapping interactions may equivalently be modeled by curve constraints. An illustration of these concepts is sketched in Fig. 1.2.2.

Parameterization over reference set

As before, suppose a reference surface $\mathcal{M}_0$ is given with a parameterization $\varphi\colon \Omega \to \mathcal{M}_0$ and the local projection and distance functions π and d, respectively. For notational convenience the interface $\mathcal{G}$ is extended locally to a hypersurface $\widetilde{\mathcal{G}} \supseteq \mathcal{G}$ such that the equality $T_x\widetilde{\mathcal{G}} = \widetilde{T}_x\mathcal{G}$ is true for all $x \in \mathcal{G}$. Given $\widetilde{\mathcal{G}}$, the *inverse local projection* Π *with respect to* $\widetilde{\mathcal{G}}$ is characterized by the property

$$\Pi(x) = x + d(\Pi(x))\nu_0(x) \in \widetilde{\mathcal{G}}$$

for x in a sufficiently small neighborhood of $\pi(\mathcal{G})$. With a unit normal vector field ν on the parameterized interface $\Gamma := \varphi^{-1}(\pi(\mathcal{G}))$ and

$$g_1 := d \circ \Pi \circ \varphi, \qquad g_2 := \nabla g_1 \cdot \nu$$

the geometric curve constraints are equivalently expressed in terms of u by

$$u|_\Gamma = g_1, \qquad \partial_\nu u|_\Gamma = g_2.$$

As for point value constraints, such a formulation needs not exist in general, especially when π is not injective on $\mathcal{G}$. Therefore, well-definedness is again an assumption, which is plausible in the case of small deformations and compatible particle orientations.

Out of scope: Averaged curve constraints

Within the Monge-gauge setting, averaged curve constraints have been introduced in [EGH+16] as a simplifying approximation of classical curve constraints with a natural connection to so-called point curvature constraints. As such they also might play a future role in linking discrete–continuum curve constraint based models and fully continuum particle–membrane models. Besides their analytic properties, it is also conceivable that averaged constraints can be a direct consequence from physical modeling steps.

The curve constraints are averaged with respect to a reference manifold $\mathcal{M}_0$. Here they are immediately stated in their parameterized formulation in order to avoid additional notation. Using the previously established notation for curve constraints, the averaged curve constraints read

$$\int_\Gamma u(x)\,\mathrm{d}x = \int_\Gamma g_1(x)\,\mathrm{d}x, \qquad \int_\Gamma \partial_\nu u(x)\,\mathrm{d}x = \int_\Gamma g_2(x)\,\mathrm{d}x.$$

This type of constraint is different from the previous ones as it is not directly expressed in terms of a subset evaluation or the trace of u, but rather applies a further linear operator to the point evaluation or trace operator. The theory developed in the subsequent chapters applies only to pure evaluation and trace type constraints and therefore averaged constraints are excluded from the following model framework. However, as these constraints are at least closely related to the pure trace operator, they can be considered a prototypical reference model for future extensions.

1.2.3 Parametric particle configurations

Besides a description of the membrane energy and the membrane–particle coupling, a fully functional model for membrane–particle interactions also needs to describe the degrees of freedom associated to each particle. The arguably most intuitive choice for particle degrees of freedom is to enable each particle to take various positions and orientations within space. However, in some other relevant situations additional parameterizations of the particle state may come into play, for example by incorporating a parameter for a particle's size or for the angle at which the membrane connects to the particle.

In order to reflect the resulting flexibility in particle modeling, it makes sense to treat interaction models in a framework that is largely able to cover such possibilities as well. To this end, the general requirement in this thesis is that there exists a smooth *particle transformation function*

$$\Psi\colon \mathbb{R}^k \times \mathbb{R}^3 \to \mathbb{R}^3, \qquad (p, x) \mapsto \Psi(p; x)$$

such that every $\Psi(p; \cdot)$ is a bijection which maps a reference particle set $\mathcal{G}_0$ (or analogously a reference hypersurface $\widetilde{\mathcal{G}}_0$ for curve constraints) to a parameterized counterpart

$$\mathcal{G}(p) := \Psi(p; \mathcal{G}_0).$$

Here $p \in \mathbb{R}^k$ denotes the *particle state*. Correspondingly, the parameterization of particles also induces p-dependent constraints, which in addition depend on the chosen membrane–particle coupling. Within the parameterized setting for curve constraints this furthermore implies p-dependent reference domains $\Omega(p) \subseteq \Omega$ since "inside" areas of the particle are excluded from the computation.

The extension of this concept to a finite and fixed number of $N \in \mathbb{N}$ particles is straight-forward. In that context, a *particle configuration* $\mathbf{p} = (\mathbf{p}_i)_{i=1,\dots,N} \in (\mathbb{R}^k)^N$ is a collection of N particle states. In general, not all particle configurations are admissible, either due to physical restrictions or limits in the model. Therefore, feasible configurations are limited to a *configuration space* $\mathcal{D} \subseteq (\mathbb{R}^k)^N$. This leads to a $\mathbf{p}$-dependent set of constraints and a *parametric reference domain* $\Omega(\mathbf{p}) := \bigcap_{i=1}^N \Omega(\mathbf{p}_i)$, if applicable.

Example 1.2.1 (Variable positions and orientations)**.** *As mentioned above, a popular choice for the particle transformation function would allow particles to move and rotate freely in space. One way of modeling this is given by*

$$\begin{aligned} \Psi\colon \mathbb{R}^6 \times \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (p, x) &\longmapsto R_3(p_6)R_2(p_5)R_1(p_4)x + \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}. \end{aligned}$$

Here the R_i are the rotation matrices around the x_i-axes, i. e.

$$R_i(\alpha)x := (e_i \otimes e_i)x + \cos(\alpha)(e_i \times x) \times e_i + \sin(\alpha)(e_i \times x)$$

where e_i is the i-th unit vector. Hence, the parameters p_1, p_2 and p_3 determine a particle's translation in space, and p_4, p_5 and p_6 set the orientation. Note that the order of the rotation matrices matters and that this choice works best if the base state $\mathcal{G}_0$ is centered around the origin as much as possible.

1.3 Overall model and the interaction potential

Up to this point, various ways to model the membrane energy and membrane–particle interactions have been discussed, including variable particle configurations. In summary, this motivates a framework for *parameterized models for membrane-mediated particle interactions* where the following *model decisions* are required:

- A *reference surface* $\mathcal{M}_0$ together with a *parameterization* $\varphi\colon \Omega \to \mathcal{M}_0$ over a *parameterization domain* $\Omega \subseteq \mathbb{R}^2$.
- A *parameterized elastic energy* $J(\Omega)$. In the context of this thesis the energy is assumed to be of the form

$$J(\Omega; u) := \int_\Omega \rho(x, u(x), Du(x), D^2u(x))\,\mathrm{d}x \tag{1.3.1}$$

  with an integrand $\rho\colon \Omega \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$. Further assumptions on ρ are stated where necessary.
- A *reference space of admissible membranes* $H^2_*(\Omega) \subseteq H^2(\Omega)$ which incorporates particle-independent Dirichlet constraints or periodic boundary conditions.
- A *reference particle geometry* $\mathcal{G}_0$ (and possibly also $\widetilde{\mathcal{G}}_0$) in the sense of point value constraints in Section 1.2.1 or curve constraints in Section 1.2.2.
- A *particle transformation* $\Psi\colon \mathbb{R}^k \times \mathbb{R}^3 \to \mathbb{R}^3$ in the sense of Section 1.2.3.
- A fixed number $N \in \mathbb{N}$ of particles together with a *configuration space* $\mathcal{D} \subseteq \mathbb{R}^{N\times k}$.

The entirety of these decisions defines a particle–membrane interaction model. Given a *particle configuration* $\mathbf{p} \in \mathcal{D}$ they induce the following quantities:

- *Parametric particles* $\mathcal{G}(\mathbf{p}_i) := \Psi(\mathbf{p}_i, \mathcal{G}_0)$, and also $\widetilde{\mathcal{G}}(\mathbf{p}_i) := \Psi(\mathbf{p}_i, \widetilde{\mathcal{G}}_0)$ in the case of curve constraints.
- *Parametric interfaces* $\Gamma_i(\mathbf{p}_i) := \varphi^{-1}(\pi(\mathcal{G}(\mathbf{p}_i)))$ together with the joint interface $\Gamma(\mathbf{p}) := \bigcup_{i=1}^N \Gamma_i(\mathbf{p}_i)$. These implicitly rely on the well-definedness of the local projection and distance functions, see also Section 1.2. It is assumed that the $\Gamma_i(\mathbf{p}_i)$ are pairwise disjoint.
- In case of curve constraints the *parametric domain* $\Omega(\mathbf{p}) := \Omega \setminus \bigcup_{i=1}^N B_i(\mathbf{p}_i)$ where $B_i(\mathbf{p}_i)$ denotes the interior of the curve $\Gamma_i(\mathbf{p}_i)$. For simplicity of notation, let $\Omega(\mathbf{p}) := \Omega$ in the case of point value constraints.
- *Parametric trace operator* $T(\mathbf{p})$. In case of point value constraints this is given by the canonical point evaluation operator

$$T(\mathbf{p})\colon H^2(\Omega(\mathbf{p})) \to \mathbb{R}^{|\Gamma(\mathbf{p})|}, \quad u \mapsto (u(x))_{x\in\Gamma(\mathbf{p})}\,,$$

  which is well-defined in view of the Sobolev embedding $H^2(\Omega(\mathbf{p})) \to C(\Omega)$, see e. g. [Ada75, Theorem 5.4]. In case of curve constraints the operator corresponds to the classical trace operator

$$T(\mathbf{p})\colon H^2(\Omega(\mathbf{p})) \to H^{3/2}(\Gamma(\mathbf{p})) \times H^{1/2}(\Gamma(\mathbf{p})), \quad u \mapsto (u|_{\Gamma(\mathbf{p})}, \partial_\nu u|_{\Gamma(\mathbf{p})})$$

  where ν is the outer unit normal vector field on $\Gamma(\mathbf{p})$. See e. g. [Gri85, Chapter 1.5] for detailed definitions and results on well-definedness for sufficiently smooth interfaces $\Gamma(\mathbf{p})$.

– *Parametric constraints* $g_i(\mathbf{p})$. For point value constraints one has $g_i(\mathbf{p}) \in \mathbb{R}^{|\Gamma_i(\mathbf{p}_i)|}$ as defined in Section 1.2.1. For curve constraints the definition is given in Section 1.2.2 and regularity $g_i(\mathbf{p}) \in H^{3/2}(\Gamma_i(\mathbf{p}_i)) \times H^{1/2}(\Gamma_i(\mathbf{p}_i))$ is assumed. The associated *joint constraints* are defined by

$$g(\mathbf{p}) := (g_i(\mathbf{p}_i))_{i=1,\dots,N} \in \mathbb{R}^{|\Gamma(\mathbf{p})|}$$

or

$$g(\mathbf{p}; x) := g_i(\mathbf{p}; x) \quad \text{for } x \in \Gamma_i(\mathbf{p}_i),$$

respectively. These quantities rely on the local projection and distance functions.

– *Parametric set of admissible membranes* $U_{\mathrm{ad}}(\mathbf{p})$, defined by

$$U_{\mathrm{ad}}(\mathbf{p}) := \left\{ v \in H^2_*(\Omega(\mathbf{p})) \mid T(\mathbf{p}; v) = g(\mathbf{p}) \right\},$$

and the corresponding *homogeneous set of admissible membranes* $U^0_{\mathrm{ad}}(\mathbf{p})$, given by

$$U^0_{\mathrm{ad}}(\mathbf{p}) := \left\{ v \in H^2_*(\Omega(\mathbf{p})) \mid T(\mathbf{p}; v) = 0 \right\}.$$

Here $H^2_*(\Omega(\mathbf{p}))$ is to be understood as the image of the canonical restriction operator from Ω to $\Omega(\mathbf{p})$ over the space $H^2_*(\Omega)$.

– *Parametric elastic energy* $J(\mathbf{p})$, given by

$$J(\mathbf{p}, u) := J(\Omega(\mathbf{p}), u) = \int_{\Omega(\mathbf{p})} \rho(x, u(x), Du(x), D^2u(x)) \,\mathrm{d}x$$

for all $u \in H^2(\Omega(\mathbf{p}))$.

– For further notational convenience it is assumed that the parametric energy $J(\mathbf{p})$ has a unique minimizer over the set of admissible membranes $U_{\mathrm{ad}}(\mathbf{p})$. Correspondingly, the *(optimal) stationary membrane* $u(\mathbf{p})$ is defined by

$$u(\mathbf{p}) := \operatorname*{arg\,min}_{v \in U_{\mathrm{ad}}(\mathbf{p})} J(\mathbf{p}; v).$$

In the special case that $J(\mathbf{p})$ induces an elliptic linear operator, the unique existence of the stationary membrane need not be assumed and is a direct consequence of the well-known theorem of Lax–Milgram instead:

Theorem 1.3.1 (Unique optimal membrane shapes for quadratic elliptic energies)**.** *Let $\boldsymbol{p} \in \mathcal{D}$ and suppose a non-empty admissible set $U_{ad}(\boldsymbol{p})$. Suppose furthermore that the first variation $\delta J(\boldsymbol{p}; u)$ of $J(\boldsymbol{p}; \cdot)$ is linear in u, and that the second variation $\delta^2 J(\boldsymbol{p})$ is elliptic over $U^0_{ad}(\boldsymbol{p})$. Then $u(\boldsymbol{p})$ is well-defined.*

Putting everything together, the *interaction potential* $\mathcal{E}$ is defined as

$$\mathcal{E}\colon \mathcal{D} \to \mathbb{R}, \qquad \mathbf{p} \mapsto J(\mathbf{p}; u(\mathbf{p})) = \inf_{v \in U_{\mathrm{ad}}(\mathbf{p})} J(\mathbf{p}; v),$$

and an element $\mathbf{p}^* \in \mathcal{D}$ is called an *optimal configuration* if and only if

$$\mathbf{p}^* \in \operatorname*{arg\,min}_{\mathbf{p} \in \mathcal{D}} \mathcal{E}(\mathbf{p})$$

holds true. For the remainder of this thesis, existence of at least one optimal configuration is postulated. Note however that uniqueness need not hold and is even ruled out in many cases due to symmetries in the configuration space.

1.3.1 Further remarks

Choice of reference surface Typical choices of reference surfaces are minimizers of the elastic energy in the absence of particles, possibly subject to particle-independent constraints. In the presence of particles, it is particularly plausible that the optimal perturbation function for such a stationary reference surface fulfills the smallness assumptions required for parameterization, at least as long as the particle count remains reasonably bounded and as long as the particles themselves induce small deformations only. In some situations, also other reference surfaces can be appropriate choices, for example if the entirety of particles induces a strong deformation of the "natural" membrane state. Compare also [Hob16] for related viewpoints.

Neglecting the Gauss curvature term In view of the Gauss–Bonnet theorem (1.1.1) and the comments from the introductory part of the chapter, it is justified to drop the Gauss curvature term if the parametric particle interfaces $\widetilde{\mathcal{G}}(\mathbf{p})$ remain constant with respect to the parameter $\mathbf{p}$ up to orthonormal transformations. Hence, the negligibility in this case depends on the choice of the transformation function Ψ and is in particular given for spatial transformations as defined in Example 1.2.1. It is furthermore also required that the $\mathbf{p}$-independent constraints in $H^2_*(\Omega)$ keep the values and normals along the boundary $\partial\Omega$ either constant or make them subject to periodic boundary conditions. In more general cases the negligibility condition demands a closer inspection of the model.

Validity of linearizing Loosely speaking, minimization of the linearized elastic energy J_{lin} yields a good approximation of the nonlinear energy's minimizer u if the associated optimal surface $\mathcal{M}(u)$ is a "small" perturbation of $\mathcal{M}_0$. This in particular imposes a smallness assumption on the perturbation function u as well as its derivatives. A quantitative error analysis of the linearization error is possible by analyzing the corresponding error term of the Taylor expansion. While such an analysis could give valuable insight into the validity of linearized interaction models in practical applications, it exceeds the scope of this work and is therefore left open for further investigations.

Beyond pure membrane-mediated interactions Within the presented framework the interaction between particles is purely mediated by the membrane's elastic energy. This is readily extended by direct particle–particle interactions which are described by an additional potential $J_{\mathrm{p–p}}\colon \mathcal{D} \to \mathbb{R}$. To this end, the definition of the interaction potential is simply augmented to

$$\mathcal{E}(\mathbf{p}) := J(\mathbf{p}, u(\mathbf{p})) + J_{\mathrm{p–p}}(\mathbf{p}).$$

Once differentiability for the classical interaction potential is shown, the differentiability of the augmented potential becomes directly accessible through smoothness properties of $J_{\mathrm{p–p}}$.

In addition, the differentiability results from the subsequent chapters generalize to some extent to a larger class of energies J, which are not necessarily of the form (1.3.1) and contain so-called soft particle–membrane interactions, such as penalty curve constraints (cf. [GK17]). For the sake of simplicity, further details in this direction are omitted, but it is noted that corresponding results are obtainable by essentially the same proofs as given in this work.

Including further constraints In this thesis the $\mathbf{p}$-independent constraints are restricted to Dirichlet and periodic boundary conditions. In view of the remarks on the choice of the reference surface and with the development of more descriptive models in mind, it may also be relevant to incorporate further constraints such as for example volume and area constraints. Within the context of particle–membrane interactions corresponding techniques for such constraints have been applied successfully in [Hob16]. However, the extension of the differentiability results below to volume and area constrained membranes is not immediate and remains an open topic.

Spontaneous curvature as a source of irregularity Closer inspection of Examples 1.1.4 and 1.1.5 shows that – at least for these examples – the linear term in the quadratic energy J_{lin} contains derivatives of u of order one or higher if and only if the spontaneous curvature c_0 is not matched with the mean curvature H_0 of the reference surface $\mathcal{M}_0$. In view of standard elliptic regularity theory (cf. [Gri85, Chapter 7] and [BR80, GGS10]), this potentially implies a loss of regularity in the stationary membrane $u(\mathbf{p})$ unless c_0 and H_0 are sufficiently smooth. Such a loss of regularity would in particular have an impact on numerical approximation schemes, which either might suffer from inefficiency due to non-optimal rates of convergence or have to account for irregularities in c_0 and H_0 explicitly. A closer treatment of these aspects is omitted in this thesis and left to special applications.

Choice of configuration space Ideally, one would like to include all those particle configurations which do not yield intersecting particle interfaces. However, such a general choice is often not feasible within the parameterized model framework as it might lead to incompatibilities with the local projection and distance functions which are required

for the definition of the parametric interfaces. Hence, a purely pragmatic choice is to limit the configuration space $\mathcal{D}$ to a subset of

$$\mathcal{D}_{\max} := \left\{ \mathbf{p} \in \mathbb{R}^{N \times k} \mid \Gamma(\mathbf{p}) \text{ and } g(\mathbf{p}) \text{ are well-defined} \right\}.$$

In many cases it is reasonable to choose a strict and possibly closed subset $\mathcal{D} \subsetneq \mathcal{D}_{\max}$. Such a step is especially justified if the interaction potential fulfills $\mathcal{E}(\mathbf{p}) \to \infty$ for $\mathbf{p} \to \partial\mathcal{D}_{\max}$, but it is also more generally valid because the existence of an optimal particle configuration on the boundary $\partial\mathcal{D}_{\max}$ would indicate a wrong model choice at some point of the framework.

2 Differentiation of the interaction potential

The differentiability of the interaction potential is crucial for a wide range of mathematical methods related to its minimization. It is tempting to argue for the differentiability of the interaction potential via the chain rule

$$\frac{\partial}{\partial \mathbf{p}} \mathcal{E}(\mathbf{p}) = \frac{\partial}{\partial \mathbf{p}} J(\mathbf{p}, u(\mathbf{p})) = J_{\mathbf{p}}(\mathbf{p}, u(\mathbf{p})) + J_u(\mathbf{p}, u(\mathbf{p})) \cdot \frac{\partial}{\partial \mathbf{p}} u(\mathbf{p}) = J_{\mathbf{p}}(\mathbf{p}, u(\mathbf{p}))$$

where the last step uses the fact that $J_u(\mathbf{p}, u(\mathbf{p})) = 0$ holds as $u(\mathbf{p})$ is a critical point of $J(\mathbf{p}, \cdot)$. However, at this point that computation only makes sense formally. This is because it implicitly relies on the existence of the derivatives $J_{\mathbf{p}}$, J_u and $\frac{\partial}{\partial \mathbf{p}} u(\mathbf{p})$, which is not yet known at all. While generally the definition and existence of the derivative J_u is relatively clear, it is less obvious what the meaning of the expressions $J_{\mathbf{p}}$ and $\frac{\partial}{\partial \mathbf{p}} u(\mathbf{p})$ is as the membrane functions $u(\mathbf{p})$ originate from different function spaces $U_{\mathrm{ad}}(\mathbf{p})$ with different domains $\Omega(\mathbf{p})$. Another potential and more practical issue is about the feasibility of the resulting expression in numerical schemes. For example, in [EGH$^+$16] a formal differentiation approach for the interaction potential is taken based on the Reynolds transport theorem [Rey03]. As a result, an expression for the gradient is derived there which consists out of boundary integrals over $\partial\Omega(\mathbf{p})$ containing derivatives of $u(\mathbf{p})$ up to third order. This is problematic in numerical schemes, because even if $u(\mathbf{p})$ is smooth, numerical approximation schemes for $u(\mathbf{p})$ are in general not able to provide error estimates for derivatives of higher than second order. As a consequence, it would not be possible to obtain suitable error estimates for approximations of the gradient of the interaction potential.

Therefore, this chapter has two objectives: The first is to establish a mathematically sound differentiability result for a sufficiently large class of interaction models, and the second one is to derive an expression for the gradient which is accessible to error estimates within the framework of finite element methods. These goals are accomplished by the following steps: In the first section the admissible sets $U_{\mathrm{ad}}(\mathbf{p})$ are related to each other via so-called trace preserving diffeomorphisms between the domains $\Omega(\mathbf{p})$. Such diffeomorphisms are constructed rather explicitly by a domain perturbation ansatz from shape calculus (see [SZ92], [DZ11]), which sets the stage for a differentiability concept similar to material derivatives in surface calculus. Afterwards, the second section transforms the elastic energies and their minimization problems onto a common reference domain and space. Then the finite dimensionality of the state space is exploited in order to reduce the matter of differentiability to an application of the implicit function theorem. Finally, the third section derives a numerically feasible representation of

the gradient in the above sense by means of known methods from matrix calculus. In this context also explicit derivative formulas for various relevant interaction models are stated.

Recall the definition of the admissible sets as

$$U_{\mathrm{ad}}(\mathbf{p}) := \left\{ u \in H^2_*(\Omega(\mathbf{p})) \mid T(\mathbf{p}; u) = g(\mathbf{p}) \right\}$$

where $H^2_*(\Omega(\mathbf{p})) \subseteq H^2(\Omega(\mathbf{p}))$ is a linear subspace encoding $\mathbf{p}$-independent constraints, $T(\mathbf{p})$ is either a point evaluation or trace operator over the parametric interface $\Gamma(\mathbf{p})$, and where $g(\mathbf{p}) \in \mathrm{range}(T(\mathbf{p}))$ is the constraint right hand side. For the sake of a simple presentation, it is assumed in the following sections that $H^2_*(\Omega) = H^2_0(\Omega)$ holds, i. e. that the $\mathbf{p}$-independent constraints are of Dirichlet-type and span the full boundary. The mild modifications that are necessary when $H^2_*(\Omega)$ incorporates periodic boundary conditions or spans only a subset of the boundary are discussed in Remark 2.3.4.

Concerning notation, throughout the whole chapter $\mathbf{p}^0 \in \mathcal{D}^\circ$ is a point in the interior of the configuration space $\mathcal{D} \subseteq \mathbb{R}^{N\times k}$ and $\mathcal{B}$ is a generic (i. e. non-fixed) convex neighborhood of $\mathbf{p}^0$ in $\mathcal{D}^\circ$. Like in the introductory formal calculation, indices of functions indicate partial derivatives, e. g. given a function $f(\mathbf{p}, x)$ one defines $f_{\mathbf{p}} := \frac{\partial}{\partial \mathbf{p}} f$. The differentiation operator D is always meant with respect to spatial variables only, *i. e.* $D := \frac{\partial}{\partial x}$. In addition, $\mathrm{Sym}(A) := A + A^T$ is defined for all matrices $A \in \mathbb{R}^{n\times n}$, and for every vector field $V \colon \mathbb{R}^2 \to \mathbb{R}^2$ and every vector $y \in \mathbb{R}^2$ the application of the Hessian is defined as $D^2V \cdot y := \sum_{k=1}^2 D^2V_k y_k$. The homogeneous admissible set is defined as

$$U^0_{\mathrm{ad}}(\mathbf{p}) := \left\{ u \in H^2_*(\Omega(\mathbf{p})) \mid T(\mathbf{p}; u) = 0 \right\}.$$

Lastly and in a slight abuse of notation, whenever a $\mathbf{p}$-dependent function $f(\mathbf{p}; x)$ is inverted, the inversion is meant for fixed $\mathbf{p}$, i. e. $f^{-1}(\mathbf{p}; x) := f(\mathbf{p}; \cdot)^{-1}(x)$.

2.1 Trace and constraint preserving maps

The following paragraphs are concerned with linking the spaces $U_{\mathrm{ad}}(\mathbf{p})$ to each other through so-called trace preserving diffeomorphisms between the domains $\Omega(\mathbf{p})$ as well as so-called constraint preserving maps.

Theorem 2.1.1. *Let $\Omega_1, \Omega_2 \subseteq \mathbb{R}^2$ and suppose a k-diffeomorphism $X \colon \Omega_1 \to \Omega_2$ where $k \in \mathbb{N}$. Then the map*

$$\begin{aligned} H^k(\Omega_1) &\longrightarrow H^k(\Omega_2) \\ u &\longmapsto u \circ X^{-1} \end{aligned}$$

is well-defined and an isomorphism.

Proof. This statement is proven in [Ada75, Theorem 3.35]. □

Lemma 2.1.2. *Suppose $p \in \mathcal{D}$ and that $X\colon \Omega(\boldsymbol{p}^0) \to \Omega(\boldsymbol{p})$ is a 2-diffeomorphism with the* trace preserving property

$$\forall u \in H^2(\Omega(\boldsymbol{p}^0))\colon\ T(\boldsymbol{p}; u \circ X^{-1}) = T(\boldsymbol{p}^0; u) \circ X^{-1}|_{\Gamma(\boldsymbol{p})} \tag{2.1.1}$$

and with Dirichlet boundary conditions

$$X|_{\partial\Omega} = \mathrm{id}_{\partial\Omega}, \qquad DX\nu|_{\partial\Omega} = \nu. \tag{2.1.2}$$

Furthermore, let $\xi \in H^2(\Omega(\boldsymbol{p}^0))$ with the constraint preserving property

$$T(\boldsymbol{p}^0; \xi) = g(\boldsymbol{p}) \circ X|_{\Gamma(\boldsymbol{p}^0)} - g(\boldsymbol{p}^0) \tag{2.1.3}$$

and with homogeneous Dirichlet boundary conditions

$$\xi|_{\partial\Omega} = 0, \qquad \partial_\nu \xi|_{\partial\Omega} = 0. \tag{2.1.4}$$

Then

$$\begin{aligned} \phi\colon U_{ad}(\boldsymbol{p}^0) &\longrightarrow U_{ad}(\boldsymbol{p}) \\ u &\longmapsto (u + \xi) \circ X^{-1} \end{aligned}$$

is a bijection.

Proof. Well-definedness of ϕ as a map $U_{\mathrm{ad}}(\mathbf{p}) \to H^2_*(\Omega(\mathbf{p}))$ is given by Theorem 2.1.1 and the Dirichlet condition (2.1.4). The linearity of T, the trace preserving property of X, and the constraint preserving property of ξ yield for all $u \in U_{\mathrm{ad}}(\mathbf{p}^0)$

$$\begin{aligned} T(\mathbf{p}; \phi(u)) &= T(\mathbf{p}; u \circ X^{-1}) + T(\mathbf{p}; \xi \circ X^{-1}) \\ &= T(\mathbf{p}^0; u) \circ X^{-1}|_{\Gamma(\mathbf{p})} + T(\mathbf{p}^0; \xi) \circ X^{-1}|_{\Gamma(\mathbf{p})} \\ &= g(\mathbf{p}^0) \circ X^{-1}|_{\Gamma(\mathbf{p})} + g(\mathbf{p}) - g(\mathbf{p}^0) \circ X^{-1}|_{\Gamma(\mathbf{p})} \\ &= g(\mathbf{p}). \end{aligned}$$

Here also the definition of $U_{\mathrm{ad}}(\mathbf{p}^0)$ and the bijectivity of X entered. Similarly, from the Dirichlet conditions (2.1.2) and (2.1.4) also the outer boundary conditions

$$(u + \xi) \circ X^{-1}|_{\partial\Omega} = (u + \xi)|_{\partial\Omega} = 0$$

and

$$\partial_\nu \left((u + \xi) \circ X^{-1}\right)|_{\partial\Omega} = \nabla(u + \xi) \cdot \left((DX)^{-1} \circ X^{-1}\right) \nu|_{\partial\Omega} = \partial_\nu u|_{\partial\Omega} + \partial_\nu \xi|_{\partial\Omega} = 0$$

are verified. Therefore, $\phi(u) \in U_{\mathrm{ad}}(\mathbf{p})$ is true. Since Theorem 2.1.1 implies injectivity of ϕ, it only remains to prove surjectivity. To this end, let $\widetilde{u} \in U_{\mathrm{ad}}(\mathbf{p})$ and define the function $u := \widetilde{u} \circ X - \xi$. Analogously to before, and again by Theorem 2.1.1 and the properties (2.1.1), (2.1.3) and (2.1.4) it is observed that $u \in H^2_0(\Omega(\mathbf{p}^0))$ holds as well as

$$T(\mathbf{p}^0; u) = T(\mathbf{p}; \widetilde{u}) \circ X|_{\Gamma(\mathbf{p}^0)} - T(\mathbf{p}^0; \xi) = g(\mathbf{p}^0),$$

implying the desired surjectivity of ϕ. □

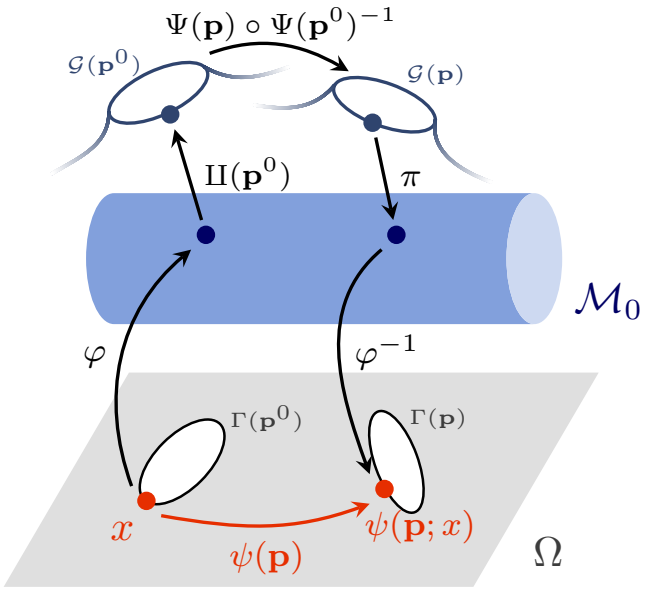

Figure 2.1.1: The red arrow shows the action of the map $\psi(\mathbf{p}; \cdot)$ on an interface point x. It is defined by the concatenation of the geometric operations indicated by the black arrows.

In the next step, a family $\mathcal{X}$ of trace preserving diffeomorphisms $\mathcal{X}(\mathbf{p})\colon \Omega(\mathbf{p}^0) \to \Omega(\mathbf{p})$ is constructed over a generic neighborhood $\mathcal{B}$ of $\mathbf{p}^0$. The construction makes use of the fact that the domains $\Omega(\mathbf{p})$ are determined by their interfaces $\Gamma(\mathbf{p})$ and that those interfaces are related to each other through a smooth family $\psi(\mathbf{p})\colon \Gamma(\mathbf{p}^0) \to \Gamma(\mathbf{p})$ of bijections. These maps $\psi(\mathbf{p})$ are induced by the geometric properties of the model and rely on the parameterization map φ, the transformation map Ψ from the definition of the parameterized particle manifolds $\mathcal{G}(\mathbf{p}_i) := \Psi(\mathbf{p}_i; \mathcal{G}_0)$, the local projection π onto the reference manifold $\mathcal{M}_0$ and its local inverse $\Pi(\mathbf{p}_i)$ with respect to $\mathcal{G}(\mathbf{p}_i)$. For $x \in \Gamma_i(\mathbf{p}_i^0)$ one then defines

$$\psi(\mathbf{p}, x) := \varphi^{-1}\left(\pi\left(\Psi\left(\mathbf{p}_i; \Psi^{-1}\left(\mathbf{p}_i^0; \Pi\left(\mathbf{p}_i^0; \varphi(x)\right)\right)\right)\right)\right).$$

In words, any point x on an interface of $\Gamma(\mathbf{p}_i^0)$ is mapped back onto the reference manifold $\mathcal{M}_0$ by φ and then mapped back onto the parametric particle manifold $\mathcal{G}(\mathbf{p}_i^0)$ through the local inverse projection Π. Then the transformation of $\Psi(\mathbf{p}_i^0)$ is undone and replaced by the transformation $\Psi(\mathbf{p}_i)$ in order to obtain a point of $\mathcal{G}(\mathbf{p}_i)$, which then is projected onto the reference manifold $\mathcal{M}_0$ and finally mapped back onto the parameterization domain Ω by φ^{-1}. A geometric interpretation of ψ is depicted in Fig. 2.1.1.

By definition, ψ admits a canonical extension to a sufficiently small neighborhood $\widetilde{\Gamma}_0$ of $\Gamma(\mathbf{p}^0)$. Since the smoothness of the involved manifolds and transformations is directly

inherited by ψ, it is in many cases justified to make generous smoothness assumptions on ψ over $\widetilde{\Gamma}_0$.

Proposition 2.1.3. *Let $m \in \mathbb{N}_{\geq 2}$ and suppose $\psi \in C^{m+1}(\mathcal{B} \times \widetilde{\Gamma}_0, \mathbb{R}^2)$. Furthermore, let $\mathcal{V} \in C^m([0,1] \times \mathcal{B} \times \mathbb{R}^2, \mathbb{R}^2)$ be a time- and configuration-dependent vector field such that the following properties are fulfilled:*

- *All partial derivatives of $\mathcal{V}$ are bounded.*
- *For all $(t, x) \in [0,1] \times \mathbb{R}^2$ holds $\mathcal{V}(t, \boldsymbol{p}^0, x) = 0$*
- *For all $(t, \boldsymbol{p}) \in [0,1] \times \mathcal{B}$ holds*

$$\mathcal{V}(t, \boldsymbol{p}, \cdot)|_{\Gamma(\boldsymbol{p})} = \psi_{\boldsymbol{p}}\left(\boldsymbol{p}(t), \psi^{-1}(\boldsymbol{p}(t), \cdot)\right)(\boldsymbol{p} - \boldsymbol{p}^0) \tag{2.1.5}$$

  where $\boldsymbol{p}(t) := \boldsymbol{p}^0 + t(\boldsymbol{p} - \boldsymbol{p}^0)$.

- *If T denotes the classical trace operator for curves and not the point evaluation operator, then for all $(t, \boldsymbol{p}) \in [0,1] \times \mathcal{B}$ also holds*

$$D\mathcal{V}(t, \boldsymbol{p}, \psi(\boldsymbol{p}(t), \cdot))|_{\Gamma(\boldsymbol{p}^0)} \cdot \nu(t) = \nu'(t) \tag{2.1.6}$$

  where $\nu(t) := \nu|_{\Gamma(\boldsymbol{p}(t))} \circ \psi(\boldsymbol{p}(t))$.

- *The Dirichlet boundary conditions*

$$\begin{aligned} \mathcal{V}(t, \boldsymbol{p}, \cdot)|_{\partial\Omega} &= 0 \\ D\mathcal{V}(t, \boldsymbol{p}, \psi(\boldsymbol{p}(t), \cdot))\nu|_{\partial\Omega} &= 0 \end{aligned} \tag{2.1.7}$$

  hold.

Define $\mathcal{X}(\boldsymbol{p}, x) := \eta(1, \boldsymbol{p}, x)$ where $\eta(\cdot, \boldsymbol{p}, x)$ solves the ordinary differential equation

$$\begin{aligned} \frac{\partial}{\partial t}\eta(t, \boldsymbol{p}, x) &= \mathcal{V}\left(t, \boldsymbol{p}, \eta(t, \boldsymbol{p}, x)\right) \\ \eta(0, \boldsymbol{p}, x) &= x. \end{aligned} \tag{2.1.8}$$

Then $\mathcal{X} \in C^m(\mathcal{B} \times \mathbb{R}^2, \mathbb{R}^2)$, and $\mathcal{X}(\boldsymbol{p}^0, \cdot) = \mathrm{id}_{\mathbb{R}^2}$, and for all $\boldsymbol{p} \in \mathcal{B}$ the restriction $\mathcal{X}(\boldsymbol{p})|_{\Omega(\boldsymbol{p}^0)}$ is an m-diffeomorphism onto $\Omega(\boldsymbol{p})$ that fulfills the trace preserving property (2.1.1) *and the Dirichlet condition* (2.1.2).

Proof. The solution function η exists and is unique in view of the Picard-Lindelöf theorem C.1 since the bounded first derivatives of $\mathcal{V}$ imply its Lipschitz-continuity. The property $\mathcal{X}(\mathbf{p}^0, \cdot) = \mathrm{id}_{\mathbb{R}^2}$ is immediate because $\mathcal{V}(\cdot, \mathbf{p}^0, \cdot) \equiv 0$ holds and therefore the solution of the ODE (2.1.8) in $(\mathbf{p}^0, x)$ is given by $\eta(t, \mathbf{p}^0, x) := x$. The smoothness $\mathcal{X} \in C^m(\mathcal{B} \times \mathbb{R}^2, \mathbb{R}^2)$ is a consequence of a standard result of ODE theory, see also Theorem C.2 in the appendix. Note that, analogously, by considering the inverse ODE of (2.1.8) the existence of $\mathcal{X}(\mathbf{p})^{-1} \in C^m(\mathbb{R}^2, \mathbb{R}^2)$ is asserted.

From $\mathcal{V}$'s boundary condition (2.1.5) and the uniqueness of η one infers for every $x \in \Gamma_0$ that $\eta(t, \mathbf{p}, x) = \psi(\mathbf{p}(t), x)$ holds. This is because of $\psi(\mathbf{p}^0, x) = x$ and

$$\begin{aligned}\frac{\partial}{\partial t}\psi(\mathbf{p}(t), x) &= \psi_{\mathbf{p}}(\mathbf{p}(t), x)(\mathbf{p} - \mathbf{p}^0) \\ &= \psi_{\mathbf{p}}\left(\mathbf{p}(t), \psi^{-1}(\mathbf{p}(t), \psi(\mathbf{p}(t), x))\right)(\mathbf{p} - \mathbf{p}^0) \\ &= \mathcal{V}(t, \mathbf{p}, \psi(\mathbf{p}(t), x)).\end{aligned}$$

In particular, $\mathcal{X}(\mathbf{p})|_{\Gamma(\mathbf{p}^0)} = \psi(\mathbf{p})$, which further implies that $\mathcal{X}(\mathbf{p})|_{\Gamma(\mathbf{p}^0)}$ is a bijection onto the interface $\Gamma(\mathbf{p})$. Analogously, the outer boundary condition (2.1.7) implies the identity $\mathcal{X}(\mathbf{p})|_{\partial\Omega} = \mathrm{id}_{\partial\Omega}$, and therefore $\mathcal{X}|_{\partial\Omega(\mathbf{p}^0)}$ is bijective onto $\partial\Omega(\mathbf{p})$. In consequence, the restriction $\mathcal{X}(\mathbf{p})|_{\Omega(\mathbf{p}^0)}$ is an m-diffeomorphism onto $\Omega(\mathbf{p})$.

If T is the point evaluation operator, then the trace preserving property (2.1.1) is immediate from the bijectivity of $\mathcal{X}(\mathbf{p})|_{\Gamma(\mathbf{p}^0)}$ onto $\Gamma(\mathbf{p})$ because

$$\begin{aligned}T(\mathbf{p}; u \circ \mathcal{X}(\mathbf{p})^{-1}) &= u \circ \mathcal{X}(\mathbf{p})^{-1}|_{\Gamma(\mathbf{p})} \\ &= u|_{\Gamma(\mathbf{p}^0)} \circ \mathcal{X}(\mathbf{p})^{-1}|_{\Gamma(\mathbf{p})} \\ &= T(\mathbf{p}^0; u) \circ \mathcal{X}(\mathbf{p})^{-1}|_{\Gamma(\mathbf{p})}.\end{aligned} \tag{2.1.9}$$

If T is the curve trace operator, then (2.1.9) analogously holds for the first component of T, but it remains to show that also the normal derivatives are preserved.

To this end, note that by another basic result of ODE theory, the derivative of $\mathcal{X}$ is again described by the solution of an ODE (see Theorem C.3 in the appendix). It holds $D\mathcal{X}(\mathbf{p}, x) = \widetilde{\eta}(1, \mathbf{p}, x)$ where $\widetilde{\eta}$ solves the initial value problem

$$\begin{aligned}\frac{\partial}{\partial t}\widetilde{\eta}(t, \mathbf{p}, x) &= D\mathcal{V}(t, \mathbf{p}, \eta(t, \mathbf{p}, x))\widetilde{\eta}(t, \mathbf{p}, x) \\ \widetilde{\eta}(0, \mathbf{p}, x) &= \mathrm{id}_{\mathbb{R}^2}.\end{aligned}$$

For $x \in \Gamma(\mathbf{p}^0)$ a solution candidate is given by

$$y(t, \mathbf{p}, x) := D\psi(\mathbf{p}(t), x)\tau \otimes \tau + \nu(t) \otimes \nu^0, \qquad \mathbf{p}(t) = \mathbf{p}^0 + t(\mathbf{p} - \mathbf{p}^0)$$

where $\nu^0 := \nu|_{\Gamma(\mathbf{p}^0)}$ is the outer unit normal field on $\Gamma(\mathbf{p}^0)$ and $\tau = (-\nu_2^0, \nu_1^0)^T$ is a unit tangent field on $\Gamma(\mathbf{p}^0)$. The solution property is verified in multiple steps: Firstly, because of $\psi(\mathbf{p}^0) = \mathrm{id}\,|_{\Gamma(\mathbf{p}^0)}$ and $\nu(0) = \nu^0$ it holds

$$y(0, \mathbf{p}, x) = \tau \otimes \tau + \nu^0 \otimes \nu^0 = \mathrm{id}_{\mathbb{R}^2}$$

and so the initial condition is met for the candidate $y(\cdot, \mathbf{p}, x)$. Moreover, the boundary condition (2.1.5) for $\mathcal{V}$ gives the identity

$$\mathcal{V}\left(t, \mathbf{p}, \psi(\mathbf{p}(t), \cdot)\right)|_{\Gamma(\mathbf{p}^0)} = \psi_{\mathbf{p}}(\mathbf{p}(t), \cdot)(\mathbf{p} - \mathbf{p}^0),$$

which again implies on $\Gamma(\mathbf{p}^0)$ by taking the tangential derivative

$$D\mathcal{V}\left(t, \mathbf{p}, \psi(\mathbf{p}(t), \cdot)\right) D\psi(\mathbf{p}(t), \cdot)\tau|_{\Gamma(\mathbf{p}^0)} = D\psi_{\mathbf{p}}(\mathbf{p}(t), \cdot)(\mathbf{p} - \mathbf{p}^0). \tag{2.1.10}$$

Combining the identity (2.1.10) with the boundary condition (2.1.6) for $D\mathcal{V}$ leads together with $\eta(t,\mathbf{p},\cdot)|_{\Gamma(\mathbf{p}^0)} = \psi(\mathbf{p}(t),\cdot)$ in the right hand side of the initial value problem to

$$\begin{aligned} D\mathcal{V}\left(t,\mathbf{p},\eta(t,\mathbf{p},x)\right) y(t,\mathbf{p},x) &= D\mathcal{V}\left(t,\mathbf{p},\psi(\mathbf{p}(t))\right)\left(D\psi(\mathbf{p}(t),x)\tau\otimes\tau + \nu(t)\otimes\nu^0\right) \\ &= D\psi_{\mathbf{p}}(\mathbf{p}(t),x)(\mathbf{p}-\mathbf{p}^0)\tau\otimes\tau + \nu'(t)\otimes\nu^0. \end{aligned}$$

On the other hand, differentiation of y in t yields

$$\frac{\partial}{\partial t} y(t,\mathbf{p},x) = D\psi_{\mathbf{p}}(\mathbf{p}(t),x)(\mathbf{p}-\mathbf{p}^0)\tau\otimes\tau + \nu'(t)\otimes\nu^0,$$

which indeed shows that $\widetilde{\eta}(t,\mathbf{p},x) = y(t,\mathbf{p},x)$ is the solution on $\Gamma(\mathbf{p}^0)$. In particular,

$$D\mathcal{X}(\mathbf{p})\nu|_{\Gamma(\mathbf{p}^0)} = \nu|_{\Gamma(\mathbf{p})}\circ\psi(\mathbf{p}) = \nu|_{\Gamma(\mathbf{p})}\circ\mathcal{X}(\mathbf{p})|_{\Gamma(\mathbf{p}^0)},$$

or equivalently,

$$\left((D\mathcal{X}(\mathbf{p}))^{-1}\circ\mathcal{X}(\mathbf{p})^{-1}\right)\nu|_{\Gamma(\mathbf{p})} = \nu|_{\Gamma(\mathbf{p}^0)}\circ\mathcal{X}(\mathbf{p})^{-1}|_{\Gamma(\mathbf{p})}$$

holds. This leads to

$$\begin{aligned} T_2(\mathbf{p}; u\circ\mathcal{X}(\mathbf{p})^{-1}) &= \partial_\nu(u\circ\mathcal{X}(\mathbf{p})^{-1})|_{\Gamma(\mathbf{p})} \\ &= \left((D\mathcal{X}(\mathbf{p}))^{-T}\circ\mathcal{X}(\mathbf{p})^{-1}\cdot\nabla u\right)\cdot\nu\big|_{\Gamma(\mathbf{p})} \\ &= \nabla u\cdot((D\mathcal{X}(\mathbf{p}))^{-1}\circ\mathcal{X}(\mathbf{p})^{-1})\nu\big|_{\Gamma(\mathbf{p})} \\ &= \partial_\nu u\circ\mathcal{X}(\mathbf{p})^{-1}|_{\Gamma(\mathbf{p})}, \end{aligned}$$

which confirms the trace preserving property (2.1.1) also for the curve trace operator.

Analogously, the Dirichlet conditions (2.1.7) of $\mathcal{V}$ imply the Dirichlet conditions (2.1.2) of $\mathcal{X}$. This completes the proof. □

The later practical derivative formulations do not require knowledge of the full family of diffeomorphisms $\mathcal{X}(\mathbf{p})$, but instead only directional derivatives $\partial_{\mathbf{q}}\mathcal{X}(\mathbf{p}^0)$ are needed. The next two results show how a single time- and $\mathbf{p}$-independent vector field may be used to induce a family of maps $\mathcal{X}(\mathbf{p})$ where the derivative $\partial_{\mathbf{q}}\mathcal{X}(\mathbf{p}^0)$ is known for one prescribed direction $\mathbf{q}$.

