Variational Discretization of Higher Order Geometric Gradient Flows Based on Phase Field Models

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von Martina Franken
aus Ratingen

Bonn, August 2012

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn
am Institut für Numerische Simulation

Erscheinungsjahr: 2013

1. Gutachter: Prof. Dr. Martin Rumpf
2. Gutachter: Prof. Dr. Holger Rauhut

Tag der Promotion: 04.02.2013

Zusammenfassung

In dieser Arbeit präsentieren wir eine Phasenfeld-basierte, geschachtelte Zeitdiskretisierung für den Willmore-Fluss. Dies ist der L^2-Gradientenfluss, der die Willmore-Energie

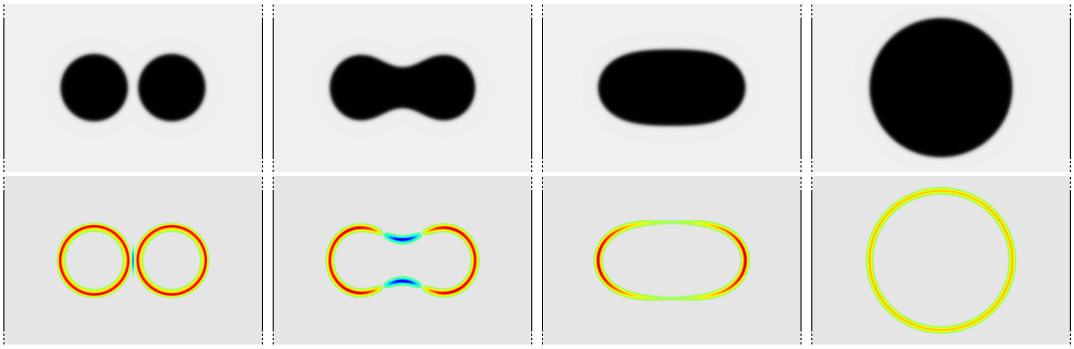
$$w[x] = \frac{1}{2} \int_\Gamma \mathbf{h}^2(x)\,\mathrm{d}\mathcal{H}^{d-1}$$

minimiert. Hierbei bezeichnen Γ eine durch die Identität x über sich selbst parametrisierte Hyperfläche, $\mathbf{h}$ die mittlere Krümmung auf dieser Hyperfläche und $\mathcal{H}^{d-1}$ das $(d-1)$-dimensionale Hausdorff-Maß.

Die grundlegende Idee unseres Ansatzes besteht darin, dass wir die mittlere Krümmung $\mathbf{h}$ durch die zeitdiskrete, approximierte Geschwindigkeit des mittleren Krümmungsflusses approximieren. Diese Geschwindigkeit wird durch das Lösen eines inneren Minimierungsproblems berechnet, das auf dem Konzept der natürlichen Zeitdiskretisierung beruht. Das äußere variationelle Problem balanciert den L^2-Abstand der Fläche zu zwei aufeinander folgenden Zeitschritten und die Abnahme der Willmore-Energie. Dies führt zu einem Optimierungsproblem mit PDG-Nebenbedingung, das uns Zeitschrittweiten bis zur Gößenordnung h erlaubt, wobei h die Gitterweite ist.