Lemma 2.1.4. *Suppose $\mathcal{X}$ and $\mathcal{V}$ as in Proposition 2.1.3. Let $\boldsymbol{q}\in\mathcal{B}-\boldsymbol{p}^0$ and $\boldsymbol{p} := \boldsymbol{p}^0+\boldsymbol{q}$. If $\mathcal{V}$ fulfills the scaling condition*

$$\forall t,\lambda\in[0,1]\colon\ \mathcal{V}(t,\boldsymbol{p}^0+\lambda\boldsymbol{q},\cdot) = \lambda\mathcal{V}(\lambda t,\boldsymbol{p}^0+\boldsymbol{q},\cdot), \tag{2.1.11}$$

then the derivative of $\mathcal{X}$ in $\boldsymbol{p}^0$ in direction $\boldsymbol{q}$ is given by

$$\partial_{\boldsymbol{q}}\mathcal{X}(\boldsymbol{p}^0) = \mathcal{V}(0,\boldsymbol{p},\cdot).$$

Proof. Again, basic ODE theory may be used to characterize $\partial_{\mathbf{q}}\mathcal{X}$ as the solution of another ODE (see Theorem C.4). It holds $\partial_{\mathbf{q}}\mathcal{X}(\mathbf{p}^0;x) = \widetilde{\eta}(1,x)$ where $\widetilde{\eta}$ solves the initial value problem

$$\begin{aligned}\frac{\partial}{\partial t}\widetilde{\eta}(t,x) &= D\mathcal{V}\left(t,\mathbf{p}^0,\eta(t,\mathbf{p}^0,x)\right)\cdot\widetilde{\eta}(t,x) + \frac{\partial}{\partial \mathbf{p}}\mathcal{V}\left(t,\mathbf{p}^0,\eta(t,\mathbf{p}^0,x)\right)\cdot\mathbf{q}\\ &= \partial_{\mathbf{q}}\mathcal{V}(t,\mathbf{p}^0,x)\end{aligned}$$

with $\widetilde{\eta}(0,x)=0$. Here the property $\mathcal{V}(\cdot,\mathbf{p}^0,\cdot)=0$, implying $D\mathcal{V}(t,\mathbf{p}^0,\cdot)=0$, as well as $\eta(t,\mathbf{p}^0,x)=x$ were used. The scaling condition (2.1.11) and $\mathcal{V}(\cdot,\mathbf{p}^0,\cdot)=0$ imply

$$\partial_{\mathbf{q}}\mathcal{V}(t,\mathbf{p}^0,x) = \lim_{\lambda\searrow 0}\frac{\mathcal{V}(t,\mathbf{p}^0+\lambda\mathbf{q},x)-\mathcal{V}(t,\mathbf{p}^0,x)}{\lambda} = \lim_{\lambda\searrow 0}\mathcal{V}(\lambda t,\mathbf{p}^0+\mathbf{q},x) = \mathcal{V}(0,\mathbf{p},x)$$

and so one arrives at

$$\partial_{\mathbf{q}}\mathcal{X}(\mathbf{p}^0,x) = \widetilde{\eta}(1,x) = \int_0^1 \partial_{\mathbf{q}}\mathcal{V}(t,\mathbf{p}^0,x)\,\mathrm{d}t = \mathcal{V}(0,\mathbf{p},x). \qquad \square$$

Lemma 2.1.5. *Let $V\in C^m(\Omega)$ and $\boldsymbol{q}\in\mathcal{B}-\boldsymbol{p}^0$. Assume that the boundary condition*

$$V|_{\Gamma(\boldsymbol{p}^0)} = \partial_{\boldsymbol{q}}\psi(\boldsymbol{p}^0) \tag{2.1.12}$$

holds and, if T is the curve trace operator, also

$$DV\nu|_{\Gamma(\boldsymbol{p}^0)} = \partial_{\boldsymbol{q}}(\nu|_{\Gamma(\boldsymbol{p})}\circ\psi(\boldsymbol{p})). \tag{2.1.13}$$

Assume that the right hand sides in (2.1.12) *and* (2.1.13) *are well-defined and admit local C^m-extensions to a neighborhood of $\{\boldsymbol{p}^0\}\times\Gamma(\boldsymbol{p}^0)$. Moreover, suppose on the outer boundary the Dirichlet boundary condition*

$$V|_{\partial\Omega}=0, \qquad DV|_{\partial\Omega}=0.$$

Then a family of trace preserving diffeomorphisms $(\mathcal{X}(\boldsymbol{p}))_{\boldsymbol{p}\in\mathcal{B}}$ exists in the sense of Proposition 2.1.3 with $\partial_{\boldsymbol{q}}\mathcal{X}(\boldsymbol{p}^0)=V$.

Proof. Extend V to a C^m-smooth vector field $\mathcal{V}(t,\mathbf{p},x)$ such that $\mathcal{V}(0,\mathbf{p}^0,\cdot)=V$ holds and the conditions (2.1.5) through (2.1.7) and (2.1.11) are fulfilled. The statement then is a direct consequence of Proposition 2.1.3 and Lemma 2.1.4. $\square$

Similarly, a single time- and $\mathbf{p}$-independent function can be used to induce a family of constraint preserving maps $\xi(\mathbf{p})$ with known directional derivative $\partial_{\mathbf{q}}\xi(\mathbf{p}^0)$. In this case the construction is even simpler because no moving interfaces need to be taken into account directly.

Lemma 2.1.6. *Let $\boldsymbol{q}\in\mathcal{B}-\boldsymbol{p}^0$. Suppose $g\circ\psi\in C^m$ in a neighborhood of $\{\boldsymbol{p}^0\}\times\Gamma(\boldsymbol{p}^0)$ and $w\in C^{m-1}(\overline{\Omega})$ with*

$$T(\boldsymbol{p}^0;w) = \partial_{\boldsymbol{q}}\left(g(\boldsymbol{p}^0)\circ\psi(\boldsymbol{p}^0)\right).$$

Moreover, assume on the outer boundary Dirichlet conditions

$$w|_{\partial\Omega} = 0, \qquad \partial_\nu w|_{\partial\Omega} = 0.$$

Then $\xi \in C^m(\mathcal{B} \times \Omega)$ *exists such that* $\xi(\boldsymbol{p}^0) = 0$, $\partial_{\boldsymbol{q}}\xi(\boldsymbol{p}^0) = w$,

$$\xi(\boldsymbol{p})|_{\partial\Omega} = 0, \qquad \partial_\nu \xi(\boldsymbol{p})|_{\partial\Omega} = 0,$$

and for all $\boldsymbol{p} \in \mathcal{B}$ *the constraint-preserving property*

$$T(\boldsymbol{p}^0; \xi) = g(\boldsymbol{p}) \circ \psi(\boldsymbol{p}) - g(\boldsymbol{p}^0)$$

is fulfilled.

Proof. The existence of an appropriate map $\xi \in C^m(\mathcal{B} \times \Omega)$ follows immediately by application of an extension theorem to the following boundary values:

- For all $x \in \Omega$ holds $\xi(\mathbf{p}^0, x) = 0$ and $\partial_{\mathbf{q}}\xi(\mathbf{p}^0, x) = w(x)$.
- For all $(\mathbf{p}, x) \in [0,1] \times \mathcal{B} \times \partial\Omega$ holds $\xi(\mathbf{p}, x) = \partial_\nu \xi(\mathbf{p}, x) = 0$.
- For all $(\mathbf{p}, x) \in [0,1] \times \mathcal{B} \times \Gamma(\mathbf{p}^0)$ holds

  $$\xi(\mathbf{p}; x) = g_1(\mathbf{p}; \psi(\mathbf{p}; x))$$

  and if T is the curve trace operator also

  $$\partial_\nu \xi(\mathbf{p}; x) = g_2(\mathbf{p}; \psi(\mathbf{p}; x)).$$

□

Combining the previous results yields:

Corollary 2.1.7. *Let* $\boldsymbol{q} \in \mathcal{B} - \boldsymbol{p}^0$, $V \in C^m(\Omega, \mathbb{R}^2)$ *and* $w \in C^{m-1}(\Omega)$ *as in Lemma 2.1.5 and Lemma 2.1.6, respectively. Then* $\mathcal{X} \in C^m(\mathcal{B} \times \Omega)$ *and* $\xi \in C^m(\mathcal{B} \times \Omega)$ *exist such that a family of bijections* $\Phi(\boldsymbol{p})\colon U_{ad}(\boldsymbol{p}^0) \to U_{ad}(\boldsymbol{p})$, $\boldsymbol{p} \in \mathcal{B}$ *is defined by*

$$\Phi(\boldsymbol{p}; u) := (u + \xi(\boldsymbol{p})) \circ \mathcal{X}(\boldsymbol{p})^{-1}$$

and such that $\partial_{\boldsymbol{q}}\mathcal{X}(\boldsymbol{p}^0) = V$ *and* $\partial_{\boldsymbol{q}}\xi(\boldsymbol{p}^0) = w$ *hold.*

Remark 2.1.8 (Whitney extension theorem)**.** *Both Lemma 2.1.5 and Lemma 2.1.6 rely on the existence of sufficiently smooth extension functions with prescribed values and derivatives on some subset of the extension domain. Under mild assumptions on the boundaries and the boundary data, the existence of such functions is asserted by* Whitney's extension theorem *(cf. [Whi34]). It implies:*

Let $m \in \mathbb{N}$, *let* $A \subset \mathbb{R}^n$ *a closed subset and suppose boundary condition functions* $f_\alpha\colon A \to \mathbb{R}$ *for multi-indices* α *with* $|\alpha| \leq m$. *If the* f_α *fulfill for all* $x, y \in A$ *the Taylor-type compatibility condition*

$$f_\alpha(x) = \sum_{|\beta| \leq m-|\alpha|} \frac{f_{\alpha+\beta}(x)}{\beta!}(x-y)^\beta + o\left(|x-y|^{m-|\alpha|}\right) \quad \text{for } |x-y| \to 0,$$

then a function $f \in C^m(\mathbb{R}^n)$ *exists such that* $D^\alpha f|_A = f_\alpha$ *for* $|\alpha| \leq m$.

Due to the fact that the verification of the compatibility conditions is rather technical and of only moderate importance in the given context, the explicit application of this theorem is omitted in this work.

2.2 Differentiability

The knowledge from the previous section on structure-preserving bijections between the admissible sets $U_{\mathrm{ad}}(\mathbf{p})$ is now used to transform the elastic energy

$$J(\mathbf{p}, u) = \int_{\Omega(\mathbf{p})} \rho\left(x, u(x), Du(x), D^2u(x)\right) \,\mathrm{d}x$$

locally onto the domain $\Omega(\mathbf{p}^0)$. This is done in such a way that an equivalent minimization problem for the transformed energy is stated over a fixed reference space $U_{\mathrm{ad}}(\mathbf{p}^0)$. Afterwards, the differentiability of the transformed energy is shown and the implicit function theorem is applied in order to reveal the differentiability of the interaction potential.

Lemma 2.2.1. *Let $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ and $X\colon \Omega_1 \to \Omega_2$ be a diffeomorphism. Suppose $u \in H^1(\Omega_2)$ and define $\widetilde{u} := u \circ X$. Then*

$$\nabla u \circ X = (DX)^{-T} \nabla \widetilde{u}. \tag{2.2.1}$$

Proof. Theorem 2.1.1 asserts $\widetilde{u} \in H^1(\Omega_1)$. Equation (2.2.1) is a direct consequence of

$$\nabla \widetilde{u} = \nabla(u \circ X) = (DX)^T \left(\nabla u \circ X\right)$$

and the invertibility of DX due to the fact that X is a diffeomorphism. □

Lemma 2.2.2. *Let $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ and $X\colon \Omega_1 \to \Omega_2$ be a 2-diffeomorphism. Suppose $u \in H^2(\Omega_2)$, let $\gamma = |\det(DX)|$, and define $\widetilde{u} := u \circ X$. Then*

$$D^2u \circ X = \frac{1}{\gamma} \sum_{k=1}^{n} \partial_k \left(\gamma (DX^{-T} \nabla \widetilde{u}) \otimes (DX^{-T} e_k)\right) \tag{2.2.2}$$

where $e_k \in \mathbb{R}^n$ denotes the k-th canonical Euclidean basis vector.

Proof. Theorem 2.1.1 implies $\widetilde{u} \in H^2(\Omega_1)$. Moreover, either $\gamma > 0$ or $\gamma < 0$ holds uniformly on Ω_1 because X is a diffeomorphism. Therefore, $|\gamma|$ is differentiable (see Lemma A.2 in the appendix). Let $\widetilde{\phi} \in C_0^\infty(\Omega_1)$ be an arbitrary smooth test function with compact support in Ω_1. Then $\widetilde{\phi} \circ X^{-1} \in H_0^2(\Omega_2)$ and application of the transformation formula yields

$$\int_{\Omega_2} D^2u \, \phi \,\mathrm{d}x = \sigma \int_{\Omega_1} \gamma (D^2u \circ X) \, \widetilde{\phi} \,\mathrm{d}x. \tag{2.2.3}$$

where $\sigma := \operatorname{sign}(\gamma)$. Integration by parts, the transformation formula and the transformation identity (2.2.1) show

$$\begin{aligned}\int_{\Omega_2} D^2u\,\phi\,\mathrm{d}x &= -\int_{\Omega_2} \nabla u \otimes \nabla\phi\,\mathrm{d}x \\ &= -\sigma\int_{\Omega_1} \gamma(DX^{-T}\nabla\widetilde{u})\otimes(DX^{-T}\nabla\widetilde{\phi})\,\mathrm{d}x \\ &= -\sigma\sum_{k=1}^{n}\int_{\Omega_1}\left(\gamma(DX^{-T}\nabla\widetilde{u})\otimes(DX^{-T}e_k)\right)\partial_k\widetilde{\phi}\,\mathrm{d}x \\ &= \sigma\sum_{k=1}^{n}\int_{\Omega_1}\partial_k\left(\gamma(DX^{-T}\nabla\widetilde{u})\otimes(DX^{-T}e_k)\right)\widetilde{\phi}\,\mathrm{d}x.\end{aligned} \tag{2.2.4}$$

Comparison of (2.2.3) with (2.2.4) yields in view of the fundamental theorem of calculus of variations the identity

$$\gamma\, D^2u\circ X = \sum_{k=1}^{n}\partial_k\left(\gamma(DX^{-T}\nabla\widetilde{u})\otimes(DX^{-T}e_k)\right),$$

which proves the transformation formula (2.2.2). □

Proposition 2.2.3. *Let $(\mathcal{X}(\boldsymbol{p}))_{\boldsymbol{p}\in\mathcal{B}}$ be a family of m-diffeomorphisms as in Proposition 2.1.3 and let $\xi \in C^m(\mathcal{B}\times\Omega)$ be as in Lemma 2.1.6. Define for every $u \in H^2(\Omega(\boldsymbol{p}^0))$ the transformed elastic energy by*

$$\widetilde{J}(\boldsymbol{p};u) := J\left(\boldsymbol{p},(u+\xi(\boldsymbol{p}))\circ\mathcal{X}^{-1}(\boldsymbol{p})\right).$$

Then:

- *For all $\boldsymbol{p}\in\mathcal{B}$ and $v\in U_{ad}(\boldsymbol{p})$ holds $J(\boldsymbol{p},v) = \widetilde{J}(\boldsymbol{p}, v\circ\mathcal{X}(\boldsymbol{p})-\xi(\boldsymbol{p}))$. In particular, every minimizer of the original energy $J(\boldsymbol{p})$ over $U_{ad}(\boldsymbol{p})$ induces a minimizer of the transformed energy $\widetilde{J}(\boldsymbol{p})$ over $U_{ad}(\boldsymbol{p}^0)$, and vice-versa.*

- *If $\rho\in C^{m-2}(\overline{\Omega}\times\mathbb{R}\times\mathbb{R}^2\times\mathbb{R}^{2\times2})$ and if the derivatives of the transformed integrand in (2.2.5) below are integrable, then $\widetilde{J}\in C^{m-2}(\mathcal{B}\times U_{ad}(\boldsymbol{p}^0))$.*

Proof. The equivalence of the energies is immediately obtained from

$$\begin{aligned}J(\mathbf{p};v) &= J\left(\mathbf{p};((v\circ\mathcal{X}(\mathbf{p})-\xi(\mathbf{p}))+\xi(\mathbf{p}))\circ\mathcal{X}^{-1}(\mathbf{p})\right)\\ &= \widetilde{J}\left(\mathbf{p};v\circ\mathcal{X}(\mathbf{p})-\xi(\mathbf{p})\right).\end{aligned}$$

As of Lemma 2.1.2, the equivalence of the minimization problems is then apparent from the bijectivity of the map $\Phi(\mathbf{p})^{-1}\colon U_{\mathrm{ad}}(\mathbf{p})\to U_{\mathrm{ad}}(\mathbf{p}^0), v\mapsto v\circ\mathcal{X}(\mathbf{p})-\xi(\mathbf{p})$.

In order to prove the differentiability of $\widetilde{J}$, let $\gamma(\mathbf{p}) := \det(D\mathcal{X}(\mathbf{p}))$, $v(\mathbf{p}) := u + \xi(\mathbf{p})$ and

$$\zeta(\mathbf{p}) := \Bigg(\mathcal{X}(\mathbf{p}),\ v(\mathbf{p}),\ D\mathcal{X}(\mathbf{p})^{-T}\nabla v(\mathbf{p}), \\ \frac{1}{\gamma(\mathbf{p})} \sum_{k=1}^{2} \partial_k \left(\gamma(\mathbf{p}) (D\mathcal{X}(\mathbf{p})^{-T}\nabla v(\mathbf{p})) \otimes (D\mathcal{X}(\mathbf{p})^{-T} e_k) \right) \Bigg).$$

In view of Lemma 2.2.1 and Lemma 2.2.2 the transformation formula yields

$$\widetilde{J}(\mathbf{p}, \widetilde{u}) = \int_{\Omega(\mathbf{p}^0)} \gamma(\mathbf{p})\, \rho\left(\zeta(\mathbf{p})\right)\, \mathrm{d}x. \tag{2.2.5}$$

Lemma A.2 implies $\gamma \in C^{m-1}$, and by assumption holds $\rho \in C^{m-2}$. Closer investigation of ζ reveals $\zeta \in C^{m-2}$ by assumptions on $\mathcal{X}$ and ξ. Altogether, this implies $\widetilde{J} \in C^{m-2}$ if the derivatives of the transformed integrand $\gamma(\mathbf{p})\rho(\zeta(\mathbf{p}))$ are integrable. □

The proof of differentiability of the interaction potential is based on the following formulation of the implicit function theorem:

Theorem 2.2.4 (Implicit function theorem)**.** *Let $\mathcal{B}$, $\mathcal{U}$, $\mathcal{W}$ be real Banach spaces, $(\boldsymbol{p}^0, v^0) \in \mathcal{B} \times \mathcal{U}$, and let $\mathcal{A} \subseteq \mathcal{B} \times \mathcal{U}$ be an open neighborhood of $(\boldsymbol{p}^0, v^0)$. Let $f \in C^k(\mathcal{A}, \mathcal{W})$, $k \geq 1$ such that $f(\boldsymbol{p}^0, v^0) = 0$ and the partial derivative $f_u(\boldsymbol{p}^0, v^0) \colon \mathcal{U} \to \mathcal{W}$ is bijective. Then there exists a neighborhood $\hat{\mathcal{B}}$ of $\boldsymbol{p}^0$ and a unique function $v \in C^k(\hat{\mathcal{B}}, \mathcal{U})$ such that $(\boldsymbol{p}, v(\boldsymbol{p})) \in \mathcal{A}$ and $f(\boldsymbol{p}, v(\boldsymbol{p})) = 0$ for all $\boldsymbol{p} \in \hat{\mathcal{B}}$.*

Proof. This is a special case of [Zei85, Theorem 4.B]. □

Theorem 2.2.5. *Suppose $\widetilde{J}$ as in Proposition 2.2.3 with $\widetilde{J} \in C^{m-2}(\mathcal{B} \times U_{ad}(\boldsymbol{p}^0))$ and $m \geq 4$. If u^0 is the unique global minimizer of $J(\boldsymbol{p}^0)$ over $U_{ad}(\boldsymbol{p}^0)$ and the second variation $J_{uu}(\boldsymbol{p}^0, u^0) \colon U^0_{ad}(\boldsymbol{p}^0) \to \left(U^0_{ad}(\boldsymbol{p}^0)\right)'$ is bijective and locally coercive, then there exists a neighborhood $\hat{\mathcal{B}} \subseteq \mathcal{B}$ of $\boldsymbol{p}^0$ such that $\mathcal{E} \in \mathbb{C}^{m-3}(\hat{\mathcal{B}})$ and*

$$\frac{\partial}{\partial \boldsymbol{p}} \mathcal{E}(\boldsymbol{p}^0) = \widetilde{J}_{\boldsymbol{p}}(\boldsymbol{p}^0, u^0).$$

Proof. Let $\mathcal{U} := U^0_{\mathrm{ad}}(\mathbf{p}^0)$ the homogeneous admissible set, define $\mathcal{W} := \mathcal{U}'$ to be its dual space, and let $\mathcal{A} := \mathcal{B} \times \mathcal{U}$. Define the corresponding shifted energy function by $F(\mathbf{p}, v) := \widetilde{J}(\mathbf{p}, u^0 + v)$ for all $v \in U^0_{\mathrm{ad}}(\mathbf{p}^0)$. Then the minimizer of $F(\mathbf{p}^0, \cdot)$ over $\mathcal{U}$ is given by $v^0 := 0 \in \mathcal{U}$, and $f := F_u \in C^{m-3}(\mathcal{A})$ by assumption on $\widetilde{J}$.

Because u^0 is a minimizer of $J(\mathbf{p}^0)$, it is in particular a critical point of $J(\mathbf{p}^0)$ and so its first variation yields $J_u(\mathbf{p}^0, u^0) = 0$. Consequently, also

$$f(\mathbf{p}^0, v^0) = F_u(\mathbf{p}^0, 0) = \widetilde{J}_u(\mathbf{p}^0, u^0) = J_u(\mathbf{p}^0, u^0) = 0.$$

Again by assumption, $J_{uu}(\mathbf{p}^0, u^0) = f_u(\mathbf{p}^0, v^0) \colon \mathcal{U} \to \mathcal{W}$ is invertible, and therefore the implicit function theorem is applicable to f for $k := m - 3$. Hence, a neighborhood $\hat{\mathcal{B}}$

of $\mathbf{p}^0$ and a function $v \in C^{m-3}(\hat{\mathcal{B}},\mathcal{U})$ exist such that $v(\mathbf{p}^0) = v^0$ and $\widetilde{J}_u(\mathbf{p}, v(\mathbf{p}) + u^0) = f(\mathbf{p}, v(\mathbf{p})) = 0$ for all $\mathbf{p} \in \hat{\mathcal{B}}$. For a sufficiently small neighborhood $\hat{\mathcal{B}}$, the uniqueness of the minimizer u^0 and the continuity properties of $\widetilde{J}$ together with its local coercivity imply that every $u(\mathbf{p}) := v(\mathbf{p}) + u_0 \in U_{\mathrm{ad}}(\mathbf{p}^0)$ is again a global minimizer of $\widetilde{J}(\mathbf{p})$ over $U_{\mathrm{ad}}(\mathbf{p}^0)$. Since the $u(\mathbf{p})$ are global minimizers it holds

$$\mathcal{E}(\mathbf{p}) = \min_{u \in U_{\mathrm{ad}}(\mathbf{p})} J(\mathbf{p}, u) = \min_{u \in U_{\mathrm{ad}}(\mathbf{p}^0)} \widetilde{J}(\mathbf{p}, u) = \widetilde{J}(\mathbf{p}, u(\mathbf{p})).$$

Because $u(\mathbf{p}) \in C^{m-3}(\hat{\mathcal{B}})$ and because $\widetilde{J} \in C^{m-2}(\mathcal{B} \times U_{\mathrm{ad}}(\mathbf{p}^0))$ is fulfilled, it follows $\mathcal{E}(\mathbf{p}) \in C^{m-3}(\hat{\mathcal{B}})$. Finally, taking the first derivative of $\mathcal{E}$ leads via the chain rule to

$$\frac{\partial}{\partial \mathbf{p}} \mathcal{E}(\mathbf{p}^0) = \widetilde{J}_{\mathbf{p}}\left(\mathbf{p}^0, u(\mathbf{p}^0)\right) + \widetilde{J}_u\left(\mathbf{p}^0, u(\mathbf{p}^0)\right) \cdot \frac{\partial}{\partial \mathbf{p}} u(\mathbf{p}^0) = \widetilde{J}_{\mathbf{p}}\left(\mathbf{p}^0, u(\mathbf{p}^0)\right) = \widetilde{J}_{\mathbf{p}}(\mathbf{p}^0, u^0).$$

and completes the proof.

□

Remark 2.2.6 (Extension to local minimizers)**.** *A similar result holds true also in the case where u^0 merely is a local and not necessarily the unique global minimizer. In that case a neighborhood $\mathcal{B}_{\boldsymbol{p}} \times \mathcal{B}_u \subseteq \mathcal{B} \times U_{ad}(\boldsymbol{p}^0)$ of the point $(\boldsymbol{p}^0, u^0)$ exists such that u^0 is the unique global minimizer of $J(\boldsymbol{p}^0)$ over $\mathcal{B}_u$. With the same arguments as above, one obtains the differentiability of the* local interaction potential

$$\mathcal{E}_{loc}(\boldsymbol{p}) := \min_{v \in \Phi(\boldsymbol{p}; \mathcal{B}_u)} J(\boldsymbol{p}, v)$$

in a neighborhood of $\boldsymbol{p}^0$ as well as

$$\frac{\partial}{\partial \boldsymbol{p}} \mathcal{E}_{loc}(\boldsymbol{p}^0) = \widetilde{J}_{\boldsymbol{p}}(\boldsymbol{p}^0, u^0).$$

Here $\Phi(\boldsymbol{p})$ is the bijection between $U_{ad}(\boldsymbol{p}^0)$ and $U_{ad}(\boldsymbol{p})$ as defined in Lemma 2.1.2.

Remark 2.2.7 (Relaxing the differentiability requirements)**.** *It is possible to relax the smoothness assumptions for $\widetilde{J}$ in Theorem 2.2.5 to a certain extent, even without major changes in the given proofs. More specifically, one observes for example that the actual proof only requires the differentiability $\widetilde{J}_u \in C^{m-3}$, but not really the stricter condition $\widetilde{J} \in C^{m-2}$. Similarly, several of the assumptions in previous results can be refined in order to reduce smoothness requirements, if necessary.*

This aspect is not treated in greater detail within this work as smoothness requirements usually do not pose any relevant restriction in the model problems that are considered later on. Therefore, a more rigorous treatment of the smoothness conditions is omitted for the sake of a simpler presentation.

2.3 Volume integral representation of the gradient

So far, the differentiability of interaction potentials induced by elastic energies of the type

$$J(\mathbf{p}, u) = \int_{\Omega(\mathbf{p})} \rho\left(x, u(x), Du(x), D^2u(x)\right) \,\mathrm{d}x$$

has been investigated. Now the results of the previous sections are combined in order to derive an explicit expression for the gradient which is feasible within the context of finite element approximations. The following calculations predominantly rely on standard methods from matrix calculus.

Lemma 2.3.1. *Let $t^* \in \mathbb{R}_{>0}$. For $t \in [0, t^*]$, suppose a family of domains $\Omega(t) \subseteq \mathbb{R}^n$ and a family of 2-diffeomorphisms $X(t)\colon \Omega(0) \to \Omega(t)$ with $X(0) = \mathrm{id}_{\Omega(0)}$. Furthermore, suppose that the derivatives $\frac{\partial}{\partial t} D^k X$ exist and are continuous for $k \leq 2$. Define the perturbation vector field $V := \frac{\partial}{\partial t} X|_{t=0}$. Then*

$$\frac{\partial}{\partial t} \det(DX(t))\bigg|_{t=0} = \operatorname{div}(V), \tag{2.3.1}$$

and

$$\frac{\partial}{\partial t} DX(t)^{-1}\bigg|_{t=0} = -DV, \tag{2.3.2}$$

and for any $\widetilde{u} \in H^2(\Omega(0))$

$$\frac{\partial}{\partial t} \frac{1}{\gamma} \sum_{k=1}^{2} \partial_k \left(\gamma (DX^{-T} \nabla \widetilde{u}) \otimes (DX^{-T} e_k)\right)\bigg|_{t=0} = -\operatorname{Sym}(D^2\widetilde{u} DV) - D^2 V \nabla \widetilde{u} \tag{2.3.3}$$

where $\gamma(t) := \det(DX(t))$ and were $e_k \in \mathbb{R}^2$ is the k-th canonical Euclidean basis vector.

Proof. For this proof, recall $\operatorname{Sym}(A) := A + A^T$ and $D^2 V \nabla \widetilde{u} := \sum_{k=1}^{2} D^2 V_k \, \partial_k \widetilde{u}$, and note $DX(0) = \mathrm{id}_{\mathbb{R}^2}$ as well as $\det(DX(0)) = 1$. Lemma A.2 in the appendix on the differentiation of the determinant implies

$$\frac{\partial}{\partial t} \det(DX(t))\bigg|_{t=0} = \det(DX(t)) \operatorname{Tr}\left(DX(t)^{-1} \frac{\partial}{\partial t} DX(t)\right)\bigg|_{t=0} = \operatorname{Tr}(DV) = \operatorname{div}(V).$$

Similarly, Lemma A.1 concerning differentiation of the matrix inverse yields

$$\frac{\partial}{\partial t} DX(t)^{-1}\bigg|_{t=0} = -DX(t)^{-1} \left(\frac{\partial}{\partial t} DX(t)\right) DX(t)^{-1}\bigg|_{t=0} = -DV.$$

This shows the identities (2.3.1) and (2.3.2). Now, note that

$$\sum_{k=1}^{2} \partial_k \nabla \widetilde{u} \otimes e_k = D^2 \widetilde{u}$$

and in particular also

$$\sum_{k=1}^{2} \partial_k \left(\gamma(DX^{-T}\nabla\widetilde{u}) \otimes (DX^{-T}e_k)\right)\Big|_{t=0} = D^2\widetilde{u}.$$

Together with $X(0) = \mathrm{id}_{\Omega(0)}$, (2.3.1), (2.3.2) and other elemental transformations this gives

$$\begin{aligned}
&\frac{\partial}{\partial t}\frac{1}{\gamma}\sum_{k=1}^{2} \partial_k \left(\gamma(DX^{-T}\nabla\widetilde{u}) \otimes (DX^{-T}e_k)\right)\Big|_{t=0} \\
&\quad = -\operatorname{div}(V)D^2\widetilde{u} + \sum_{k=1}^{2} \partial_k \operatorname{div}(V)\nabla\widetilde{u}\otimes e_k - \sum_{k=1}^{2} \partial_k DV^T\nabla\widetilde{u}\otimes e_k \\
&\qquad - \sum_{k=1}^{2} \nabla\widetilde{u}\otimes\left(\partial_k DV^T e_k\right) + \operatorname{div}(V)D^2\widetilde{u} - DV^T D^2\widetilde{u} - D^2\widetilde{u}DV \\
&\quad = -\sum_{k=1}^{2} D^2V_k\partial_k\widetilde{u} - DV^T D^2\widetilde{u} - D^2\widetilde{u}DV,
\end{aligned}$$

which shows equation (2.3.3) and completes the proof. □

Proposition 2.3.2. *Let $\mathcal{X}(\boldsymbol{p})$ be a family of trace-preserving 3-diffeomorphisms induced by a parameter-dependent vector field $\mathcal{V}$ as in Proposition 2.1.3, and let $\xi(\boldsymbol{p}) \in C^3(\mathcal{B}\times\Omega)$ be a family of constraint-preserving maps as in Lemma 2.1.6. Furthermore, let $u \in U_{ad}(\boldsymbol{p}^0)$, $\boldsymbol{q} \in \mathcal{B} - \boldsymbol{p}^0$, $V := \mathcal{V}(0, \boldsymbol{p}^0 + \boldsymbol{q}, \cdot)$ and assume that $\mathcal{V}$ fulfills the scaling condition (2.1.11). If the transformed energy $\widetilde{J}$ from Proposition 2.2.3 is differentiable, then the derivative in direction $\boldsymbol{q}$ is given by*

$$\begin{aligned}
\partial_{\boldsymbol{q}}\widetilde{J}(\boldsymbol{p}^0, u) = &\int_{\Omega(\boldsymbol{p}^0)} \operatorname{div}(V)\rho(\zeta_0) + \rho_x(\zeta_0)V + \rho_y(\zeta_0)\xi_0' + \rho_z(\zeta_0)\cdot(\nabla\xi_0' - DV\nabla u)\,\mathrm{d}x \\
&+ \int_{\Omega(\boldsymbol{p}^0)} \rho_Z(\zeta_0) : \left(D^2\xi_0' - D^2V\nabla u - \operatorname{Sym}(D^2u\cdot DV)\right)\,\mathrm{d}x
\end{aligned} \tag{2.3.4}$$

where $\zeta_0(x) := (x, u(x), \nabla u(x), D^2u(x))$, $\xi_0' := \partial_q\xi(\boldsymbol{p}^0)$, and where ρ_x, ρ_y, ρ_z and ρ_Z are the partial derivatives of ρ with respect to its first, second, third and fourth component, respectively.

Proof. As in the proof of Proposition 2.2.3, define $\gamma(\mathbf{p}) := \det(D\mathcal{X}(\mathbf{p}))$, $v(\mathbf{p}) := u + \xi(\mathbf{p})$ and

$$\begin{aligned}
\zeta(\mathbf{p}) := \Bigg(&\mathcal{X}(\mathbf{p}),\ v(\mathbf{p}),\ D\mathcal{X}(\mathbf{p})^{-T}\nabla v(\mathbf{p}), \\
&\frac{1}{\gamma(\mathbf{p})}\sum_{k=1}^{2} \partial_k \left(\gamma(\mathbf{p})(D\mathcal{X}(\mathbf{p})^{-T}\nabla v(\mathbf{p}))\otimes(D\mathcal{X}(\mathbf{p})^{-T}e_k)\right)\Bigg)
\end{aligned}$$

to obtain the transformed energy

$$\widetilde{J}(\mathbf{p}, u) := J(\mathbf{p}, v(\mathbf{p}) \circ \mathcal{X}(\mathbf{p})) = \int_{\Omega(\mathbf{p}^0)} \gamma(\mathbf{p})\, \rho\left(\zeta(\mathbf{p})\right) \,\mathrm{d}x.$$

Let $X(t) := \mathcal{X}(\mathbf{p}^0 + t(\mathbf{p} - \mathbf{p}_0))$ and note $\frac{\partial}{\partial t} X|_{t=0} = V = \mathcal{V}(0, \mathbf{p}, \cdot)$ as of Lemma 2.1.4. As of $X \in C^3$ and $\mathcal{X}(\mathbf{p}^0) = \mathrm{id}_{\Omega(\mathbf{p}^0)}$, the identities from Lemma 2.3.1 are applicable and one obtains via the product and the chain rules

$$\begin{aligned}
\partial_{\mathbf{q}} \widetilde{J}(\mathbf{p}^0, u) &= \frac{\partial}{\partial t} \widetilde{J}(\mathbf{p}^0 + t\mathbf{q}, u)\Big|_{t=0} \\
&= \int_{\Omega(\mathbf{p}^0)} \mathrm{div}(V)\rho(\zeta_0) + \rho_x(\zeta_0)V - \rho_z(\zeta_0) \cdot (DV\nabla u)\,\mathrm{d}x \\
&\quad - \int_{\Omega(\mathbf{p}^0)} \rho_Z(\zeta_0) : \left(D^2 V \nabla u + \mathrm{Sym}(D^2 u \cdot DV)\right) \,\mathrm{d}x \\
&\quad + \int_{\Omega(\mathbf{p}^0)} \rho_y(\zeta_0)\xi_0' + \rho_z(\zeta_0)\nabla\xi_0' + \rho_Z(\zeta_0) : D^2\xi_0' \,\mathrm{d}x,
\end{aligned}$$

which is equivalent to (2.3.4). □

In summary, the previous paragraphs lead to the following differentiability result:

Theorem 2.3.3. *Assume that $u^0 \in U_{ad}(\boldsymbol{p}^0)$ is the unique minimizer of $J(\boldsymbol{p}^0)$ over $U_{ad}(\boldsymbol{p}^0)$ and that the second variation $J_{uu}(\boldsymbol{p}^0, u^0)\colon U^0_{ad}(\boldsymbol{p}^0) \to (U^0_{ad}(\boldsymbol{p}^0))'$ is invertible. Let $\boldsymbol{q} \in \mathbb{R}^{N\times k}$ and suppose $V \in H^2(\Omega(\boldsymbol{p}^0), \mathbb{R}^2)$ such that*

$$V|_{\Gamma(\boldsymbol{p}^0)} = \partial_{\boldsymbol{q}} \psi(\boldsymbol{p}^0) \tag{2.3.5}$$

and, if T is the curve trace operator,

$$DV\nu|_{\Gamma(\boldsymbol{p}^0)} = \partial_{\boldsymbol{q}} \left(\nu|_{\Gamma(\boldsymbol{p}^0)} \circ \psi(\boldsymbol{p}^0)\right). \tag{2.3.6}$$

Also suppose $w \in H^2(\Omega(\boldsymbol{p}^0))$ with

$$T(\boldsymbol{p}^0; w) = \partial_{\boldsymbol{q}} \left(g(\boldsymbol{p}^0) \circ \psi(\boldsymbol{p}^0)\right). \tag{2.3.7}$$

Furthermore, have V and w fulfill the Dirichlet boundary conditions

$$\begin{aligned}
V|_{\partial\Omega} &= 0 & \qquad w|_{\partial\Omega} &= 0 \\
DV\nu|_{\partial\Omega} &= 0 & \qquad \partial_\nu w|_{\partial\Omega} &= 0.
\end{aligned}$$

Finally, assume that the transformed energy $\widetilde{J}$ as in Proposition 2.2.3 is differentiable twice and that the right hand sides in equations (2.3.5), (2.3.6) and (2.3.7) are well-defined and admit C^4-extensions to a neighborhood of $\boldsymbol{p}^0 \times \Gamma(\boldsymbol{p}^0)$. Then the interaction potential $\mathcal{E}$ is differentiable in $\boldsymbol{p}^0$ and the value of the directional derivative $\partial_{\boldsymbol{q}}\mathcal{E}(\boldsymbol{p}^0)$ is given by equation (2.3.4) with $\xi_0' = w$.

Proof. Assume at first $V \in C^4$ and $w \in C^3$ as well as $\mathbf{q} \in \mathcal{B} - \mathbf{p}^0$. Then by Lemma 2.1.5 and Lemma 2.1.6 the existence of C^4-smooth constraint-preserving families of maps $\mathcal{X}(\mathbf{p})$ and $\xi(\mathbf{p})$ is ensured. In particular, Proposition 2.3.2 is applicable and $\partial_{\mathbf{q}}\widetilde{J}(\mathbf{p}^0, u)$ may be computed as stated in equation (2.3.4). Moreover, by the assumptions on J and $\widetilde{J}$, Theorem 2.2.5 is applicable and hence the interaction potential $\mathcal{E}$ is differentiable in $\mathbf{p}^0$ and the directional derivative $\partial_{\mathbf{q}}\mathcal{E}(\mathbf{p}^0)$ is given by equation (2.3.4) with $u = u(\mathbf{p}^0)$ and $\xi_0' = w$.

For arbitrary $\mathbf{q} \in \mathbb{R}^{N\times k}$ note that there exists an $\varepsilon \in \mathbb{R}_{>0}$ such that $\mathbf{q}_\varepsilon := \varepsilon\mathbf{q} \in \mathcal{B} - \mathbf{p}^0$. Then the above argumentation is applicable with the scaled functions $V_\varepsilon := \varepsilon V$ and $w_\varepsilon := \varepsilon w$. This yields that $\partial_{\mathbf{q}_\varepsilon}\mathcal{E}(\mathbf{p}^0)$ is given by the formula (2.3.4) with $(V_\varepsilon, w_\varepsilon)$ instead of (V, w). However, because of the linearity $\partial_{\mathbf{q}_\varepsilon}\mathcal{E}(\mathbf{p}^0) = \varepsilon\partial_{\mathbf{q}}\mathcal{E}(\mathbf{p}^0)$ and because the formula (2.3.4) is linear in (V, w), it follows that $\partial_{\mathbf{q}}\mathcal{E}(\mathbf{p}^0)$ still may be computed by equation (2.3.4) with just (V, w).

Finally, a classical density argument extends this result to the case $V \in H^2(\Omega(\mathbf{p}^0))$ and $\xi_0' \in H^2(\Omega(\mathbf{p}^0))$. This completes the proof.

□

Remark 2.3.4 (Periodic boundary conditions)**.** *As mentioned earlier, all of the above results are stated for the case where the $\boldsymbol{p}$-independent boundary $\partial\Omega$ is equipped with Dirichlet-type boundary conditions. In many applications, however, also periodic boundary conditions are of interest, either because it is required by the parameterization (e. g. when perturbations of tubular membranes are considered) or because it is a pure model decision.*

Suppose $\partial\Omega \supseteq \Gamma_D \dot\cup \Gamma_P$ where Γ_D is the Dirichlet-boundary and Γ_P is the periodic boundary, which again is split into its counter-parts $\Gamma_P = \Gamma_{P,1} \dot\cup \Gamma_{P,2}$. In that situation the analogue of Theorem 2.3.3 is simply obtained by replacing the boundary conditions for V and w by

$$\begin{aligned} V|_{\Gamma_D} &= 0 & w|_{\Gamma_D} &= 0 \\ DV\nu|_{\Gamma_D} &= 0 & \partial_\nu w|_{\Gamma_D} &= 0 \\ V|_{\Gamma_{P,1}} &= V|_{\Gamma_{P,2}} & w|_{\Gamma_{P,1}} &= w|_{\Gamma_{P,2}} \\ DV\nu|_{\Gamma_{P,1}} &= -DV\nu|_{\Gamma_{P,2}} & \partial_\nu w|_{\Gamma_{P,1}} &= -\partial_\nu w|_{\Gamma_{P,2}}. \end{aligned}$$

The general proof remains identical once the corresponding analogues of Lemma 2.1.5 and Lemma 2.1.6 are established, which again imply the existence of families of suitable bijections $\Phi(\boldsymbol{p})\colon U_{ad}(\boldsymbol{p}^0) \to U_{ad}(\boldsymbol{p})$ induced by the functions V and w.

The following examples discuss the derivative formula for some special interaction models. It is assumed that all quantities are given in the sense of Theorem 2.3.3, unless it is stated otherwise.

Example 2.3.5 (Monge-gauge)**.** *Recall Example 1.1.4 where the linearization of the elastic energy without Gaussian curvature over a flat reference domain is stated, also*

known as Monge-gauge formulation. In that case the energy is up to a constant given by

$$J(\boldsymbol{p}, u) = \int_{\Omega(\boldsymbol{p})} \frac{\kappa}{2}(\Delta u)^2 + \frac{\sigma}{2} \|\nabla u\|^2 \, \mathrm{d}x$$

and the homogeneous admissible spaces are $U_{ad}^0(\boldsymbol{p}) = H_0^2(\Omega(\boldsymbol{p}))$.

From the theory of linear elliptic operators it is well-known that a minimizer of $J(\boldsymbol{p}, \cdot)$ *over the closed affine subspace* $U_{ad}(\boldsymbol{p})$ *exists and is unique because the second variation*

$$J_{uu}(\boldsymbol{p}, u; v_1, v_2) = \int_{\Omega(\boldsymbol{p})} \kappa \Delta v_1 \, \Delta v_2 + \sigma \nabla v_1 \cdot \nabla v_2 \, \mathrm{d}x$$

is elliptic over $H_0^2(\Omega(\boldsymbol{p}))$ *for any* $u \in H^2(\Omega(\boldsymbol{p}))$ *(see also [EGH+16]). In particular,* $J_{uu} \colon U_{ad}^0(\boldsymbol{p}) \to (U_{ad}^0(\boldsymbol{p}))'$ *is always invertible.*

Given $\mathcal{X} \in C^m$ *and* $\xi \in C^m$*, the transformed energy* $\widetilde{J}$ *in* $\boldsymbol{p}^0$ *reads*

$$\widetilde{J}(\boldsymbol{p}, u) = \frac{1}{2} \int_{\Omega(\boldsymbol{p})} \frac{\kappa \operatorname{div}(A \nabla u)^2}{\det(D\mathcal{X}(\boldsymbol{p}))} + \sigma \nabla u \cdot A \nabla u \, \mathrm{d}x$$

where $A = \det(D\mathcal{X}(\boldsymbol{p})) D\mathcal{X}(\boldsymbol{p})^{-1} D\mathcal{X}(\boldsymbol{p})^{-T}$*. Hence,* $\widetilde{J}$ *is* ∞*-smooth in* u *and* m*-smooth in* $\boldsymbol{p}$*. Putting all these properties together and given that the boundary data is sufficiently smooth, the differentiability result in Theorem 2.3.3 is applicable for* $m \geq 2$*,*

The integrand in $J(\boldsymbol{p}, \cdot)$ *is described by*

$$\rho(x, y, z, Z) = \frac{\kappa}{2}(Z_{11} + Z_{22})^2 + \frac{\sigma}{2}(z_1^2 + z_2^2)$$

and its derivatives in a point $\zeta = (x, u, Du, D^2u)$ *are given by*

$$\rho_x(\zeta) = 0, \qquad \rho_y(\zeta) = 0, \qquad \rho_z(\zeta) = \sigma Du, \qquad \rho_Z(\zeta) = \kappa \Delta u \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore, the derivative formula (2.3.4) *reads*

$$\begin{aligned} \partial_{\boldsymbol{q}} \mathcal{E}(\boldsymbol{p}^0) = \int_{\Omega(\boldsymbol{p}^0)} \frac{1}{2} \operatorname{div}(V) \left(\kappa (\Delta u)^2 + \sigma \|\nabla u\|^2 \right) + \sigma \nabla u \cdot (\nabla \xi_0' - DV \nabla u) \, \mathrm{d}x \\ + \int_{\Omega(\boldsymbol{p}^0)} \kappa \Delta u \, \left(\Delta \xi_0' - \Delta V \cdot \nabla u - 2DV : D^2 u \right) \, \mathrm{d}x. \end{aligned}$$

Example 2.3.6 (Willmore energy in graph case)**.** *Consider the nonlinear elastic bending energy parameterized over a flat reference surface, as shown in Example 1.1.1, but with* $c_0 = \kappa_G = 0$*. Then*

$$J(\boldsymbol{p}, u) = \int_{\Omega(\boldsymbol{p})} \frac{\kappa}{2} \cdot \frac{\left(\Delta u + \nabla u \cdot A(u) \nabla u \right)^2}{(1 + \|\nabla u\|^2)^{5/2}} + \sigma \sqrt{1 + \|\nabla u\|^2} \, \mathrm{d}x$$

where

$$A(u) = \begin{pmatrix} u_{22} & -u_{12} \\ -u_{12} & u_{11} \end{pmatrix}.$$

In this setting it is mathematically challenging to establish whether a global minimizer of $J(\boldsymbol{p})$ exists in $U_{ad}(\boldsymbol{p})$ and if it is unique. Moreover, solution methods might only be able to find local minimums. Therefore, it is simply assumed here that the differentiability result Theorem 2.3.3 is applicable and, in the spirit of Remark 2.2.6, that the derivative formula is considered for a local minimizer u.

Let $\hat{H} := \hat{H}(z, Z) = Z_{11}(1 + z_2^2) + Z_{22}(1 + z_1^2) - z_1 z_2 Z_{12} - z_1 z_2 Z_{21}$ the numerator of the mean curvature term. Then the integrand of the energy functional is described by

$$\rho(x, y, z, Z) = \frac{\kappa}{2} \cdot \frac{\hat{H}^2}{(1 + z_1^2 + z_2^2)^{5/2}} + \sigma\sqrt{1 + z_1^2 + z_2^2}$$

and the relevant derivatives in a point $\zeta = (x, y, z, Z)$ are given by

$$\rho_x(\zeta) = 0, \qquad \rho_y(\zeta) = 0, \qquad \rho_Z(\zeta) = \frac{\kappa\hat{H}}{(1 + z_1^2 + z_2^2)^{5/2}} \begin{pmatrix} 1 + z_2^2 & -z_1 z_2 \\ -z_1 z_2 & 1 + z_1^2 \end{pmatrix}.$$

and

$$\begin{aligned} \rho_z(\zeta) = {} & \frac{\kappa\hat{H}}{(1 + z_1^2 + z_2^2)^{5/2}} \begin{pmatrix} 2z_1 Z_{22} - z_2(Z_{12} + Z_{21}) \\ 2z_2 Z_{11} - z_1(Z_{12} + Z_{21}) \end{pmatrix} \\ & - \frac{5\kappa\hat{H}^2}{2(1 + z_1^2 + z_2^2)^{7/2}} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \frac{\sigma}{\sqrt{1 + z_1^2 + z_2^2}} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}. \end{aligned}$$

Using $Q := \sqrt{1 + \|\nabla u\|^2}$ and $\hat{H} := \hat{H}(Du, D^2u)$, the derivative formula (2.3.4) reads

$$\begin{aligned} \partial_{\boldsymbol{q}}\mathcal{E}(\boldsymbol{p}^0) = {} & \int_{\Omega(\boldsymbol{p}^0)} \operatorname{div}(V) \left(\frac{\kappa\hat{H}^2}{2Q^5} + \sigma Q \right) \mathrm{d}x \\ & + \int_{\Omega(\boldsymbol{p}^0)} \left(\frac{2\kappa\hat{H}}{Q^5} \begin{pmatrix} u_1 u_{22} - u_2 u_{12} \\ u_2 u_{11} - u_1 u_{12} \end{pmatrix} + \left(\frac{\sigma}{Q} - \frac{5\kappa\hat{H}^2}{2Q^7} \right) \nabla u \right) \cdot (\nabla\xi_0' - DV\nabla u) \, \mathrm{d}x \\ & + \int_{\Omega(\boldsymbol{p}^0)} \frac{\kappa\hat{H}}{Q^5} \begin{pmatrix} 1 + u_2^2 & -u_1 u_2 \\ -u_1 u_2 & 1 + u_1^2 \end{pmatrix} : \left(D^2\xi_0' - D^2V\nabla u - \operatorname{Sym}(D^2 u DV) \right) \mathrm{d}x. \end{aligned}$$

Example 2.3.7 (linearized energy over tube)**.** *Given a tubular reference manifold, the linearized elastic energy parameterized over a domain $\Omega = [0, L] \times [0, 2\pi]$ is stated in Example 1.1.5. For the sake of simplicity the non-physical assumption is made that the mean curvature of the tube is matched with the spontaneous curvature, i. e. let $c_0 = \frac{1}{r}$. After neglecting the Gaussian curvature term, the resulting linearized energy is up to a constant given by*

$$J(\boldsymbol{p}, u) = \frac{r}{2} \int_{\Omega(\boldsymbol{p})} \kappa \left(u_{11} + \frac{u + u_{22}}{r^2} \right)^2 + \sigma \left(u_1^2 + \frac{u_2^2}{r^2} \right) + \frac{2\sigma u}{r} \, \mathrm{d}x.$$

Correspondingly,

$$\rho(x,y,z,Z) = \frac{r\kappa}{2}\left(Z_{11} + \frac{y+Z_{22}}{r^2}\right)^2 + \frac{r\sigma}{2}\left(z_1^2 + \frac{2y}{r} + \frac{z_2^2}{r^2}\right)$$

with derivatives

$$\rho_x(\zeta) = 0, \qquad \rho_y(\zeta) = r\kappa\left(Z_{11} + \frac{y+Z_{22}}{r^2}\right) + \sigma,$$
$$\rho_z(\zeta) = r\sigma\begin{pmatrix} z_1 \\ \frac{z_2}{r^2}\end{pmatrix}, \qquad \rho_Z(\zeta) = r\kappa\left(Z_{11} + \frac{y+Z_{22}}{r^2}\right)\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2}\end{pmatrix}.$$

Hence, if $U_{ad}(\boldsymbol{p})$ is chosen such that $J(\boldsymbol{p})$ is elliptic, the interaction potential is differentiable and directional derivatives may be computed by formula (2.3.4)*. With the shortcuts $\nabla_r := (\partial_1, \partial_2/r)$, $\Delta_r := \partial_{11} + \partial_{22}/r^2$ and $\mathrm{Tr}_r(A) := A_{11} + A_{22}/r^2$ it reads*

$$\begin{aligned}\partial_{\boldsymbol{q}}\mathcal{E}(\boldsymbol{p}^0) = \frac{r}{2}\int_{\Omega(\boldsymbol{p}^0)} \mathrm{div}(V)\left(\kappa\left(\Delta_r u + \frac{u}{r^2}\right)^2 + \sigma\left\|\nabla_r u\right\|^2 + \frac{2\sigma u}{r}\right)\mathrm{d}x \\ + \int_{\Omega(\boldsymbol{p}^0)}\left(r\kappa\left(\Delta_r u + \frac{u}{r^2}\right)\sigma\right)\xi_0' + \sigma\nabla_r u\cdot\left(\nabla\xi_0' - DV\nabla u\right)\mathrm{d}x \\ + \int_{\Omega(\boldsymbol{p}^0)} r\kappa\left(\Delta_r u + \frac{u}{r^2}\right)\left(\Delta_r\xi_0' - \Delta_r V\cdot\nabla u - 2\,\mathrm{Tr}_r(D^2 u DV)\right)\mathrm{d}x.\end{aligned}$$

Note that the functions V and w here need to fulfill periodic boundary conditions along the segments $\Gamma_{P,1} = [0,L]\times\{0\}$ and $\Gamma_{P,2} = [0,L]\times\{2\pi\}$ in the spirit of Remark 2.3.4.

3 Discretization for linear models

Gradient-type minimization methods for the interaction potential

$$\mathcal{E}(\mathbf{p}) := \min_{v \in U_{\mathrm{ad}}(\mathbf{p})} J(\mathbf{p}, v) = J(\mathbf{p}, u(\mathbf{p})), \qquad \mathbf{p} \in \mathcal{D}$$

generally require a large number of evaluations of the gradient $\nabla\mathcal{E}$ and possibly also of the function $\mathcal{E}$ itself. In view of the definition of $\mathcal{E}$ and the derivative formula in Theorem 2.3.3, this means in a practical algorithm that the stationary membranes $u(\mathbf{p}) \in H^2(\Omega(\mathbf{p}))$ need to be computed frequently and for many different particle configurations $\mathbf{p} \in \mathcal{D}$. Since the membrane function $u(\mathbf{p})$ is characterized by the solution of a partial differential equation on a possibly complicated domain, in general no analytical representations are available and instead one needs to rely on numerical approximation techniques.

This chapter directs one's attention towards the approximation of the membrane function $u(\mathbf{p})$ for one single configuration $\mathbf{p} \in \mathcal{D}^\circ$ in the case where the elastic energy $J(\mathbf{p}, \cdot)$ is of a special class of elliptic quadratic energies that is specified in detail later on. Both point value and curve constraints are considered. Because the latter constraint type is numerically more challenging than the former, the focus lies mostly on curve constraints. The goal is to develop a stable, optimal-order finite element scheme that avoids potentially expensive re-meshing steps for updates in the configuration parameter $\mathbf{p}$ by using a configuration-independent finite element grid.

3.1 Preliminaries

Since only the approximation of $u(\mathbf{p})$ for one fixed $\mathbf{p} \in \mathcal{D}^\circ$ is considered here, the configuration parameter $\mathbf{p}$ is dropped entirely for the remainder of this chapter. For notational clarity $\Omega_0 := \Omega(\mathbf{p})$ is defined and $\Omega \supseteq \Omega_0$ denotes a *bulk domain* which does not necessarily have to coincide with the set Ω from the parameterization domain in Chapter 1. Similarly, also $J_0 := J(\mathbf{p}, \cdot)$ and $U_{\mathrm{ad}} := U_{\mathrm{ad}}(\mathbf{p})$ are defined. All estimates in the subsequent sections may have an implicit dependence on the configuration parameter $\mathbf{p}$ without further notice.

For now, the attention is on curve constraints. Departing from previous chapters, the trace operator T in the following sections is defined to incorporate not only the $\mathbf{p}$-dependent but also the $\mathbf{p}$-independent Dirichlet boundary conditions. Hence, let the set $\Gamma \subseteq \partial\Omega_0$ denote the portion of the boundary $\partial\Omega_0$ on which Dirichlet boundary conditions are prescribed, and extend the constraint right hand side g to all of Γ by zero. It is assumed that Γ has a non-trivial one-dimensional measure.

Recall the definition of the admissible set

$$U_{\mathrm{ad}} := \left\{ v \in H^2_*(\Omega_0) \mid Tv = g \right\}.$$

Here, T is the trace operator on the Dirichlet boundary Γ, i. e.

$$T\colon H^2(\Omega_0) \to H^{3/2}(\Gamma) \times H^{1/2}(\Gamma), \qquad v \mapsto (v|_\Gamma, \partial_\nu v|_\Gamma).$$

Furthermore, $H^2_*(\Omega_0) \subseteq H^2(\Omega_0)$ is a closed linear subspace which possibly encodes periodic or homogeneous natural boundary conditions, and the function $g \in \mathrm{range}(T)$ is the Dirichlet constraint right hand side.

As before, the homogeneous admissible set is analogously defined by

$$U^0_{\mathrm{ad}} := \left\{ v \in H^2_*(\Omega_0) \mid Tv = 0 \right\}.$$

Finally, for later use it is also convenient to define the *extended trace operator*

$$\widetilde{T}\colon H^2(\Omega_0) \to H^{3/2}(\Gamma) \times H^{1/2}(\Gamma)^2, \qquad v \mapsto (v|_\Gamma, \nabla v|_\Gamma).$$

Structure of the membrane energy

The optimal membrane shape is determined by the function $u \in U_{\mathrm{ad}}$ which solves the energy minimization problem

$$\min_{v \in U_{\mathrm{ad}}} J_0(v).$$

As mentioned earlier, in this chapter the energy J_0 is assumed to be quadratic and elliptic, i. e. up to a v-independent constant holds

$$J_0(v) = \frac{1}{2} a_{\Omega_0}(v, v) - \ell_0(v)$$

where a_{Ω_0} is an elliptic bilinear form on the linear subspace U^0_{ad}. and $\ell_0 \in H^2(\Omega_0)'$ is a bounded linear form. Moreover, it is assumed that the bilinear form a_{Ω_0} has a special structure:

Suppose

$$a_{\Omega_0}(v, w) = \sum_{i=1}^{n} \langle \delta_i v, \delta_i w \rangle_{L^2(\Omega_0)}$$

where the δ_i define differential operators (A1)

$$\delta_i v := \sum_{|\alpha| \le 2} c^i_\alpha \partial^\alpha v$$

with bounded coefficient functions $c^i_\alpha \in C^2(\Omega)$.

It is remarked that the quadratic parts in the Monge-gauge energy from Example 1.1.4 as well as those in the linearized energy for tubular reference manifolds from Example 1.1.5 are of this form.

Unfitted grid of quadrilaterals

The upcoming discretization works with an unfitted grid on an arbitrarily simple background domain. Therefore, it is assumed that the bulk domain Ω is rectangular with its border lines parallel to the coordinate axes. The grid $\mathcal{T}$ itself is assumed to have no hanging nodes and to consist out of quadrilateral elements E such that $\Omega = \bigcup_{E\in\mathcal{T}} E$. For simplicity of notation the elements E are assumed to be closed sets with pairwise disjoint interiors.

The set of *elements that intersect the domain* Ω_0 *non-trivially* is denoted by

$$\mathcal{T}_0 := \{E \in \mathcal{T} \mid E^\circ \cap \Omega_0 \neq \emptyset\}$$

where E° denotes the interior of E. This again induces the *grid-dependent bulk domain*

$$\widetilde{\Omega} := \bigcup_{E\in\mathcal{T}_0} E.$$

Similarly, the set of *Dirichlet boundary elements* is defined as

$$\mathcal{T}_\Gamma := \{E \in \mathcal{T} \mid E^\circ \cap \Gamma \neq \emptyset\},$$

which induces the *grid-dependent neighborhood of the Dirichlet boundary*

$$\Omega_\Gamma := \bigcup_{E\in\mathcal{T}_\Gamma} E.$$

Moreover, the discretization also requires the set of *boundary faces*

$$\mathcal{F} := \{E_1 \cap E_2 \mid E_1 \in \mathcal{T}_\Gamma,\ E_2 \in \mathcal{T}_0,\ E_1 \neq E_2,\ |E_1 \cap E_2|_1 > 0\}$$

where $|E_1 \cap E_2|_1$ is the one-dimensional measure of the set $E_1 \cap E_2$. For these faces a geometric assumption is made which ensures a grid-independent upper bound on the connectivity of boundary elements with non-boundary elements:

> Suppose there exists an h-independent $d \in \mathbb{N}$ with the following property:
> For all boundary elements $E_0 \in \mathcal{T}_\Gamma$ exist $d' \leq d$ elements $E_1, \dots, E_{d'} \in \mathcal{T}_0$ (A2)
> with $E_i \cap E_{i+1} \in \mathcal{F}$ and where $E_i \in \mathcal{T}_\Gamma$ for $i < d'$ and $E_{d'} \in \mathcal{T}_0 \setminus \mathcal{T}_\Gamma$.

Lastly, the *grid size function* $h\colon \Omega \to \mathbb{R}$ *with respect to the grid* $\mathcal{T}$ is defined almost-everywhere on Ω by the following rule:

$$h(x) = \begin{cases} \operatorname{diam}(E)/\sqrt{2}, & \text{if } x \in E^\circ \text{ for an } E \in \mathcal{T}, \\ \operatorname{diam}(E_1 \cap E_2), & \text{if } x \in E_1 \cap E_2 \text{ for } E_1, E_2 \in \mathcal{T}. \end{cases}$$

Here, $\operatorname{diam}(A) := \sup_{x,y\in A} \|x - y\|$ denotes the diameter of a set A. This definition is well-posed everywhere except in those interface points x where also an $E_3 \in \mathcal{T}$ exists such that $x \in E_1 \cap E_2 \cap E_3$. The function h shall remain undefined in those points. For

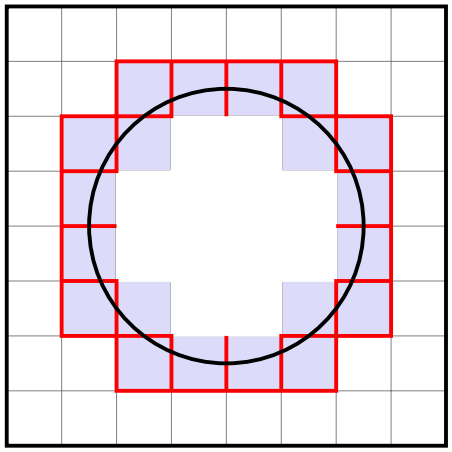

Figure 3.1.1: Unfitted uniform grid of quadrilaterals. Boundary elements are colored in gray and boundary interfaces are marked in red.

later convenience it is assumed that $h \leq 1$ holds uniformly. Also the weight $\sqrt{2}$ is due to reasons of convenience.
Finally, the following non-degeneracy condition is imposed:

> Suppose $\sigma \in \mathbb{R}_{\geq 1}$ such that for all adjacent elements $E_1, E_2 \in \mathcal{T}_0$ the inequality $\frac{|E_1|}{|E_2|} \leq \sigma$ holds h-independently where $|E_i|$ is the area of E_i. (A3)

Figure 3.1.1 shows an illustration of a typical example grid for a square domain with a circular cut-out. More specifically, a discretization of $\Omega = \{x \in [-1.6, 1.6]^2 \mid \|x\| \geq 1\}$ by uniform squares with side length 0.4 is depicted. In particular, the grid size function is constant with $h \equiv 0.4$, and $\sigma = 1$. All quadratic squares together, i. e. including the white and gray ones, define the set $\mathcal{T}_0$, whose union in turn spans the domain $\widetilde{\Omega}$. The boundary grid $\mathcal{T}_\Gamma$ corresponds to the union of all gray squares, the union of which yields the discrete boundary domain $\widetilde{\Omega}_\Gamma$. The boundary faces $\mathcal{F}$ are given by the red element faces. Assumption A2 holds with $d = 2$.

Piecewise $\mathcal{Q}_k$ finite element space

The base finite element space for the discretization is assumed to be given by

$$\mathcal{S} := \left\{ v \in C^{(k-1)/2}(\widetilde{\Omega}) \mid \forall E \in \mathcal{T}_0 \colon \ v|_E \in \mathcal{Q}_k \right\}$$

with an odd $k \in \mathbb{N}_{\geq 3}$. Here, $\mathcal{Q}_k$ is a subset of the polynomials up to order $2k$ in $\mathbb{R}^2$ which is defined by

$$\mathcal{Q}_k := \left\{ \sum_{i,j=0}^{k} a_{ij} x_1^i x_2^j \;\middle|\; a \in \mathbb{R}^{k \times k} \right\}.$$

Note that $\mathcal{S}$ indeed is non-empty and defines a conforming finite element space in $H^2(\widetilde{\Omega})$. A local basis of $\mathcal{S}$ is readily constructed via the solution of suitable Hermite interpolation problems. The special and most simple case $k = 3$ corresponds to the space spanned by Bogner–Fox–Schmit finite elements [BFS65].

The finite element space is assumed to fulfill the typical approximation properties:

> There exists a grid-independent constant $c \in \mathbb{R}_{>0}$ such that for all $s \in [2, k+1]$ and all $v \in H^s(\widetilde{\Omega}) \cap H^2_*(\widetilde{\Omega})$ exists an approximation $\overline{v} \in \mathcal{S}_*$ such that the *volume approximation property*
>
> $$\forall E \in \mathcal{T}_0 \; \forall |\alpha| \leq s\colon \quad \|\partial^\alpha (v - \overline{v})\|_{L^2(E)} \leq c \left\| h^{s-|\alpha|} v \right\|_{L^2(E)}$$
>
> and the *trace approximation property* for the boundary Γ
>
> $$\forall E \in \mathcal{T}_\Gamma \; \forall |\alpha| \leq s - \tfrac{1}{2}\colon \quad \|\partial^\alpha (v - \overline{v})\|_{L^2(\Gamma \cap E)} \leq c \left\| h^{s-1/2-|\alpha|} v \right\|_{L^2(E)} \tag{A4}$$
>
> as well as for the element faces
>
> $$\forall E \in \mathcal{T}_\Gamma \; \forall |\alpha| \leq s - \tfrac{1}{2}\colon \quad \|\partial^\alpha (v - \overline{v})\|_{L^2(\partial E)} \leq c \left\| h^{s-1/2-|\alpha|} v \right\|_{L^2(E)}$$
>
> hold.

Also the inverse inequalities are required:

> There exists a grid-independent constant $c_{\mathrm{inv}} \in \mathbb{R}_{\geq 1}$ such that for all $v \in \mathcal{S}_*$ the *volume inverse inequality*
>
> $$\forall E \in \mathcal{T}_0 \; \forall |\alpha| \leq k \; \forall \beta \leq \alpha\colon \quad \|\partial^\alpha v\|_{L^2(E)} \leq c_{\mathrm{inv}} \left\| h^{|\beta|-|\alpha|} \partial^\beta v \right\|_{L^2(E)}$$
>
> and the *trace inverse inequality* (A5)
>
> $$\forall E \in \mathcal{T}_\Gamma \; \forall |\alpha| \leq k \; \forall \beta \leq \alpha\colon \quad \|\partial^\alpha v\|_{L^2(E \cap \Gamma)} \leq c_{\mathrm{inv}} \left\| h^{|\beta|-|\alpha|-1/2} \partial^\beta v \right\|_{L^2(E)}$$
>
> are fulfilled. Here the notation $\beta \leq \alpha$ means that the multi-index β is a subset of the multi-index α.

Under slight assumptions on the interface Γ, both Assumption A4 and Assumption A5 are usually fulfilled, see also [Cia78] for details.

Regularity, bulk and extension assumptions

The existence of a unique membrane function $u \in H^2(\Omega_0)$ is readily established by virtue of the well-known Lax–Milgram theorem. Additional regularity in u is assumed

in order to prove a-priori discretization error estimates.

Suppose $u \in H^{k+1}(\Omega_0)$ where $k \in \mathbb{N}$ is the order of the $\mathcal{Q}_k$ finite element space. (A6)

Note that in many cases this assumption is fulfilled naturally in view of elliptic regularity theory (cf. [Gri85, Chapter 7] and [BR80, GGS10]) if the boundary data is sufficiently smooth.

The following assumption extends the coercivity of a_{Ω_0} to the domain $\widetilde{\Omega}$:

Suppose that the bilinear form

$$a_{\widetilde{\Omega}}(v,w) := \sum_{i=1}^{n} \langle \delta_i v, \delta_i w \rangle_{L^2(\widetilde{\Omega})} \tag{A7}$$

is coercive on the space $H^2_0(\widetilde{\Omega})$ with a grid-independent coercivity constant $\gamma \in \mathbb{R}_{>0}$.

Furthermore, since a bulk discretization is considered, it is also assumed that the solution function u admits a suitable extension:

Suppose there exists a grid-independent constant $c \in \mathbb{R}_{>0}$ and an extension $\widetilde{u} \in H^{k+1}(\widetilde{\Omega})$ such that $\widetilde{u}|_{\Omega_0} = u$ and

$$\forall |\alpha| \leq k+1 \colon \quad \|\partial^\alpha \widetilde{u}\|_{L^2(\widetilde{\Omega})} \leq c \, \|\partial^\alpha u\|_{L^2(\Omega_0)} \,. \tag{A8}$$

3.2 Fictitious domain stabilized Nitsche method for curve constraints

3.2.1 Discretization scheme

The discretization scheme requires three further ingredients besides the bilinear form a_{Ω_0} and the linear operator ℓ_0: A *penalty term*, a *Nitsche term*, and a so-called *stabilization term* augment the elastic energy J_0.

Penalty term

As the finite element space $\mathcal{S}_*$ does not incorporate the constraints $Tv = g$ explicitly they are enforced approximately via a penalty term. For later convenience this is done with help of the extended trace $\widetilde{T}v = (v|_\Gamma, \nabla v|_\Gamma)$. In this sense, note that the condition $Tv = g$ is equivalent to $\widetilde{T}v = (g_0, g_1\nu + \partial_\tau g_0 \tau) =: \widetilde{g}$ where τ is an arbitrary unit tangent field on the Dirichlet boundary Γ. The scheme uses a weighted L^2-penalization of this constraint,

$$\frac{\lambda}{2} \left(h^{-3} \, \|u - \widetilde{g}_0\|^2_{L^2(\Gamma)} + h^{-1} \, \|\nabla u - \widetilde{g}_1\|^2_{L^2(\Gamma)} \right),$$

where $\lambda \in \mathbb{R}_{>0}$ is a "large enough" constant that is specified in greater detail later on. Taking the first variation of this quadratic energy term naturally leads to the symmetric positive semi-definite bilinear form

$$a_{\mathrm{pen}}(v,w) := \lambda \sum_{i=0}^{1} \langle h^{i-3/2}\, \widetilde{T}_i v, h^{i-3/2}\, \widetilde{T}_i w\rangle_{L^2(\Gamma)}$$

and the linear functional

$$\ell_{\mathrm{pen}}(v) := \lambda \sum_{i=0}^{1} \langle h^{i-3/2}\, \widetilde{g}_i, h^{i-3/2}\, \widetilde{T}_i v\rangle_{L^2(\Gamma)}$$

where the components of the extended trace operator are indexed by $\widetilde{T} = (\widetilde{T}_0, \widetilde{T}_1)$.

Nitsche term

The Nitsche term ensures consistency of a minimizer of the augmented elastic energy. This term is essentially obtained by modifying the bilinear form a_{Ω_0} by a Nitsche term a_{Nit} such that the optimal membrane function u fulfills the characterizing variational equation over the larger space $H^2_*(\Omega_0) \supseteq U^0_{\mathrm{ad}}$, i. e.

$$\forall v \in H^2_*(\Omega_0)\colon\ a_{\Omega_0}(u,v) + a_{\mathrm{Nit}}(u,v) = \ell_0(v). \tag{3.2.1}$$

It is possible to derive a boundary representation of a_{Nit} by transitioning to the strong formulation for the membrane function u. To this end, let $c^i_{\alpha\beta} := c^i_\alpha \cdot c^i_\beta$ where the c^i_α are the coefficients in the definitions of the differential operators δ_i as stated in Assumption A1. Then

$$a_{\Omega_0}(v,w) = \sum_{i=1}^{n} \langle \delta_i v, \delta_i w\rangle_{L^2(\Omega_0)} = \sum_{i=1}^{n} \sum_{|\alpha|,|\beta|\le 2} \int_{\Omega_0} c^i_{\alpha\beta}\, \partial^\alpha v\, \partial^\beta w \,\mathrm{d}x$$

and integration by parts yields for $v \in H^2_*(\Omega_0)$

$$\begin{aligned} a_{\Omega_0}(u,v) = &\sum_{i=1}^{n} \sum_{|\alpha|,|\beta|\le 2} (-1)^{|\beta|} \int_{\Omega_0} \partial^\beta \left(c^i_{\alpha\beta}\partial^\alpha u\right) v \,\mathrm{d}x \\ &+ \sum_{i=1}^{n} \sum_{|\alpha|,|\beta|\le 2} \sum_{j=1}^{|\beta|} (-1)^{j+1} \int_{\Gamma} \partial_{\underline{\ }}^{\underline{\beta}^{j-1}} \left(c^i_{\alpha\beta}\partial^\alpha u\right)\, \partial^{\overline{\beta}^{j+1}} v\, \nu_{\beta_j} \,\mathrm{d}x. \end{aligned}$$

where $\underline{\beta}^j = (\beta_1, \dots, \beta_j)$ and $\overline{\beta}^j = (\beta_j, \dots, \beta_{|\beta|})$. Here also the periodic or natural boundary conditions on $\partial\Omega \setminus \Gamma$ play an important role in order to reduce the occurring boundary integral to just the Dirichlet boundary Γ. Upon collecting terms and defining coefficient functions $c^\Gamma_{\alpha\beta}$ such that

$$\sum_{\substack{|\alpha|\le 3,|\beta|\le 1,\\ |\alpha|+|\beta|\le 3}} \int_\Gamma c^\Gamma_{\alpha\beta}\partial^\alpha u \partial^\beta v \,\mathrm{d}x = \sum_{i=1}^{n} \sum_{|\alpha|,|\beta|\le 2} \sum_{j=1}^{|\beta|} (-1)^j \int_\Gamma \partial_{\underline{\ }}^{\underline{\beta}^{j-1}} \left(c^i_{\alpha\beta}\partial^\alpha u\right)\, \partial^{\overline{\beta}^{j+1}} v\, \nu_{\beta_j} \,\mathrm{d}x$$

holds, the Nitsche bilinear form is given by

$$a_{\mathrm{Nit}}(v,w) := \sum_{\substack{|\alpha|\leq 3,|\beta|\leq 1,\\ |\alpha|+|\beta|\leq 3}} \int_\Gamma c^\Gamma_{\alpha\beta} \partial^\alpha v \partial^\beta w \,\mathrm{d}x.$$

Based on the previous computations it is readily verified that the consistency property (3.2.1) is indeed fulfilled.

The actual discretization requires also the linear functional that is induced by the Nitsche bilinear form applied to the optimal membrane u. For $v \in H^4(\Omega_0)$ it is defined by

$$\ell_{\mathrm{Nit}}(v) := a_{\mathrm{Nit}}(v,u) = \sum_{\substack{|\alpha|\leq 3,|\beta|\leq 1,\\ |\alpha|+|\beta|\leq 3}} \int_\Gamma c^\Gamma_{\alpha\beta}\, \partial^\alpha v\, g_{|\beta|} \,\mathrm{d}x.$$

Stabilization term

The stabilization term is sometimes also referred to as ghost penalty term. It is purely quadratic and is defined on the discrete space $\mathcal{S}$ of piecewise $\mathcal{Q}_k$-elements by

$$a_{\mathrm{stab}}(v,w) = \sum_{F\in\mathcal{F}} \sum_{j=2}^{k} \frac{1}{(j!)^2(2j+1)} \left\langle h^{j-3/2} \left[\!\left[\partial^j_\nu v\right]\!\right], h^{j-3/2} \left[\!\left[\partial^j_\nu w\right]\!\right] \right\rangle_{L^2(F)}.$$

Here, ν is a unit normal field on the respective face $F = E_1 \cap E_2$. Moreover, with the restrictions $v_i := v|_{E_i}$, the expressions

$$\left[\!\left[\partial^j_\nu v\right]\!\right] := \sum_{|\alpha|=j} (\partial^\alpha v_1 - \partial^\alpha v_2)\, \nu^\alpha$$

denote the jumps of the normal derivatives along the edge F. The terms in a_{stab} are independent of the orientation of ν due to their quadratic appearance.

The full discretization scheme

For all $v, w \in \mathcal{S}$ let

$$a(v,w) := a_{\Omega_0}(v,w) + a_{\mathrm{pen}}(v,w) + a_{\mathrm{Nit}}(v,w) + a_{\mathrm{Nit}}(v,w) + a_{\mathrm{stab}}(v,w) \tag{3.2.2}$$

and

$$\ell(v) := \ell_0(v) + \ell_{\mathrm{pen}}(v) + \ell_{\mathrm{Nit}}(v). \tag{3.2.3}$$

The finite element approximation $u_h \in \mathcal{S}_*$ of the membrane function u is defined as the solution of the minimization problem

$$\min_{v\in\mathcal{S}_*} \frac{1}{2} a(v,v) - \ell(v)$$

or, equivalently, as the solution of the variational equation

$$\forall v \in \mathcal{S}_* \colon \; a(u_h, v) = \ell(v).$$

Of crucial importance for the later error estimates is the fact that the continuous solution function u fulfills the consistency property

$$\forall v \in H^2_*(\Omega_0) \colon \; a(u, v) = \ell(v). \tag{3.2.4}$$

This is a direct consequence of the construction of the discretization scheme, and in particular relies on the definition of the Nitsche term a_{Nit}. An immediate benefit of this variational equality is that the discrete approximation u_h preserves a Galerkin orthogonality property with respect to the bilinear form a over the space $\mathcal{S}_*$, i. e. it holds

$$\forall v \in \mathcal{S}_* \colon \; a(u, v) = a(u_h, v). \tag{3.2.5}$$

In Appendix D this discretization scheme is motivated from a more pragmatic viewpoint where it is put in the context of the better known Galerkin, penalty and Nitsche schemes. The subsequent sections, however, are independent of those considerations.

3.2.2 Existence and uniqueness of the discrete solution

Before it is possible to prove existence and uniqueness of the discrete approximation u_h within the proposed scheme, it is necessary to show some technical results first.

For the sake of a more concise notation later on, define the *triple-norm on* $\mathcal{S}_*$ by

$$|||v||| := \|v\|_{a_{\Omega_0}+a_{\mathrm{pen}}+a_{\mathrm{stab}}} = \sqrt{a_{\Omega_0}(v,v) + a_{\mathrm{pen}}(v,v) + a_{\mathrm{stab}}(v,v)}.$$

Furthermore, let the *trace norm on* $H^2(\Omega_0)$ be defined by

$$\|v\|_\Gamma := \sqrt{\sum_{i=0}^{1} \|T_i v\|^2_{L^2(\Gamma)}}$$

and define the *weighted extended trace norm* by

$$\|v\|_{\Gamma,h} := \sqrt{\sum_{i=0}^{1} \left\| h^{i-3/2} \widetilde{T}_i v \right\|^2_{L^2(\Gamma)}}.$$

One very crucial component in the upcoming proofs is an estimate for piecewise polynomials that relates the L^2-norms on adjacent elements to each other via the stabilization jump terms along their joint interface. The following result is an adaption of [MLLR14, Lemma 5.1] to the special case at hand.

Lemma 3.2.1. *Suppose $k \in \mathbb{N}$, $\sigma \in \mathbb{R}_{\geq 1}$ and $h, h_1, h_2 \in \mathbb{R}_{>0}$ with $\frac{h_1}{h_2} \leq \sigma$. Let $E_1 := [-h_1, 0] \times [0, h]$ and $E_2 := [0, h_2] \times [0, h]$ be two connected elements with the joint face $F := \{0\} \times [0, h]$. There exists an h- and σ-independent constant $c \in \mathbb{R}_{>0}$ such that for all $v \colon E_1 \cup E_2 \to \mathbb{R}$ with $v|_{E_i} \in \mathcal{Q}_k$ holds*

$$\|v\|^2_{L^2(E_1)} \leq c\sigma \left(\|v\|^2_{L^2(E_2)} + \sum_{j=0}^{k} \frac{h^{2j+1}}{(j!)^2(2j+1)} \left\| \left[\!\left[\partial_\nu^j v \right]\!\right] \right\|^2_{L^2(F)} \right).$$

Proof. Let

$$\pi \colon E_1 \cup E_2 \longrightarrow F, \qquad x \longmapsto \begin{pmatrix} 0 \\ x_2 \end{pmatrix}$$

the orthogonal projection from $E_1 \cup E_2$ onto the joint face F. Furthermore, denote by $v_i \colon \mathbb{R}^2 \to \mathbb{R} \in \mathcal{Q}_k$ the canonical extension of the restriction $v|_{E_i}$ to $\mathbb{R}^2$. Given $x \in E_1$, a Taylor expansion of $v_i(x)$ in $\pi(x)$ leads to

$$v_i(x) = \sum_{j=0}^{\infty} \frac{1}{j!} \frac{\partial^j}{\partial t^j} v_i \left(\pi(x) + t(x - \pi(x)) \right) \Big|_{t=0} = \sum_{j=0}^{k} \frac{x_1^j}{j!} \frac{\partial^j}{\partial x_1^j} v_i(\pi(x)).$$

This implies

$$v_1(x) = v_2(x) + \sum_{j=0}^{k} \frac{x_1^j}{j!} \left(\partial_\nu^j v_1(\pi(x)) - \partial_\nu^j v_2(\pi(x)) \right)$$

and squaring both sides and applying Cauchy–Schwarz inequality yields

$$v_1(x)^2 \leq (k+2) \left(v_2(x)^2 + \sum_{j=0}^{k} \frac{x_1^{2j}}{(j!)^2} \left[\!\left[\partial_\nu^j v(\pi(x)) \right]\!\right]^2 \right).$$

Now, integration of both sides over E_1 leads to

$$\begin{aligned} \frac{1}{k+2} \int_{E_1} v_1(x)^2 \,\mathrm{d}x &\leq \int_{E_1} v_2(x)^2 \,\mathrm{d}x + \sum_{j=0}^{k} \int_{E_1} \frac{x_1^{2j}}{(j!)^2} \left[\!\left[\partial_\nu^j v(\pi(x)) \right]\!\right]^2 \,\mathrm{d}x \\ &= \int_{E_1} v_2(x)^2 \,\mathrm{d}x + \sum_{j=0}^{k} \frac{h^{2j+1}}{(j!)^2(2j+1)} \int_F \left[\!\left[\partial_\nu^j v(\pi(x)) \right]\!\right]^2 \,\mathrm{d}x. \end{aligned}$$

Let $T_1 := [-1, 0] \times [0, 1]$ and $T_2 := [0, 1] \times [0, 1]$ be reference elements associated to E_1 and E_2, respectively. As of the finite-dimensionality of $\mathcal{Q}_k$, the norms $\|\cdot\|_{L^2(T_1)}$ and $\|\cdot\|_{L^2(T_2)}$ are equivalent over $\mathcal{Q}_k$, and in particular there exists a constant $\hat{c} \in \mathbb{R}_{\geq 1}$ such that

$\|w\|^2_{L^2(T_1)} \leq \hat{c}\,\|w\|^2_{L^2(T_2)}$ holds true for all $w \in \mathcal{Q}_k$. By applying this to the transformed function $\hat{v}_2(x_1, x_2) := v_2(h_1 x_1, h x_2) \in \mathcal{Q}_k$ one derives the estimate

$$\begin{aligned}\int_{E_1} v_2(x)^2 \,\mathrm{d}x = h_1 h \int_{T_1} v_2(h_1 x_1, h_2 x_2)^2 \,\mathrm{d}x = h_1 h \,\|\hat{v}_2\|^2_{L^2(T_1)} \\ \leq \hat{c} h_1 h \,\|\hat{v}_2\|^2_{L^2(T_2)} = \hat{c}\sigma \int_{E_2} v_2(x)^2 \,\mathrm{d}x.\end{aligned}$$

Combination of this inequality with the previous one gives the desired statement with the h- and σ-independent constant $c = (k+2)\hat{c}$. □

Lemma 3.2.2. *Let for all $v, w \in H^2(\widetilde{\Omega})$*

$$a_{\widetilde{\Omega}}(v, w) := \sum_{i=1}^{n} \langle \delta_i v, \delta_i w \rangle_{L^2(\widetilde{\Omega})}.$$

There exists an h-independent constant $c_1 \in \mathbb{R}_{>0}$ such that for all $v \in \mathcal{S}$

$$\|v\|_{a_{\widetilde{\Omega}}} \leq c_1 \left(\|v\|_{a_{\Omega_0}} + \|v\|_{a_{stab}} \right).$$

Proof. Let $E_0 \in \mathcal{T}_\Gamma$ and $c_0 := c\sigma \geq 1$ the constant from Lemma 3.2.1 where σ coincides with the shape regularity constant given in Assumption A3. The geometric assumption A2 states the existence of $d' \leq d$ pairwise different elements $E_i \in \mathcal{T}_0$ such that the last element is no boundary element, i. e. $E_{d'} \in \mathcal{T}_0 \setminus \mathcal{T}_\Gamma$, and such that two consecutive elements are always connected by a joint boundary face, i. e. $E_i \cap E_{i+1} \in \mathcal{F}$. Up to d-fold iterative application of Lemma 3.2.1 yields

$$\sum_{i=1}^{n} \|\delta_i v\|^2_{L^2(E_0)} \leq c_0^d \left(\sum_{i=1}^{n} \|\delta_i v\|^2_{L^2(E_{d'})} + n \sum_{i=1}^{d'} \sum_{j=0}^{k} \frac{h^{2j+1}}{(j!)^2(2j+1)} \left\| \left[\!\left[\partial_\nu^j v \right]\!\right] \right\|^2_{L^2(E_{i-1} \cap E_i)} \right).$$

When summing the above inequality over all elements $E \in \mathcal{T}_\Gamma$ Assumption A2 also ensures that each interior element $E \in \mathcal{T}_0 \setminus \mathcal{T}_\Gamma$ and each boundary face $F \in \mathcal{F}$ is visited at most $4d^2$ times. Therefore holds for all $v \in \mathcal{S}$

$$\begin{aligned} a_{\widetilde{\Omega}}(v, v) &= \sum_{E \in \mathcal{T}_0} \sum_{i=1}^{n} \|\delta_i v\|^2_{L^2(E)} \\ &= \sum_{E \in \mathcal{T}_\Gamma} \sum_{i=1}^{n} \|\delta_i v\|^2_{L^2(E)} + \sum_{E \in \mathcal{T}_0 \setminus \mathcal{T}_\Gamma} \sum_{i=1}^{n} \|\delta_i v\|^2_{L^2(E)} \\ &\leq (4c_0^d d^2 + 1) \sum_{E \in \mathcal{T}_0 \setminus \mathcal{T}_\Gamma} \sum_{i=1}^{n} \|\delta_i v\|^2_{L^2(E)} + 4c_0^d d^2 n \sum_{F \in \mathcal{F}} \sum_{j=0}^{k} \frac{h^{2j+1}}{(j!)^2(2j+1)} \left\| \left[\!\left[\partial_\nu^j v \right]\!\right] \right\|^2_{L^2(F)} \\ &\leq (4c_0^d d^2 n + 1) \left(a_{\Omega_0}(v, v) + a_{\mathrm{stab}}(v, v) \right). \end{aligned}$$

In particular, one obtains with $c_1 := \sqrt{4c_0^d d^2 n + 1}$ the inequality

$$\|v\|_{a_{\widetilde{\Omega}}} = \sqrt{a_{\widetilde{\Omega}}(v, v)} \leq \sqrt{4c_0^d d^2 n + 1} \cdot \sqrt{\|v\|^2_{a_{\Omega_0}} + \|v\|^2_{a_{\mathrm{stab}}}} \leq c_1 \left(\|v\|_{a_{\Omega_0}} + \|v\|_{a_{\mathrm{stab}}} \right). \square$$

Lemma 3.2.3. *There exists an h-independent constant $c_2 \in \mathbb{R}_{\geq 1}$ such that for all $v \in \mathcal{S}$ holds*

$$|a_{Nit}(v,v)| \leq c_2 \|v\|_{H^2(\widetilde{\Omega})} \|v\|_{\Gamma,h}$$

Proof. Recall

$$a_{\mathrm{Nit}}(v,w) = \sum_{\substack{|\alpha|\leq 3, |\beta|\leq 1,\\ |\alpha|+|\beta|\leq 3}} \int_\Gamma c_{\alpha\beta}^\Gamma \, \partial^\alpha v \, \partial^\beta w \, \mathrm{d}x$$

and

$$\|v\|_{\Gamma,h}^2 = \sum_{i=0}^{1} \left\| h^{i-3/2} \widetilde{T}_i v \right\|_{L^2(\Gamma)}^2 = \sum_{|\beta|\leq 1} \left\| h^{|\beta|-3/2} \partial^\beta v \right\|_{L^2(\Gamma)}^2 .$$

Let $c_\infty := \max_{\alpha,\beta} \|c_{\alpha\beta}^\Gamma\|_\infty$. Then Cauchy–Schwarz and the inverse inequalities from Assumption A5 lead to the desired estimate

$$\begin{aligned}
|a_{\mathrm{Nit}}(v,v)| &\leq c_\infty \sum_{\substack{|\alpha|\leq 3, |\beta|\leq 1,\\ |\alpha|+|\beta|\leq 3}} \sum_{E\in\mathcal{T}_\Gamma} \left\| h^{\max(0,|\alpha|-3/2)} \partial^\alpha v \right\|_{L^2(\Gamma\cap E)} \left\| h^{\min(0,3/2-|\alpha|)} \partial^\beta v \right\|_{L^2(\Gamma\cap E)} \\
&\leq c_\infty \sum_{\substack{|\alpha|\leq 3, |\beta|\leq 1,\\ |\alpha|+|\beta|\leq 3}} \sum_{E\in\mathcal{T}_\Gamma} c_{\mathrm{inv}} \|v\|_{H^2(E)} \left\| h^{|\beta|-3/2} \partial^\beta v \right\|_{L^2(\Gamma\cap E)} \\
&\leq 29 c_\infty c_{\mathrm{inv}} \|v\|_{H^2(\widetilde{\Omega})} \|v\|_{\Gamma,h} .
\end{aligned}$$

□

Lemma 3.2.4. *Let H be a Hilbert space and $a_1, a_2 \colon H \times H \to \mathbb{R}$ continuous symmetric bilinear forms. Suppose that a_1 has a finite dimensional kernel $\ker a_1$ and that a_1 is coercive on its orthogonal complement $(\ker a_1)^\perp$. If furthermore a_2 is positive semi-definite on H and positive definite on $\ker a_1$, then $a_1 + a_2$ is coercive on H.*

Proof. See [Grä15]. □

Lemma 3.2.5. *There exists an h-independent constant $c_3 \in \mathbb{R}_{\geq 1}$ such that for all functions $v \in H^2(\widetilde{\Omega})$ holds*

$$\|v\|_{H^2(\widetilde{\Omega})} \leq c_3 \left(\|v\|_{a_{\widetilde{\Omega}}} + \|v\|_\Gamma \right)$$

Proof. By Assumption A7 the bilinear form $a_{\widetilde{\Omega}}$ is coercive on $H_0^2(\widetilde{\Omega})$ with an h-independent coercivity constant $\gamma \in \mathbb{R}_{>0}$, i. e.

$$\gamma \|v\|_{H^2(\widetilde{\Omega})}^2 \leq a_{\widetilde{\Omega}}(v,v)$$

is true for all $v \in H_0^2(\widetilde{\Omega})$. From the definition of $a_{\widetilde{\Omega}}$ it is also apparent that its kernel is finite-dimensional with

$$\ker a_{\widetilde{\Omega}} \subseteq \operatorname{span}\{1, x_1, x_2\}$$

whereas the bilinear form inducing the norm $\|\cdot\|_\Gamma$

$$a_\Gamma(v, w) := \langle v, w\rangle_{L^2(\Gamma)} + \langle \partial_\nu v, \partial_\nu w\rangle_{L^2(\Gamma)}$$

is positive semi-definite on $H^2(\widetilde{\Omega})$ and strictly positive on $\ker a_{\widetilde{\Omega}}$.

Altogether, this means that the conditions for Lemma 3.2.4 are met and therefore the bilinear form $a_{\widetilde{\Omega}} + a_\Gamma$ is coercive on $H^2(\widetilde{\Omega})$ with some coercivity constant $\widetilde{\gamma} \in \mathbb{R}_{>0}$. Hence, the desired estimate holds with $c_3 := \widetilde{\gamma}^{-1/2}$.

□

The previous results allow to conclude the following discrete ellipticity property for the bilinear form a with respect to the triple-norm $|||\cdot|||$:

Proposition 3.2.6. *Suppose $\lambda \geq 12(c_1c_2c_3)^2$ where the c_i are the constants from Lemma 3.2.2, Lemma 3.2.3 and Lemma 3.2.5, respectively. Then for all $v \in \mathcal{S}$*

$$\frac{1}{2}|||v|||^2 \leq a(v, v) \leq \frac{3}{2}|||v|||^2.$$

Proof. Application of Lemma 3.2.3, then Lemma 3.2.5, and then Lemma 3.2.2 yields the estimate

$$\begin{aligned}
|a_{\mathrm{Nit}}(v, v)| &\leq c_2 \|v\|_{H^2(\widetilde{\Omega})} \|v\|_{\Gamma,h} \\
&\leq c_2c_3 \left(\|v\|_{a_{\widetilde{\Omega}}} + \|v\|_\Gamma\right) \|v\|_{\Gamma,h} \\
&\leq c_2c_3 \left(c_1 \left(\|v\|_{a_{\Omega_0}} + \|v\|_{a_{\mathrm{stab}}}\right) + \|v\|_\Gamma\right) \|v\|_{\Gamma,h}\,.
\end{aligned}$$

By application of Young's inequality, $h \leq 1$, and from the assumption $\lambda \geq 12(c_1c_2c_3)^2$ one infers furthermore

$$|a_{\mathrm{Nit}}(v, v)| \leq \frac{1}{4}\|v\|_{a_{\Omega_0}}^2 + \frac{1}{4}\|v\|_{a_{\mathrm{stab}}}^2 + 3(c_1c_2c_3)^2\|v\|_{\Gamma,h}^2 \leq \frac{1}{4}|||v|||^2.$$

Plugging this into the definition of the bilinear form a readily gives

$$a(v, v) = a_{\Omega_0}(v, v) + a_{\mathrm{pen}}(v, v) + a_{\mathrm{stab}}(v, v) + a_{\mathrm{Nit}}(v, v) = |||v|||^2 + a_{\mathrm{Nit}}(v, v) \leq \frac{3}{2}|||v|||$$

and analogously also

$$a(v, v) = |||v|||^2 + a_{\mathrm{Nit}}(v, v) \geq \frac{1}{2}|||v|||$$

holds true. This completes the proof. □

A direct consequence of the ellipticity result Proposition 3.2.6 is the existence and uniqueness of the discrete approximation u_h.

Corollary 3.2.7. *Suppose $\lambda \geq 12(c_1c_2c_3)^2$. There exists a unique finite element function $u_h \in \mathcal{S}_*$ such that*

$$\forall v \in \mathcal{S}_* \colon a(u_h, v) = \ell(v).$$

Proof. Apply Lax–Milgram's theorem to the bilinear form a over the discrete space $\mathcal{S}_*$ equipped with the norm $|||\cdot|||$. □

3.2.3 Discretization error estimate

In the following, let $\overline{u} \in \mathcal{S}_*$ be the approximation of the membrane function u in the sense of Assumption A4, and let $\widetilde{u} \in H^r(\widetilde{\Omega})$ the extension of u in the sense of Assumption A8.

From now on it is assumed that λ is large enough for the discrete ellipticity result Proposition 3.2.6 to hold true. All constants $c \in \mathbb{R}_{>0}$ in the results below are stated with the understanding that they are grid-independent.

Lemma 3.2.8. *There exists $c \in \mathbb{R}_{>0}$ such that for all $u \in H^2(\Omega_0)$*

$$\|u\|_{H^2(\Omega_0)} \leq c \, \|u\|_{a_{\Omega_0} + a_{pen}} .$$

Proof. The proof is the same as for Lemma 3.2.5. □

Lemma 3.2.9. *There exists $c \in \mathbb{R}_{>0}$ such that*

$$\|u - \overline{u}\|_{a_{\Omega_0}} + \|u - \overline{u}\|_{a_{pen}} \leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})} .$$

Proof. Let $c_\infty := \sum_{i=1}^n \max_{|\alpha|,|\beta| \leq 2} \|c^i_{\alpha\beta}\|_{L^\infty(\widetilde{\Omega})}$. Then the volume approximation inequality from Assumption A4 and other basic estimates yield

$$\begin{aligned}
\|u - \overline{u}\|^2_{a_{\Omega_0}} &\leq \|\widetilde{u} - \overline{u}\|^2_{a_{\widetilde{\Omega}}} \\
&= \sum_{i=1}^n \sum_{|\alpha|,|\beta| \leq 2} \sum_{E \in \mathcal{T}_0} \int_E c^i_{\alpha,\beta} \partial^\alpha (\widetilde{u} - \overline{u}) \, \partial^\beta (\widetilde{u} - \overline{u}) \, \mathrm{d}x \\
&\leq c_\infty \sum_{|\alpha|,|\beta| \leq 2} \sum_{E \in \mathcal{T}_0} \|\partial^\alpha (\widetilde{u} - \overline{u})\|_{L^2(E)} \, \|\partial^\beta (\widetilde{u} - \overline{u})\|_{L^2(E)} \\
&\leq c' \sum_{E \in \widetilde{\mathcal{T}}} \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|^2_{L^2(E)} \\
&= c' \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|^2_{L^2(\widetilde{\Omega})}
\end{aligned}$$

where $c' \in \mathbb{R}_{>0}$ is a generic grid-independent constant. Similarly, for a suitable constant $c'' \in \mathbb{R}_{>0}$, the trace approximation formula gives

$$\begin{aligned}\|u-\overline{u}\|_{a_{\mathrm{pen}}}^2 &= \|\widetilde{u}-\overline{u}\|_{a_{\mathrm{pen}}}^2 \\ &= \lambda \sum_{|\alpha|\leq 1} \sum_{E\in\mathcal{T}_\Gamma} \|\partial^\alpha(\widetilde{u}-\overline{u})\|_{L^2(E\cap\Gamma)}^2 \\ &\leq c'' \sum_{E\in\mathcal{T}_\Gamma} \left\|h^{k-1}D^{k+1}\widetilde{u}\right\|_{L^2(E)}^2 \\ &= c'' \left\|h^{k-1}D^{k+1}\widetilde{u}\right\|_{L^2(\widetilde{\Omega})}^2 .\end{aligned}$$

Combining these two estimates yields the desired result. □

Lemma 3.2.10. *There exists $c \in \mathbb{R}_{>0}$ such that for all $v \in \mathcal{S}$*

$$a_{Nit}(u-\overline{u},v) + a_{Nit}(v,u-\overline{u}) \leq c \left\|h^{k-1}D^{k+1}\widetilde{u}\right\|_{L^2(\widetilde{\Omega})} |||v|||.$$

Proof. Let

$$c_\infty := \max_{\substack{|\alpha|\leq 3,|\beta|\leq 1\\ |\alpha|+|\beta|\leq 3}} \|c_{\alpha\beta}^\Gamma\|_{L^\infty(\Gamma_0)}.$$

With the trace approximation from Assumption A4, the inequality $|\beta| - \frac{3}{2} \leq \frac{3}{2} - |\alpha|$ for the multi-indices occurring below, and the Cauchy-Schwarz inequality one has for appropriate generic constants $c', c'' \in \mathbb{R}_{>0}$ the chain of estimates

$$\begin{aligned}a_{\mathrm{Nit}}(\widetilde{u}-\overline{u},v) &\leq c_\infty \sum_{E\in\mathcal{T}_\Gamma} \sum_{\substack{|\alpha|\leq 3,|\beta|\leq 1\\ |\alpha|+|\beta|\leq 3}} \|\partial^\alpha(\widetilde{u}-\overline{u})\|_{L^2(\Gamma\cap E)} \left\|\partial^\beta v\right\|_{L^2(\Gamma\cap E)} \\ &\leq c_\infty c' \sum_{E\in\mathcal{T}_\Gamma} \sum_{\substack{|\alpha|\leq 3,|\beta|\leq 1\\ |\alpha|+|\beta|\leq 3}} \left\|h^{\max(|\alpha|-\frac{3}{2},0)+s-|\alpha|-\frac{1}{2}}D^s\widetilde{u}\right\|_{L^2(E)} \left\|h^{\min(\frac{3}{2}-|\alpha|,0)}\partial^\beta v\right\|_{L^2(\Gamma\cap E)} \\ &\leq c_\infty c' \sum_{E\in\mathcal{T}_\Gamma} \sum_{\substack{|\alpha|\leq 3,|\beta|\leq 1\\ |\alpha|+|\beta|\leq 3}} \left\|h^{k-1}D^{k+1}\widetilde{u}\right\|_{L^2(E)} \left\|h^{|\beta|-\frac{3}{2}}\partial^\beta v\right\|_{L^2(\Gamma_0\cap E)} \\ &\leq c'' \left\|h^{k-1}D^{k+1}\widetilde{u}\right\|_{L^2(\widetilde{\Omega})} \|v\|_{a_{\mathrm{pen}}} .\end{aligned}$$

Similarly, one infers with suitable generic constants $c', c'' \in \mathbb{R}_{>0}$ from the inverse trace estimate in Assumption A5, the trace approximation property in Assumption A4, and

the Cauchy-Schwarz inequality

$$\begin{aligned}
a_{\mathrm{Nit}}(v, \widetilde{u} - \overline{u}) &\leq c_\infty \sum_{E \in \mathcal{T}_\Gamma} \sum_{\substack{|\alpha| \leq 3, |\beta| \leq 1 \\ |\alpha| + |\beta| \leq 3}} \|\partial^\alpha v\|_{L^2(\Gamma \cap E)} \left\| \partial^\beta (\widetilde{u} - \overline{u}) \right\|_{L^2(\Gamma \cap E)} \\
&\leq c' \sum_{E \in \mathcal{T}_\Gamma} \sum_{\substack{|\alpha| \leq 3, |\beta| \leq 1 \\ |\alpha| + |\beta| \leq 3}} \left\| h^{\min\left(0, \frac{3}{2} - |\alpha|\right)} v \right\|_{H^2(E)} \left\| h^{k-1-|\beta|-\frac{1}{2}} D^{k+1} \widetilde{u} \right\|_{L^2(E)} \\
&\leq c'' \sum_{E \in \mathcal{T}_\Gamma} \|v\|_{H^2(E)} \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(E)} \\
&\leq c'' \|v\|_{H^2(\widetilde{\Omega})} \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}.
\end{aligned}$$

The $H^2(\widetilde{\Omega})$ coercivity result from Lemma 3.2.5 and the stabilization estimate in Lemma 3.2.2 imply

$$\begin{aligned}
\|v\|_{H^2(\widetilde{\Omega})} &\leq c_3 \left(\|v\|_{a_{\widetilde{\Omega}}} + \|v\|_{a_{\mathrm{pen}}} \right) \\
&\leq c_1 c_3 \left(\|v\|_{a_{\Omega_0}} + \|v\|_{a_{\mathrm{pen}}} + \|v\|_{a_{\mathrm{stab}}} \right) \\
&\leq 3 c_1 c_3 |||v|||.
\end{aligned}$$

Finally, putting all of the above inequalities together proves the assertion. □

Lemma 3.2.11. *There exists $c \in \mathbb{R}_{>0}$ such that*

$$\|\overline{u}\|_{a_{stab}} \leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}.$$

Proof. As of the regularity assumption $\widetilde{u} \in H^{k+1}(\widetilde{\Omega})$, the jump term $[\![\partial_\nu^j \widetilde{u}]\!]$ is well-defined for every order $j \leq k$ and vanishes for every boundary element face $F \in \mathcal{F}$. Using this and the trace approximation error estimate for element faces in Assumption A4 one directly obtains for a suitable constant $c \in \mathbb{R}_{>0}$

$$\begin{aligned}
a_{\mathrm{stab}}(\overline{u}, \overline{u}) &= a_{\mathrm{stab}}(\widetilde{u} - \overline{u}, \widetilde{u} - \overline{u}) \\
&= \sum_{F \in \mathcal{F}} \sum_{j=2}^{k} \frac{1}{(j!)^2 (2j+1)} \left\| h^{j-3/2} [\![\partial_\nu^j (\widetilde{u} - \overline{u})]\!] \right\|_{L^2(F)}^2 \\
&\leq \sum_{E \in \mathcal{T}_0} \sum_{j=2}^{k} \frac{1}{(j!)^2 (2j+1)} \left\| h^{j-3/2} \partial_\nu^j (\widetilde{u} - \overline{u})|_E \right\|_{L^2(\partial E)}^2 \\
&\leq c \sum_{E \in \mathcal{T}_0} \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(E)}^2 \\
&= c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}^2.
\end{aligned}$$

□

Proposition 3.2.12. *There exists $c \in \mathbb{R}_{>0}$ such that*

$$|||\overline{u} - u_h||| \leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}.$$

Proof. For the sake of a simpler notation, let $\overline{e} := u - \overline{u}$ be the interpolation error and $v := u_h - \overline{u} \in \mathcal{S}_*$ be the error between the finite element approximation and the solution's interpolation. Consecutive use of the $|||\cdot|||$-coercivity result in Proposition 3.2.6, the Galerkin orthogonality (3.2.5) of the solution function u and the discrete approximation u_h, the definition of the bilinear form a, and the approximation error estimates Lemma 3.2.9, Lemma 3.2.10 and Lemma 3.2.11 allows one to conclude for a sufficiently large grid-independent constant $c \in \mathbb{R}_{>0}$

$$\begin{aligned}
\frac{1}{2}|||u_h - \overline{u}|||^2 &\leq a(u_h - \overline{u}, u_h - \overline{u}) \\
&= a(u - \overline{u}, u_h - \overline{u}) \\
&= a_{\Omega_0}(\overline{e}, v) + a_{\mathrm{pen}}(\overline{e}, v) + a_{\mathrm{Nit}}(\overline{e}, v) + a_{\mathrm{Nit}}(v, \overline{e}) + a_{\mathrm{stab}}(\overline{u}, v) \\
&\leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})} \left(\|v\|_{a_{\Omega_0} + a_{\mathrm{pen}}} + |||v||| + \|v\|_{a_{\mathrm{stab}}} \right) \\
&\leq 3c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})} |||v|||,
\end{aligned}$$

which immediately implies the claimed statement. □

Theorem 3.2.13. *There exists a grid-independent constant $c \in \mathbb{R}_{>0}$ such that*

$$\|u - u_h\|_{H^2(\Omega_0)} \leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}.$$

Proof. In the following proof let $c \in \mathbb{R}_{>0}$ be a generic constant. The triangle inequality allows to split the discretization error into

$$\|u - u_h\|_{H^2(\Omega_0)} \leq \|u - \overline{u}\|_{H^2(\Omega_0)} + \|\overline{u} - u_h\|_{H^2(\Omega_0)}.$$

From the volume approximation property in Assumption A4 the estimate

$$\|u - \overline{u}\|_{H^2(\Omega_0)} \leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}$$

is immediately apparent. From the coercivity estimate in Lemma 3.2.8 and the approximation property in Proposition 3.2.12 one infers the estimate

$$\|\overline{u} - u_h\|_{H^2(\Omega_0)} \leq c|||\overline{u} - u_h||| \leq c \left\| h^{k-1} D^{k+1} \widetilde{u} \right\|_{L^2(\widetilde{\Omega})}.$$

Combination of all three estimates completes the proof. □

Altogether, the above results prove under the given assumptions the optimal asymptotic convergence behavior

$$\|u - u_h\|_{H^2(\Omega_0)} \in \mathcal{O}(h^{k-1}) \text{ for } h \to 0$$

for quasi-uniform grids.

Remark 3.2.14. *Further important properties in general numerical applications that are related to the above results concern error estimates in the sense of Aubin-Nitsche and Strang as well as bounds on condition numbers for the resulting discrete systems. These aspects play a subordinate role in the remainder of this thesis and therefore their precise treatment is left open. Various results in these directions are readily obtained either via standard techniques or along the lines of the corresponding proofs presented in [BH12].*

3.3 Conforming discretization for point value constraints

This section derives a simple discretization scheme for point value constraints when the membrane parameterization domain Ω_0 is discretized by a matched grid but the evaluation points do not necessarily coincide with nodes in the grid. No new grid types or finite element spaces are introduced in the following paragraphs for the sake of a briefer notation. Instead, it is simply assumed that those are given as before and that the domain Ω_0 is rectangular and resolved exactly. However, the technique described below readily generalizes to rather arbitrary finite element schemes and discrete spaces.

Let the membrane–particle interaction be described by point value constraints and denote by $N \in \mathbb{N}$ the total number of evaluation points $x_i \in \Omega_0$. The membrane function $u \in H^2_*(\Omega_0)$ is then characterized by the solution of the constrained minimization problem

$$\min_{v \in H^2_*(\Omega_0)} \frac{1}{2} a_{\Omega_0}(v,v) - \ell_0(v), \qquad Tv = g$$

where $T\colon H^2(\Omega_0) \to \mathbb{R}^N$ is the evaluation operator in the points $(x_i)_{i=1,\dots,N}$ and $g \in \mathbb{R}^N$ is the vector of corresponding point values.

Given the conforming subspace $\mathcal{S}_* \subseteq H^2_*(\Omega_0)$, the discrete approximation $u_h \in \mathcal{S}_*$ of the solution function u is analogously defined as the solution of the problem

$$\min_{v \in \mathcal{S}_*} \frac{1}{2} a_{\Omega_0}(v,v) - \ell_0(v), \qquad Tv = g$$

or, equivalently due to the ellipticity of a_{Ω_0}, by the variational equation

$$\forall v \in \mathcal{S}_*\colon \; a_{\Omega_0}(u_h, v) = \ell_0(v).$$

Let $M \in \mathbb{N}$ be the dimension of $\mathcal{S}_*$ and let $(\psi_i)_{i=1,\dots,M}$ be a nodal basis of $\mathcal{S}_*$. Then the discrete approximation u_h is expressed via a coefficient vector $\mathbf{u} \in \mathbb{R}^M$ as

$$u_h = \sum_{i=1}^{M} \mathbf{u}_i \psi_i.$$

In the spirit of the standard Galerkin method, define discretizations $A \in \mathbb{R}^{M\times M}$, $b \in \mathbb{R}^M$, $E \in \mathbb{R}^{N\times M}$ of the operators a_{Ω_0}, ℓ_0 and T by

$$A_{ij} := a_{\Omega_0}(\psi_i, \psi_j), \qquad b_i := \ell_0(\psi_i), \qquad E_{ij} := \psi_j(x_i).$$

Assuming that E has full rank $N \leq M$, the coefficient vector $\mathbf{u}$ is uniquely determined by the solution of the finite-dimensional quadratic optimization problem

$$\min_{\mathbf{v}\in\mathbb{R}^M} \frac{1}{2}\mathbf{v}^T A\mathbf{v} - b^T\mathbf{v}, \qquad E\mathbf{v} = g \tag{3.3.1}$$

or, equivalently, as the solution of the linear system of equations

$$\begin{pmatrix} A & E^T \\ E & 0 \end{pmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} b \\ g \end{pmatrix}$$

where $\lambda \in \mathbb{R}^N$. Because in general the constraints are low-dimensional in comparison to the full finite element space, i. e. $N \ll M$, and because the evaluation property is local, i. e. the matrix E is sparse, it is possible to avoid the reformulation of this equation as saddle point problem and instead one may incorporate the constraints directly into the finite element space $\mathcal{S}_*$ in order to obtain a sparse conforming discretization.

To this end, assume without loss of generality that the basis indices are sorted such that the evaluation matrix E and the right hand side g take the block form

$$E = \begin{pmatrix} E_1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & E_l & 0 & \cdots & 0 \end{pmatrix}, \qquad g = \begin{pmatrix} g_1 \\ \vdots \\ g_l \end{pmatrix}$$

where $E_i \in \mathbb{R}^{n_i \times m_i}$ and $g_i \in \mathbb{R}^{n_i}$. Note that as of $N \ll M$ and due to the the local support of the basis functions ψ_i, also $\sum_{i=1}^{l} m_i \ll M$ holds true.

For each $i \in \{1, \ldots, l\}$, let

$$E_i^T = Q^{(i)} R^{(i)} \tag{3.3.2}$$

a QR-decomposition of the transposed block $E_i^T \in \mathbb{R}^{m_i \times n_i}$ with an orthogonal matrix $Q^{(i)} \in \mathbb{R}^{m_i \times m_i}$ and an upper triangular matrix $R^{(i)} \in \mathbb{R}^{m_i \times n_i}$.

The full rank assumption on E implies that the blocks E_i also have full rank $n_i \leq m_i$ such that the R-matrices take the block form

$$R^{(i)} = \begin{pmatrix} \widetilde{R}^{(i)} \\ 0 \end{pmatrix}$$

with invertible upper triangular matrices $\widetilde{R}^{(i)} \in \mathbb{R}^{n_i \times n_i}$. Let $Q_j^{(i)}$ denote the columns of the matrix $Q^{(i)}$ and split it into the matrices

$$\hat{Q}_i := \begin{pmatrix} Q_1^{(i)} & \cdots & Q_{n_i}^{(i)} \end{pmatrix} \in \mathbb{R}^{m_i \times n_i}, \qquad \widetilde{Q}_i = \begin{pmatrix} Q_{n_i+1}^{(i)} & \cdots & Q_{m_i}^{(i)} \end{pmatrix} \in \mathbb{R}^{m_i \times (m_i - n_i)}.$$

Since

$$\begin{aligned} E_i\mathbf{v}_i = g_i &\iff \begin{pmatrix} \left(\widetilde{R}^{(i)}\right)^T & 0 \end{pmatrix} \left(Q^{(i)}\right)^T \mathbf{v}_i = g_i \\ &\iff \begin{pmatrix} I & 0 \end{pmatrix} \left(Q^{(i)}\right)^T \mathbf{v}_i = \left(\widetilde{R}^{(i)}\right)^{-1} g_i \end{aligned}$$

one has

$$E_i \mathbf{v}_i = g_i \iff \exists c \in \mathbb{R}^{m_i - n_i} \colon \ \mathbf{v}_i = \hat{Q}_i \big(\widetilde{R}^{(i)}\big)^{-T} g_i + \widetilde{Q}_i c.$$

In particular, a feasible solution vector $\mathbf{u}^0 \in \mathbb{R}^M$ with $\sum_{i=1}^M \mathbf{u}_i^0 \psi_i \in U_{\mathrm{ad}}$ is given by

$$\mathbf{u}^0 := \begin{pmatrix} \hat{Q}_1 \big(\widetilde{R}^{(1)}\big)^{-T} g_1 \\ \vdots \\ \hat{Q}_l \big(\widetilde{R}^{(l)}\big)^{-T} g_l \end{pmatrix}, \tag{3.3.3}$$

and the kernel of E is spanned by the columns of

$$\widetilde{Q} := \begin{pmatrix} \widetilde{Q}_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & \widetilde{Q}_l & 0 \\ 0 & \cdots & 0 & I \end{pmatrix} \in \mathbb{R}^{M \times (M-N)}. \tag{3.3.4}$$

Therefore, the optimization problem (3.3.1) is equivalent to the reduced problem

$$\min_{\mathbf{v} \in \mathbb{R}^{M-N}} \frac{1}{2} (\widetilde{Q}\mathbf{v} + \mathbf{u}^0)^T A (\widetilde{Q}\mathbf{v} + \mathbf{u}^0) - b^T (\widetilde{Q}\mathbf{v} + \mathbf{u}^0).$$

Analogously to the previous formulations, the minimizer $\widetilde{\mathbf{u}} \in \mathbb{R}^{M-N}$ of this problem is equivalently given by the solution of the linear system

$$\widetilde{A}\widetilde{\mathbf{u}} = \widetilde{b} \tag{3.3.5}$$

where

$$\widetilde{A} := \widetilde{Q}^T A \widetilde{Q}, \qquad \widetilde{b} := \widetilde{Q}^T \left(b - A\mathbf{u}^0\right).$$

The discrete solution vector of the original problem $\mathbf{u}$ is related to the vector $\widetilde{\mathbf{u}}$ via $\mathbf{u} = \widetilde{Q}\widetilde{\mathbf{u}} + \mathbf{u}^0$.

It is readily verified that the matrix $\widetilde{A}$ is again symmetric positive definite. Moreover, it is also sparse as of the sparsity of A and E, and the assembling effort is comparable to the one for the original system matrix A. Hence, computation of $\mathbf{u}$ by solving the system (3.3.5) is numerically feasible and admits the use of typical solvers for symmetric positive definite matrices.

Since one is dealing with an equivalent reformulation of the original Galerkin scheme, the well-known Céa-Lemma is still applicable and yields the estimate

$$\|u - u_h\|_{H^2(\Omega_0)} \leq c \inf_{v \in \mathcal{S}_* \cap U_{\mathrm{ad}}} \|u - v\|_{H^2(\Omega_0)} \leq c \, \|u - \overline{u}\|_{H^2(\Omega_0)}$$

for an h-independent constant $c \in \mathbb{R}_{>0}$ and an approximation function $\overline{u} \in \mathcal{S}_* \cap U_{\mathrm{ad}}$ of the solution u. Therefore, given the regularity assumption $u \in H^r(\Omega_0)$ with $r \in [2, k+1]$, one expects for quasi-uniform grids the asymptotic convergence behavior

$$\|u - u_h\|_{H^2(\Omega_0)} \in \mathcal{O}(h^{r-2}) \text{ for } h \to 0.$$

3.4 Discretization error in the gradient of the interaction potential

The greater goal behind the preceding finite element analysis is to substitute the membrane function u in the derivative formula from Theorem 2.3.3 by the discrete function u_h. This section investigates how this step affects the approximation error in the evaluation of the gradient $\nabla\mathcal{E}(\mathbf{p})$.

To this end, suppose $u \in H^r(\Omega(\mathbf{p}))$ as in Assumption A6 with $r > 2$ and suppose a family of discrete approximations $u_h \in \mathcal{S}_*$ fulfilling the discretization error estimate

$$\|u - u_h\|_{H^2(\Omega(\mathbf{p}))} \leq c \left\|h^{r-2}u\right\|_{H^r(\Omega(\mathbf{p}))} \tag{3.4.1}$$

with a grid-independent constant $c \in \mathbb{R}_{>0}$. Let a direction vector $\mathbf{q} \in \mathbb{R}^{n\times k}$, and functions $V \in H^2(\Omega(\mathbf{p}), \mathbb{R}^2)$, $\xi_0' \in H^2(\Omega(\mathbf{p}))$ be given as required in the derivative result Theorem 2.3.3. For the sake of a simpler notation, further define the joint vector field $W := (V, \xi_0')\colon \Omega(\mathbf{p}) \to \mathbb{R}^3$. As before, a_{Ω_0} is assumed to fulfill Assumption A1 on the structure of the elastic energy. The function $\widetilde{J}$ is the transformed energy as defined in Proposition 2.2.3.

Theorem 3.4.1. *Suppose $h \leq 1$ and $W = (V, \xi_0') \in H^{2,\infty}(\Omega(\boldsymbol{p}), \mathbb{R}^3)$. Moreover, suppose the partial derivatives $\partial_{\boldsymbol{q}}\widetilde{J}(\boldsymbol{p}; \cdot)$ are evaluated via equation (2.3.4) with the given functions V and ξ_0'. Then for a grid- and W-independent constant $C \in \mathbb{R}_{>0}$ holds*

$$\left|\partial_{\boldsymbol{q}}\widetilde{J}(\boldsymbol{p}; u) - \partial_{\boldsymbol{q}}\widetilde{J}(\boldsymbol{p}; u_h)\right| \leq C \|W\|_{H^{2,\infty}(\Omega(\boldsymbol{p}))} \left\|h^{r-2}u\right\|_{H^r(\Omega(\boldsymbol{p}))}.$$

Proof. Suppose $v \in H^2(\Omega(\mathbf{p}))$. By Assumption A1 the integrand ρ associated to the energy J is of the form

$$\rho(x, v(x), Dv(x), D^2v(x)) = \sum_{|\alpha|,|\beta|\leq 2} c_{\alpha\beta}(x)\partial^\alpha v(x)\partial^\beta v(x)$$

with coefficient functions $c_{\alpha\beta} \in C^2(\Omega)$. As a consequence, closer inspection of the derivative formula (2.3.4) shows that the directional derivative $\partial_{\mathbf{q}}\widetilde{J}(\mathbf{p}; v)$ takes the form

$$\partial_{\mathbf{q}}\widetilde{J}(\mathbf{p}; v) = \sum_{i=1}^{3} \sum_{|\alpha|,|\beta|,|\gamma|\leq 2} \int_{\Omega(\mathbf{p})} c^{(i)}_{\alpha\beta\gamma}\partial^\alpha v\, \partial^\beta v\, \partial^\gamma W_i \,\mathrm{d}x$$

with coefficient functions $c^{(i)}_{\alpha\beta\gamma} \in C^2(\Omega)$. In particular, there exist a continuous bilinear form $\widetilde{a}\colon \left(H^2(\Omega(\mathbf{p}))\right)^2 \to \mathbb{R}$, a continuous linear map $\widetilde{\ell} \in \left(H^2(\Omega(\mathbf{p}))\right)'$ and a constant $\widetilde{c} \in \mathbb{R}$ such that

$$\partial_{\mathbf{q}}\widetilde{J}(\mathbf{p}; v) = \widetilde{a}(v, v) + \widetilde{\ell}(v) + \widetilde{c},$$

and for the operators $\widetilde{a}$, $\widetilde{\ell}$ holds

$$\max\left\{\|\widetilde{a}\|, \|\widetilde{\ell}\|\right\} \leq C' \|W\|_{H^{2,\infty}(\Omega(\mathbf{p}))}$$

with a grid- and W-independent constant $C' \in \mathbb{R}_{>0}$. Given that and the approximation property (3.4.1), this admits the estimate

$$\begin{aligned}\left|\partial_{\mathbf{q}}\widetilde{J}(\mathbf{p};u) - \partial_{\mathbf{q}}\widetilde{J}(\mathbf{p};u_h)\right| &= |\widetilde{a}(u+u_h, u-u_h)| + \left|\widetilde{\ell}(u-u_h)\right| \\ &\leq \|\widetilde{a}\| \, \|u+u_h\|_{H^2(\Omega(\mathbf{p}))} \, \|u-u_h\|_{H^2(\Omega(\mathbf{p}))} + \|\widetilde{\ell}\| \, \|u-u_h\|_{H^2(\Omega(\mathbf{p})} \\ &\leq C'\left(1 + \|u+u_h\|_{H^2(\Omega(\mathbf{p}))}\right) \cdot c \left\|h^{r-2}u\right\|_{H^r(\Omega(\mathbf{p}))}.\end{aligned}$$

Noting

$$\begin{aligned}\|u+u_h\|_{H^2(\Omega(\mathbf{p}))} &\leq 2\|u\|_{H^2(\Omega(\mathbf{p}))} + \|u-u_h\|_{H^2(\Omega(\mathbf{p}))} \\ &\leq (2+c)\left\|h^{r-2}u\right\|_{H^r(\Omega(\mathbf{p}))} \\ &\leq (2+c)\|u\|_{H^r(\Omega(\mathbf{p}))}\end{aligned}$$

this proves the claimed assertion with $C := (3+c)c \cdot C'$. □

Consider now the approximation $G \in \mathbb{R}^{n\times k}$ of the gradient $\nabla\mathcal{E}(\mathbf{p})$ whose components are defined as

$$G^h_{(i,j)} := \partial_{(i,j)}\widetilde{J}^{(i,j)}(\mathbf{p};u_h), \qquad i = 1,\dots,n, \quad j = 1,\dots,k..$$

Here, $\partial_{(i,j)}$ denotes differentiation with respect to the (i,j)-th standard unit vector in the configuration space $\mathbb{R}^{n\times k}$ and $\widetilde{J}^{(i,j)}$ is the transformed energy generated by an associated vector field $V^{(i,j)}$. The evaluation of the partial derivatives $\partial_{(i,j)}\widetilde{J}^{(i,j)}(\mathbf{p};u_h)$ is carried out via the derivative formula (2.3.4) with an associated constraint function $(\xi_0')^{(i,j)}$. Application of Theorem 3.4.1 directly shows the following error estimate for quasi-uniform grids:

Corollary 3.4.2. *Suppose $V^{(i,j)} \in H^{2,\infty}(\Omega(\boldsymbol{p}),\mathbb{R}^2)$ and $(\xi_0')^{(i,j)} \in H^{2,\infty}(\Omega(\boldsymbol{p}))$ for every pair (i,j). If $V^{(i,j)} \in \mathcal{O}(1)$ and $(\xi_0')^{(i,j)} \in \mathcal{O}(1)$ for $h \to 0$, then also*

$$\left\|\nabla\mathcal{E}(\boldsymbol{p}) - G^h\right\| \in \mathcal{O}(h^{r-2}) \text{ for } h \to 0.$$

4 Implementation and numerical experiments

This chapter is divided into two parts: The first one sketches an implementation for the computation of stationary membranes in linearized particle–membrane interaction models based on the established finite element discretizations for linearized energies in Chapter 3. In addition, it introduces a level set ansatz for the description of the configuration-dependent domains and interfaces. Such an implementation is then used in the second part in order to illustrate the previously derived discretization error estimates for finite element approximations of stationary membranes within various different models.

4.1 Implementation remarks

4.1.1 Level set description of parametric interfaces and constraints

While the overall grid in the discretization is chosen independently of the particle configuration $\mathbf{p}$, it is still necessary to keep track of the interface $\Gamma(\mathbf{p})$ and the boundary conditions $g(\mathbf{p})$ for assembling purposes. In the case of point value constraints this is done relatively easily, however for curve constraints more work is required, which this section elaborates on. The core idea is to represent the $\Gamma(\mathbf{p})$ interfaces via level sets. For simplicity of notation, only one single particle is considered, i. e. $\mathbf{p} \in \mathbb{R}^{n\times k}$ with $n = 1$. The generalization to multiple particles is straight-forward.

First, recall and re-state the relevant definitions in the spirit of Sections 1.2.2 and 1.2.3. For the particle model, suppose a particle reference hypersurface $\widetilde{\mathcal{G}}_0 \subseteq \mathbb{R}^3$, a reference interface $\Gamma_0 \subseteq \mathbb{R}^2$ in the plane and a diffeomorphism $\gamma\colon \widetilde{\Gamma}_0 \to \operatorname{im}(\gamma) \subseteq \widetilde{\mathcal{G}}_0$ that is defined in a neighborhood $\widetilde{\Gamma}_0$ of Γ_0. The set $\gamma(\Gamma_0)$ describes the interface along which the membrane connects to the particle and therefore induces the later constraints and interfaces. A smooth particle transformation map $\Psi\colon \mathbb{R}^k \times \mathbb{R}^3 \to \mathbb{R}^3$ is postulated such that $\Psi(\mathbf{p};\cdot)$ is bijective and the transformed particle hypersurface is defined by $\widetilde{\mathcal{G}}(\mathbf{p}) := \Psi(\mathbf{p};\widetilde{\mathcal{G}}_0)$. In particular, the transformed membrane–particle interface is given by

$$\mathcal{G}(\mathbf{p}) := \Psi\left(\mathbf{p};\begin{pmatrix}\Gamma_0\\ \gamma(\Gamma_0)\end{pmatrix}\right).$$

Concerning the membrane model, assume that a reference surface $\mathcal{M}_0$, a parameterization $\varphi\colon \Omega \to \mathcal{M}_0$, a local projection π and a local distance function d are given. The inverse local projection onto $\widetilde{\mathcal{G}}(\mathbf{p})$ is again denoted by $\Pi(\mathbf{p})$, and it is assumed that all of

the following expressions containing these local functions are well-defined in sufficiently large neighborhoods of the interfaces.

The transformed particle-interface on the reference domain Ω is defined by

$$\Gamma(\mathbf{p}) := \varphi^{-1}\left(\pi\left(\Psi\left(\mathbf{p};\begin{pmatrix}\Gamma_0\\ \gamma(\Gamma_0)\end{pmatrix}\right)\right)\right),$$

i. e. the transformed membrane–particle interface $\mathcal{G}(\mathbf{p})$ is projected onto the reference manifold $\mathcal{M}_0$ and then pulled back to the parameterization domain Ω.

In order to express the interface $\Gamma(\mathbf{p})$ via a level set function, suppose first a level set function $f_0\colon \mathbb{R}^2 \to \mathbb{R}$ such that its zero level set describes the reference interface Γ_0, i. e.

$$\Gamma_0 = \left\{x \in \mathbb{R}^2 \mid f_0(x) = 0\right\}$$

holds. Moreover, it is assumed for later that $f_0 > 0$ holds in the interior of Γ_0 and $f_0 < 0$ on the outside.

By considering the inverse of the function in the definition of $\Gamma(\mathbf{p})$, a zero level set function for $\Gamma(\mathbf{p})$ is given by

$$f(\mathbf{p}; x) := f_0(\widetilde{y}(\mathbf{p}; x))$$

where

$$y(\mathbf{p}; x) := \Psi^{-1}\left(\mathbf{p}; \amalg\left(\mathbf{p}; \varphi(x)\right)\right), \qquad \widetilde{y}(\mathbf{p}; x) = \begin{pmatrix} y_1(\mathbf{p}; x)\\ y_2(\mathbf{p}; x)\end{pmatrix}.$$

If no such y exists, for example because x is "too far away" from the interface $\Gamma(\mathbf{p})$, then let $f(\mathbf{p}; x)$ stay undefined, i. e. $f(\mathbf{p}; x) = \mathrm{NaN}$ in programming terms.

The transformed constraint function $g(\mathbf{p}) = (g_1(\mathbf{p}), g_2(\mathbf{p}))$ is defined on the interface $\Gamma(\mathbf{p})$ and extended to the full domain Ω by

$$g_1(\mathbf{p}) := d \circ \amalg(\pi) \circ \varphi, \qquad g_2(\mathbf{p}) := \partial_{\widetilde{\nu}} g_1(\mathbf{p})$$

where $\widetilde{\nu}$ is an arbitrary extension of the outer unit normal field ν on $\partial\Omega(\mathbf{p})$ to a vector field in all of $\Omega(\mathbf{p})$. Hence, $g_1(x)$ simply is the distance of the point $\varphi(x) \in \mathcal{M}_0$ in direction of the normal $\nu_{\mathcal{M}_0}$ on $\mathcal{M}_0$ to the transformed surface $\widetilde{\mathcal{G}}(\mathbf{p})$. This means that if $g_1(x)$ is defined it has the property

$$\Psi\left(\mathbf{p}; \begin{pmatrix}\widetilde{y}(\mathbf{p}; x)\\ \gamma(\widetilde{y}(\mathbf{p}; x))\end{pmatrix}\right) = \varphi(x) + g_1(x)\nu_0(\varphi(x)). \tag{4.1.1}$$

Again, if $g(\mathbf{p}; x)$ is not well-defined in a point $x \in \Omega$, let $g(\mathbf{p}; x) = \mathrm{NaN}$.

Within a practical implementation all relevant quantities are readily accessible once it is clear how to evaluate $\widetilde{y}$ and g_1 and an extension vector field $\widetilde{\nu}$ is constructed. Under the assumption that the equation (4.1.1) may be used to characterize $\widetilde{y}(\mathbf{p}; x)$ and $g_1(\mathbf{p}; x)$, Algorithm 1 is proposed.

Algorithm 1 Evaluation of $\widetilde{y}(\mathbf{p};x)$ and $g_1(\mathbf{p};x)$

1: Let $\mathrm{tol} \in \mathbb{R}_{>0}$ and

$$F(\widetilde{x},t) := \frac{1}{2}\left\|\Psi\left(\mathbf{p};\begin{pmatrix}\widetilde{x}\\ \gamma(\widetilde{x})\end{pmatrix}\right) - \varphi(x) - t\nu_{\mathcal{M}_0}(\varphi(x))\right\|^2.$$

2: Minimize F (e. g. using Newton's method) yielding final iterates $(\widetilde{x}^*, t^*)$.

3: **if** $F(\widetilde{x}^*, t^*) < \mathrm{tol}$ **then**
4: **return** $(\widetilde{x}^*, t^*)$
5: **else**
6: **return** NaN
7: **end if**

With this algorithm at hand the evaluation of $f(\mathbf{p})$ is possible as well, which enables the approximation of $\Gamma(\mathbf{p})$ and the computation of the sub-grids $\mathcal{T}_0$ and $\mathcal{T}_\Gamma$ by standard level set methods. Furthermore, an extension of the unit normal vector field is given in a neighborhood of the interface $\Gamma(\mathbf{p})$ by

$$\widetilde{\nu}(\mathbf{p};x) := \frac{\nabla f(\mathbf{p};x)}{\|\nabla f(\mathbf{p};x)\|}.$$

Finally, whenever needed, differentiation of $\widetilde{y}$ and g_1 can for example be carried out through a sensitivity analysis of the corresponding implicit equation (4.1.1) or also only approximatively via classical difference methods.

4.1.2 Approximation algorithm for stationary membranes

Suppose a configuration $\mathbf{p} \in \mathcal{D}^\circ$ and a grid size parameter $h \in (0,1]$ as well as a rectangular bulk domain $\Omega \supseteq \Omega(\mathbf{p})$. A prototypical algorithm for the discrete approximation of the stationary membrane $u(\mathbf{p})$ subject to curve constraints is described in Algorithm 2. The analogous prototypical algorithm for point constraints in the sense of Section 3.3 is shown in Algorithm 3.

Corresponding implementations are done in DUNE, a C++ toolbox for solving partial differential equations [BBD+08b, BBD+08a, BBD+16]. The grids are created as uniform grids of rectangles using the DUNE-UGGRID module (cf. [BBJ+97]), and are equipped with Bogner–Fox–Schmit finite elements. Interface-dependent quantities are computed by a level set ansatz in the sense of Section 4.1.1. The assembly of the system itself and the incorporation of remaining boundary conditions is then realized by standard techniques. The line integrals along $\Gamma(\mathbf{p})$ are approximated by means of the techniques described in [AV93]. For the volume quadrature on $\mathcal{T}_\Gamma$ element-specific quadrature rules for implicitly defined domains in the sense of [OS16] are applied, and on $\mathcal{T}_0 \setminus \mathcal{T}_\Gamma$ quadrature is performed exactly using classical Gauss quadrature rules. Finally, the linear system is solved using the direct solver SUPERLU [Li05].

Algorithm 2 Approximation of a stationary membrane for curve constraints

1: Create an h-dependent grid $\mathcal{T}$ of rectangles covering the bulk domain Ω.
2: Equip the grid with a $\mathcal{Q}_k$ finite element space and basis functions φ_i, $i = 1, \dots, M$.
3: Compute the required $\mathbf{p}$-dependent geometric information, such as $\mathcal{T}_0$, $\mathcal{T}_\Gamma$, quadrature points on $\mathcal{T}_\Gamma$ and values of the constraint right hand sides $g(\mathbf{p})$.
4: Assemble the system matrix $A \in \mathbb{R}^{M \times M}$ and right hand side vector $b \in \mathbb{R}^M$ over the sub-grid $\mathcal{T}_0$ where

$$A_{ij} := a(\varphi_i, \varphi_j), \qquad b_i := \ell(\varphi_i)$$

is defined in the sense of the fictitious domain stabilized Nitsche method (3.2.2) and (3.2.3).
5: Modify A and b in order to incorporate periodic and/or homogeneous Dirichlet boundary conditions along the outer boundary $\partial\Omega$.
6: Solve the linear system $A\mathbf{u} = b$.

Algorithm 3 Approximation of a stationary membrane for point constraints

1: Create an h-dependent grid $\mathcal{T}$ of rectangles covering the bulk domain Ω.
2: Equip the grid with a $\mathcal{Q}_k$ finite element space and basis functions φ_i, $i = 1, \dots, M$.
3: Compute the required $\mathbf{p}$-dependent geometric information, such as the constraint right hand side $g(\mathbf{p}) \in \mathbb{R}^N$.
4: Compute a transformed basis $\widetilde{\varphi}_i$, $i = 1, \dots, M - N$ based on local QR-decompositions (3.3.2) and set up the basis transformation matrix $\widetilde{Q} \in \mathbb{R}^{M \times (M-N)}$ as in (3.3.4).
5: Compute a feasible vector $\mathbf{u}^0 \in \mathbb{R}^M$ by formula (3.3.3).
6: Assemble the system matrix $\widetilde{A} \in \mathbb{R}^{(M-N)\times(M-N)}$ and right hand side vector $\widetilde{b} \in \mathbb{R}^{M-N}$ over the sub-grid $\mathcal{T}_0$ where

$$\widetilde{A}_{ij} := a_{\Omega(\mathbf{p})}(\widetilde{\varphi}_i, \widetilde{\varphi}_j), \qquad \widetilde{b}_i := \ell(\mathbf{p}; \widetilde{\varphi}_i) - \sum_{j=1}^{M} a_{\Omega(\mathbf{p})}(\mathbf{u}_j^0 \varphi_j, \varphi_i)$$

is defined in the sense of equation (3.3.5)
7: If necessary, modify $\widetilde{A}$ and $\widetilde{b}$ in order to incorporate periodic and/or homogeneous Dirichlet boundary conditions along the outer boundary $\partial\Omega$.
8: Solve the linear system $\widetilde{A}\widetilde{\mathbf{u}} = \widetilde{b}$ and set $\mathbf{u} := \widetilde{Q}\widetilde{\mathbf{u}} + \mathbf{u}^0$.

It is noted that the implementation restricts itself to a simple level set algorithm. Its precise treatment is omitted for the sake of brevity, but it is emphasized that it works sufficiently well for the subsequent examples and applications. However, in rare situations this implementation fails to reconstruct the full interfaces, leading to disproportionate discretization errors. Such issues will be addressed appropriately when they occur. More sophisticated methods and algorithms for improved level set computations are presented for example in [Set99].

4.1.3 Evaluation of gradients of the interaction potential

In view of the derivative formula in Proposition 2.3.2 for the computation of directional derivatives, the evaluation of the gradient $\nabla\mathcal{E}(\mathbf{p})$ of the interaction potential requires the construction of one vector field V and one constraint-preserving function ξ_0' for each degree of freedom in the configuration parameter $\mathbf{p} \in \mathbb{R}^{n\times k}$, i. e. $n \cdot k$ constructions are necessary.

For the sake of a shorter notation those functions are taken together into a single vector field $W := (V, \xi)\colon \Omega(\mathbf{p}) \to \mathbb{R}^3$. Following Theorem 2.3.3, the evaluation of $\partial_{(i,j)}\mathcal{E}(\mathbf{p})$, $i = 1, \dots, n$, $j = 1, \dots, k$ requires the vector field W to fulfill the boundary conditions on $\Gamma(\mathbf{p})$

$$W|_{\Gamma(\mathbf{p})} = \begin{pmatrix} \partial_i\psi(\mathbf{p}) \\ \partial_i(g_1(\mathbf{p}) \circ \psi(\mathbf{p})) \end{pmatrix}, \qquad DW\nu|_{\Gamma(\mathbf{p})} = \begin{pmatrix} \partial_i(\widetilde{\nu}(\mathbf{p}) \circ \psi(\mathbf{p}) \\ \partial_i(g_2(\mathbf{p}) \circ \psi(\mathbf{p})) \end{pmatrix} \tag{4.1.2}$$

where the partial derivatives $\partial_{(i,j)}$ here are always understood with respect to the $\mathbf{p}$-variable. In the spirit of Remark 2.3.4 the conditions on the $\mathbf{p}$-independent boundary $\partial\Omega$ read

$$\begin{aligned} W|_{\Gamma_D} &= 0 & W|_{\Gamma_{P,1}} &= W|_{\Gamma_{P,2}} \\ DW\nu|_{\Gamma_D} &= 0 & DW\nu|_{\Gamma_{P,1}} &= -DW\nu|_{\Gamma_{P,2}} \end{aligned} \tag{4.1.3}$$

where Γ_D denotes the $\mathbf{p}$-independent Dirichlet boundary and $\Gamma_P = \Gamma_{P,1}\dot\cup\Gamma_{P,2}$ is the periodic boundary.

In general, one is free in the choice of extension for W, however the error result Corollary 3.4.2 on the approximation of the gradient suggests to choose W such that the derivatives up to second order are bounded grid-independently. This may be achieved for example if the components W_l, $l = 1, 2, 3$ each solve the minimization problem

$$\min_{v\in H^2(\Omega(\mathbf{p}))} \frac{1}{2} a_{\Omega(\mathbf{p})}(v, v), \qquad \text{s. t. } v \text{ fulfills the } l\text{-th boundary condition in (4.1.2) and (4.1.3).}$$

This problem is essentially identical to the minimization problem associated to the stationary membrane $u(\mathbf{p})$, but with different right hand sides in the constraints. In particular, the same discretization schemes are applicable for the approximative computation of the components $W_{l,h} \approx W_l$ as for the computation of the approximation $u_h \approx u(\mathbf{p})$. Within the setting of curve constraints, the required smoothness

property $W \in H^{2,\infty}(\Omega_0, \mathbb{R}^3)$ in Corollary 3.4.2 is fulfilled under the typical smoothness conditions for the boundary data from elliptic regularity theory, which generally hold true in the upcoming example computations. For point constraints, however, this regularity property is not obtained as easily and hence the behavior of this extension approach technically requires further investigation.

In conclusion, the evaluation of the gradient may for example be implemented along the lines of Algorithm 4.

Algorithm 4 Approximation G of the gradient of the interaction potential $\nabla\mathcal{E}(\mathbf{p})$

1: Suppose an approximation vector $\mathbf{u}$ is known from Algorithm 2 (or Algorithm 3).
2: **for** $i = 1, \dots, n$ and $j = 1, \dots, k$ **do**
3: **for** $l = 1, 2, 3$ **do**
4: Assemble a right hand side $b^{(l)}$ as in Algorithm 2 (or Algorithm 3) incorporating the boundary conditions from W_l.
5: Approximate W_l by solving the system $A\mathbf{W}^{(l)} = b^{(l)}$ (or $\widetilde{A}\mathbf{W}^{(l)} = b^{(l)}$).
6: **end for**
7: Compute $G_{(i,j)} \approx \partial_{(i,j)}\mathcal{E}(\mathbf{p})$ by carrying out the integration in the volume formula (2.3.4) with the approximations $u_h \approx u$ and $W_h \approx (V, \xi)$.
8: **end for**
9: **return** G

Within the DUNE implementation of this algorithm the solution of the involved linear systems is comparatively inexpensive because a factorization of A (or $\widetilde{A}$) is already known from the computation of the approximation vector $\mathbf{u}$ via the direct solver SUPERLU, and a similar efficiency should be achievable by using e. g. iterative solvers based on geometric multigrid methods. Moreover, the quantities required for the assembly of the right hand sides are indirectly available from the considerations in Section 4.1.1, and the integration is done by similar quadrature rules as for the assembly of the system matrices.

4.2 Curve constraints: Optimal orders of convergence

4.2.1 Radially symmetric example

This example considers the Monge-gauge parameterization with bending rigidity $\kappa = 1$ and vanishing surface tension $\sigma = 0$ coupled to curve constraints. The reference domain is assumed to be a two-dimensional R-sphere, i. e. such that $\Omega = B_R \subseteq \mathbb{R}^2$ holds, and the projection of the particle–membrane-interface onto Ω is supposed to be circular and centered around the origin, i. e. such that $\Gamma = \partial B_r$ holds. In addition, it is assumed that the outer boundary $\partial\Omega$ fulfills zero Dirichlet boundary conditions and that the particle induces radially symmetric Dirichlet boundary conditions along the interface Γ. Effectively, this leaves the domain of computation $\Omega_0 := B_R \setminus B_r$. Altogether, this

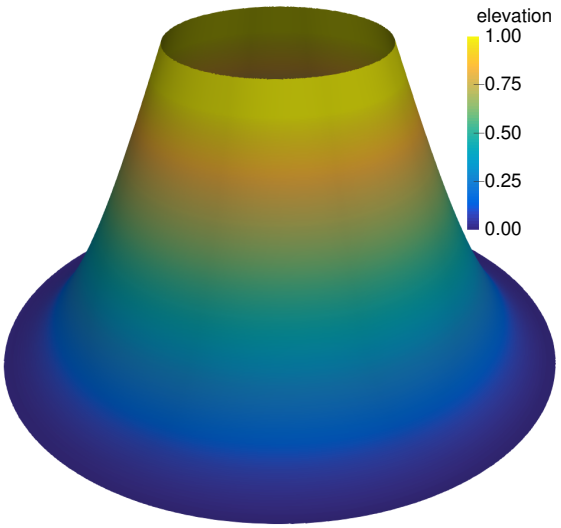

Figure 4.2.1: Exact solution of the example problem stated in Section 4.2.1. The coloring encodes the function's elevation.

reduces the computation of the stationary membrane to the solution of the fourth order elliptic partial differential equation with Dirichlet boundary conditions

$$\Delta^2 u = 0 \text{ on } \Omega_0, \qquad Tu = \begin{cases} (0,0) & \text{on } \partial B_R \\ (c_1, c_2) & \text{on } \partial B_r \end{cases}$$

for some constants $c_1, c_2 \in \mathbb{R}$.

Due to radial symmetry the solution takes the form $u(x) = u_0(\|x\|)$ where u_0 solves the differential equation

$$u_0''''(t) + \frac{2}{t} u_0'''(t) - \frac{1}{t^2} u_0''(t) + \frac{1}{t^3} u_0'(t) = 0$$

subject to the boundary conditions

$$u_0(r) = c_1, \qquad u_0'(r) = c_2, \qquad u(R) = 0, \qquad u_0'(R) = 0.$$

For the numerical computations the special choices $r = \frac{1}{e}$, $R = 1$, $c_1 = \frac{e^2-3}{4}$, and $c_2 = -e$ are made. The solution u_0 is therefore given by

$$u_0(t) = \frac{e^2}{4}\left(1 - t^2 + 2t^2 \log(t)\right).$$

See Fig. 4.2.1 for a visualization of the resulting stationary membrane u.

The numerical approximation of the stationary membrane is performed using the fictitious domain stabilized Nitsche method applied to a uniform background mesh of

square elements over the domain $[-1,1]^2$ with varying grid sizes $h \in \mathbb{R}_{>0}$. In view of the error results from Section 3.2.3, the smoothness of u over Ω_0, and the approximation properties of Bogner–Fox–Schmit finite elements, one expects within the given DUNE implementation the asymptotic convergence

$$\|u - u_h\|_{H^2(\Omega_0)} \in \mathcal{O}(h^2) \text{ for } h \to 0.$$

This convergence behavior is explored numerically in Fig. 4.2.2 where the discretization errors $\left\|D^i(u-u_h)\right\|_{L^2(\Omega_0)}$ are plotted over the range of grid sizes $h = \frac{2}{N}$ for $i \in \{0,1,2\}$ and $N \in \{16,\dots,512\}$. Each error was computed via a high order quadrature rule with respect to the given grid, and the penalty weighting parameter $\lambda = 1024$ was used within the discretization scheme.

The H^2-norm of the error generally is in good agreement with the reference line and thus confirms the predicted convergence order 2. Similarly, the plots of the errors in the L^2-norm and the H^1-norm closely follow their respective reference lines and so show the convergence orders 4 and 3, respectively. This is the typical expected behavior where every order less of differentiation in the error norm yields an extra order of convergence. Altogether, this corresponds to an optimal convergence behavior.

However, there are outliers within the error plots that seemingly disagree with the observed convergence behavior, especially in the H^2-norm. As the theory states grid-independence of the constants in the approximation estimates for the given discretization scheme, these outliers are assumed to be attributable to instabilities in the used implementation of the level set description and the resulting errors in the quadrature rules. Similarly, the divergence of the L^2-error around the threshold $2^{-31} \approx 5 \cdot 10^{-10}$ might be explained by rounding errors and numerical instabilities in the quadrature rules within the computation of the L^2-norm.

4.2.2 Non-symmetric example

This example investigates the rates of convergence in a more complicated, non-symmetric case with two particles. As before, the Monge-gauge parameterization with bending rigidity $\kappa = 1$ and vanishing surface tension $\sigma = 0$ as well as zero Dirichlet boundary conditions on the outer boundary $\partial\Omega$ is considered. However, this time the parameterization domain is given by $\Omega = [-1,1]^2$ and the particle–membrane interface $\Gamma(\mathbf{p})$ is induced by the following elliptic particle model: Suppose the semi axes are $a = \frac{1}{3}$ and $b = \frac{2}{9}$ and the slope is $s = 1$. The reference particle manifold is given by

$$\widetilde{\mathcal{G}}_0 := \left\{ x \in \mathbb{R}^3 \;\middle|\; \frac{2x_3}{as} = 1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} \right\}$$

and the reference particle–membrane interface is defined as

$$\mathcal{G}_0 := \left\{ x \in \mathbb{R}^2 \times \{0\} \;\middle|\; \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \right\} \subseteq \widetilde{\mathcal{G}}_0.$$

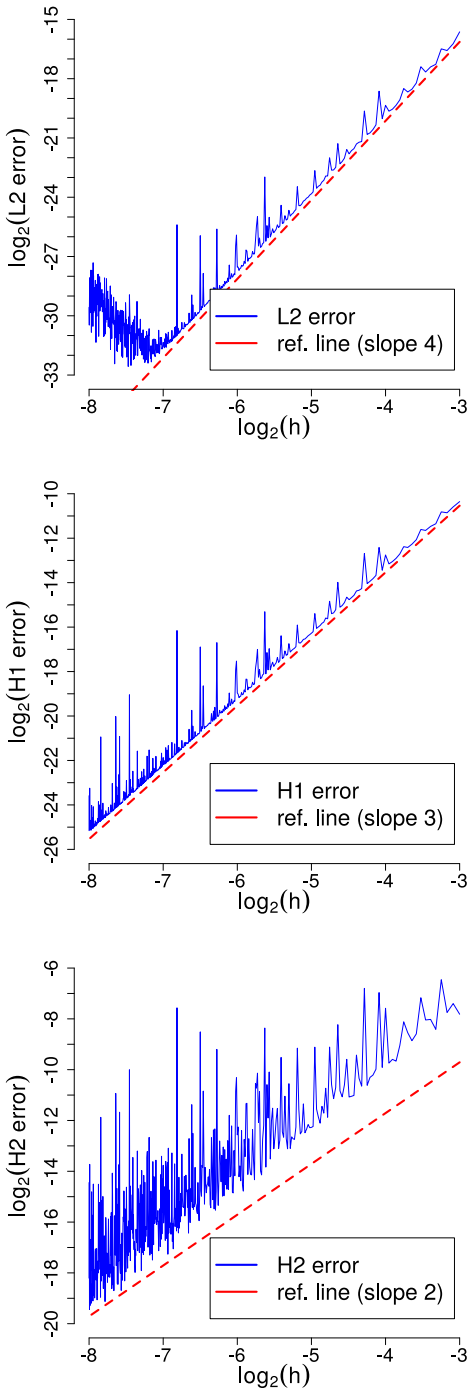

Figure 4.2.2: Double logarithmic plot of the L^2-, H^1- and H^2 discretization error (from top to bottom) over the range $h \in [1/256, 1/8]$ of grid sizes for the example problem in Section 4.2.1.

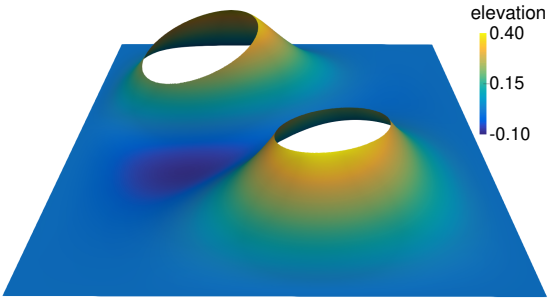

Figure 4.2.3: Fine grid solution of the example problem stated in Section 4.2.2. The coloring encodes the function's elevation.

Furthermore, each particle is equipped with six degrees of freedom describing both its position and the orientation. Hence, following Example 1.2.1 the transformation map may be defined by

$$\Psi(p;x) := R_3(p_6)R_2(p_5)R_1(p_4)x + \begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix}^T$$

where $R_i(\alpha)$ denotes the counter-clockwise rotation around the x_i-axis by the angle α. The particle configuration parameter chosen in the example is

$$\mathbf{p} := \begin{pmatrix} -0.4 & 0.4 & 0.3 & 10^\circ & 15^\circ & 30^\circ \\ 0.25 & -0.35 & 0.3 & -10^\circ & -15^\circ & 60^\circ \end{pmatrix}.$$

Altogether, the complete problem reads

$$\Delta^2 u = 0 \text{ on } \Omega(\mathbf{p}), \qquad u = \partial_\nu u = 0 \text{ on } \partial\Omega, \qquad \begin{pmatrix} u \\ \partial_\nu u \end{pmatrix} = g(\mathbf{p}) \text{ on } \Gamma(\mathbf{p})$$

with an implicit interface $\Gamma(\mathbf{p})$ and boundary conditions $g(\mathbf{p})$. A discrete solution of this problem is depicted in Fig. 4.2.3.

As for the symmetric example, the solution is approximated numerically via the fictitious domain stabilized Nitsche method applied to a uniform mesh of squares. From regularity theory also the smoothness of u is apparent and thus again quadratic convergence of the approximation in the H^2-norm is expected. As no analytical solution u is available, the error computations approximate it by a fine grid solution u_H with $H = 1/128$. Corresponding plots of the approximate errors $\left\|D^i(u_H - u_h)\right\|_{L^2(\Omega_0)}$, $i \in \{0,1,2\}$ are shown in Fig. 4.2.4 for the grid sizes $h = \frac{1}{N}$, $N \in \{16,\dots,127\}$. Each error was computed by a high order quadrature rule with respect to the given fine grid. The penalty weighting parameter was set to $\lambda = 1024$.

Like in the symmetric case, the error plots follow the given reference lines relatively well and thus exhibit the convergence orders 4, 3 and 2 for the L^2-, H^1- and H^2-norm of the error, respectively. Thus, optimal convergence is observed in this example as well.

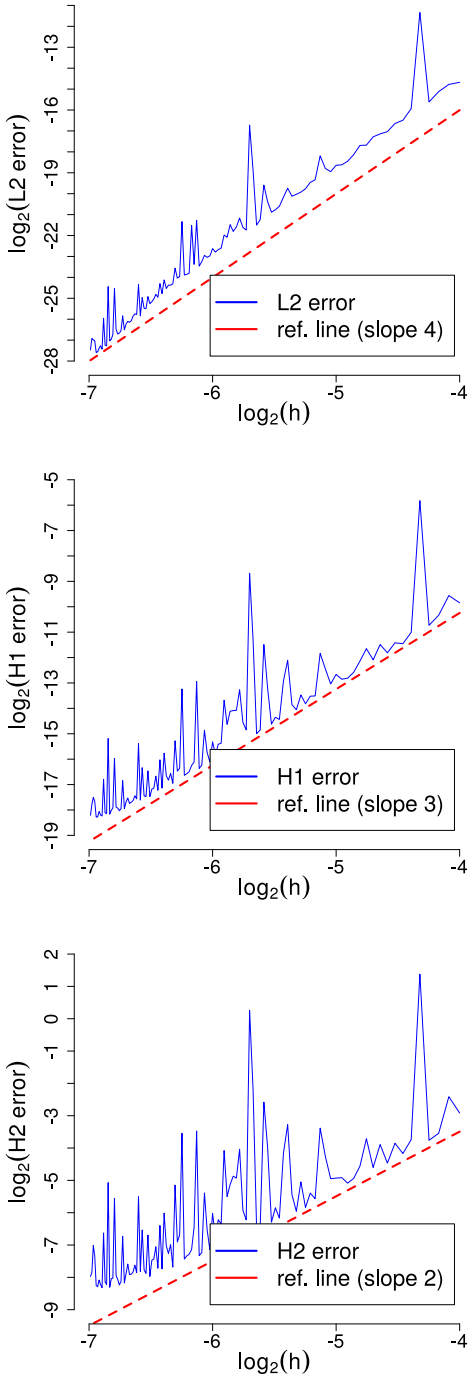

Figure 4.2.4: Double logarithmic plot of the approximate L^2-, H^1- and H^2 discretization error (from top to bottom) over the range $h \in [1/128, 1/16]$ of grid sizes for the example problem in Section 4.2.2.

Again, fluctuations in the error plots might be explained by grid-dependent instabilities within the given implementation. As is to be expected, the error curves degenerate and lose their significance for $h \to H$.

4.3 Point value constraints: Non-optimal regularity

4.3.1 Radially symmetric example

This example uses the analogue to the problem in Section 4.2.1 but with point value constraints instead of curve constraints. Let $\Omega := B_R \subseteq \mathbb{R}^2$ bean R-sphere and suppose a single point value constraint in the origin, while maintaining zero Dirichlet boundary conditions. Again with the Monge-gauge parameterization and the parameters $\kappa = 1$ and $\sigma = 0$ this corresponds to the equation

$$\Delta^2 u = 0 \text{ on } \Omega, \qquad u = \partial_\nu u = 0 \text{ on } \partial\Omega, \qquad u(0) = g$$

for some constant $g \in \mathbb{R}$.

As in Section 4.2.1, due to radial symmetry the solution of this problem takes the form $u(x) = u_0(\|x\|)$ where u_0 solves the boundary value problem

$$u_0''''(t) + \frac{2}{t} u_0'''(t) - \frac{1}{t^2} u_0''(t) + \frac{1}{t^3} u_0'(t) = 0, \qquad u_0(0) = g, \qquad u_0(R) = 0.$$

For the numerical computations the radius is set to $R = 1$ and the point constraint value is set to $g = 1$. Hence, one has

$$u_0(t) = 1 - t^2 + t^2 \log(t).$$

The resulting solution function u is depicted in Fig. 4.3.1.

Because of the logarithmic term $t^2 \log(t)$ in the definition of u_0, the solution u exhibits a singular behavior around the constraint point $x = 0$. This leads to a loss of regularity, and so instead of H^4-regularity only $u \in H^{3-\varepsilon}(\Omega)$ is obtained for arbitrary $\varepsilon \in \mathbb{R}_{>0}$.

The continuous solution is approximated numerically by the conforming transformation method as presented in Section 3.3 with a uniform background grid of squares and grid size $h \in \mathbb{R}_{>0}$. The outer Dirichlet boundary conditions on the curved boundary $\partial\Omega$ are enforced weakly through the fictitious domain stabilized Nitsche method. As of $u \notin H^4(\Omega)$ the error estimates derived in Section 3.2.3 for the stabilized Nitsche method are technically not applicable. Still, in view of the typical approximation properties as mentioned in Section 3.3, one expects at best that the H^2-norm of the discretization error converges with order $1 - \varepsilon \approx 1$. Plots of the error norms $\left\|D^i(u - u_h)\right\|_{L^2(\Omega)}$ for $i \in \{0, 1, 2\}$ and $h = \frac{2}{N}$ with $N = \{16, \ldots, 512\}$ are shown in Fig. 4.3.2. The error was computed using high order quadrature rules on the respective grid.

The graph for the H^2-error is in excellent agreement with the drawn reference line, which supports the claim for a convergence rate of order 1. Similarly, the H^1-error and the L^2-error strictly follow their respective reference lines and so each suggest convergence rates of order 2. The quadratic convergence of the H^1-error is expected due to the

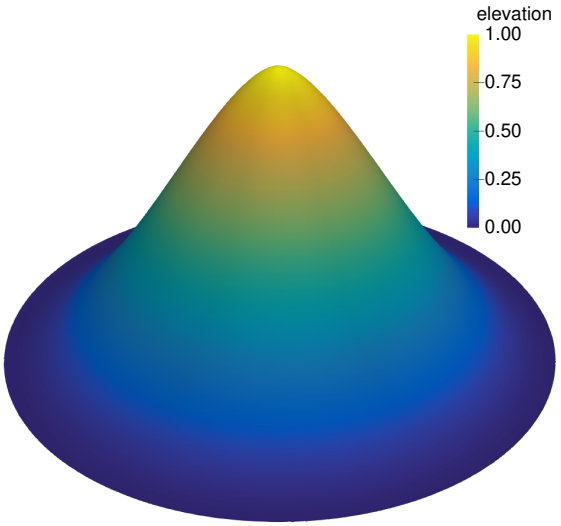

Figure 4.3.1: Exact solution of the example problem stated in Section 4.3.1. The coloring encodes the function's elevation.

linear H^2-error and the typically observed gain of one convergence order per order of differentiation less. In that sense it may seem surprising that the L^2-error only exhibits a quadratic convergence behavior as well. However, in view of the Aubin–Nitsche trick the convergence order of the L^2-error can at most be expected to be twice of the order of the H^2-error. Therefore, the limitation to quadratic convergence in the L^2-norm is explained by the reduced regularity of the solution. Interestingly, the plots show a fluctuating behavior, mostly depending on whether the N in $h = 2/N$ is even or odd, which again decides whether the constraint point $x = 0$ is resolved by the grid exactly or not. Although such a behavior hints at a grid-dependence of the employed method, this does not impact the overall observed rate of convergence.

In summary, the realistically expectable convergence speed is reduced in comparison to the curve constraint examples because of the non-optimal regularity properties of the continuous solution, and so the convergence properties of the discretization appear to be optimal with respect to the regularity of u. But they are certainly non-optimal with respect to the given finite element space as the use of Bogner–Fox–Schmit elements is inefficient when no quadratic convergence in the H^2-norm is obtained.

4.3.2 Non-symmetric example

This example considers a more generic point value constraint problem for two point-value type particles whose solution is not radially symmetric and which admits a conforming discretization without the use of the fictitious domain stabilized Nitsche method. As usual, it uses the Monge-gauge parameterization with $\kappa = 1$ and $\sigma = 0$ over some square domain $\Omega := [-4, 4]^2$. However, this time the outer boundary $\partial\Omega$ is to be equipped with

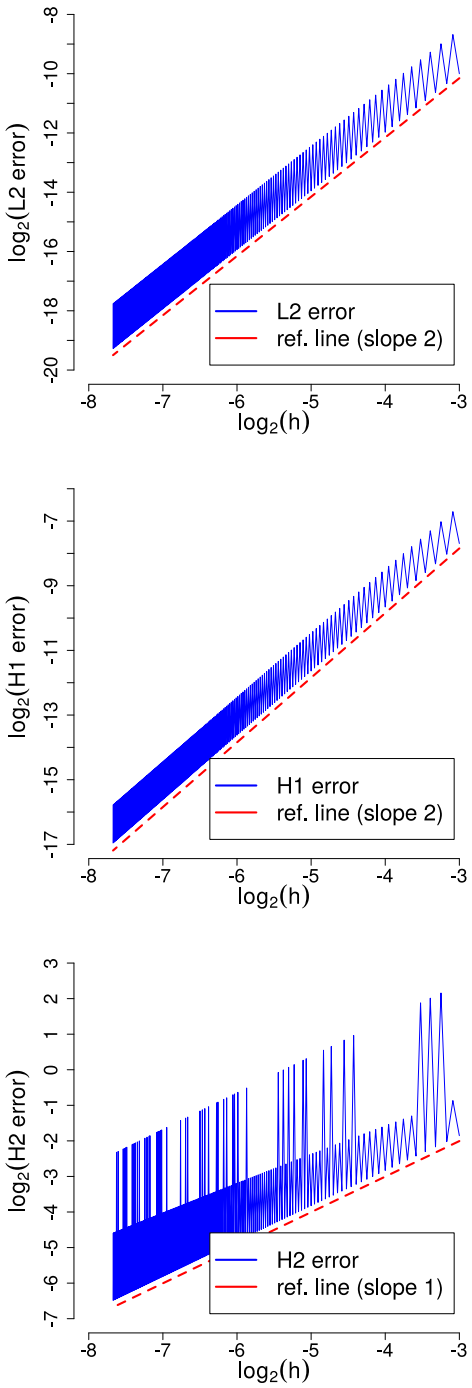

Figure 4.3.2: Double logarithmic plot of the L^2-, H^1- and H^2 discretization error (from top to bottom) over the range $h \in [1/256, 1/8]$ of grid sizes for the example problem in Section 4.3.1.

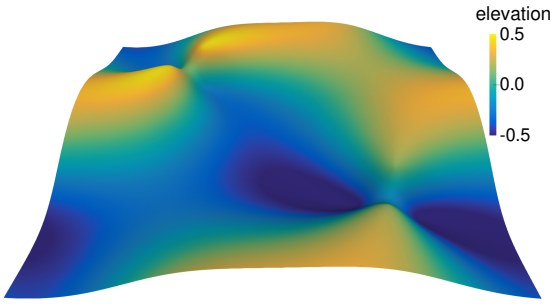

Figure 4.3.3: Discrete solution of the example problem stated in Section 4.3.2. The coloring encodes the function's elevation.

periodic boundary conditions. For the particle model the reference point set

$$\mathcal{G}_0 := \left\{ \begin{pmatrix} -0.35 \\ -0.35 \\ 0.3 \end{pmatrix}, \begin{pmatrix} 0.35 \\ -0.35 \\ -0.3 \end{pmatrix}, \begin{pmatrix} 0.35 \\ 0.35 \\ 0.3 \end{pmatrix}, \begin{pmatrix} -0.35 \\ 0.35 \\ -0.3 \end{pmatrix} \right\}$$

is defined and each particle is equipped with six degrees of freedom to describe its position and rotation in space, i. e. as in Example 1.2.1 and Section 4.2.2 one has

$$\Psi(p; x) := R_3(p_6)R_2(p_5)R_1(p_4)x + \begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix}^T.$$

The concrete particle configuration chosen in this example is

$$\mathbf{p} := \begin{pmatrix} 2 & -2 & -0.2 & 0 & 0 & 30^\circ \\ -2 & 2 & 0.1 & 0 & 0 & 10^\circ \end{pmatrix}.$$

Altogether, this results in the point-constraint problem

$$\Delta^2 u = 0 \text{ on } \Omega, \qquad u(x_i(\mathbf{p})) = g_i(\mathbf{p}) \text{ for } i = 1, \dots, 8$$

with periodic boundary conditions

$$\begin{aligned} u(-2, \cdot) &= u(2, \cdot), & u(\cdot, -2) &= u(\cdot, 2) \\ \nabla u(-2, \cdot) &= \nabla u(2, \cdot), & \nabla u(\cdot, -2) &= \nabla u(\cdot, 2) \end{aligned}$$

and parametric points $x_i(\mathbf{p}) \in \Omega$, $i = 1, \dots, 8$ and constraints $g(\mathbf{p}) \in \mathbb{R}^8$. A discrete solution of this problem is depicted in Fig. 4.3.3.

A standard finite element discretization is applied which is made conforming by restriction to a subspace which fulfills the periodic boundary conditions on $\partial\Omega$ and the point value constraints over the $x_i(\mathbf{p})$. As before, this space is constructed using the transformation technique described in Section 3.3.

It appears reasonable to expect that the point value constraints would again lead to singularities that behave like $r^2 \log(r)$ around the constraint points $x_i(\mathbf{p})$. If that is

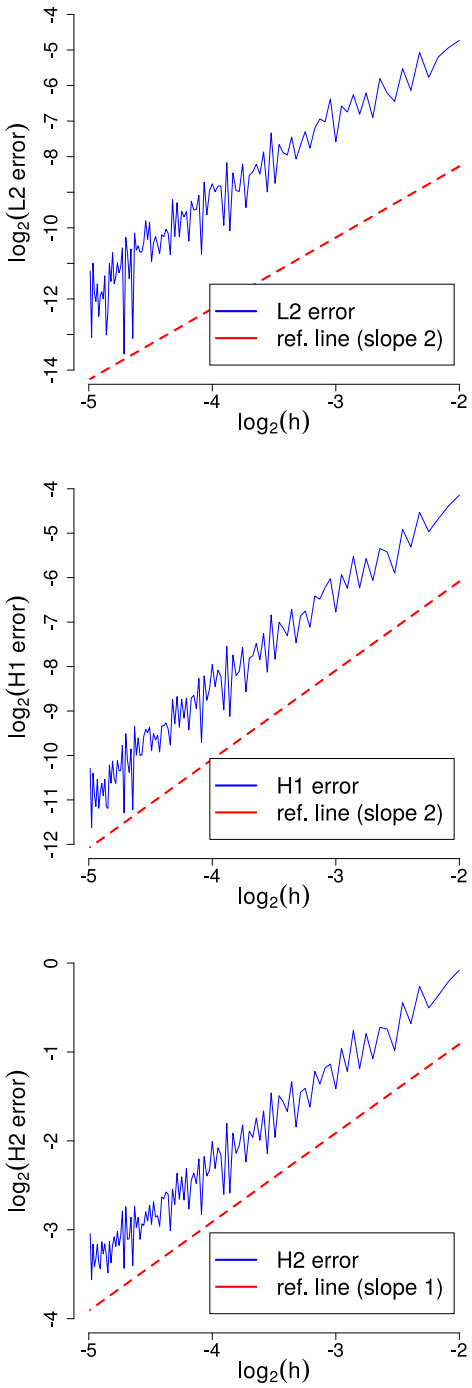

Figure 4.3.4: Double logarithmic plot of the L^2-, H^1- and H^2 discretization error (from top to bottom) over the range $h \in [\frac{1}{32}, \frac{1}{4}]$ of grid sizes for the example problem in Section 4.3.2.

indeed true, then the regularity $u \in H^{3-\varepsilon}(\Omega)$ is to be expected again together with the convergence orders 2, 2 and 1 for the errors in the L^2-, H^1- and H^2-norm, respectively.

The approximate error plots of the $\|D^i(u_H - u_h)\|_{L^2(\Omega)}$, $i \in \{0,1,2\}$ as obtained for the above mentioned discretization are depicted in Fig. 4.3.4 for the grid sizes $h = \frac{4}{N}$, $N \in \{16, \dots, 127\}$. Here, the fine grid approximation u_H corresponds to the discrete solution obtained with $H = \frac{1}{32}$. The errors were computed by a high order quadrature method with respect to the grid associated to u_H.

The error curves in the plots closely follow their corresponding reference lines and thus confirm the anticipated orders of convergence. This may be seen as an indicator for the claimed regularity $u \in H^{3-\varepsilon}(\Omega)$ which then justifies an identical argumentation for the observed convergence rates as in Section 4.3.1. Altogether, this again hints at an optimal convergence with respect to the regularity of the solution u, but also to sub-optimal efficiency of the chosen discretization scheme.

4.4 Gradient evaluation: Non-optimal error estimates

4.4.1 Curve constraints

This example numerically assesses the rate of convergence for the discrete approximation of the gradient of the interaction potential. To this end, the model in Section 4.2.2 for two ellipsoidal particles with free translation and rotation is re-used and again the configuration parameter is set to

$$\mathbf{p} := \begin{pmatrix} -0.4 & 0.4 & 0.3 & 10^\circ & 15^\circ & 30^\circ \\ 0.25 & -0.35 & 0.3 & -10^\circ & -15^\circ & 60^\circ \end{pmatrix} \in \mathbb{R}^{2\times 6}.$$

Approximations G_h of the gradient $\nabla\mathcal{E}(\mathbf{p})$ are obtained by Algorithm 4 using the same discrete approximations u_h of the membrane function as in Section 4.2.2. As of the earlier observed quadratic convergence $\|u - u_h\|_{H^2(\Omega(\mathbf{p}))} \in \mathcal{O}(h^2)$ for $h \to 0$ and the error estimate in Corollary 3.4.2, one expects also quadratic convergence in the gradient errors, i. e. $\|\nabla\mathcal{E}(\mathbf{p}) - G_h\| \in \mathcal{O}(h^2)$ for $h \to 0$.

In Fig. 4.4.1 approximate gradient errors are shown for grid sizes $h = \frac{2}{N}$, $N \in \{16, \dots, 256\}$. There the unknown gradient $\nabla\mathcal{E}(\mathbf{p})$ is substituted by the fine grid approximation G_H with $H = \frac{1}{128}$. On the left hand side, the approximate gradient errors $\|G_H - G_h\|$ are shown, whereas the right hand side shows the errors $\|G_H - \widetilde{G}_h\|$ where the $\widetilde{G}_h$ stem from a central different quotient approximation of $\nabla\mathcal{E}(\mathbf{p})$. More precisely, the components of $\widetilde{G}_h$ are defined by

$$\widetilde{G}_{h,i} := \frac{1}{2h}\left(J(\mathbf{p} + he_i; u_h(\mathbf{p} + he_i)) - J(\mathbf{p} - he_i; u_h(\mathbf{p} - he_i))\right) \tag{4.4.1}$$

where the e_i are the canonical Euclidean basis vectors and the $u_h(\mathbf{p} \pm te_i)$ are the discrete solutions with respect to the perturbed configurations $\mathbf{p} \pm he_i$.

Somewhat surprisingly, the errors $\|G_H - G_h\|$ actually indicate a quartic rate of convergence, which is twice what is predicted by the analytic results. Moreover, the

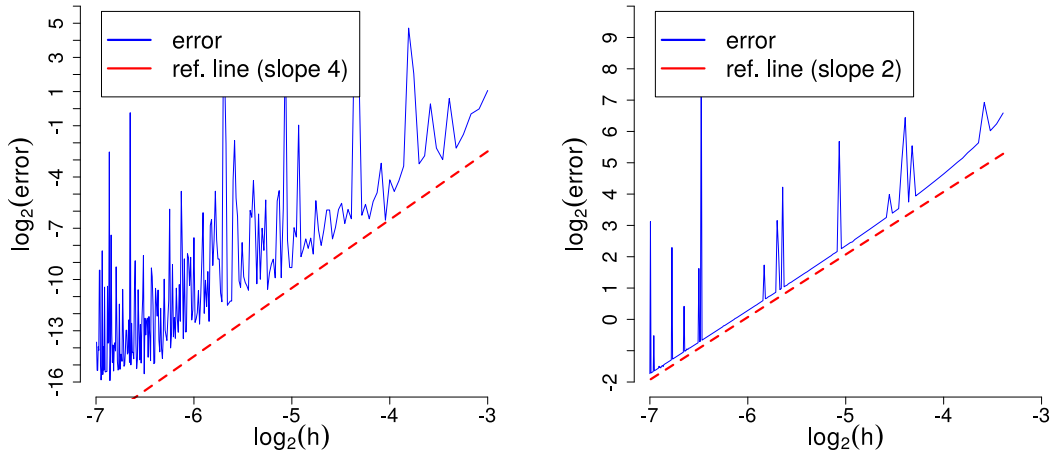

Figure 4.4.1: Double logarithmic error plots for the approximate gradient $G_H \approx \nabla\mathcal{E}(\mathbf{p})$ over grid sizes $h \in [\frac{1}{128}, \frac{1}{8}]$ with $H = \frac{1}{128}$ for the curve constraint problem considered in Section 4.4.1 Left: Plot of $\|G_H - G_h\|$. Right: Plot of $\|G_H - \widetilde{G}_h\|$ with difference quotient approximations $\widetilde{G}_h$ of $\nabla\mathcal{E}(\mathbf{p})$.

quadratic convergence of the errors $\|G_H - \widetilde{G}_h\|$ confirms that the G_h actually converge towards the gradient $\nabla\mathcal{E}(\mathbf{p})$. As in the previous examples, the fluctuations and some of the outliers in the error plots are likely to be explained by instabilities within the implementation of the level set method introduced in Section 4.1.1. The degeneration of the convergence rate for $h \to H$ is seen to be typical. Altogether, it may be concluded that the G_h indeed converge at least quadratically towards the gradient $\nabla\mathcal{E}(\mathbf{p})$. However, at this point it is unclear whether really fourth order convergence is obtained within the given approximation scheme. Investigation of this question requires a more refined error analysis in order to rule out discretization artifacts in the observations. It is emphasized that during the numerical computations indeed the 2-norms of the gradients were calculated and not their squares.

4.4.2 Point value constraints

This example considers again the non-symmetric point model from Section 4.3.2 and the configuration

$$\mathbf{p} := \begin{pmatrix} 2 & -2 & -0.2 & 0 & 0 & 30^\circ \\ -2 & 2 & 0.1 & 0 & 0 & 10^\circ \end{pmatrix} \in \mathbb{R}^{2\times 6}.$$

Just as in the previous example, the gradient $\nabla\mathcal{E}(\mathbf{p})$ is approximated by vectors G_h as obtained by Algorithm 4 using the discrete approximations u_h from Section 4.3.2. This time, in view of the observation $\|u - u_h\|_{H^2(\Omega)} \in \mathcal{O}(h)$ for $h \to 0$ and Corollary 3.4.2, it is justified to expect linear convergence $\|\nabla\mathcal{E}(\mathbf{p}) - G_h\| \in \mathcal{O}(h)$ for $h \to 0$.

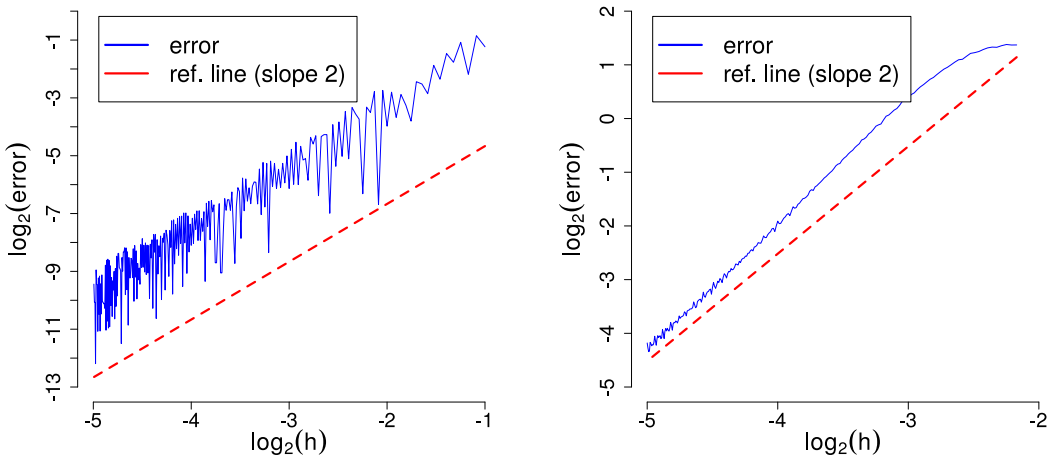

Figure 4.4.2: Double logarithmic error plots for the approximate gradient $G_H \approx \nabla\mathcal{E}(\mathbf{p})$ over grid sizes $h \in [\frac{1}{32}, \frac{1}{2}]$ with $H = \frac{1}{32}$ for the point constraint problem considered in Section 4.4.2 Left: Plot of $\|G_H - G_h\|$. Right: Plot of $\|G_H - \widetilde{G}_h\|$ with difference quotient approximations $\widetilde{G}_h$ of $\nabla\mathcal{E}(\mathbf{p})$.

In Fig. 4.4.2 approximate gradient errors are shown for grid sizes $h = \frac{8}{N}$, $N \in \{16, \dots, 256\}$ with $G_H \approx \nabla\mathcal{E}(\mathbf{p})$ and $H := \frac{1}{32}$. The plot on the left side shows the errors $\|G_H - G_h\|$, and the right hand plot depicts the progress of $\|G_H - \widetilde{G}_h\|$ where the $\widetilde{G}_h$ are the central difference quotient approximations of the gradient $\nabla\mathcal{E}(\mathbf{p})$ as defined in (4.4.1).

Like in Section 4.4.1, the convergence order observed for $\|G_H - G_h\|$ is twice the expected magnitude, i. e. quadratic in this case. The errors $\|G_H - \widetilde{G}_h\|$ again suggest quadratic convergence. Therefore, it appears justified to assume that the G_h indeed converge quadratically towards the gradient $\nabla\mathcal{E}(\mathbf{p})$. However, in view of the previous analysis this claim still remains to be proven, and hence only linear convergence is certain. Also for this experiment it is emphasized that indeed the 2-norms of the gradients were computed and not their squares.

5 Applications

The results from the previous chapters open a wide array of possible applications for gradient-type methods in conjunction with a plethora of different particle–membrane models. This chapter illustrates some of the capabilities of the established interaction framework by means of several example applications.

5.1 Perturbed gradient method

All of the following application examples consider at some point an optimization problem for an interaction potential $\mathcal{E}$ over some closed configuration space $\mathcal{D}$, reading

$$\min_{\mathbf{p}\in\mathcal{D}} \mathcal{E}(\mathbf{p}). \tag{5.1.1}$$

Because it is known at this point that the interaction potential is differentiable under general assumptions, it is in principle feasible to approach this problem using e. g. gradient descent methods. However, special care needs to be taken as of the inexact evaluation of the gradient $\nabla\mathcal{E}(\mathbf{p})$. A method which is able to cope with this circumstance of non-vanishing perturbations of the gradient is the so-called *perturbed gradient method* as introduced in [SS97]. The pseudocode in Algorithm 5 shows a specialization of the algorithm presented therein. In that context the convention $\min\emptyset := \infty$ is declared for convenience. The method performs line searches in the perturbed steepest descent directions $d^i = -G^i \approx -\nabla\mathcal{E}(\mathbf{p}^i)$, similar to the classical gradient descent method, and uses an Armijo-type step size rule. However, it omits the classical stopping criteria such as "$\|\mathbf{p}^{i+1} - \mathbf{p}^i\| < \varepsilon$" or "$\|\nabla\mathcal{E}(\mathbf{p}^i)\| < \varepsilon$" because they are no longer reasonable choices for termination in the presence of non-vanishing gradient perturbations. Instead, the algorithm terminates if within an iteration step no valid Armijo step size 2^{-k_i} with $k_i \in \{0, \dots, K\}$ is found where $K \in \mathbb{N}$ is a pre-defined threshold.

Under slight additional regularity assumptions the following convergence result holds for large enough K:

Theorem 5.1.1. *Suppose Lipschitz-continuity of $\nabla\mathcal{E}$, i. e. $\mathcal{E} \in C^{1,1}$, and denote by $L \in \mathbb{R}_{>0}$ the Lipschitz-constant of $\nabla\mathcal{E}$. In addition, assume*

$$K \geq \log_2\left(\frac{L}{2(1-\omega)}\right)$$

where $\omega \in (0,1)$ is the Armijo-constant in Algorithm 5. Let $(\boldsymbol{p}^i)$ be a sequence generated by Algorithm 5 for $N := \infty$. If this sequence is finite with last element $\boldsymbol{p}^*$, then

$$\boldsymbol{p}^* \in \left\{ \boldsymbol{p} \in \mathbb{R}^{n\times k} \;\middle|\; \|\nabla\mathcal{E}(\boldsymbol{p})\| \leq \frac{2\,\|\nabla\mathcal{E}(\boldsymbol{p}^*) - G^*\|}{\sqrt{8\omega+1} - 1 - 2\omega} \right\}.$$

Algorithm 5 Perturbed gradient method for $\mathcal{E}$ over $\mathbb{R}^{n\times k}$

Suppose $\mathbf{p}^0 \in \mathcal{D}$, $K \in \mathbb{N}$, $N \in \mathbb{N}$ and $\omega \in (0,1)$.
for $i \in 0, \dots, N-1$ **do**
  Compute $d^i := -G^i \approx -\nabla\mathcal{E}(\mathbf{p}^i)$, using e. g. Algorithm 4.
  Let $k_i := \min\left\{k \in \{0,\dots,K\} \mid \mathcal{E}(\mathbf{p}^i) - \mathcal{E}(\mathbf{p}^i + 2^{-k}d^i) \geq \omega 2^{-k}\left\|d^i\right\|^2\right\}$.
  if $k_i = \infty$ **then**
    return $\mathbf{p}^i$
  end if
  Let $\mathbf{p}^{i+1} := \mathbf{p}^i + 2^{-k_i}d^i$.
end for
return $\mathbf{p}^N$

If the sequence is infinite, then either $\mathcal{E}$ is unbounded from below or the sequence $(\mathcal{E}(\boldsymbol{p}^i))_{i\in\mathbb{N}}$ converges and every accumulation point $\boldsymbol{p}^*$ of $(\boldsymbol{p}^i)_{i\in\mathbb{N}}$ fulfills

$$\boldsymbol{p}^* \in \left\{\boldsymbol{p} \in \mathbb{R}^{n\times k} \;\middle|\; \|\nabla\mathcal{E}(\boldsymbol{p})\| \leq \limsup_{i\to\infty}\left\|\nabla\mathcal{E}(\boldsymbol{p}^i) - G^i\right\|\right\}.$$

Moreover, if $\left\|\nabla\mathcal{E}(\boldsymbol{p}^i)\right\| \geq \frac{2\|\nabla\mathcal{E}(\boldsymbol{p}^*)-G^*\|}{\sqrt{8\omega+1}-1-2\omega}$ for all $i \in \mathbb{N}$, then $\lim_{i\to\infty}\nabla\mathcal{E}(\boldsymbol{p}^i) \to 0$ and in particular every accumulation point of the sequence $(\boldsymbol{p}^i)_{i\in\mathbb{N}}$ is a stationary point.

Proof. The proof of an even more general statement is given by [SS97, Theorem 2.2]. □

A readily apparent issue with this algorithm is that it is designed for unconstrained problems, while the problem (5.1.1) of interest is in general constrained. In the setting where the admissible set $\mathcal{D}$ is closed and convex, perturbed gradient projection methods have been studied as in [Sol97] and similar properties as for the perturbed gradient method have been shown. Unfortunately, usually the configuration space $\mathcal{D}$ resulting from particle–membrane interaction models – including the ones that are to follow in this chapter – is non-convex, and so more care needs to be taken for finding suitable optimization methods for (5.1.1). However, it also is the case that the numerical applications of Algorithm 5 to the examples below do exclusively generate iterates within the interior of the feasible domain. This likely is because the energies are such that $\mathcal{E}(\mathbf{p}) \to \infty$ holds for $\mathbf{p} \to \partial\mathcal{D}$, and hence the problems behave like unconstrained ones. As a consequence, the use of Algorithm 5 is as an exception justified for the upcoming problems, but still more investigations are required in order to develop a stable optimization method for these problems in general.

5.2 Pattern formation of isotropic conical inclusions in the Monge-gauge

A popular particle model in the study of elastic particle–membrane interactions are isotropic conical inclusions. As the name suggests, those are transmembrane particles with a conical shape and in particular circular cross sections. The particle–membrane interface is hence given by a circle, and a membrane connects to it with a slope that is induced by the cone. Usually, such a conical particle is free in space and therefore it is justified to equip it with the corresponding translational and rotational degrees of freedom.

The membrane is assumed to be "almost flat" and the particles are assumed to only induce small deformation in order to justify a Monge-gauge parameterization later on. Moreover, in order to avoid artificial boundary effects the membrane is extended periodically.

In this setting the overall goal is to investigate the pattern formation of conical inclusions on the membrane. It is first necessary to formulate the above objects mathematically within the interaction framework. This requires the specification of a reference particle manifold $\widetilde{\mathcal{G}}_0$ and an interface submanifold $\mathcal{G}_0 \subseteq \widetilde{\mathcal{G}}_0$ such that $\mathcal{G}_0$ induces suitable constraints. For the level set approach as in Section 4.1.1 also a planar reference interface $\Gamma_0 \subseteq \mathbb{R}^2$ and a height function $\gamma\colon \mathbb{R}^2 \to \mathbb{R}$ are needed such that $(x, \gamma(x)) \in \widetilde{\mathcal{G}}_0$ for all $x \in \mathbb{R}^2$ and such that the map $\Gamma_0 \ni x \mapsto (x, \gamma(x)) \in \mathbb{R}^3$ is bijective onto $\mathcal{G}_0$. In addition also a level set function $f\colon \mathbb{R}^2 \to \mathbb{R}$ for Γ_0 must be defined. Moreover, a transformation function Ψ has to be chosen such that the degrees of freedom of a particle are represented as intended. Finally, also a reference manifold $\mathcal{M}_0$ with an associated bending energy J is to be defined together with a parameterization $\varphi\colon \Omega \to \mathcal{M}_0$ and a parameterization domain $\Omega \subseteq \mathbb{R}^2$. Then the associated function space $H^2(\Omega)$ needs to be equipped with the appropriate configuration-independent boundary conditions.

As stated before, the particle object induces a circular interface. Without loss of generality, this interface is embedded into the two-dimensional Euclidean plane, i. e. it holds

$$\mathcal{G}_0 := \left\{ x \in \mathbb{R}^2 \times \{0\} \mid x_1^2 + x_2^2 = r^2 \right\}$$

for some pre-defined radius $r \in \mathbb{R}_{>0}$. Correspondingly, the planar interface is simply given by

$$\Gamma_0 := \left\{ x \in \mathbb{R}^2 \;\middle|\; \|x\|^2 = r^2 \right\}$$

and a feasible level set representation of Γ_0 is immediately defined by $f(x) := r^2 - \|x\|^2$.

In addition to the constant height of the circular interface, the particle is seen to also induce a constant slope. Hence, by prescribing the radially symmetric height function

$$\gamma(x) := \frac{s}{2r} \left(r^2 - \|x\|^2 \right)$$

where $s \in \mathbb{R}$ is a slope parameter, a suitable reference particle manifold which extends the interface $\mathcal{G}_0$ and which has a constant slope is defined via

$$\widetilde{\mathcal{G}}_0 := \left\{(x, \gamma(x)) \in \mathbb{R}^3 \mid x \in \mathbb{R}^2\right\}.$$

Concerning the transformation Ψ, it is reasonable to equip every particle with the typical six degrees of freedom for the translations and rotations and thus define for every state $p \in \mathbb{R}^6$ and position $x \in \mathbb{R}^3$

$$\Psi(p, x) := R_3(p_6)R_2(p_5)R_1(p_4)x + \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

as in Example 1.2.1 where the $R_i(\alpha)$ are the counter-clockwise rotations around the x_i axis by the angle α. The membrane model and the elastic energy are given by the well-known Monge-gauge as described in Example 1.1.4. In particular, a domain $\Omega \subseteq \mathbb{R}^2$ may be chosen freely and the reference surface is then defined by $\mathcal{M}_0 := \Omega \times \{0\}$. The prototypical membrane energy upon perturbations of $\mathcal{M}_0$ is defined as

$$J(u) := \frac{1}{2} \int_\Omega \kappa(\Delta u)^2 + \sigma \left\|\nabla u\right\|^2 \, \mathrm{d}x.$$

In order to easily incorporate the periodicity of the surface, it is assumed that $\Omega = [-L, L]^2$ is a square domain. This leads to the function space

$$\begin{aligned} H^2_*(\Omega) := \big\{u \in H^2(\Omega) \mid\ & u(L, \cdot) = u(-L, \cdot),\ u(\cdot, L) = u(\cdot, -L) \\ & \partial_{x_1} u(L, \cdot) = \partial_{x_1} u(-L, \cdot),\ \partial_{x_2} u(\cdot, L) = \partial_{x_2} u(\cdot, -L)\big\} \end{aligned}$$

of feasible periodic membrane shapes. Parameter-dependent constraints are going to be realized via curve constraints in the sense of Section 1.2.2.

Given $n \in \mathbb{N}$ particles and a configuration $\mathbf{p} \in \mathbb{R}^{n \times 6}$, the optimal membrane shape is in this setting uniquely defined by

$$u(\mathbf{p}) := \operatorname*{arg\,min}_{v \in U_{\mathrm{ad}}(\mathbf{p})} \frac{1}{2} \int_{\Omega(\mathbf{p})} \kappa(\Delta v)^2 + \sigma \left\|\nabla v\right\|^2 \, \mathrm{d}x$$

and so the interaction potential takes the value

$$\mathcal{E}(\mathbf{p}) := \frac{1}{2} \int_{\Omega(\mathbf{p})} \kappa(\Delta u(\mathbf{p}))^2 + \sigma \left\|\nabla u(\mathbf{p})\right\|^2 \, \mathrm{d}x.$$

Here the involved configuration-dependent interfaces $\Gamma(\mathbf{p})$ and domains $\Omega(\mathbf{p})$ as well as the admissible sets $U_{\mathrm{ad}}(\mathbf{p})$ may be defined and represented via a level set approach as in Section 4.1.1.

Altogether, the problem of finding stationary particle configurations is of the form (5.1.1), i. e. it reads

$$\min_{\mathbf{p} \in \mathcal{D}} \mathcal{E}(\mathbf{p})$$

where the space of admissible configurations is $\mathcal{D} := \{\mathbf{p} \in \mathbb{R}^{n\times 6} \mid U_{\mathrm{ad}}(\mathbf{p}) \neq \emptyset\}$. Differentiability of the interaction potential $\mathcal{E}$ is known from Example 2.3.5.

This problem is tackled numerically for the case of $n = 9$ particles with the radius $r = \frac{3}{13}$ and slope $s = 1$ over the domain $\Omega = [-2, 2]^2$ by the perturbed gradient method from Algorithm 5. In an attempt to speed up the convergence, the rescaled interaction energy

$$\widetilde{\mathcal{E}}(\mathbf{p}) := \mathcal{E}(R\mathbf{p})$$

is minimized instead of the original interaction energy. It uses the purely heuristically motivated rescaling function

$$(R\mathbf{p})_{ij} := \begin{cases} \mathbf{p}_{ij} & , \ j \in \{1, 2\} \\ \frac{\mathbf{p}_{ij}}{10} & , \ j = 3 \\ \frac{\mathbf{p}_{ij}}{20r\pi} & , \ j \in \{4, 5, 6\}. \end{cases}$$

As initial iterate the configuration $\widetilde{\mathbf{p}}^0 := R\mathbf{p}^0$ with

$$\mathbf{p}^0 := \begin{pmatrix} -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 - \frac{1}{7} & -1 + \frac{1}{7} & 0 & 0 & 0 & 0 \\ -1 - \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 + \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ -1 - \frac{1}{7} & 1 - \frac{1}{7} & 0 & 0 & 0 & 0 \\ \frac{1}{7} & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 + \frac{1}{7} & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{9\times 6}$$

is chosen. The gradients within the perturbed gradient method are evaluated using Algorithm 4 and the discrete approximations $u_h(\mathbf{p})$ of $u(\mathbf{p})$ are computed by the fictitious domain stabilized Nitsche method in Algorithm 2 with grid size $h = \frac{1}{32}$. The bending rigidity and the surface tension are set to $\kappa := 1$ and $\sigma := 0$, respectively, and the parameters in the perturbed gradient method are set to $K := 11$, $N := 10^4$ and $\omega := 0.01$. Due to symmetries in the problem, the position of the central particle $\mathbf{p}_5^i := \mathbf{p}_5^0$ remains fixed and is not changed during the iterations.

In Fig. 5.2.1 the resulting iterates are shown. The method required a total of 9 iteration steps and 29 function evaluations. In particular, the method terminated before reaching the maximum iteration count. The energy decreased from 8.44 to 5.88.

In this simulation the particles quickly arrange into a highly structured pattern which resembles a uniform grid of quadrilaterals. This finding appears to be plausible because of the periodic boundary conditions and because it is known that the pairwise interaction between conical inclusions is strictly repulsive in case of Monge-gauge parameterizations with vanishing surface tension (see [WKH98, EGH$^+$16]). While results are also known where a pairwise repulsive interaction can still lead to stable clustering for a many-particle interaction (see e. g. [KNO98]), it appears that in this concrete example the repulsion dominates, which again leads to the formation of such a regular pattern.

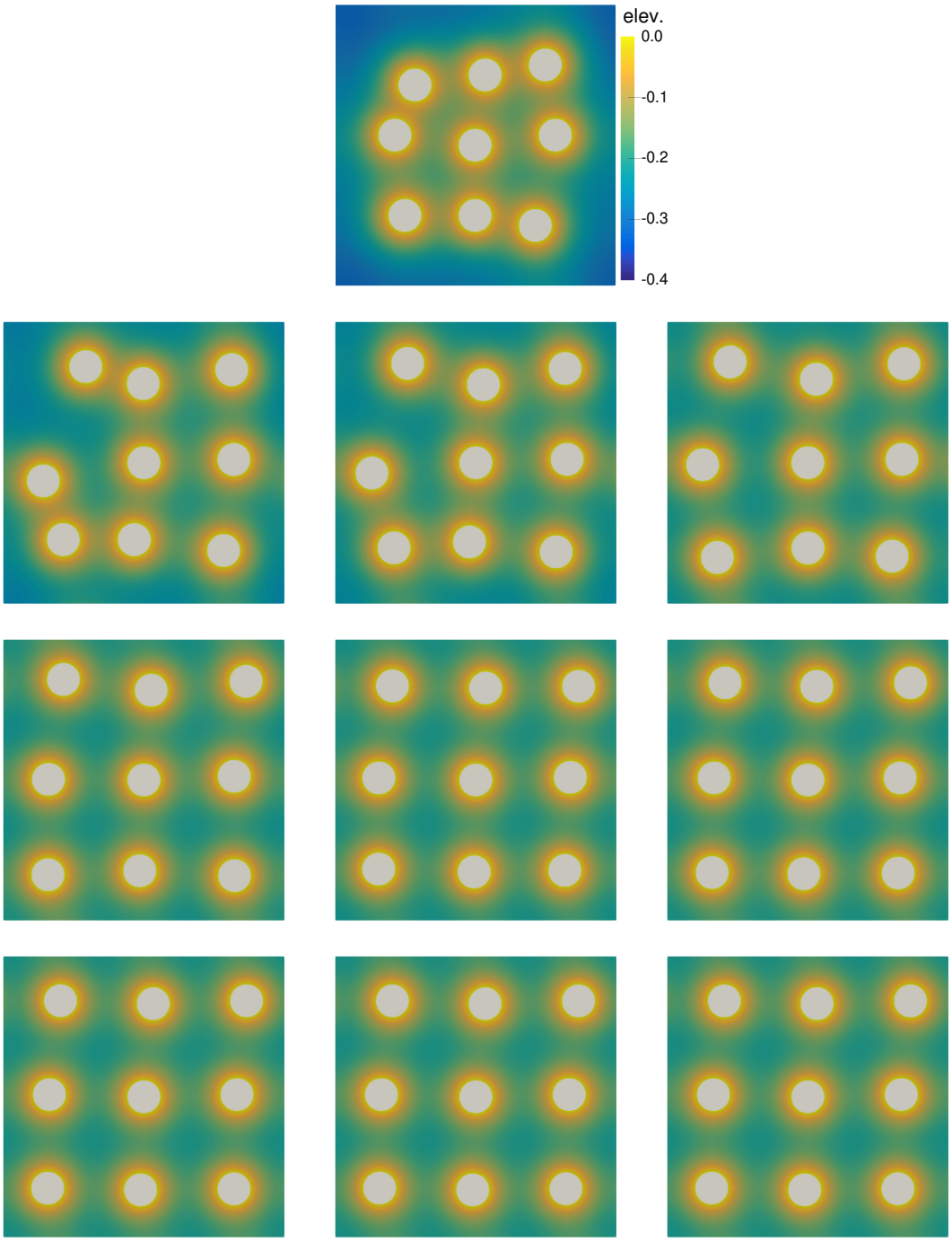

Figure 5.2.1: Top view of the optimal membrane shapes associated to the iterates $\mathbf{p}^i$ for $i = 0, \ldots, 9$ (from left to right, from top to bottom) as generated by the perturbed gradient method applied to the problem in Section 5.2. The coloring encodes the elevation of the membrane deformation function.

5.3 Aggregation of anisotropic ellipsoidal protein scaffolds in the Monge-gauge

The following example model is identical to the previous one except that it considers anisotropic scaffolding particles instead of isotropic inclusions, which are here chosen to have an elliptic shape. The concept behind a scaffolding particle is that it is seen to lie on the membrane, and while doing so it forces the membrane to take the particle's shape, hence the name "scaffolding particle." Therefore the membrane shape within the domain that is covered by the particle always remains constant, independently of the particle configuration. In particular, the energy contribution from the affected membrane section remains constant as well and is in consequence negligible during energy minimization. As a result, only the influence of the particle on the membrane's shape and slope along its boundary matters, which leads to the same concept for the particle–membrane interface as inclusion particles.

In consequence, the essential mathematical framework and the discretization remain identical to Section 5.2, only that now the height function γ and the planar reference interface Γ_0 are different. For this example, let $r := \frac{4}{13}$, $s = 0.2$ and

$$\gamma(x) := \frac{sr}{2}\left(\frac{x_1^2}{r^2} - \frac{x_2^2}{4r^2}\right)$$

as well as

$$f(x) := 1 - \frac{x_1^2}{r^2} - \frac{x_2^2}{4r^2}$$

such that $\Gamma_0 := f^{-1}(\{0\})$ forms an ellipse and $\mathcal{G}_0 := (\Gamma_0, \gamma(\Gamma_0))$ a saddle. The remaining model definitions stay the same as in the previous section.

Also the parameters in the optimization method and those in the discretization schemes are kept the same. The only differences are that now $n = 6$ particles are to be considered, that the heuristic rescaling for the rescaled interaction potential $\widetilde{\mathcal{E}}(\mathbf{p}) := \mathcal{E}(R\mathbf{p})$ is now defined by

$$(R\mathbf{p})_{ij} := \begin{cases} \mathbf{p}_{ij} & , \ j \in \{1,2\} \\ \frac{\mathbf{p}_{ij}}{10} & , \ j = 3 \\ \frac{\mathbf{p}_{ij}}{10r\pi} & , \ j = 4 \\ \frac{\mathbf{p}_{ij}}{20r\pi} & , \ j = 5 \\ 5\mathbf{p}_{ij} & , \ j = 6 \end{cases}$$

and that the initial configuration is set to $\widetilde{\mathbf{p}}^0 = R\mathbf{p}^0$ with

$$\mathbf{p}^0 :\approx \begin{pmatrix} -0.96 & -0.62 & 0 & 0 & 0 & 25^\circ \\ 0.04 & -0.54 & 0 & 0 & 0 & 20^\circ \\ 0.88 & -0.62 & 0 & 0 & 0 & 20^\circ \\ -1.12 & 0.54 & 0 & 0 & 0 & -10^\circ \\ 0 & 0.5 & 0 & 0 & 0 & 0^\circ \\ 1.04 & 0.38 & 0 & 0 & 0 & -20^\circ \end{pmatrix} \in \mathbb{R}^{6\times 6}.$$

The particle position $\mathbf{p}_5^i := \mathbf{p}_5^0$ is again set to stay constant during minimization.

A selection of the iterates resulting from multiple consecutive applications of the perturbed gradient method is shown in Fig. 5.3.1. Those iterates were obtained by repeatedly running the algorithm until it terminated due to violation of the Armijo step length rule and then restarting it with a perturbation of the last iterate in direction $-10^{-3}G$, i. e. "$\mathbf{p}_0^{\text{new}} = \mathbf{p}^{\text{old}} - 10^{-3}G^{\text{old}}$." This essentially corresponds to a combination of the original method with a standard gradient flow algorithm as a backup, and by doing so it was possible to circumvent apparent dead ends in the algorithms due to numerical instabilities. As such an approach is not fully satisfactory from a numerical viewpoint, this once more emphasizes the need for further investigations for robust discretization and optimization algorithms. In total 300 iteration steps were computed, which required 898 function evaluations. The method was terminated manually, and the energy decreased from 7.41 to 5.05.

Starting from their ladder-like arrangement (iterates 0, 10 and 25), the particles begin to assemble in a ellipsoidal shape (iterates 50, 75 and 100) and ultimately end up aggregating in form of a circle (iterates 150 and above). Roughly speaking, it appears to be energetically most favorable to limit the strong curvature generated at the long ends of the particles by linear alignment. Hence, depending on the particle number linear structures such as simple lines or circles might be expected, or also the union of many lines and circles in the case of large particle counts. Of course, the present data can not verify if the obtained circular configuration is indeed optimal. To this end it would arguably be necessary to apply global optimization methods to the given problem. Alternatively, a first insight might be gained from considering further variations in particle numbers and initial configurations.

Altogether, the given anisotropic particles appear to have the potential ability to build superstructures solely based on the membrane-mediated interaction force. In particular, the exhibited behavior is vastly different from the isotropic case in Monge-gauge where there was no attractive collaboration of particles. These results fit well with a range of other results where similar interactions between various types of anisotropic particles have been studied [DF99, KNO00, DF02, SK15, KKKH16a].

5.4 Interaction of isotropic conical inclusions on a tube

This application investigates the interaction of two radially symmetric conical inclusions on a tubular membrane. As in Section 5.2, the particle model is defined through the quantities

$$\begin{aligned}
\mathcal{G}_0 &:= \left\{x \in \mathbb{R}^2 \times \{0\} \mid x_1^2 + x_2^2 = r^2\right\} \\
\Gamma_0 &:= \left\{x \in \mathbb{R}^2 \mid \|x\|^2 = r^2\right\} \\
\gamma(x) &:= \frac{s}{2r}\left(r^2 - \|x\|^2\right) \\
\widetilde{\mathcal{G}}_0 &:= \left\{(x, \gamma(x)) \in \mathbb{R}^3 \mid x \in \mathbb{R}^2\right\}.
\end{aligned}$$

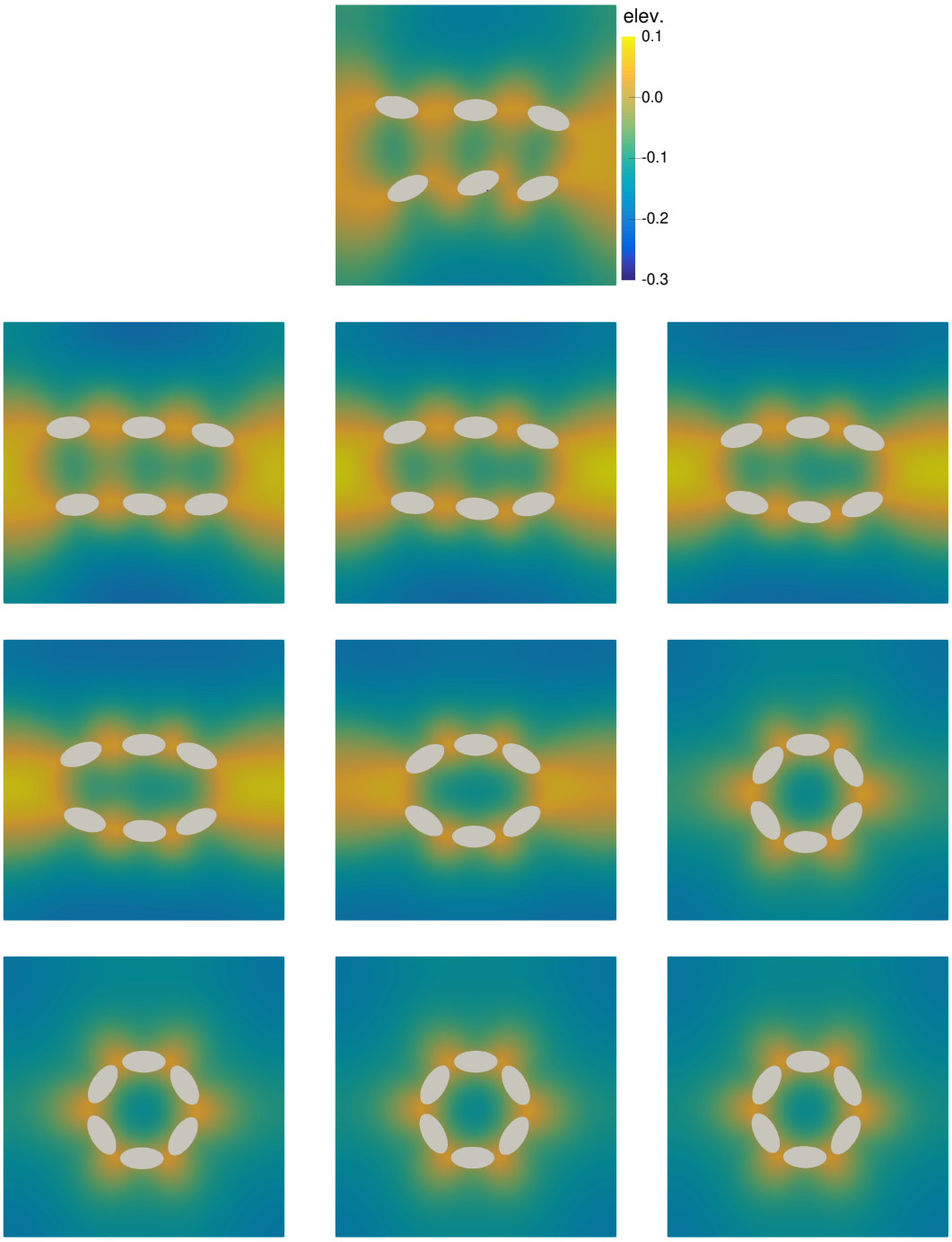

Figure 5.3.1: Top view of the optimal membrane shapes associated to the iterates $\mathbf{p}^i$ for $i = 0, 10, 25, 50, 75, 100, 150, 200, 250, 300$ (from left to right, from top to bottom) as generated by the perturbed gradient method applied to the problem in Section 5.3. The coloring encodes the elevation of the membrane deformation function.

Here $r \in \mathbb{R}_{>0}$ is the particle radius and $s \in \mathbb{R}$ is the particle slope. As before, the particles are allowed to move and rotate freely in space. In particular, the transformation function Ψ is identical to those in Sections 5.2 and 5.3.

The reference membrane is given by a tube of length $L \in \mathbb{R}_{>0}$ and radius $R \in \mathbb{R}_{>0}$, i. e. by

$$\mathcal{M}_0 := \left\{ x \in [0, L] \times \mathbb{R}^2 \mid x_2^2 + x_3^2 = R^2 \right\},$$

which is again parameterized over the domain $\Omega := [0, L] \times [0, 2\pi]$ via the map

$$\varphi(x) := (x_1, R\cos(x_2), R\sin(x_2))^T.$$

Under the simplifying (but non-physical) assumption that the membrane's spontaneous curvature is matched with its mean curvature and upon neglecting both surface tension and Gaussian curvature, the elastic energy function is defined by

$$J(\mathbf{p}; v) := \frac{R\kappa}{2} \int_{\Omega(\mathbf{p})} \left(v_{11} + \frac{v + v_{22}}{R^2} \right)^2 \mathrm{d}x,$$

in accordance to Example 1.1.5. The reference function space mixes periodic and Dirichlet boundary conditions and is given by

$$H^2_*(\Omega) := \left\{ v \in H^2(\Omega) \mid v|_{\Gamma_{P,1}} = v|_{\Gamma_{P,2}},\ \partial_\nu v|_{\Gamma_{P,1}} = -\partial_\nu v|_{\Gamma_{P,2}},\ v|_{\Gamma_D} = \partial_\nu v|_{\Gamma_D} = 0 \right\}$$

where $\Gamma_{P,1} := [0, L] \times \{0\}$, $\Gamma_{P,2} := [0, L] \times \{2\pi\}$, and $\Gamma_D := \{0, L\} \times [0, 2\pi]$.

This application restricts the attention towards the interaction of the particles along a single circular line segment on the tube. Hence, the configuration space is defined as

$$\mathcal{D} := \left\{ \mathbf{p} \in (\{L/2\} \times \mathbb{R}^5)^2 \mid U_{\mathrm{ad}}(\mathbf{p}) \neq \emptyset \right\},$$

i. e. both particles are required to position themselves at center height of the tube. The derivative of the interaction potential is given by the formula in Example 2.3.7.

Within the numerical computations the introduced parameters are set to $r := 0.2$, $s := -0.3$, $R := 1$, and $L := 6$. In this application minimization is applied to the interaction potential $\mathcal{E}$ directly, without any rescaling. The initial iterate is set to

$$\mathbf{p}^0 := \begin{pmatrix} L/2 & 0 & R & 0 & 0 & 0 \\ L/2 & R\sin(\alpha) & R\cos(\alpha) & -\alpha & 0 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 6}$$

with $\alpha := 45°$. Gradients in the perturbed gradient method are again computed according to Algorithm 4 and discrete approximations $u_h(\mathbf{p})$ of $u(\mathbf{p})$ are obtained via Algorithm 2 with a uniform 123-by-128 grid over Ω, yielding $h \leq \frac{2\pi}{128}$. The bending rigidity is again set to $\kappa := 1$ and the parameters in the perturbed gradient method are $K := 11$, $N := 10^4$ and $\omega := 0.01$. As in Section 5.3, the method again performs a gradient flow step if no valid Armijo step size is found.

Numerical results are presented in Figs. 5.4.1 and 5.4.2. The method computed 250 iteration steps, which required 578 function evaluations. The iteration was terminated

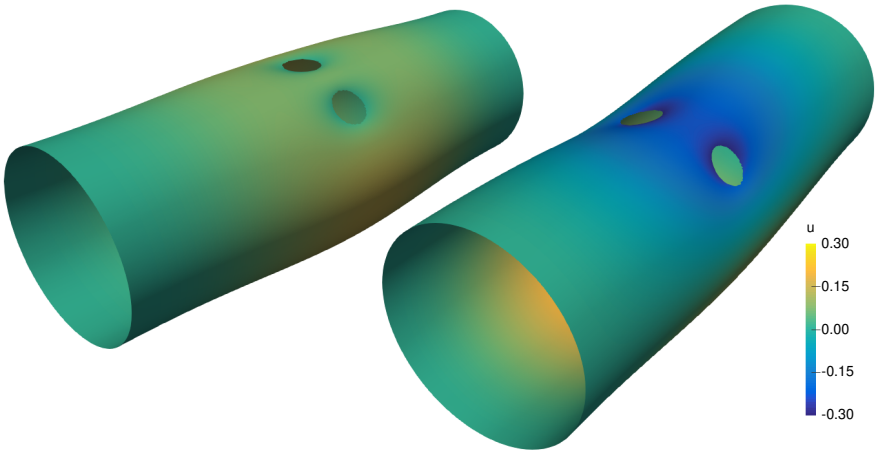

Figure 5.4.1: Rendered 3D views of the first iterate (left) and the last iterate (right) as obtained by the perturbed gradienet method applied to the problem in Section 5.4. The coloring indicates the magnitude of the perturbation function u.

manually, during the course of which the energy decreased from 0.582 to 0.209. The cross sections in Fig. 5.4.2 show how the iterates quickly change the height of the particles relative to the center of the tube. Subsequent iterations then proceed with a radial repositioning of the particles, until a stationary state is found.

Most remarkably, it turns out that the particles in optimal configuration are not diametrically opposed to each other. This in turn indicates that the interaction of identical isotropic particles in a tubular membrane is not strictly repulsive, like it would be for almost flat Membranes in Monge-gauge. These results are in strong accordance with those obtained e. g. in [BLW12] and emphasize the impact of the membrane shape on the overall particle interaction.

5.5 Towards simulating superstructures induced by FCHo2 F-BAR domains

The FCHo2 F-BAR domain protein is a scaffolding particle and was first studied in [HKF$^+$07]. Since its discovery, those domains became known to play a role in connection with the early stages of clathrin-mediated endocytosis, possibly by slightly bending the membrane in a circular area and thus preparing the boundary for the clathrin-coated pit from which the vesicle later buds [HBM$^+$10]. The precise bending behavior of these proteins remains an unanswered question and is thus still open to debate.

This section proposes a model for the interaction of FCHo2 F-BAR domains in terms of the previously developed framework. To this end, Fig. 5.5.1 shows a graphical representation of the protein as obtained under the protein data base (PDB) identifier *2V0O*.

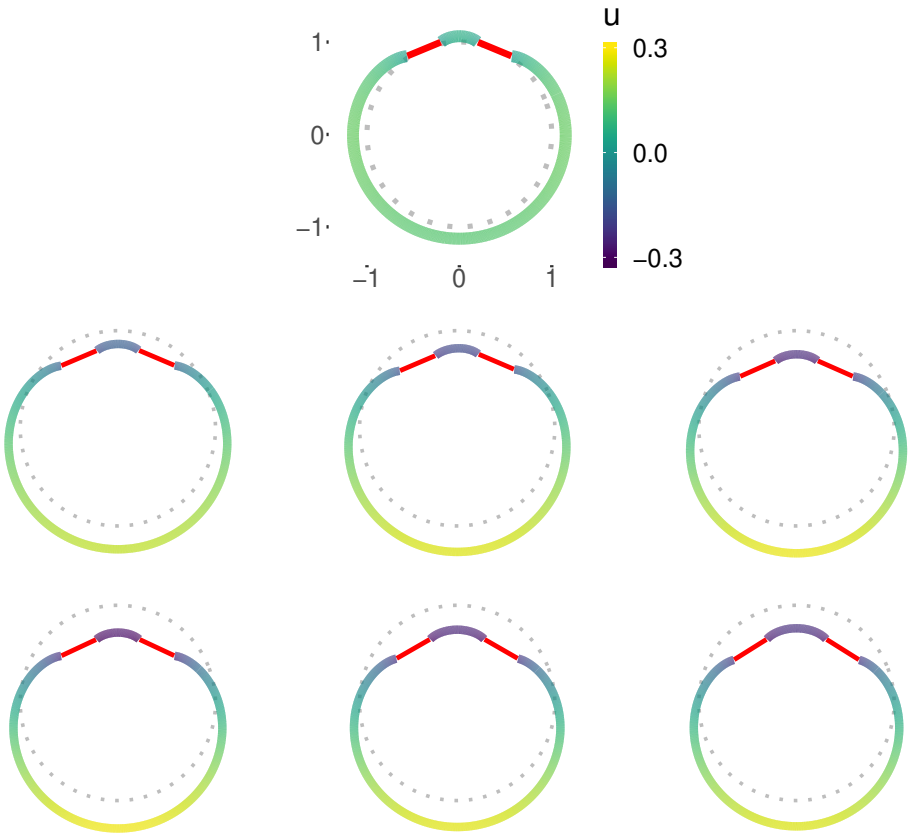

Figure 5.4.2: Cross sections of the optimal membrane shapes along the plane $x_1 = 3$ at the iterates $\mathbf{p}^i$ for $i = 0, 5, 10, 25, 150, 250$ (from left to right, from top to bottom) as generated by the perturbed gradient method applied to the problem in Section 5.4. The dotted line shows the unperturbed reference tube, and red lines indicate particle positions. The remaining coloring encodes the elevation of the membrane perturbation function u.

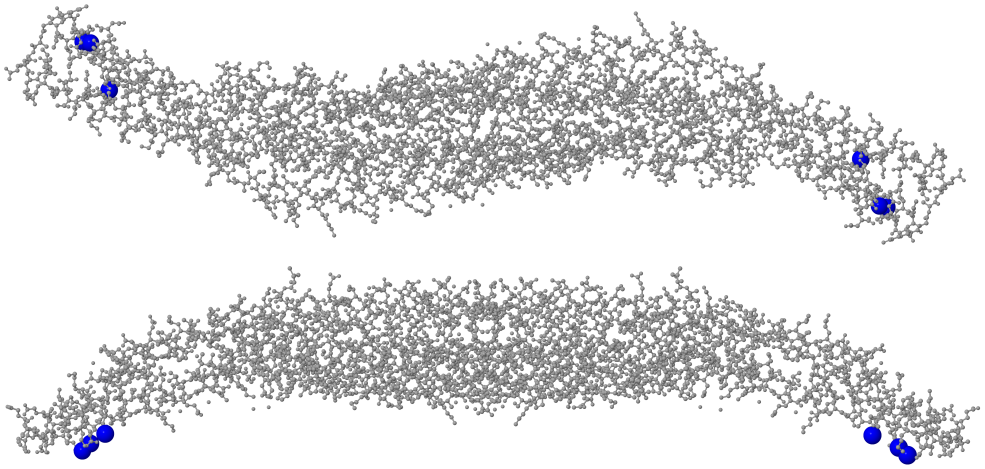

Figure 5.5.1: Visualization of the FCHo2 F-BAR domain (PDB ID *2V0O*). Blue spheres indicate the positions of the charged residue tips which bind to the membrane. Top: Protein as seen from above. Bottom: Side view of the particle.

It depicts an atomistic representation of the two monomers of FCHo2 with a total of 4478 atoms.

Of special significance are the six points which are depicted as large blue spheres. Those are the charged ends of the residues Lys146 and Arg152 of each monomer, which are given by the NZ, NH1 and NH2 atoms, respectively. The biomembrane connects to those points preferentially, which makes them a relevant object for the development of a particle model in the sense of the interaction framework. After an orthogonal transformation of the coordinates for improving the alignment of the atom positions with the two-dimensional Euclidean plane, the coordinates of these six points are approximately given in Ångströms (Å) by

$$\mathcal{G}_0 := \left\{ \begin{pmatrix} 78.40 \\ 7.45 \\ 3.71 \end{pmatrix}, \begin{pmatrix} 81.54 \\ -3.08 \\ -1.17 \end{pmatrix}, \begin{pmatrix} 83.13 \\ -4.01 \\ -2.54 \end{pmatrix}, \begin{pmatrix} -78.53 \\ -7.74 \\ 3.54 \end{pmatrix}, \begin{pmatrix} -81.67 \\ 2.96 \\ -1.12 \end{pmatrix}, \begin{pmatrix} -82.87 \\ 4.41 \\ -2.43 \end{pmatrix} \right\}.$$

The approach taken here is to use exactly these positions as point value constraints in the sense of Section 1.2.1, which essentially defines the particle model.

Concerning the membrane model, it is justified to again use the Monge-gauge parameterization of the elastic energy. This is because the FCHo2 particles are investigated for an early phase of the clathrin-mediated endocytosis where the membrane is not bent significantly yet and the particles themselves are expected to only induce a relatively small deformation. Furthermore, periodic boundary positions are imposed on the boundary of the parameterization domain in order to soften boundary effects.

For simplicity, the transformation function Ψ is defined to only incorporate lateral displacements and rotations in the x_1–x_2-plane. I. e., given a particle state $p \in \mathbb{R}^4$ one

has for all $x \in \mathbb{R}^3$

$$\Psi(p;x) := \begin{pmatrix} \cos(p_4) & -\sin(p_4) & 0 \\ \sin(p_4) & \cos(p_4) & 0 \\ 0 & 0 & 1 \end{pmatrix} x + \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}.$$

The configuration space is set to $\mathcal{D} := \mathbb{R}^4$ by periodic extension.

In this application the elastic membrane energy alone is not sufficient for a reasonable description of the system's particle interactions, and the reason behind this is that a simple mathematical consideration quickly reveals that an optimal particle configuration $\mathbf{p} \in \mathbb{R}^{n\times 4}$ would satisfy the condition $\mathbf{p}_i = \mathbf{p}_j$ for all $i, j \in \{1, \dots, n\}$. This again would imply perfect overlapping of all particles, which of course should be infeasible. Instead of forcing a separation of particles by restriction of the configuration space, the approach taken here is to augment the interaction potential by a direct particle–particle interaction $\mathcal{E}_{\text{particle}}$ which is sufficiently repulsive and thus prevents overlapping of particles. The interaction potential is hence defined as

$$\mathcal{E}(\mathbf{p}) := J(\mathbf{p}; u(\mathbf{p})) + \mathcal{E}_{\text{particle}}(\mathbf{p})$$

where J is the standard elastic membrane energy. The precise choice of $\mathcal{E}_{\text{particle}}$ in this application example is based on a smoothing of the repulsive term in the Lennard–Jones interaction between each particle's atoms, see Appendix E for further details. In this context, while physically motivated, the particle–particle interaction term should rather be understood as a regularization than a physical extension of the model.

Altogether, the task of finding stationary particle configurations corresponds again to a minimization problem of the type (5.1.1). The gradient perturbation method with restarts is applied to this problems for $n = 5$ particles over the quadratic reference domain $\Omega = [-500, 500]^2$. The discrete membrane function u_h is computed by the discretization for point constraints as introduced in Section 3.3 using a locally refined grid with element sizes $h \leq \frac{500}{32}$. As in the previous applications, a heuristic rescaling is applied. Here it is defined by

$$(R\mathbf{p})_{ij} := \begin{cases} 100\,\mathbf{p}_{ij} & , \ j \in \{1,2\} \\ 10\,\mathbf{p}_{ij} & , \ j = 3 \\ \frac{10}{13\pi}\,\mathbf{p}_{ij} & , \ j = 4 \end{cases}$$

and the initial position is defined as $\widetilde{\mathbf{p}}^0 := R\mathbf{p}^0$ with

$$\mathbf{p}^0 := \begin{pmatrix} -300 & 10 & 0 & 110^\circ \\ -150 & 20 & 0 & 100^\circ \\ 0 & 0 & 0 & 90^\circ \\ 150 & 30 & 0 & 70^\circ \\ 300 & -20 & 0 & 80^\circ \end{pmatrix} \in \mathbb{R}^{5\times 4}.$$

The remaining parameters are defined as in the previous sections.

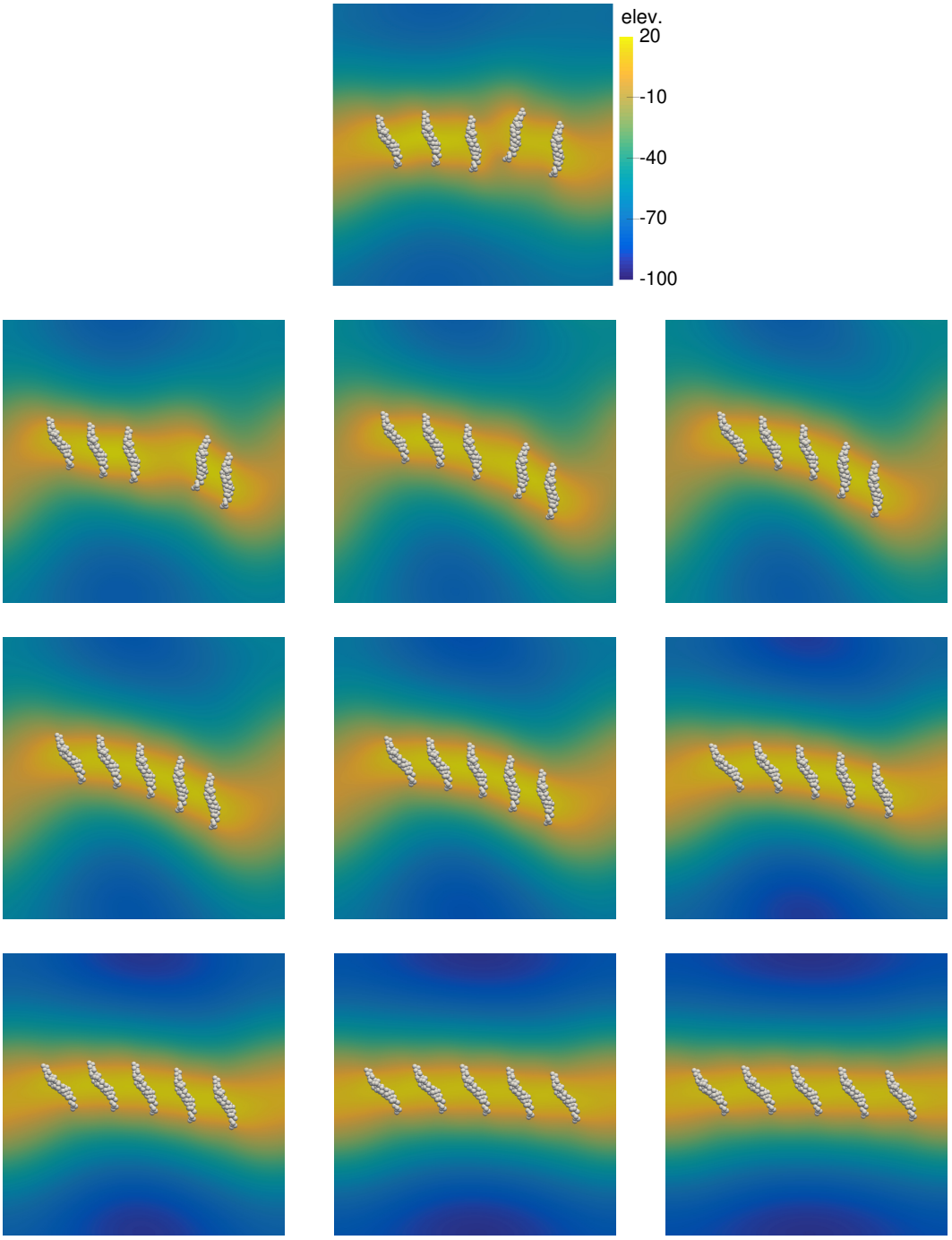

Figure 5.5.2: Top view of the optimal membrane shapes associated to the iterates $\mathbf{p}^i$ for $i = 0, 10, 25, 50, 75, 100, 150, 200, 300, 400$ (from left to right, from top to bottom) as generated by the perturbed gradient method applied to the problem in Section 5.5. The coloring encodes the elevation of the membrane deformation function in Å. The space-filling model of the protein is shown to aid visualization of the particle positions and geometries.

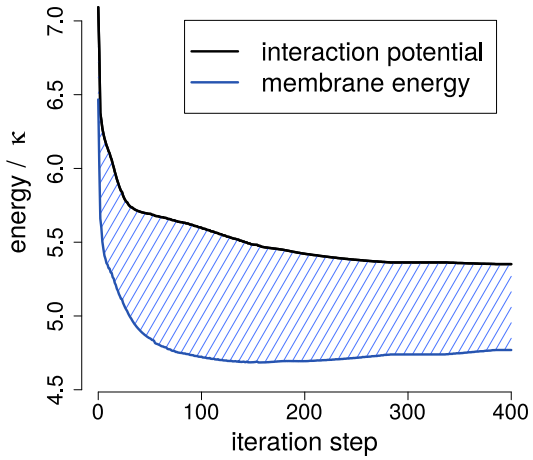

Figure 5.5.3: Energy decrease obtained by application of the perturbed gradient method to the FCHo2 aggregation problem in Section 5.5. The black line shows the values of the interaction potential $\mathcal{E}(\mathbf{p}^i) = J(\mathbf{p}^i, u(\mathbf{p}^i)) + \mathcal{E}_{\text{particle}}(\mathbf{p}^i)$ and the blue line corresponds to the pure membrane energy J. The shaded area indicates the development of the direct particle–particle interaction $\mathcal{E}_{\text{particle}}$.

At this point, special attention needs to be given to the value of the bending rigidity κ. A physical parameter choice for the bending rigidity lies for example around $\kappa = 20\,k_BT \approx 2.8 \times 10^{-24}\,\mathrm{J\,K^{-1}}$ for room temperature (cf. [BHD04]). However, because the surface tension $\sigma = 0$ vanishes, it is readily seen that the interaction potential $\mathcal{E}$ scales linearly with κ, where it is assumed for simplicity that the direct particle–particle interaction is chosen proportionally to κ as well. Therefore, it is sufficient to carry out the numerical computations for $\kappa = 1$. The obtained energies may then be interpreted to be given in the unit $[\kappa]$.

Numerical results are shown in Figs. 5.5.2 and 5.5.3. A total of 400 iteration steps was performed, which required 1531 function evaluations. The first restart of the method happened at iteration 270, which was reached after only 530 function evaluations, and a total of 70 restarts were executed from there on.

As depicted in Fig. 5.5.2, the iterations start with five particles that are roughly, but not perfectly, assembled into a linear configuration. In particular, between the third and fourth particle appears a kink, which could be energetically unfavorable. Over the course of iterations it is observed that this kink is resolved, which yields a slightly curved area of deformation with uniformly distributed particles around iteration 75. This deformation area is slightly rotated until iteration 200, while still maintaining some of its curved shape. The remaining iterations up to step 400 then proceed with straightening the deformation shape.

As shown in Fig. 5.5.3, the initial energy of the interaction potential starts at 7.09, then quickly decreases to 5.60 until iteration step 100, and ultimately reaches the value 5.35 at step 400. Simultaneously, the membrane energy is reduced from 6.46 to 4.72 at step 100 and to 4.76 at step 400. The particle–particle interaction starts at 0.63, increasing to 0.87 at step 100 and finally decreasing to 0.58 at step 400.

Since the direct particle–particle interaction is supposed to act mostly as a regularization, the main attention is on the development of the pure elastic membrane energy. While ensuring that no particle overlap occurs, the regularization is sufficiently weak such that the energetic behavior of the system is still dominated by the membrane-mediated interactions. The optimal membrane energies are obtained between the iterations 100 and 200, which corresponds to a slightly curved line formation of the F-BAR proteins. Due to the non-physical nature of the regularization it is not clear from the data if this state is stationary in the physical sense. However, due to the particle–particle interaction being purely repulsive, it may be concluded that such a configuration is energetically favorable in comparison to the initial line setup with a kink. There the membrane energy is roughly decreased by 1.78κ, which for $\kappa = 20\,k_BT$ corresponds to an energy decrease of $7.12\,k_BT$ per particle.

In conclusion, the obtained results support the theory that FCHo2 F-BAR domains favor linear aggregation. In view of the slight bending of those aggregates it furthermore appears plausible that large numbers of such particles are able to form circular pits. While the implementation and resources used for the computations did not admit significantly larger domains and particle counts, appropriately sized simulations might in principle be accessible through state-of-the-art solvers for partial differential equations

and by investing a larger amount of computing power. Moreover, additional research may be advised in order to investigate the feasibility of the chosen constraint model and to improve the chosen particle–particle interaction model for further justification in a physical sense.

6 Outlook: Free energy computations

Up to this point, all of the previous models were stated in the so-called zero temperature regime, i. e. in the absence of any thermal fluctuations and where stationary states are simply characterized by minimizers of the system's potential. From a biophysical standpoint, however, it often is desirable to further assess the stability of stationary states in a thermodynamic context. This leads to the subject of free energy computations, which is motivated in the following paragraphs.

In order to avoid confusion, it is worth emphasizing that the notation in the subsequent presentation is essentially based on the one established in the previous chapters. In particular, some of the notation is unavoidably in disagreement with the conventions from standard literature on statistical mechanics and thermodynamics.

First, let $\mathcal{D} \subseteq \mathbb{R}^{n\times k}$ again be a space of feasible particle configurations, and suppose that the system is described by the *separable Hamiltonian*

$$H(\mathbf{p}, \mathbf{v}) := \mathcal{E}(\mathbf{p}) + E_{\mathrm{kin}}(\mathbf{v})$$

where $\mathcal{E}$ is an interaction potential, E_{kin} describes the kinetic energy and $\mathbf{v} \in \mathbb{R}^{n\times k}$ is a velocity vector. Moreover, the so-called phase space is defined as the set of all feasible position–velocity pairs and coincides here with the tangent bundle $\mathrm{T}\mathcal{D} = \mathcal{D} \times \mathbb{R}^{n\times k}$. This leads to the deterministic *Hamiltonian dynamics*

$$\begin{aligned} \mathbf{p}'(t) &= \nabla E_{\mathrm{kin}}(\mathbf{v}) \\ \mathbf{v}'(t) &= -\nabla \mathcal{E}(\mathbf{p}) \end{aligned}$$

of the particle system, whose stochastic counterpart is the *Langevin process*

$$\begin{aligned} \mathrm{d}\mathbf{p}(t) &= \nabla E_{\mathrm{kin}}(\mathbf{v})\,\mathrm{d}t \\ \mathrm{d}\mathbf{v}(t) &= -\nabla \mathcal{E}(\mathbf{p})\,\mathrm{d}t - \gamma(\mathbf{v})\nabla E_{\mathrm{kin}}(\mathbf{v})\,\mathrm{d}t + \sigma(\mathbf{v})\,\mathrm{d}W_t. \end{aligned} \tag{6.1}$$

Here $\sigma(\mathbf{v})\,\mathrm{d}W_t$ is fluctuation by a $(n \times k)$-dimensional Brownian motion, which adds energy to the system and which is dissipated by the viscous friction term $-\gamma(\mathbf{v})\nabla E_{\mathrm{kin}}(\mathbf{v})$. This equation aims at describing a particle's motion that essentially follows the above Hamiltonian dynamics while also being embedded into an infinite heat bath which keeps the temperature constant. While the connection between the Hamiltonian dynamics and the Langevin dynamics is relatively well understood in certain models, see for example [KSTT02], it is less established in many other cases.

It is important to emphasize that the purpose of the Langevin dynamics here is not to reproduce the physical behavior in time truthfully but rather to serve as a tool for sampling statistical quantities from time-independent equilibria later on. In that sense

the time parameter t is always meant in a non-physical way and it is understood that further enhancements of the model may be required.

It is also worth noting that the above model neglects any fluctuations within the biomembrane, which can be interpreted as the implicit assumption that fluctuations of the membrane happen on a much faster timescale than those of the remaining particles. Moreover, the fluctuations in the particle parameters $\mathbf{p}_i$ are modeled by Brownian motion, which is arguably well-established when a $\mathbf{p}_i$ describes a lateral position, but may be a less canonical choice if $\mathbf{p}_i$ describes other degrees of freedom.

As a technicality, it shall be assumed that $H(\mathbf{p}, \mathbf{v}) \to \infty$ sufficiently fast for $\mathbf{p} \to \partial\mathcal{D}$ such that trajectories $(\mathbf{p}(t), \mathbf{v}(t))_{t\geq 0}$ stay inside of the phase space $\mathrm{T}\mathcal{D}$ almost-surely.

In order to relate the microscopic phase space to a macroscopic state of the system, a *stochastic ensemble* is required, which induces a probability density μ on the phase space $\mathrm{T}\mathcal{D}$. One standard choice is the *canonical ensemble* which describes the system as embedded into an infinite heat bath and with a fixed number of particles, volume, and also temperature. Under the condition that the *fluctuation–dissipation theorem* holds with the consequential relation $\gamma = \frac{1}{2}\sigma^2\beta$, where $\beta = (k_B T)^{-1}$ denotes *inverse temperature*, it leads to the *canonical measure*

$$\mu(\mathbf{p}, \mathbf{v}) := Z_\mu^{-1} \exp(-\beta H(\mathbf{p}, \mathbf{v})), \qquad Z_\mu := \int_{\mathrm{T}\mathcal{D}} \exp(-\beta H(\mathbf{p}, \mathbf{v}))\, \mathrm{d}(\mathbf{p}, \mathbf{v}).$$

Then, given a *microscopic observable* $A\colon \mathrm{T}\mathcal{D} \to \mathbb{R}$, the *macroscopic quantity associated with* A is defined by its expectational value with respect to the canonical measure, i. e.

$$\mathbb{E}_\mu(A) := \int_{\mathrm{T}\mathcal{D}} A(\mathbf{p}, \mathbf{v})\, \mathrm{d}\mu(\mathbf{p}, v).$$

The dynamics $(\mathbf{p}(t), \mathbf{v}(t))_{t\geq 0}$ is called *ergodic*, if for every observable holds

$$\mathbb{E}_\mu(A) = \lim_{T\to\infty} \frac{1}{T} \int_0^T A(\mathbf{p}(t), \mathbf{v}(t)). \tag{6.2}$$

This essentially means that the trajectory $(\mathbf{p}(t), \mathbf{v}(t))_{t\in\mathbb{R}}$ is able to visit all points in the phase space $\mathrm{T}\mathcal{D}$ while replicating the canonical density μ.

As a related quantity, the *absolute Helmholtz free energy* is defined up to a constant by

$$F := -\frac{\ln(Z_\mu)}{\beta}$$

and can be interpreted as a measure of stability in the system where a lower free energy means a more stable state of the system. In general, not the absolute free energy itself but rather the *free energy difference* $\Delta F := F_1 - F_0$ resulting from different Hamiltonians H_0 and H_1 with the joint phase space $\mathrm{T}\mathcal{D}$ is of interest. However, the evaluation of the integrals Z_{μ_0} and Z_{μ_1} by standard quadrature methods usually is prohibitive due to the potentially large dimension of the space $\mathrm{T}\mathcal{D}$ as well as the possibly complicated

geometry in $\mathcal{D}$. A well-established and numerically feasible way to compute such free energy differences is the *free energy perturbation* method due to Zwanzig [Zwa54], which relates ΔF to the observable $A(\mathbf{p}, \mathbf{v}) := \exp(-\beta(H_1(\mathbf{p}, \mathbf{v}) - H_0(\mathbf{p}, \mathbf{v})))$ via

$$e^{-\beta \Delta F} = \frac{Z_{\mu_1}}{Z_{\mu_0}} = Z_{\mu_0}^{-1} \int_{\mathrm{T}\mathcal{D}} e^{-\beta H_1(\mathbf{p},\mathbf{v})} \,\mathrm{d}(\mathbf{p}, \mathbf{v}) = \int_{\mathrm{T}\mathcal{D}} e^{-\beta(H_1(\mathbf{p},\mathbf{v}) - H_0(\mathbf{p},\mathbf{v}))} \,\mathrm{d}\mu_0 = \mathbb{E}_{\mu_0}(A).$$

Therefore, if $(\mathbf{p}(t), \mathbf{v}(t))_{t \geq 0}$ is a realization of the dynamics (6.1) and ergodic in the sense of (6.2), and if furthermore $(\mathbf{p}_{\tau,i}, \mathbf{v}_{\tau,i})_{i \in \{0,\dots,N\}}$ is a discrete approximation with step size $\tau \in \mathbb{R}_{>0}$, then one may approximate

$$\begin{aligned} e^{-\beta \Delta F} = \mathbb{E}_{\mu_0}(A) &= \lim_{T \to \infty} \frac{1}{T} \int_0^T e^{-\beta(H_1(\mathbf{p}(t),\mathbf{v}(t)) - H_0(\mathbf{p}(t),\mathbf{v}(t)))} \,\mathrm{d}t \\ &\approx \frac{1}{N} \sum_{i=0}^{N} e^{-\beta(H_1(\mathbf{p}_{\tau,i},\mathbf{v}_{\tau,i}) - H_0(\mathbf{p}_{\tau,i},\mathbf{v}_{\tau,i}))}. \end{aligned}$$

If both $H_0(\mathbf{p}, \mathbf{v}) = V_0(\mathbf{p}) + E_{\mathrm{kin}}(\mathbf{v})$ and $H_1(\mathbf{p}, \mathbf{v}) = V_1(\mathbf{p}) + E_{\mathrm{kin}}(\mathbf{v})$ are separable Hamiltonians with the same kinetic energy term E_{kin}, then the formula reduces further to

$$e^{-\beta \Delta F} = \lim_{T \to \infty} \frac{1}{T} \int_0^T e^{-\beta(V_1(\mathbf{p}(t)) - V_0(\mathbf{p}(t))} \,\mathrm{d}t \approx \frac{1}{N} \sum_{i=0}^{N} e^{-\beta(V_1(\mathbf{p}_{\tau,i}) - V_0(\mathbf{p}_{\tau,i}))}$$

where the trajectory $(p(t))_{t \in \mathbb{R}_{>0}}$ needs to be ergodic with respect to the probability density ν_0 over the configuration space $\mathcal{D}$. It is defined by

$$\nu_0(\mathbf{p}) := Z_{\nu_0}^{-1} \exp(-\beta V_0(\mathbf{p})), \qquad Z_{\nu_0} := \int_{\mathcal{D}} \exp(-\beta V_0(\mathbf{p})) \,\mathrm{d}\mathbf{p}$$

Under certain conditions, such a trajectory $(\mathbf{p}(t))_{t \geq 0}$ is obtained from samples of the *overdamped Langevin equation*

$$\mathrm{d}\mathbf{p}(t) = -\nabla \mathcal{E}(\mathbf{p}(t)) + \sqrt{\frac{2}{\beta}} \,\mathrm{d}W_t$$

where W_t is again a $(k \times n)$-dimensional Brownian motion. This equation is arguably the most easily discretized via the *explicit Euler–Maruyama-scheme*

$$\mathbf{p}_{\tau,i+1} = \mathbf{p}_{\tau,i} - \tau \nabla \mathcal{E}(\mathbf{p}_{\tau,i}) + \sqrt{\frac{2\tau}{\beta}} G_i \tag{6.3}$$

where the entries of the $G_i \in \mathbb{R}^{n \times k}$ are independent and identically distributed centered Gaussian random variables with unit variance. Since, mathematically speaking, this is very closely related to a (systematically) perturbed gradient flow method, the computation of free energy differences is immediately accessible by means of the implementation introduced in the earlier chapters. A numerical illustration is shown in Example 6.1 and Fig. 6.1.

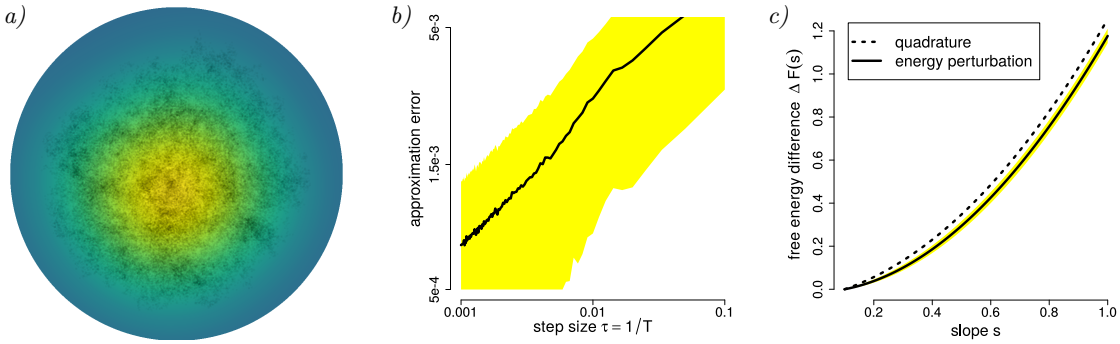

Figure 6.1: Numerical results as obtained from application of the free energy perturbation method applied to Example 6.1. From left to right: a) Positions in a trajectory generated by the Euler–Maruyama scheme (6.3). b) Convergence of the energy perturbation method to $\Delta F(0.2)$ for $\tau \to 0$ and $T \to \infty$. c) Free energy differences $\Delta F(s)$ for $s \in [0.1, 1]$.

Example 6.1. *Consider a single conical inclusion over a flat circular domain in Monge-gauge, $\mathcal{M}_0 := \Omega := \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$, and let the particle induce a circular particle–membrane interface of radius 0.1, i. e. $\mathcal{G}_0 := \Gamma_0 := \{x \in \mathbb{R}^2 \mid \|x\| = 0.1\}$. The bending rigidity is set to $\kappa = 1$ and surface tension is neglected, $\sigma = 0$. It is assumed that the particle is free to move in space laterally, but no vertical movement is admitted. Hence, the particle state is described by a two-dimensional position vector $\boldsymbol{p} \in \mathcal{D} \subseteq \mathbb{R}^2$ where the configuration space is set to $\mathcal{D} := \{\boldsymbol{p} \in \mathbb{R}^2 \mid \|p\| < 0.9\}$ in order to avoid collision of the particle with the outer boundary $\partial\Omega$. The parametric interfaces are given by $\Gamma(\boldsymbol{p}) := \{x \in \mathbb{R}^2 \mid \|x - \boldsymbol{p}\| = 0.1\}$ and the induced Dirichlet boundary conditions read $u|_{\Gamma(\boldsymbol{p})} = 0$, $\partial_\nu u|_{\Gamma(\boldsymbol{p})} = s$ where $s \in \mathbb{R}$ is a fixed slope. Homogeneous Dirichlet conditions are imposed $\boldsymbol{p}$-independently on $\partial\Omega$.*

The aim is to investigate how the free energy of this system changes upon changes in the slope s. To this end, the energy perturbation method is applied to the system for computing the free energy differences $\Delta F(s) := F(s) - F(s_0)$. The reference slope is chosen to be $s_0 := 0.1$ and the inverse temperature is set to $\beta = 25$.

Numerical results are presented in Fig. 6.1. The black points in plot a) show the positions of an example trajectory generated by the Euler–Maruyama scheme (6.3) with $\tau = 0.01$ and $T = 1000$. The circular domain depicts the membrane domain Ω and bright colors indicate the position density of the trajectory as obtained from a kernel density estimate. There it is visible how the dynamics explore the position space $\mathcal{D}$ while residing preferentially in the central high probability areas of the canonical measure. In b) the convergence behavior of the energy perturbation method is explored at $s = 0.2$ for $\tau \to 0$ and $T \to \infty$ in a double logarithmic plot. To this end, the exact value $\Delta F(s)$ was approximated by standard quadrature schemes and the energy perturbation method was applied for various step sizes $\tau \in [10^{-3}, 0.1]$ and with $T = \tau^{-1}$. At each point the solid line shows the mean error of the values obtained from 2500 runs of the energy

perturbation method given the corresponding τ-value, and the adjacent yellow shaded area depicts this mean error plus or minus one standard deviation. One observes that the method converges roughly with rate $\frac{1}{2}$. In order to assess the optimality of this behavior a more extensive study of parameter choices or the consideration of enhanced methods may be advised. Finally, figure c) shows a plot of the free energy differences $\Delta F(s)$ for $s \in [s_0, 1]$. The dotted curve shows the exact values as computed by a numerical quadrature scheme and the solid line shows the mean of results obtained from 100 trajectories in the energy perturbation method with $\tau = 0.1$ and $T = 1$ where the initial position always is $\boldsymbol{p}^0 = 0$. The surrounding yellow shaded area indicates the mean value plus or minus two standard deviations. While this approximation is rather simple, it is already sufficient to replicate the qualitative behavior of the exact values. Indeed further computations showed that the difference between the two curves would quickly become indistinguishable for more expensive parameter choices. While this is a promising first result in favor of the applicability of the free energy perturbation method in the context of hybrid particle–membrane interaction models, further investigations are still required in order to ensure the applicability of the method in more challenging applications as well.

Altogether, these considerations merely open a first perspective onto the possibilities of the application of stochastic gradient-type methods in the context of membrane-mediated particle interactions. A comprehensive introduction to free energy computations, including further approximation methods and convergence results is found in [LSR10]. Here future challenges may in particular lie in the further development and statistical justification of the model and the more efficient computation of free energy differences. Also, uses of Langevin-type computations beyond free energies might be of interest, for example in the context of global optimization via so-called *simulated annealing methods*, see e. g. [HP95, HPR02] for an introduction.

Appendix

A Matrix calculus: Differentiation of determinants and inverse matrices

In what follows, let $\mathcal{B} \subseteq \mathbb{R}^d$ an open domain, $A \in C^1(\mathcal{B}, \mathbb{R}^{n\times n})$ and $x \in \mathcal{B}$. The symbol ∂_i denotes the partial derivative with respect to the i-th component.

Lemma A.1. *If $A(x)$ is regular, then*

$$\partial_i \left(A(x)^{-1}\right) = -A(x)^{-1} \left(\partial_i A(x)\right) A(x)^{-1}.$$

Proof. [MN88, Chapter 8] □

Lemma A.2 (Jacobi's formula)**.** *It holds*

$$\partial_i \det(A(x)) = \operatorname{Tr}\left(\operatorname{adj}(A(x))\partial_i A(x)\right).$$

If $A(x)$ is regular, then

$$\partial_i \det(A(x)) = \det(A(x)) \operatorname{Tr}\left(A(x)^{-1}\partial_i A(x)\right).$$

Proof. [MN88, Chapter 8] □

B Differential geometry for perturbed surfaces

This part is concerned with the computation of an expression for the linearized elastic energy J_{lin} in the sense of Section 1.1.2 which does not rely explicitly on geometric quantities of the perturbed surfaces $\mathcal{M}(t)$, but instead only uses terms in the perturbation function u and the reference surface $\mathcal{M}_0$. The following derivation of such a representation of J_{lin} relies crucially on results for the analysis of surface PDEs and surface differentials in the sense of [DE13]. Moreover, the general approach is also based on ideas from [Hob16] where first and second variations of the Willmore functional are derived with the help of such methods. In the spirit of developing a widely applicable particle–membrane interaction framework, those results are now extended to more general energies which include spontaneous curvature and Gaussian curvature. In addition, the analysis is carried out differently in order to avoid non-symmetric terms entering the expressions in case of a non-vanishing boundary $\partial\mathcal{M}_0$, and for the sake of easier accessibility of the formulas the presentation is carried out on the basis of matrix calculus as far as possible. Similar results for even more general energies but only for variations up to first order have been derived in [DN12].

First, the necessary notation and some basic results are introduced. To this end, let $\varphi\colon \Omega \to \mathcal{M}_0$ a parameterization of $\mathcal{M}_0$ as before and let a map of t-dependent parameterizations of the perturbed surfaces $\mathcal{M}(t)$ be given by

$$\begin{aligned}\phi\colon [0,1]\times\Omega &\longrightarrow \mathbb{R}^3\\ (t,x) &\longmapsto \varphi(x) + t\,u(\varphi(x))\,\nu_0(\varphi(x)).\end{aligned}$$

The space–time manifold of perturbed surfaces is defined by

$$\mathcal{M}_T := \bigcup_{t\in[-1,1]} \{t\}\times\mathcal{M}(t).$$

In a slight abuse of notation, the maps u and ν_0 are extended to $\mathcal{M}_T$ via the definition $u(t,x) := u(\varphi(\phi(t)^{-1}(x)))$ and $\nu_0(t,x) := \nu_0(\varphi(\phi(t)^{-1}(x)))$.

By $\nu\colon \mathcal{M}_T \to \mathbb{R}^3$ the family of oriented unit normal vector fields on $\mathcal{M}(t)$ is denoted which is continuous with respect to t and which fulfills $\nu(0) = \nu_0$. Note that generally $\nu_0(t) \neq \nu(t)$ holds for $t \neq 0$.

The *metric tensor on* $\mathcal{M}(t)$ is given by $G(t) := D\phi(t)^T D\phi(t)$, where here and in the following the letter D always denotes the Jacobian with respect to the spatial variables only, i. e. excluding the time-component.

For any function $f\colon \mathcal{M}_T \to \mathbb{R}^k$, define the associated *parameterized function of* f by $\widetilde{f}(t,x) := f(t,\phi(t,x))$ for all $(t,x)\in[-1,1]\times\Omega$. The *material derivative of* f is defined implicitly by

$$\dot\partial f\circ\phi(t) := \frac{\partial}{\partial t}\widetilde{f}. \tag{B.1}$$

The *surface derivative of* $f(t)$ is defined by

$$D_{\mathcal{M}(t)} f(t)\circ\phi(t) := D\widetilde{f}(t)\,G(t)^{-1}D\phi(t)^T \tag{B.2}$$

and the *surface gradient* is $\nabla_{\mathcal{M}(t)} f(t) := (D_{\mathcal{M}(t)} f(t))^T$. For $k = 1$ the *surface Hessian of* $f(t)$ is

$$D^2_{\mathcal{M}(t)} f(t) := \nabla_{\mathcal{M}(t)} \nabla_{\mathcal{M}(t)} f(t).$$

Both the material derivatives and the surface derivatives are independent of the choice of the parameterization φ.

The families $H(t)$ and $K(t)$ denote the mean and Gaussian curvature on $\mathcal{M}(t)$, respectively, and $\kappa_1(t)$ and $\kappa_2(t)$ are the principal curvatures with principal directions $\mu_1(t)$ and $\mu_2(t)$, respectively. In case $\kappa_1(t,x) = \kappa_2(t,x)$ holds for some $(t,x) \in \mathcal{M}_T$, it is assumed that the μ_i are chosen such that the $\kappa_i(t)$ and $\mu_i(t)$ are differentiable almost-everywhere.

The *extended Weingarten map on* $\mathcal{M}$ is defined by $\mathcal{H}(t) := D_{\mathcal{M}(t)} \nu(t)$. The family $\mathcal{H}(t)$ defines a field of pointwise symmetric matrices and it holds $\mathcal{H}(t)\nu(t) = 0$ as well as $\mathcal{H}(t)\mu_i(t) = \kappa_i(t)\mu_i(t)$. In addition, a regularization of the extended Weingarten map $\hat{\mathcal{H}}(t) := \mathcal{H}(t) + \nu \otimes \nu$ is used, which fulfills $K = \det(\hat{\mathcal{H}}(t))$. Here, $a \otimes b := ab^T$ denotes the outer vector product.

In this context the identity (cf. [DE13, Lemma 2.7])

$$D^2_{\mathcal{M}(t)} f(t) - D^2_{\mathcal{M}(t)} f(t)^T = \nu(t) \otimes (\mathcal{H} \nabla_{\mathcal{M}(t)} f(t)) - (\mathcal{H} \nabla_{\mathcal{M}(t)} f(t)) \otimes \nu(t) \tag{B.3}$$

is noteworthy. Most importantly, it reveals that, unlike in the Euclidean case, the surface Hessian $D^2_{\mathcal{M}(t)} f(t)$ needs not be symmetric, but $D^2_{\mathcal{M}(t)} f(t) - \nu(t) \otimes (\mathcal{H} \nabla_{\mathcal{M}(t)} f(t))$ is. From this it is readily shown that $D^2_{\mathcal{M}(t)} f(t)$ induces a symmetric bilinear form on the tangent space of $\mathcal{M}(t)$.

The parameter t is dropped whenever the dependence and implied value are clear from the context. Furthermore, for convenience and clarity of presentation the shortcut $H_0 := H(0)$ is defined. This completes the preparations for the actual linearization of the elastic bending energy.

B.1 Material derivatives of common quantities

Lemma B.1. *Let* $\widetilde{u} := u \circ \phi$ *and* $\widetilde{\nu}_0 := \nu_0 \circ \phi$. *It holds*

$$\frac{\partial}{\partial t} \phi = \widetilde{u}\, \widetilde{\nu}_0 \tag{B.4}$$

$$\frac{\partial}{\partial t} D\phi = \widetilde{\nu}_0 \otimes \nabla \widetilde{u} + \widetilde{u} D \widetilde{\nu}_0 \tag{B.5}$$

$$\frac{\partial}{\partial t} G = D\phi^T \left(\widetilde{\nu}_0 \otimes \nabla \widetilde{u} + \widetilde{u} D \widetilde{\nu}_0 \right) + \left(\nabla \widetilde{u} \otimes \widetilde{\nu}_0 + \widetilde{u} D \widetilde{\nu}_0^T \right) D\phi \tag{B.6}$$

$$\frac{\partial}{\partial t} G^{-1} = -G^{-1} \left(\frac{\partial}{\partial t} G \right) G^{-1} \tag{B.7}$$

Proof. The identities (B.4), (B.5) and (B.6) are immediately clear from the definition of the involved quantities. Equation (B.7) corresponds to the result from Lemma A.1 on differentiation of matrix inverses. □

Lemma B.2. *Let*

$$\begin{aligned}\mathcal{C} &:= \nabla_{\mathcal{M}(t)} u \otimes (\nu\nu^T\nu_0) - \nu_0 \otimes \nabla_{\mathcal{M}(t)} u - u D_{\mathcal{M}(t)}\nu_0\\ \mathcal{C}_0 &:= \nabla_{\mathcal{M}_0} u \otimes \nu_0 - u\mathcal{H}_0.\end{aligned}$$

Then

$$\dot\partial D_{\mathcal{M}(t)} f = \left(D_{\mathcal{M}(t)} f\right)\mathcal{C} + D_{\mathcal{M}(t)}\dot\partial f$$

and

$$\dot\partial D_{\mathcal{M}(t)} f\Big|_{t=0} = \left(D_{\mathcal{M}_0} f\right)\mathcal{C}_0 + D_{\mathcal{M}_0}\dot\partial f\Big|_{t=0}$$

Proof. Application of Lemma B.1 yields

$$\begin{aligned}D\widetilde{f}G^{-1}\frac{\partial}{\partial t}D\phi^T &= D\widetilde{f}G^{-1}\left(\nabla\widetilde{u}\otimes\widetilde{\nu}_0 + \widetilde{u}D\widetilde{\nu}_0^T\right)\\ &= D\widetilde{f}G^{-1}D\phi^T\left(D\phi G^{-1}\widetilde{u}\otimes\widetilde{\nu}_0 + \widetilde{u}D\phi G^{-1}D\widetilde{\nu}_0^T\right)\\ &= \left(D_{\mathcal{M}(t)}f\left(\nabla_{\mathcal{M}(t)}u\otimes\nu_0 + u(D_{\mathcal{M}(t)}\nu_0)^T\right)\right)\circ\phi.\end{aligned} \tag{B.8}$$

The fact that $\Pi := D\phi G^{-1}D\phi^{-1}$ is an orthogonal projection onto the tangent space implies $D\phi G^{-1}D\phi^T\left(D_{\mathcal{M}(t)}\nu_0\circ\phi\right) = D_{\mathcal{M}(t)}\nu_0\circ\phi$. Therefore,

$$\begin{aligned}D\widetilde{f}\left(\frac{\partial}{\partial t}G^{-1}\right)D\phi^T &= -D\widetilde{f}G^{-1}D\phi^T\left(\widetilde{\nu}_0\otimes\nabla\widetilde{u} + \widetilde{u}D\widetilde{\nu}_0\right)G^{-1}D\phi^T\\ &\quad - D\widetilde{f}G^{-1}D\phi^T\left(D\phi G^{-1}\nabla\widetilde{u}\otimes\widetilde{\nu}_0 + \widetilde{u}D\phi G^{-1}D\widetilde{\nu}_0^T\right)D\phi G^{-1}D\phi^T\\ &= -\left(D_{\mathcal{M}(t)}f\left(\nu_0\otimes\nabla_{\mathcal{M}(t)}u + uD_{\mathcal{M}(t)}\nu_0\right)\right)\circ\phi\\ &\quad - \left(D_{\mathcal{M}(t)}f\left(\nabla_{\mathcal{M}(t)}u\otimes(\Pi\nu_0) + u(D_{\mathcal{M}(t)}\nu_0)^T\right)\right)\circ\phi.\end{aligned} \tag{B.9}$$

Note that the orthogonal complement of the projection Π fulfills $\nu_0 - \Pi\nu_0 = \nu\nu^T\nu_0$. Altogether, it is concluded from the definitions in (B.1) and (B.2) and by combining (B.8) and (B.9) that

$$\begin{aligned}\dot\partial D_{\mathcal{M}(t)} f\circ\phi &= \frac{\partial}{\partial t}\left(D\widetilde{f}G^{-1}D\phi^T\right)\\ &= D\left(\frac{\partial}{\partial t}\widetilde{f}\right)G^{-1}D\phi^T + D\widetilde{f}\,\frac{\partial}{\partial t}\left(G^{-1}D\phi\right)\\ &= \left(D_{\mathcal{M}(t)}\dot\partial f + \left(D_{\mathcal{M}(t)}f\right)\mathcal{C}\right)\circ\phi,\end{aligned}$$

which proves the first equation. The second equation for $t = 0$ is an immediate consequence of $\nu\nu^T\nu_0\big|_{t=0} = \nu_0$ and $\left(D_{\mathcal{M}(t)}f\right)\nu_0\big|_{t=0} = 0$. □

Lemma B.3. *It holds*

$$\dot\partial u = 0 \qquad \dot\partial \nu_0 = 0$$
$$\dot\partial \nabla_{\mathcal{M}(t)} u = \mathcal{C}^T \nabla_{\mathcal{M}(t)} u \qquad \dot\partial D_{\mathcal{M}(t)} \nu_0 = D_{\mathcal{M}(t)} \nu_0 \, \mathcal{C}$$

Proof. This is a direct consequence of Lemma B.2 and the definition of u and ν_0. □

Lemma B.4. *It holds*

$$\dot\partial \nu = -\left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u D_{\mathcal{M}(t)} \nu_0\right) \nu \tag{B.10}$$
$$\dot\partial^2 \nu = \left(\left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u D_{\mathcal{M}(t)} \nu_0\right)^2 - \mathcal{C}^T \left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u D_{\mathcal{M}(t)} \nu_0^T\right)\right) \nu \tag{B.11}$$

and

$$\dot\partial \nu\Big|_{t=0} = -\nabla_{\mathcal{M}_0} u \tag{B.12}$$
$$\dot\partial^2 \nu\Big|_{t=0} = 2u\mathcal{H}_0 \nabla_{\mathcal{M}_0} u - \|\nabla_{\mathcal{M}_0} u\|^2 \nu_0. \tag{B.13}$$

Proof. Equation (B.5) implies

$$0 = \frac{\partial}{\partial t}\left(D\phi^T \widetilde\nu\right) = \left(\nabla \widetilde u \otimes \widetilde\nu_0 + \widetilde u (D\widetilde\nu_0)^T\right) \widetilde\nu + D\phi^T \frac{\partial}{\partial t} \widetilde\nu$$

and therefore

$$D\phi^T \left(\dot\partial \nu \circ \phi\right) = -\left(\nabla \widetilde u \otimes \widetilde\nu_0 + \widetilde u (D\widetilde\nu_0)^T\right) \widetilde\nu. \tag{B.14}$$

In addition,

$$0 = \dot\partial(\nu^T \nu) = 2\nu^T \dot\partial \nu,$$

from which it is inferred that $\dot\partial\nu$ is an element of the tangent space. Consequently, $\Pi \dot\partial \nu = \dot\partial \nu$ holds where $\Pi = D\phi G^{-1} D\phi^T$ is the orthogonal projection onto the tangent space. In combination with (B.14) equation (B.10) is obtained by

$$\begin{aligned}
\dot\partial \nu \circ \phi &= D\phi G^{-1} D\phi^T \left(\dot\partial \nu \circ \phi\right) \\
&= -D\phi G^{-1} \left(\nabla \widetilde u \otimes \widetilde\nu_0 + \widetilde u (D\widetilde\nu_0)^T\right) \widetilde\nu \\
&= -\left(\left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u (D_{\mathcal{M}(t)} \nu_0)^T\right) \nu\right) \circ \phi.
\end{aligned}$$

Application of Lemma B.2 and Lemma B.3 yields (B.11) by

$$\begin{aligned}
\dot\partial\left(\dot\partial \nu\right) &= -\dot\partial \left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u (D_{\mathcal{M}(t)} \nu_0)^T\right) \nu - \left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u (D_{\mathcal{M}(t)} \nu_0)^T\right) \dot\partial \nu \\
&= -\mathcal{C}^T \left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u (D_{\mathcal{M}(t)} \nu_0)^T\right) \nu + \left(\nabla_{\mathcal{M}(t)} u \otimes \nu_0 + u (D_{\mathcal{M}(t)} \nu_0)^T\right)^2 \nu.
\end{aligned}$$

Equation (B.12) is immediately apparent from $\nu|_{t=0} = \nu_0$ and $D_{\mathcal{M}_0}\nu_0 \cdot \nu_0 = 0$. Finally, for (B.13) the symmetry of $\mathcal{H}_0 = D_{\mathcal{M}_0}\nu_0$ is used to conclude

$$\begin{aligned}\dot{\partial}^2\nu\Big|_{t=0} &= \left(\left(\nabla_{\mathcal{M}_0}u \otimes \nu_0 + uD_{\mathcal{M}_0}\nu_0\right)^2 - \mathcal{C}_0^T\left(\nabla_{\mathcal{M}_0}u \otimes \nu_0 + uD_{\mathcal{M}_0}\nu_0\right)\right)\nu_0 \\ &= uD_{\mathcal{M}_0}\nu_0\nabla_{\mathcal{M}_0}u - \mathcal{C}_0^T\nabla_{\mathcal{M}_0}u \\ &= 2u\mathcal{H}_0\nabla_{\mathcal{M}_0}u - \|\nabla_{\mathcal{M}_0}u\|^2\nu_0.\end{aligned}$$

□

Lemma B.5. *Let*

$$\dot{\mathcal{C}}_0 := u^2\mathcal{H}_0^2 - 2u\mathcal{H}_0\nabla_{\mathcal{M}_0}u \otimes \nu_0 - \nabla_{\mathcal{M}_0}u \otimes \nabla_{\mathcal{M}_0}u.$$

It holds

$$D_{\mathcal{M}_0}f\,\dot{\partial}\mathcal{C}|_{t=0} = D_{\mathcal{M}_0}f\,\dot{\mathcal{C}}_0.$$

Proof. In view of Lemma B.3 and Lemma B.4 holds $\dot{\partial}\left(\nu\nu^T\nu_0\right)|_{t=0} = -\nabla_{\mathcal{M}_0}u$. Therefore,

$$\begin{aligned}\dot{\partial}\mathcal{C}|_{t=0} &= \dot{\partial}\left(\nabla_{\mathcal{M}(t)}u \otimes (\nu\nu^T\nu_0) - \nu_0 \otimes \nabla_{\mathcal{M}(t)}u - uD_{\mathcal{M}(t)}\nu_0\right)\Big|_{t=0} \\ &= \mathcal{C}_0^T\nabla_{\mathcal{M}_0}u \otimes \nu_0 - \nabla_{\mathcal{M}_0}u \otimes \nabla_{\mathcal{M}_0}u - \nu_0 \otimes (\mathcal{C}_0^T\nabla_{\mathcal{M}_0}u) - u\mathcal{H}_0\mathcal{C}_0.\end{aligned}$$

As of

$$\mathcal{C}_0 = \nabla_{\mathcal{M}_0}u \otimes \nu_0 - u\mathcal{H}_0$$

and $D_{\mathcal{M}_0}f\nu_0 = 0$ the identity

$$D_{\mathcal{M}_0}f\,\dot{\partial}\mathcal{C}|_{t=0} = D_{\mathcal{M}_0}f\left(-u\mathcal{H}_0\nabla_{\mathcal{M}_0}u \otimes \nu_0 - \nabla_{\mathcal{M}_0}u \otimes \nabla_{\mathcal{M}_0}u - u\mathcal{H}_0\nabla_{\mathcal{M}_0}u \otimes \nu_0 + u^2\mathcal{H}_0^2\right)$$

is true and completes the proof. □

Lemma B.6. *It holds*

$$\dot{\partial}^2 D_{\mathcal{M}(t)}f = D_{\mathcal{M}(t)}f\left(\mathcal{C}^2 + \dot{\partial}\mathcal{C}\right) + 2\left(D_{\mathcal{M}(t)}\dot{\partial}f\right)\mathcal{C} + D_{\mathcal{M}(t)}\dot{\partial}^2 f$$

and

$$\dot{\partial}^2 D_{\mathcal{M}(t)}f|_{t=0} = D_{\mathcal{M}_0}f\left(\mathcal{C}_0^2 + \dot{\mathcal{C}}_0\right) + 2\left(D_{\mathcal{M}_0}\dot{\partial}f\right)\mathcal{C}_0 + D_{\mathcal{M}_0}\dot{\partial}^2 f$$

where

$$\mathcal{C}_0^2 + \dot{\mathcal{C}}_0 = 2u^2\mathcal{H}_0^2 - 3u\mathcal{H}_0\nabla_{\mathcal{M}_0}u \otimes \nu_0 - \nabla_{\mathcal{M}_0}u \otimes \nabla_{\mathcal{M}_0}u.$$

Proof. Repeated application of Lemma B.2 yields

$$\begin{aligned}\dot\partial\left(\dot\partial D_{\mathcal{M}(t)}f\right) &= \dot\partial\left(D_{\mathcal{M}(t)}f\,\mathcal{C}+D_{\mathcal{M}(t)}\dot\partial f\right)\\ &= \left(\dot\partial D_{\mathcal{M}(t)}f\right)\mathcal{C}+\left(D_{\mathcal{M}(t)}f\right)\dot\partial\mathcal{C}+\left(D_{\mathcal{M}(t)}\dot\partial f\right)\mathcal{C}+D_{\mathcal{M}(t)}\dot\partial^2 f\\ &= D_{\mathcal{M}(t)}f\left(\mathcal{C}^2+\dot\partial\mathcal{C}\right)+2\left(D_{\mathcal{M}(t)}\dot\partial f\right)\mathcal{C}+D_{\mathcal{M}(t)}\dot\partial^2 f.\end{aligned}$$

The remainder of the statement is quickly verified by invoking Lemma B.5. □

Lemma B.7. *It holds*

$$\dot\partial\mathcal{H}|_{t=0}=\mathcal{H}_0\nabla_{\mathcal{M}_0}u\otimes\nu_0-u\mathcal{H}_0^2-(D^2_{\mathcal{M}_0}u)^T$$

and

$$\begin{aligned}\dot\partial^2\mathcal{H}|_{t=0} &= 2u^2\mathcal{H}_0^3-3u\mathcal{H}_0^2\nabla_{\mathcal{M}_0}u\otimes\nu_0\\ &\quad+\mathcal{H}_0\left(\nabla_{\mathcal{M}_0}u\otimes\nabla_{\mathcal{M}_0}u+2u(D^2_{\mathcal{M}_0}u)^T-\|\nabla_{\mathcal{M}_0}u\|^2 I\right)\\ &\quad+2(D^2_{\mathcal{M}_0}u)^T\left(u\mathcal{H}_0-\nabla_{\mathcal{M}_0}u\otimes\nu_0\right)+2uD_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0}u.\end{aligned}$$

Proof. From Lemma B.2, Lemma B.4 and $\nu|_{t=0}=\nu_0$ the equation

$$\begin{aligned}\dot\partial D_{\mathcal{M}(t)}\nu|_{t=0} &= D_{\mathcal{M}_0}\nu_0\,\mathcal{C}_0-D_{\mathcal{M}_0}\nabla_{\mathcal{M}_0}u\\ &= \mathcal{H}_0\nabla_{\mathcal{M}_0}u\otimes\nu_0-u\mathcal{H}_0^2-(D^2_{\mathcal{M}_0}u)^T\end{aligned}$$

is obtained. Similarly, application of Lemma B.6 yields

$$\begin{aligned}\dot\partial^2 D_{\mathcal{M}(t)}\nu|_{t=0} &= \mathcal{H}_0\left(\mathcal{C}_0^2+\dot{\mathcal{C}}_0\right)-2(D^2_{\mathcal{M}_0}u)^T\mathcal{C}_0+D_{\mathcal{M}_0}\left(2u\mathcal{H}_0\nabla_{\mathcal{M}_0}u-\|\nabla_{\mathcal{M}_0}u\|^2\nu_0\right)\\ &= 2u^2\mathcal{H}_0^3-3u\mathcal{H}_0^2\nabla_{\mathcal{M}_0}u\otimes\nu_0-\mathcal{H}_0\nabla_{\mathcal{M}_0}u\otimes\nabla_{\mathcal{M}_0}u\\ &\quad+2(D^2_{\mathcal{M}_0}u)^T\left(u\mathcal{H}_0-\nabla_{\mathcal{M}_0}u\otimes\nu_0\right)\\ &\quad+2\mathcal{H}_0\nabla_{\mathcal{M}_0}u\otimes\nabla_{\mathcal{M}_0}u+2u\mathcal{H}_0(D^2_{\mathcal{M}_0}u)^T+2uD_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0}u\\ &\quad-2\nabla_{\mathcal{M}_0}u^T(D^2_{\mathcal{M}_0}u)^T\nu_0-\|\nabla_{\mathcal{M}_0}u\|^2\mathcal{H}_0\end{aligned}$$

where $D_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0}u=\sum_{k=1}^2 D^2_{\mathcal{M}_0}\nu_{0,k}\nabla_{\mathcal{M}_0}u_k$. The proof is completed by collecting terms and by making use of $(D^2_{\mathcal{M}_0}u)^T\nu_0=0$. □

Lemma B.8. *It holds*

$$\dot\partial\hat{\mathcal{H}}|_{t=0}=\mathcal{H}_0\nabla_{\mathcal{M}_0}u\otimes\nu_0-u\mathcal{H}_0^2-(D^2_{\mathcal{M}_0}u)^T-\nabla_{\mathcal{M}_0}u\otimes\nu_0-\nu_0\otimes\nabla_{\mathcal{M}_0}u$$

and

$$\begin{aligned}\dot\partial^2\hat{\mathcal{H}}|_{t=0} &= 2u^2\mathcal{H}_0^3-3u\mathcal{H}_0^2\nabla_{\mathcal{M}_0}u\otimes\nu_0\\ &\quad+\mathcal{H}_0\left(\nabla_{\mathcal{M}_0}u\otimes\nabla_{\mathcal{M}_0}u+2u(D^2_{\mathcal{M}_0}u)^T-\|\nabla_{\mathcal{M}_0}u\|^2 I\right)\\ &\quad+2(D^2_{\mathcal{M}_0}u)^T\left(u\mathcal{H}_0-\nabla_{\mathcal{M}_0}u\otimes\nu_0\right)+2uD_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0}u\\ &\quad+\left(2u\mathcal{H}_0\nabla_{\mathcal{M}_0}u-\|\nabla_{\mathcal{M}_0}u\|^2\nu_0\right)\otimes\nu_0+\nu_0\otimes\left(2u\mathcal{H}_0\nabla_{\mathcal{M}_0}u-\|\nabla_{\mathcal{M}_0}u\|^2\nu_0\right)\\ &\quad+2\nabla_{\mathcal{M}_0}u\otimes\nabla_{\mathcal{M}_0}u.\end{aligned}$$

Proof. The statement is a combination of Lemma B.4 and Lemma B.7. □

Proposition B.9. *It holds*

$$\begin{aligned}\dot\partial H|_{t=0} &= -u\,\|\mathcal{H}_0\|_F^2 - \Delta_{\mathcal{M}_0} u\\ \dot\partial^2 H|_{t=0} &= 2u^2\,\mathrm{Tr}(\mathcal{H}_0^3) - \nabla_{\mathcal{M}_0} u^T \mathcal{H}_0 \nabla_{\mathcal{M}_0} u + 4u\mathcal{H}_0 : D^2_{\mathcal{M}_0} u\\ &\quad - \|\nabla_{\mathcal{M}_0} u\|^2 H_0 + 2u\Delta_{\mathcal{M}_0}\nu_0 \cdot \nabla_{\mathcal{M}_0} u\end{aligned}$$

where $\Delta_{\mathcal{M}_0}$ is the Laplace-Beltrami operator on $\mathcal{M}_0$ and $\|\cdot\|_F$ denotes the Frobenius norm with $:$ as the inducing matrix scalar product.

Proof. The proof makes use of the identities

$$\mathrm{Tr}(a\otimes b) = a^T b, \qquad \mathrm{Tr}(AB^T) = A:B, \qquad A:A = \|A\|_F^2$$

for vectors a, b and matrices A, B. From the properties of the extended Weingarten map the equality $H = \mathrm{Tr}(\mathcal{H})$ and the symmetry of $\mathcal{H}$ are known. With the help of Lemma B.7 one concludes

$$\begin{aligned}\dot\partial H|_{t=0} &= \mathrm{Tr}\left(\dot\partial \mathcal{H}|_{t=0}\right)\\ &= \mathrm{Tr}\left(\mathcal{H}_0\nabla_{\mathcal{M}_0} u\otimes\nu_0 - u\mathcal{H}_0^2 - D^2_{\mathcal{M}_0} u\right)\\ &= \nu_0^T\mathcal{H}_0\nabla_{\mathcal{M}_0} u - u\,\mathrm{Tr}(\mathcal{H}_0^2) - \mathrm{Tr}(D^2_{\mathcal{M}_0} u)\\ &= -u\,\|\mathcal{H}_0\|_F^2 - \Delta_{\mathcal{M}_0} u\end{aligned}$$

as well as

$$\begin{aligned}\dot\partial^2 H|_{t=0} &= \mathrm{Tr}\left(\dot\partial^2 \mathcal{H}|_{t=0}\right)\\ &= \mathrm{Tr}\left(2u^2\mathcal{H}_0^3 - 3u\mathcal{H}_0^2\nabla_{\mathcal{M}_0} u\otimes\nu_0\right)\\ &\quad + \mathrm{Tr}\left(\mathcal{H}_0\left(\nabla_{\mathcal{M}_0} u\otimes\nabla_{\mathcal{M}_0} u + 2u(D^2_{\mathcal{M}_0} u)^T - \|\nabla_{\mathcal{M}_0} u\|^2 I\right)\right)\\ &\quad + \mathrm{Tr}\left(2(D^2_{\mathcal{M}_0} u)^T(u\mathcal{H}_0 - \nabla_{\mathcal{M}_0} u\otimes\nu_0) + 2uD_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0} u\right)\\ &= 2u^2\,\mathrm{Tr}(\mathcal{H}_0^3) + \nabla_{\mathcal{M}_0} u^T\mathcal{H}_0\nabla_{\mathcal{M}_0} u + 2u\mathcal{H}_0 : D^2_{\mathcal{M}_0} u - \|\nabla_{\mathcal{M}_0} u\|^2\,\mathrm{Tr}(\mathcal{H}_0)\\ &\quad + 2u\mathcal{H}_0 : D^2_{\mathcal{M}_0} u - 2\nu_0^T(D^2_{\mathcal{M}_0} u)^T\nabla_{\mathcal{M}_0} u + 2u\,\mathrm{Tr}(D_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0} u).\end{aligned}$$

One further recognizes

$$\nu_0^T(D^2_{\mathcal{M}_0} u)^T\nabla_{\mathcal{M}_0} u = \nabla_{\mathcal{M}_0} u^T\mathcal{H}_0\nabla_{\mathcal{M}_0} u$$

by virtue of (B.3) as well as

$$\mathrm{Tr}(D_{\mathcal{M}_0}\mathcal{H}_0\nabla_{\mathcal{M}_0} u) = \sum_{i,k=1}^{2}(D^2_{\mathcal{M}_0}\nu_{0,k})_{ii}\nabla_{\mathcal{M}_0} u_k = \sum_{k=1}^{2}\Delta_{\mathcal{M}_0}\nu_{0,k}\nabla_{\mathcal{M}_0} u_k = \Delta_{\mathcal{M}_0}\nu_0\cdot\nabla_{\mathcal{M}_0} u.$$

Plugging these identities into the above formula completes the proof. □

Proposition B.10. *Let* $\operatorname{adj}(\hat{\mathcal{H}})$ *denote the adjugate matrix of* $\hat{\mathcal{H}}$*. It holds*

$$\begin{aligned}\dot{\partial} K|_{t=0} &= \operatorname{Tr}(\operatorname{adj}(\hat{\mathcal{H}})\dot{\partial}\hat{\mathcal{H}})|_{t=0}\\ \dot{\partial}^2 K|_{t=0} &= \operatorname{Tr}(\dot{\partial}\operatorname{adj}(\hat{\mathcal{H}})\,\dot{\partial}\hat{\mathcal{H}})|_{t=0} + \operatorname{Tr}(\operatorname{adj}(\hat{\mathcal{H}})\dot{\partial}^2\hat{\mathcal{H}})|_{t=0}\end{aligned}$$

where expressions for $\dot{\partial}\hat{\mathcal{H}}|_{t=0}$ *and* $\dot{\partial}^2\hat{\mathcal{H}}|_{t=0}$ *are given in Lemma B.8. If furthermore* $K(0)\neq 0$*, let* $A := \hat{\mathcal{H}}^{-1}\dot{\partial}\mathcal{H}^{-1}|_{t=0}$*. Then*

$$\begin{aligned}\dot{\partial} K|_{t=0} &= K\operatorname{Tr}(A)\\ \dot{\partial}^2 K|_{t=0} &= K\left(\operatorname{Tr}(A)^2 - \|A\|_F^2 + \operatorname{Tr}\left(\hat{\mathcal{H}}^{-1}\dot{\partial}^2\hat{\mathcal{H}}\right)|_{t=0}\right).\end{aligned}$$

Proof. This is a direct consequence of Jacobi's formula in Lemma A.2, the product rule, and Lemma A.1 for the differentiation of inverse matrices. □

B.2 Derivatives of integrals over geometric quantities

In this section integration of a hypersurface $\mathcal{M}(t)$ is always taken with respect to the corresponding Hausdorff-measure $\mathbb{H}(\mathcal{M}(t))$. In the following statements this is only indicated by the shorthand $\mathrm{d}\mathbb{H}$ for ease of notation.

Theorem B.11 (Transport formula)**.** *Let* V *such that* $\frac{\partial}{\partial t}\phi = V\circ\phi$*. It holds*

$$\frac{\partial}{\partial t}\int_{\mathcal{M}(t)} f\,\mathrm{d}\mathbb{H} = \int_{\mathcal{M}(t)} \dot{\partial} f + f\operatorname{div}_{\mathcal{M}(t)} V\,\mathrm{d}\mathbb{H}$$

and

$$\frac{\partial^2}{\partial t^2}\int_{\mathcal{M}(t)} f\,\mathrm{d}\mathbb{H} = \int_{\mathcal{M}(t)} \dot{\partial}^2 f + 2\dot{\partial} f\operatorname{div}_{\mathcal{M}(t)} V + f\left(\left(\operatorname{div}_{\mathcal{M}(t)} V\right)^2 + \dot{\partial}\operatorname{div}_{\mathcal{M}(t)} V\right)\mathrm{d}\mathbb{H}.$$

Proof. The proof for the first equation is found in [DE13, Theorem 5.1]. The second equation follows directly by repeated application of the first. □

Lemma B.12. *Let* V *such that* $\frac{\partial}{\partial t}\phi = V\circ\phi$*. Then*

$$\operatorname{div}_{\mathcal{M}(0)} V = uH_0$$

and

$$\dot{\partial}\operatorname{div}_{\mathcal{M}(t)} V|_{t=0} = \|\nabla_{\mathcal{M}_0} u\|^2 - u^2\|\mathcal{H}_0\|_F^2$$

Proof. The definition of V and Lemma B.1 give for $\widetilde{V} := V\circ\phi$

$$D_{\mathcal{M}(t)}V\circ\phi = D\widetilde{V}G^{-1}D\phi^T = \left(\frac{\partial}{\partial t}D\phi\right)G^{-1}D\phi^T = \left(\nu_0\otimes\nabla_{\mathcal{M}(t)}u + uD_{\mathcal{M}(t)}\nu_0\right)\circ\phi.$$

Consequently,

$$\operatorname{div}_{\mathcal{M}(0)} V = \operatorname{Tr}(D_{\mathcal{M}_0} V) = \operatorname{Tr}(\nu_0 \otimes \nabla_{\mathcal{M}_0} u + u\mathcal{H}_0) = \nabla_{\mathcal{M}_0} u^T \nu_0 + u \operatorname{Tr}(\mathcal{H}_0) = uH_0.$$

As of

$$\dot{\partial} V \circ \phi = \frac{\partial}{\partial t} \widetilde{V} = \frac{\partial^2}{\partial t^2} \phi = \frac{\partial}{\partial t} (\widetilde{u}\widetilde{\nu_0}) = 0$$

Lemma B.2 yields

$$\begin{aligned}
\dot{\partial} \operatorname{div}_{\mathcal{M}(t)} V|_{t=0} &= \operatorname{Tr}\left(\dot{\partial} D_{\mathcal{M}(t)} V|_{t=0}\right) \\
&= \operatorname{Tr}\left(D_{\mathcal{M}_0} V \, \mathcal{C}_0\right) \\
&= \operatorname{Tr}\left((\nu_0 \otimes \nabla_{\mathcal{M}_0} u + u\mathcal{H}_0)\left(\nabla_{\mathcal{M}_0} u \otimes \nu_0 - u\mathcal{H}_0\right)\right) \\
&= \operatorname{Tr}\left(\|\nabla_{\mathcal{M}_0} u\|^2 \nu_0 \otimes \nu_0 + u\mathcal{H}_0 \nabla_{\mathcal{M}_0} u \otimes \nu_0 - \nu_0 \otimes (\mathcal{H}_0 \nabla_{\mathcal{M}_0} u) - u^2 \mathcal{H}_0^2\right) \\
&= \|\nabla_{\mathcal{M}_0} u\|^2 - u^2 \|\mathcal{H}_0\|_F^2 .
\end{aligned}$$

□

Theorem B.13. *It holds*

$$\frac{\partial}{\partial t} \int_{\mathcal{M}(t)} 1 \, \mathrm{d}\mathbb{H} \bigg|_{t=0} = \int_{\mathcal{M}_0} uH_0 \, \mathrm{d}\mathbb{H}$$

and

$$\frac{\partial^2}{\partial t^2} \int_{\mathcal{M}(t)} 1 \, \mathrm{d}\mathbb{H} \bigg|_{t=0} = \int_{\mathcal{M}_0} \|\nabla_{\mathcal{M}_0} u\|^2 + 2u^2 K_0 \, \mathrm{d}\mathbb{H}.$$

Proof. The first equation is a direct consequence of $\dot{\partial} 1 = \dot{\partial}^2 1 = 0$, $D_{\mathcal{M}_0} 1 = 0$, Theorem B.11 and Lemma B.12. Similarly, one also obtains for the second equation

$$\frac{\partial^2}{\partial t^2} \int_{\mathcal{M}(t)} 1 \, \mathrm{d}\mathbb{H} \bigg|_{t=0} = \int_{\mathcal{M}_0} u^2 H_0^2 + \|\nabla_{\mathcal{M}_0} u\|^2 - u^2 \|\mathcal{H}_0\|_F^2 \, \mathrm{d}\mathbb{H}.$$

Since the Weingarten map $\mathcal{H}_0$ has the eigenvalues 0, κ_1 and κ_2 with corresponding eigenvalues ν_0, μ_1, μ_2, respectively, the matrix $\mathcal{H}_0^2$ has the eigenvalues 0, κ_1^2 and κ_2^2. Together with the symmetry of $\mathcal{H}_0$ this implies

$$\|\mathcal{H}_0\|_F^2 = \mathcal{H}_0 : \mathcal{H}_0 = \operatorname{Tr}(\mathcal{H}_0^2) = \kappa_1^2 + \kappa_2^2.$$

As of $H_0 = \kappa_1 + \kappa_2$ one further has

$$H_0^2 - \|\mathcal{H}_0\|_F^2 = (\kappa_1 + \kappa_2)^2 - (\kappa_1^2 + \kappa_2^2) = 2\kappa_1\kappa_2 = 2K_0.$$

Plugging this into the above identity completes the proof. □

Theorem B.14. *It holds*

$$\frac{\partial}{\partial t}\frac{1}{2}\int_{\mathcal{M}(t)}(H-c_0)^2\,\mathrm{d}\mathbb{H}\bigg|_{t=0} = \int_{\mathcal{M}_0} -(H_0-c_0)\left(u\,\|\mathcal{H}_0\|_F^2+\Delta_{\mathcal{M}_0}u+\dot\partial c_0\right) + \frac{1}{2}uH_0(H_0-c_0)^2\,\mathrm{d}\mathbb{H}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial t^2}\frac{1}{2}\int_{\mathcal{M}(t)}(H-c_0)^2\,\mathrm{d}\mathbb{H}\bigg|_{t=0} &= \int_{\mathcal{M}_0}\left(u\,\|\mathcal{H}_0\|_F^2+\Delta_{\mathcal{M}_0}u+\dot\partial c_0|_{t=0}\right)^2\,\mathrm{d}\mathbb{H}\\ &\quad+\int_{\mathcal{M}_0}(H_0-c_0)(\dot\partial^2 H|_{t=0}-\dot\partial^2 c_0|_{t=0})\,\mathrm{d}\mathbb{H}\\ &\quad-\int_{\mathcal{M}_0}2uH_0(H_0-c_0)\left(u\,\|\mathcal{H}_0\|_F^2+\Delta_{\mathcal{M}_0}u+\dot\partial c_0|_{t=0}\right)\,\mathrm{d}\mathbb{H}\\ &\quad+\int_{\mathcal{M}_0}\frac{1}{2}(H_0-c_0)^2\left(\|\nabla_{\mathcal{M}_0}u\|^2+u^2K_0\right)\,\mathrm{d}\mathbb{H}\end{aligned}$$

where $\dot\partial^2 H|_{t=0}$ is as in Proposition B.9.

Proof. Those equations are a direct consequence of Proposition B.9, Theorem B.11 and Lemma B.12. □

Theorem B.15. *It holds*

$$\frac{\partial}{\partial t}\int_{\mathcal{M}(t)}K\,\mathrm{d}\mathbb{H}\bigg|_{t=0} = \int_{\mathcal{M}_0}\dot\partial K|_{t=0}+uK_0H_0\,\mathrm{d}\mathbb{H}$$

and

$$\frac{\partial^2}{\partial t^2}\int_{\mathcal{M}(t)}K\,\mathrm{d}\mathbb{H}\bigg|_{t=0} = \int_{\mathcal{M}_0}\dot\partial^2 K|_{t=0}+2uH_0\dot\partial K|_{t=0}+K_0\left(\|\nabla_{\mathcal{M}_0}u\|^2+u^2K_0\right)\,\mathrm{d}\mathbb{H}$$

where $\dot\partial K|_{t=0}$ and $\dot\partial^2 K|_{t=0}$ are given in Proposition B.10.

Proof. Those equations are a direct consequence of Proposition B.10, Theorem B.11 and Lemma B.12. □

C Ordinary differential equations: Existence and sensitivity

Theorem C.1 (Picard–Lindelöf)**.** *Suppose $n \in \mathbb{N}$, $y_0 \in \mathbb{R}^n$, $\tau, r \in \mathbb{R}_{>0}$, and define*

$$I = [0, \tau], \qquad \Omega = \{x \in \mathbb{R}^n \mid \|x - y_0\|_\infty \leq r\}.$$

Let $f \in C(I \times \Omega, \mathbb{R}^n)$ be uniformly Lipschitz continuous with respect to $x \in \Omega$ and define $J := [0, \min(\tau, r/\|f\|_{C(I\times\Omega)})]$. Then there exists a unique function $y\colon J \to \mathbb{R}^n$ such that for all $t \in J$

$$y'(t) = f(t, y(t)), \qquad y(0) = y_0.$$

Proof. [Har02, Chapter II, Theorem 1.1] □

Theorem C.2. *Let $k, n \in \mathbb{N}$, $\tau \in \mathbb{R}_{>0}$ and let $A \subseteq \mathbb{R}^k \times \mathbb{R}^n$ be open. Suppose $m \in \mathbb{N}_{\geq 1}$ and $f \in C^m([0,\tau] \times \overline{A}, \mathbb{R}^n)$ such that all partial derivatives of f are bounded. Then there exists a unique function $\eta \in C^m([0,\tau] \times A, \mathbb{R}^n)$ such that for all $t \in [0,\tau]$ and $(x, y_0) \in A$ holds*

$$\frac{\partial}{\partial t}\eta(t, x, y_0) = f\left(t, x, \eta(t, x, y_0)\right), \qquad \eta(0, x, y_0) = y_0.$$

Proof. [Har02, Chapter V, Theorem 4.1] □

Theorem C.3. *Let $\eta(t, x, y_0)$ as in Theorem C.2 and suppose $\eta \in C^1([0,\tau] \times A, \mathbb{R}^n)$. Then $\frac{\partial}{\partial y_0}\eta = \widetilde{\eta}$ where $\widetilde{\eta}$ solves the ODE*

$$\frac{\partial}{\partial t}\widetilde{\eta}(t, x, y_0) = \frac{\partial}{\partial y_0} f\left(t, x, \eta(t, x, y_0)\right) \cdot \widetilde{\eta}(t, x, y_0), \qquad \widetilde{\eta}(0, x, y_0) = \mathrm{id}\,.$$

Proof. [Har02, Chapter V, Theorem 3.1] □

Theorem C.4. *Let $\eta(t, x, y_0)$ as in Theorem C.2 and suppose $\eta \in C^1([0,\tau] \times A, \mathbb{R}^n)$. Then $\frac{\partial}{\partial x}\eta = \widetilde{\eta}$ where $\widetilde{\eta}$ solves the ODE*

$$\begin{aligned}\frac{\partial}{\partial t}\widetilde{\eta}(t, x, y_0) &= \frac{\partial}{\partial y_0} f\left(t, x, \eta(t, x, y_0)\right) \cdot \widetilde{\eta}(t, x, y_0) + \frac{\partial}{\partial x} f\left(t, x, \eta(t, x, y_0)\right) \\ \widetilde{\eta}(0, x, y_0) &= 0.\end{aligned}$$

Proof. [Har02, Chapter V, Theorem 3.1] □

D Schemes related to the fictitious domain stabilized Nitsche method

In order to put the fictitious domain stabilized Nitsche method as presented in Section 3.2.1 in the context of simpler existing schemes, the following paragraphs briefly review a hierarchy of approximation schemes for linear elliptic PDEs that lead to this method.

Standard Galerkin scheme

A standard and arguably the most known approach for the approximation of elliptic PDEs is to construct a conforming discrete subspace $\mathcal{S}' \subseteq U_{\mathrm{ad}}$ (or to approximate such a space sufficiently well in the sense of Strang) and define the discrete approximation function via the discrete variational equation

$$\forall v \in \mathcal{S}' \colon \; a_{\Omega_0}(u_h, v - u_h) = \ell_0(v - u_h),$$

which is equivalent to the solution of a linear equation.

Within this framework, approximation estimates are readily derived, for example by virtue of the Céa-Lemma which implies the asymptotic order of convergence $k - 1$ for $h \to 0$ if $u \in H^{k+1}(\Omega_0)$. Moreover, one also finds that the condition number of the discrete system behaves for fourth order problems like $\mathcal{O}(h^4)$ for $h \to 0$.

The main disadvantage of this scheme in the given setting is that it cannot work with unfitted grids, which is however desirable in order to avoid re-meshing whenever the configuration parameter $\mathbf{p}$ changes in the greater scheme. The reason behind this is the possibility that there could exist elements with vanishing relative measure $|E \cap \Omega_0| \,/\, |E| \to 0$ for $h \to 0$, which in turn would lead to a blow-up in the condition number of the system matrix and would hence make this method numerically infeasible. Moreover, it may be relatively expensive to incorporate the boundary constraints (approximately) into the discrete space $\mathcal{S}'$ if the boundary is not resolved exactly.

Penalty formulation

Penalty formulations avoid the construction of a conforming subspace $\mathcal{S}'$ by incorporating the Dirichlet boundary conditions inexactly and implicitly via a penalization term in the overall energy. The augmented minimization problem reads

$$\min_{v \in \mathcal{S}} \frac{1}{2}\Big(a_{\Omega_0}(v, v) + a_{\mathrm{pen}}(v, v)\Big) - \ell_0(v) - \ell_{\mathrm{pen}}(v)$$

and is equivalent to the variational equation

$$\forall v \in \mathcal{S} \colon \; a_{\Omega_0}(u_h, v) + a_{\mathrm{pen}}(u_h, v) = \ell_0(v) + \ell_{\mathrm{pen}}(v).$$

Within the setting of elliptic PDEs analogous and related penalty approaches and related ones have been studied for example in [BS70], [Bab73], [BE86], [GK17] both for fitted and unfitted grids. The main disadvantage of these methods generally is that they usually lead to sub-optimal rates of convergence due to a violation of the consistency property (3.2.4) or, just like the standard Galerkin scheme, they may suffer from a blow-up in the condition number if there are elements with vanishing relative measure $|E \cap \Omega_0| \,/\, |E| \to 0$ for $h \to 0$.

Nitsche formulation

A remedy for the problem of reduced rates of convergence in penalty formulations was found by Nitsche in [Nit71] for the Laplace problem with Dirichlet boundary conditions on matched grids. In that paper, an analogous term to a_{Nit} in Section 3.2.1 is introduced which preserves the consistency (3.2.4) over the full unconstrained function space, and a discrete minimization problem analogous to

$$\min_{v \in \mathcal{S}_*} \frac{1}{2}\widetilde{a}(v,v) - \widetilde{\ell}(v)$$

with

$$\begin{aligned} \widetilde{a}(v,w) &= a_{\Omega_0}(v,w) + a_{\mathrm{pen}}(v,w) + a_{\mathrm{Nit}}(v,w) + a_{\mathrm{Nit}}(w,v) \\ \widetilde{\ell}(v) &= \ell_0(v) + \ell_{\mathrm{pen}}(v) + a_{\mathrm{Nit}}(v,u) \end{aligned}$$

is considered.

While this modification preserves symmetry in $\widetilde{a}$, it in general breaks the ellipticity because a_{Nit} is indefinite. However, Nitsche shows for fitted grids that the bilinear form $\widetilde{a}$ is positive definite at least on the discrete subspace $\mathcal{S}_*$ if the h-independent penalty coefficient λ is large enough, depending on the constant c_{inv} from the inverse estimates as in Assumption A5. Under these conditions optimal rates of convergence are also proven.

The disadvantage that remains, however, is that this method is not suitable for the application to unfitted grids. This is because the constant c_{inv} in the inverse estimates for a boundary segment $\Gamma \cap E$ and a cut element $E \cap \Omega_0$ is in general unbounded with respect to the grid size parameter h if the cut volume vanishes, i. e. if $|E \cap \Omega_0| \,/\, |E| \to 0$ holds for $h \to 0$. Then also the penalty weighting parameter λ is h-dependent and unbounded, which leads to a deterioration of the condition number of the discrete system and numerical infeasibility of the method.

Stabilized Nitsche formulation

A solution to the above mentioned stability problem for spurious elements is called the stabilized Nitsche formulation, or also the cut finite element method, and is presented in [HH02] and [BH12]. For the Laplace problem and linear finite elements it is shown there

that the analogue of the above stabilized discretization scheme is able to preserve both optimal rates of convergence while also maintaining the original behavior of the condition number, independently of how the mesh $\mathcal{T}$ is located relative to the domain Ω_0. Since then various other works applied this type of discretization to a plethora of different problems and derived corresponding results. An extensive review of those is out of the scope of this thesis and therefore omitted.

The key ingredient in these schemes usually is a stability property in the spirit of Lemma 3.2.1 that allows the relation of L^2-norms of piecewise polynomials on neighboring elements to each other up to terms from a stabilization bilinear form a_{stab}, which is defined on the discrete space $\mathcal{S}$ only. With the help of this it is possible to circumvent the application of inverse inequalities with respect to restricted elements $E \cap \Omega_0$ as in the standard Nitsche method. Instead, the inverse inequalities are applied to the full elements E and, if necessary, related to a fully contained interior element $E' \subseteq \Omega_0$ through said stability property. As a result, the h-independent penalty weight λ may be chosen grid-independently such that the condition number remains stable and optimal convergence rates are achieved.

E A smooth repulsion potential for FCHo2 particle–particle interactions

As mentioned in Section 5.5, a particle interaction potential for FCHo2 F-BAR domains that is purely driven by membrane elasticity is infeasible within a point value constraint model due to the possibility of overlapping. One way of solving this issue is by augmenting said interaction potential by a repulsive direct particle–particle interaction. The idea in this section to introduce such a particle–particle interaction potential based on pairwise atomic forces.

To this end, suppose two FCHo2 particles as well as a particle transformation function $\Psi\colon \mathbb{R}^4 \times \mathbb{R}^3 \to \mathbb{R}^3$ with some particle configuration $\mathbf{p} \in \mathbb{R}^{2\times 4}$ as introduced in Section 5.5. The reference particle has $N := 4478$ reference atoms $\widetilde{a}_i \in \mathbb{R}^3$, $i \in \{1, \dots, N\}$, and correspondingly the transformed atom positions are $\widetilde{a}_i^k := \Psi(\mathbf{p}_k; a_i)$ for $k \in \{1, 2\}$ and $i \in \{1, \dots, N\}$. For reasons of simplicity and in view of the particle model in Section 5.5, the following paragraphs only consider the projections $a_i := (\widetilde{a}_{i,1}, \widetilde{a}_{i,2}) \in \mathbb{R}^2$ and $a_i^k := (\widetilde{a}_{i,1}^k, \widetilde{a}_{i,2}^k) \in \mathbb{R}^2$ of these atoms onto the Euclidean plane. However, generalizations of the subsequent methods to the original three-dimensional space are readily available.

The *Lennard–Jones potential* [Jon24] is used to describe the interaction of two atoms at a given distance $r \in \mathbb{R}_{>0}$. It is defined by

$$E_{\mathrm{LJ}}^0(r) := \frac{A}{r^{12}} - \frac{B}{r^6}$$

where $A, B \in \mathbb{R}_{>0}$ are constants depending on the type of the atoms. Here, the A-term is a heuristic description of the repulsion of electrons at short ranges, and the B-term corresponds to the long-range attractive van der Waals force. As attention is restricted to repulsive interactions only, the B-term is neglected for the remainder of this section by defining $B := 0$.

Based on the Lennard–Jones potential, a repulsive interaction potential between the two given FCHo2 particles is therefore defined by

$$\mathcal{E}_{\mathrm{LJ}} := \sum_{i=1}^{N} \sum_{j=1}^{N} E_{\mathrm{LJ}}^0(\|a_i^1 - a_j^2\|).$$

While being physically justified, this potential is not necessarily suitable for being used as a particle–particle repulsion in conjunction with an elasticity-driven membrane-mediated interaction. The reason behind this is of numerical nature and is explained with the help of Fig. E.1 a) where a level set view of the function

$$E_{\mathrm{LJ}}\colon \mathbb{R}^2 \longrightarrow \mathbb{R}, \qquad x \longmapsto \sum_{i=1}^{n} E_{\mathrm{LJ}}^0\left(\|x - a_i\|\right)$$

is plotted over the domain $[-110, 110] \times [-40, 40]$. There the constant A in E_{LJ}^0 is set to $A := 28 \cdot 3.5^{12}$, which is roughly in the range of physically relevant parameter

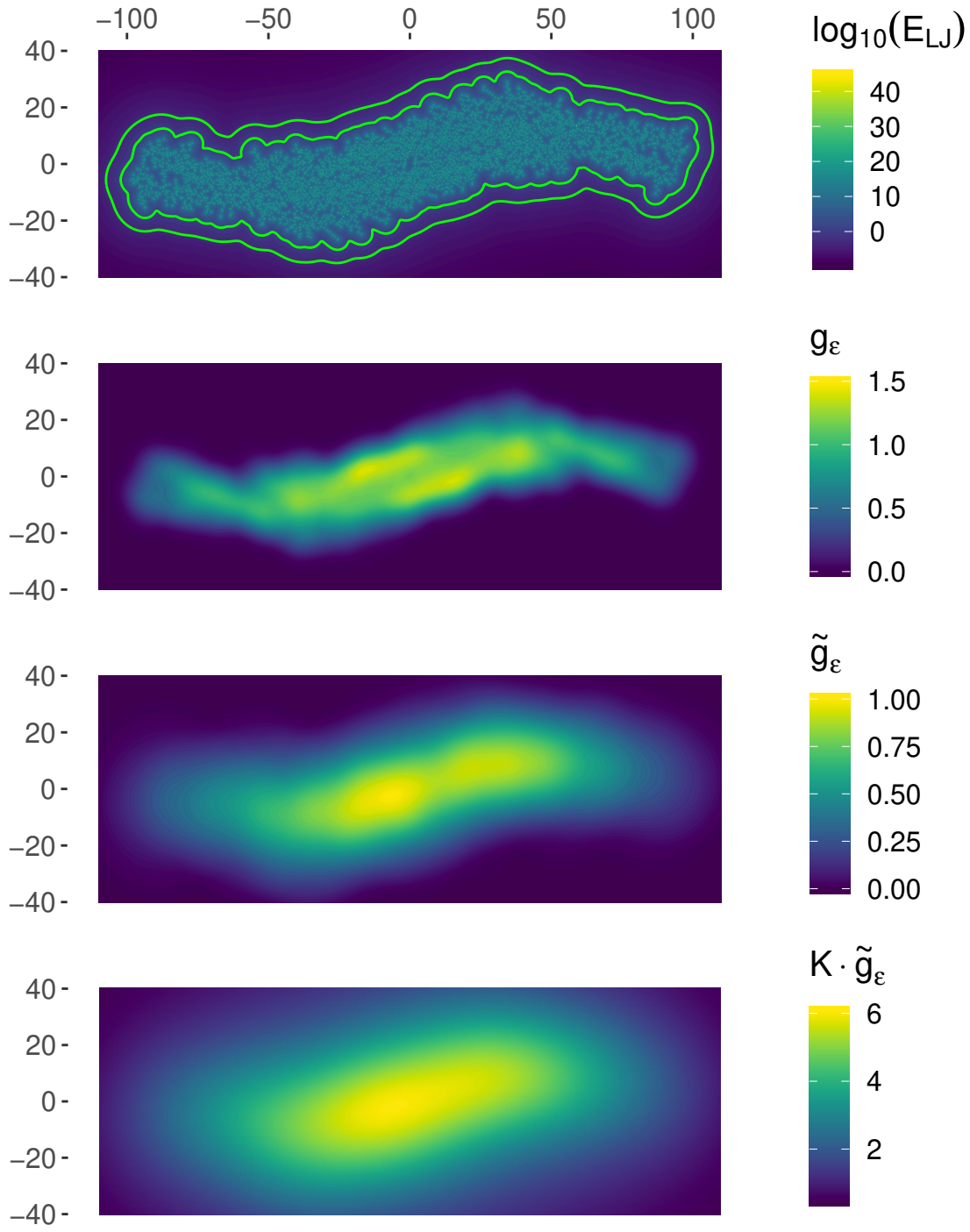

Figure E.1: Level set plots of various functions. From top to bottom: a) Lennard–Jones interaction potential E_{LJ} (logarithmic). b) Dirac approximation g_ε of the delta distributions used in the Lennard–Jones interaction potential. c) Low-dimensional approximation $\widetilde{g}_\varepsilon$ of g_ε. d) Kernel used in the definition of the final particle–particle interaction potential $\mathcal{E}_{\mathrm{particle}}$.

choices but not justified here any further. Within the plot the outer green curve emphasizes the level set $\{E_{\mathrm{LJ}} = 10^{-3}\}$, and the inner green curve corresponds to the level set $\{E_{\mathrm{LJ}} = 10\}$. It is observed that E_{LJ} exhibits a highly singular structure with geometrically complicated level sets. Considering that the elastic membrane interaction energy in practical applications usually spans only one order of magnitude, these vastly different behaviors in contrast imply a considerable challenge for a gradient-type descent methods. This is because once two particles come close enough to each other, the gradient of the Lennard–Jones interaction potential would disproportionately dominate the elastic membrane interaction and either lead to a particle configuration which is largely unrelated to the previous one or restrict the line search to step lengths which would effectively make the gradient method come to a halt. Ideally, an approximation of the Lennard–Jones interaction potential is wanted whose gradients are of lower or comparable order to those originating from the elastic membrane interaction and which nonetheless leads to a similar effective system behavior.

Using a formal calculation and two-dimensional delta-distributions $\delta_i^k := \delta(\cdot - a_i^k)$, the Lennard–Jones interaction may be rewritten as

$$\begin{aligned}
\mathcal{E}_{\mathrm{LJ}} &= \sum_{i=1}^{N}\sum_{j=1}^{N} E_{\mathrm{LJ}}^0(\|a_i^1 - a_j^2\|) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \int_{\mathbb{R}^2\times\mathbb{R}^2} E_{\mathrm{LJ}}^0(\|x-y\|)\, \delta_i^1(x)\, \delta_j^2(y)\, \mathrm{d}(x,y) \\
&= \int_{\mathbb{R}^2\times\mathbb{R}^2} E_{\mathrm{LJ}}^0\left(\|\Psi(\mathbf{p}_1;x) - \Psi(\mathbf{p}_2;y)\|\right)\, g(x)\, g(y)\, \mathrm{d}(x,y)
\end{aligned}$$

where $g := \sum_{i=1}^{N} \delta_i$ with $\delta_i := \delta(\cdot - a_i)$. Note how this step implicitly uses the fact that the chosen transformation function Ψ is orthonormal and compatible with the projection of the atoms onto the Euclidean plane. In more general cases the integration would necessarily need to be carried out over the full space $\mathbb{R}^3 \times \mathbb{R}^3$.

Define furthermore a C^1-smooth two-dimensional Dirac sequence by $\phi_\varepsilon(x) := \frac{1}{\varepsilon^2}\phi(\|x\|/\varepsilon)$ for $\varepsilon \in \mathbb{R}_{>0}$ where

$$\phi(r) := \begin{cases} \frac{3}{\pi}(r^2-1)^2 & \text{for } r \le 1, \\ 0 & \text{for } r > 1. \end{cases}$$

The sum of delta distributions g is then further approximated by

$$g_\varepsilon(x) := \sum_{i=1}^{N} \phi_\varepsilon(x - a_i).$$

A level set view plot of g_ε for $\varepsilon = 10$ is shown in Fig. E.1 b). There it becomes apparent that also this approximation is not numerically feasible yet due to the presence of still relatively large fluctuations in g_ε and because of the large number N of summands,

which would imply a considerable computational expense when carrying out a numerical quadrature scheme using g_ε in the integral above. Therefore, this function is further approximated by the sparse representation

$$\widetilde{g}_\varepsilon(x) := \frac{N}{|I|} \sum_{i \in I} \phi_\varepsilon(x - a_i)$$

where $I \subseteq \{1, \dots, N\}$. In Fig. E.1 c) the level set view of $\widetilde{g}_\varepsilon$ resulting from $\varepsilon := 30$ and

$$I := \{1 + 100k \mid k \in \mathbb{N},\, 1 + 100k \leq N\}$$

is depicted and shows a significantly smoother structure. Of course, at this step also any other low-dimensional approximation of g_ε might be considered as well and the given choice is merely motivated by its simplicity.

Finally, the relatively singular behavior of E^0_{LJ} is softened by transitioning to the kernel $K(r) := (100 + r^2)^{-1}$ instead. Altogether, given $n \in \mathbb{N}$ particles and a particle configuration $\mathbf{p} \in \mathbb{R}^{n \times 4}$, a numerically feasible repulsive particle–particle interaction potential is defined by

$$\mathcal{E}_{\mathrm{particle}}(\mathbf{p}) := \frac{1}{2} \sum_{i,j=1}^{n} \int_{\mathbb{R}^2 \times \mathbb{R}^2} K(\left\|\Psi(\mathbf{p}_i; x) - \Psi(\mathbf{p}_j; y)\right\|)\, \widetilde{g}_\varepsilon(x)\, \widetilde{g}_\varepsilon(y)\, \mathrm{d}(x, y).$$

This is the interaction chosen for the application example in Section 5.5. It is emphasized that the integration in $\mathcal{E}_{\mathrm{particle}}$ may be carried out approximatively using standard quadrature schemes as of the local support of $\widetilde{g}_\varepsilon$.

The choice of K is motivated both by the desire to have an upper bound on the gradient in the interaction potential and to admit as coarse a quadrature scheme as reasonably possible. In Fig. E.1 d) a plot of the function $x \mapsto \int_{\mathbb{R}^2} K(\|x - y\|)\widetilde{g}_\varepsilon(y)\, \mathrm{d}x$ is shown. It illustrates how the the integrand in $\mathcal{E}_{\mathrm{particle}}$ is smooth enough to avoid any disproportional domination of the elastic membrane interaction which could lead to numerical instabilities while still being able to penalize particle overlapping sufficiently strongly and decaying reasonably fast to zero in a neighborhood of the particle. This makes the repulsive interaction above not only a numerically feasible but also a potentially physically plausible choice.

Bibliography

[Ada75] R. Adams. *Sobolev Spaces*. Pure and applied mathematics. Academic Press, 1975.

[AS07] E. Atilgan and S. Sun. Shape transitions in lipid membranes and protein mediated vesicle fusion and fission. *Journal of Chemical Physics*, 126(9), 2007.

[AV93] K. Atkinson and E. Venturino. Numerical Evaluation of Line Integrals. *SIAM Journal on Numerical Analysis*, 30(3):882–888, 1993.

[Bab73] I. Babuška. The finite element method with penalty. *Mathematics of computation*, 27(122):221–228, 1973.

[BBD+08a] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. Part II: implementation and tests in DUNE. *Computing*, 82(2):121–138, 2008.

[BBD+08b] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. Part I: abstract framework. *Computing*, 82(2):103–119, 2008.

[BBD+16] M. Blatt, A. Burchardt, A. Dedner, C. Engwer, J. Fahlke, B. Flemisch, C. Gersbacher, C. Gräser, F. Gruber, C. Grüninger, D. Kempf, R. Klöfkorn, T. Malkmus, S. Müthing, M. Nolte, M. Piatkowski, and O. Sander. The Distributed and Unified Numerics Environment, Version 2.4. *Archive of Numerical Software*, 4(100):13–29, 2016.

[BBJ+97] P. Bastian, K. Birken, K. Johannsen, S. Lang, N. Neuß, H. Rentz-Reichert, and C. Wieners. UG–a flexible software toolbox for solving partial differential equations. *Computing and Visualization in Science*, 1(1):27–40, 1997.

[BE86] J. W. Barrett and C. M. Elliott. Finite element approximation of the Dirichlet problem using the boundary penalty method. *Numerische Mathematikm*, 49(4):343–366, 1986.

[BF03] D. Bartolo and J.-B. Fournier. Elastic interaction between "hard" or "soft" pointwise inclusions on biological membranes. *European Physical Journal E*, 11(2):141–146, 2003.

[BFS65] F. K. Bogner, L. Fox, and L. A. Schmit. The Generation of Interelement Compatible Stiffness and Mass Matrices by the Use of Interpolation Formulas. *Proceedings of the Conference on Matrix Methods in Structural Mechanics*, pages 397–444, 1965.

[BGN08] J. W. Barrett, H. Garcke, and R. Nürnberg. Parametric approximation of Willmore flow and related geometric evolution equations. *SIAM Journal on Scientific Computing*, 31(1):225–253, 2008.

[BH12] E. Burman and P. Hansbo. Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method. *Applied Numerical Mathematics*, 62(4):328–341, 2012.

[BHD04] H. Bermúdez, D. Hammer, and D. Discher. Effect of bilayer thickness on membrane bending rigidity. *Langmuir*, 20(3):540–543, 2004.

[BLW12] A. H. Bahrami, R. Lipowsky, and T. R. Weikl. Tubulation and aggregation of spherical nanoparticles adsorbed on vesicles. *Physical review letters*, 109(18):188102, 2012.

[BR80] H. Blum and R. Rannacher. *On the Boundary Value Problem of the Biharmonic Operator on Domains with Angular Corners*. Preprint: Sonderforschungsbereich Approximation und Mathematische Optimierung in einer Anwendungsbezogenen Mathematik. SFB 72, 1980.

[Bro08] F. L. Brown. Elastic modeling of biomembranes and lipid bilayers. *Annu. Rev. Phys. Chem.*, 59:685–712, 2008.

[Bro11] F. Brown. Continuum simulations of biomembrane dynamics and the importance of hydrodynamic effects. *Quarterly Reviews of Biophysics*, 44(4):391–432, 2011.

[BS70] J. H. Bramble and A. H. Schatz. Rayleigh-Ritz-Galerkin methods for dirichlet's problem using subspaces without boundary conditions. *Communications on Pure and Applied Mathematics*, 23(4):653–675, 1970.

[Can70] P. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *Journal of Theoretical Biology*, 26(1):61–81, 1970.

[Cia78] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.

[DD06] K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces and free boundaries*, 8(1):21–46, 2006.

[DE13] G. Dziuk and C. M. Elliott. Finite element methods for surface PDEs. *Acta Numerica*, 22:289–396, 2013.

[Des15] M. Deserno. Fluid lipid membranes: From differential geometry to curvature stresses. *Chemistry and Physics of Lipids*, 185:11–45, 2015.

[DF99] P. Dommersnes and J.-B. Fournier. N-body study of anisotropic membrane inclusions: Membrane mediated interactions and ordered aggregation. *European Physical Journal B*, 12(1):9–12, 1999.

[DF02] P. Dommersnes and J.-B. Fournier. The many-body problem for anisotropic membrane inclusions and the self-assembly of "saddle" defects into an "egg carton". *Biophysical Journal*, 83(6):2898–2905, 2002.

[DGR17] K. Deckelnick, H.-C. Grunau, and M. Röger. Minimising a relaxed Willmore functional for graphs subject to boundary conditions. *Interfaces and Free Boundaries*, 19(1):109–140, 2017.

[DKS15] K. Deckelnick, J. Katz, and F. Schieweck. A C^1-finite element method for the Willmore flow of two-dimensional graphs. *Mathematics of Computation*, 84(296):2617–2643, 2015.

[DN12] G. Doan and R. H. Nochetto. First variation of the general curvature-dependent surface energy. *ESAIM: Mathematical Modelling and Numerical Analysis - Modlisation Mathmatique et Analyse Numrique*, 46(1):59–79, 2012.

[DZ91] M. C. Delfour and J.-P. Zolésio. Velocity method and Lagrangian formulation for the computation of the shape Hessian. *SIAM journal on control and optimization*, 29(6):1414–1442, 1991.

[DZ11] M. Delfour and J. Zolésio. *Shapes and Geometries*. Society for Industrial and Applied Mathematics, second edition, 2011.

[Dzi08] G. Dziuk. Computational parametric Willmore flow. *Numerische Mathematik*, 111(1):55, 2008.

[EGH+16] C. M. Elliott, C. Gräser, G. Hobbs, R. Kornhuber, and M.-W. Wolf. A Variational Approach to Particles in Lipid Membranes. *Archive for Rational Mechanics and Analysis*, 222(2):1011–1075, 2016.

[FG15] J.-B. Fournier and P. Galatola. High-order power series expansion of the elastic interaction between conical membrane inclusions. *The European Physical Journal E*, 38(8):86, 2015.

[GBP93] M. Goulian, R. Bruinsma, and P. Pincus. Long-Range Forces in Heterogeneous Fluid Membranes. *EPL (Europhysics Letters)*, 22(2):145, 1993.

[GGS10] F. Gazzola, H.-C. Grunau, and G. Sweers. *Polyharmonic boundary value problems: positivity preserving and nonlinear higher order elliptic equations in bounded domains*. Springer Science & Business Media, 2010.

[GK17] C. Gräser and T. Kies. Discretization error estimates for penalty formulations of a linearized Canham–Helfrich-type energy. *IMA Journal of Numerical Analysis*, 2017.

[Grä15] C. Gräser. A note on Poincaré- and Friedrichs-type inequalities. *ArXiv e-prints*, 2015.

[Gri85] P. Grisvard. *Elliptic problems in nonsmooth domains*. Monographs and studies in mathematics. Pitman Advanced Pub. Program, Boston, London, Melbourne, 1985.

[Har02] P. Hartman. *Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, second edition, 2002.

[HBM+10] W. M. Henne, E. Boucrot, M. Meinecke, E. Evergren, Y. Vallis, R. Mittal, and H. T. McMahon. FCHo proteins are nucleators of clathrin-mediated endocytosis. *Science*, 328(5983):1281–1284, 2010.

[Hel73] W. Helfrich. Elastic properties of lipid bilayers: theory and possible experiments. *Zeitschrift für Naturforschung C*, 28(11-12):693–703, 1973.

[HH02] A. Hansbo and P. Hansbo. An unfitted finite element method, based on Nitsches method, for elliptic interface problems. *Computer methods in applied mechanics and engineering*, 191(47-48):5537–5552, 2002.

[HKF+07] W. M. Henne, H. M. Kent, M. G. Ford, B. G. Hegde, O. Daumke, P. J. G. Butler, R. Mittal, R. Langen, P. R. Evans, and H. T. McMahon. Structure and analysis of FCHo2 F-BAR domain: a dimerizing and membrane recruitment module that effects membrane curvature. *Structure*, 15(7):839–852, 2007.

[Hob16] G. Hobbs. *Particles and biomembranes: a variational PDE approach*. Ph.D. thesis, University of Warwick, 2016.

[HP95] R. Horst and P. Pardalos. *Handbook of Global Optimization*. Nonconvex Optimization and Its Applications. Springer US, 1995.

[HPR02] R. Horst, P. Pardalos, and H. Romeijn. *Handbook of Global Optimization*. Number Bd. 2 in Handbook of Global Optimization. Springer US, 2002.

[ISSS10] T. Idema, S. Semrau, C. Storm, and T. Schmidt. Membrane mediated sorting. *Physical Review Letters*, 104(19), 2010.

[Jon24] J. E. Jones. On the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 106(738):463–477, 1924.

[KKKH16a] O. Kahraman, P. Koch, W. Klug, and C. Haselwandter. Architecture and function of mechanosensitive membrane protein lattices. *Scientific Reports*, 6, 2016.

[KKKH16b] O. Kahraman, P. Koch, W. Klug, and C. Haselwandter. Bilayer-thickness-mediated interactions between integral membrane proteins. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 93(4), 2016.

[KNO98] S. Kim, K. J. Neu, and G. Oster. Curvature-Mediated Interactions Between Membrane Proteins. *Biophysical Journal*, 75(5):2274–2291, 1998.

[KNO00] K. Kim, J. Neu, and G. Oster. Effect of protein shape on multibody interactions between membrane inclusions. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 61(4 B):4281–4285, 2000.

[KSTT02] R. Kupferman, A. Stuart, J. Terry, and P. Tupper. Long-term behaviour of large mechanical systems with random initial data. *Stochastics and Dynamics*, 2(04):533–562, 2002.

[LGR18] T. Ludescher, S. Gross, and A. Reusken. A Multigrid Method for Unfitted Finite Element Discretizations of Elliptic Interface Problems. *ArXiv e-prints*, 2018.

[Li05] X. S. Li. An overview of SuperLU: Algorithms, implementation, and user interface. *ACM Transactions on Mathematical Software (TOMS)*, 31(3):302–325, 2005.

[LSR10] T. Lelièvre, G. Stoltz, and M. Rousset. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, 2010.

[MD10] M. M. Müller and M. Deserno. Cell model approach to membrane mediated protein interactions. *Progress of Theoretical Physics Supplement*, 184:351–363, 2010.

[MDG05a] M. Müller, M. Deserno, and J. Guven. Geometry of surface-mediated interactions. *Europhysics Letters*, 69(3):482–488, 2005.

[MDG05b] M. Müller, M. Deserno, and J. Guven. Interface-mediated interactions between particles: a geometrical approach. *Physical Review E*, 72(6):061407, 2005.

[MG05] H. T. McMahon and J. L. Gallop. Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature*, 438(7068):590, 2005.

[MLLR14] A. Massing, M. G. Larson, A. Logg, and M. E. Rognes. A stabilized Nitsche fictitious domain method for the Stokes problem. *Journal of Scientific Computing*, 61(3):604–628, 2014.

[MN88] J. R. Magnus and H. Neudecker. Matrix differential calculus with applications in statistics and econometrics. *Wiley series in probability and mathematical statistics*, 1988.

[Nit71] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 36(1):9–15, 1971.

[OS16] M. A. Olshanskii and D. Safin. Numerical integration over implicitly defined domains for higher order unfitted finite element methods. *Lobachevskii Journal of Mathematics*, 37(5):582–596, 2016.

[RD11] B. Reynwar and M. Deserno. Membrane-mediated interactions between circular particles in the strongly curved regime. *Soft Matter*, 7(18):8567–8575, 2011.

[Rey03] O. Reynolds. *Scientific papers on mechanical and physical subjects*. Cambridge University Press, 1903.

[Set99] J. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 1999.

[SK15] Y. Schweitzer and M. Kozlov. Membrane-Mediated Interaction between Strongly Anisotropic Protein Scaffolds. *PLoS Computational Biology*, 11(2), 2015.

[Sol97] M. Solodov. Convergence analysis of perturbed feasible descent methods. *Journal of Optimization Theory and Applications*, 93(2):337–353, 1997.

[SS97] M. Solodov and B. Svaiter. Descent methods with linesearch in the presence of perturbations. *Journal of Computational and Applied Mathematics*, 80(2):265–275, 1997.

[SZ92] J. Sokolowski and J.-P. Zolesio. *Introduction to shape optimization*. Springer, 1992.

[Wei02] T. Weikl. Dynamic phase separation of fluid membranes with rigid inclusions. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 66(6):061915/1–061915/6, 2002.

[Whi34] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Am. Math. Soc.*, 36(1):63–89, 1934.

[Wil13] T. Willmore. *An Introduction to Differential Geometry.* Dover Books on Mathematics. Dover Publications, 2013.

[WKH98] T. R. Weikl, M. M. Kozlov, and W. Helfrich. Interaction of conical membrane inclusions: Effect of lateral tension. *Phys. Rev. E*, 57:6988–6995, 1998.

[Zei85] E. Zeidler. *Nonlinear Functional Analysis and Its Applications: Fixed point theorems.* Nonlinear Functional Analysis and Its Applications. Springer-Verlag, 1985.

[Zwa54] R. W. Zwanzig. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *The Journal of Chemical Physics*, 22(8):1420–1426, 1954.

Zusammenfassung

Diskret-kontinuierliche hybride Modelle sind ein populäres Mittel zur Beschreibung elastisch Membran-mediierter Interaktionen von Partikeln in und auf Doppellipidschichten. Der kontinuierliche Anteil ist hierbei üblicherweise durch eine Näherung der Lipidmembran als unendlich dünne und hinreichend glatte Hyperfläche gegeben, deren elastische Energie durch ein Canham–Helfrich-artiges Funktional bestimmt ist. Die diskrete Komponente ergibt sich aus der Modellierung von Nicht-Membran-Partikeln als diskrete und starre Objekte, die in Abhängigkeit ihrer Konfiguration lokale Restriktionen der Membran entlang der Membran-Partikel-Schnittstellen induzieren.

In diesem Zusammenhang beschreibt das Interaktionspotential die optimale elastische Energie eines solchen hybriden Systems in Abhängigkeit einer festen Partikelkonfiguration. Dementsprechend folgt aus dem Energieminimierungsprinzip, dass stationäre Partikelkonfigurationen durch die lokalen Minima des Interaktionspotentials gegeben sind.

Das Hauptziel dieser Arbeit ist der Beweis der Differenzierbarkeit des Interaktionspotentials für eine ausgewählte Klasse von Modellen im Sinne des in [EGH$^+$16] eingeführten variationellen Ansatzes zum Zwecke der Entwicklung und Anwendung robuster numerischer Optimierungsverfahren zur Berechnung stationärer Partikelkonfigurationen. Entsprechend steht darüber hinaus ebenso die Herleitung einer numerisch zugänglichen Darstellung des Gradienten im Mittelpunkt, inklusive deren Diskretisierung und der relevanten numerischen Analysis.

Der Differenzierbarkeitsbeweis wird erbracht durch eine Anwendung des Satzes über implizite Funktionen. Die Grundlage hierfür sind sogenannte randwerterhaltende Gebietstransformationen, die von geeigneten Familien von Vektorfeldern induziert werden und die lokal die Umformulierung der im Interaktionspotential implizit definierten Minimierungsprobleme bezüglich einer festen Referenzkonfiguration erlauben. Dies ermöglicht im Anschluss die Darstellung des Gradienten als Volumenintegral mittels Methoden der Matrix-Analysis.

Die Diskretisierung der Differentialgleichungen zur Beschreibung optimaler Membranformen erfolgt über Finite-Elemente-Methoden. Für Partikelmodelle mit sogenannten Kurvenrestriktionen wird ein stabilisiertes Nitscheverfahren mit fiktiven Gebietsanteilen entwickelt, und für Modelle mit Punktwertrestriktionen wird durch lokale QR-Transformationen der nodalen Finite-Element-Basis eine konforme Galerkin-Diskretisierung ermöglicht. In beiden Fällen werden geeignete a priori Fehlerabschätzungen bewiesen, und ebenso werden in diesem Rahmen auch Fehlerabschätzungen für die Volumendarstellung des Gradienten gezeigt.

Durch die entwickelten Verfahren wird die effiziente Simulation von Makrostrukturen durch isotrope und anisotrope Partikel erschlossen, was anhand verschiedener Beispielanwendungen und mit Hilfe gestörter Gradientenverfahren illustriert wird.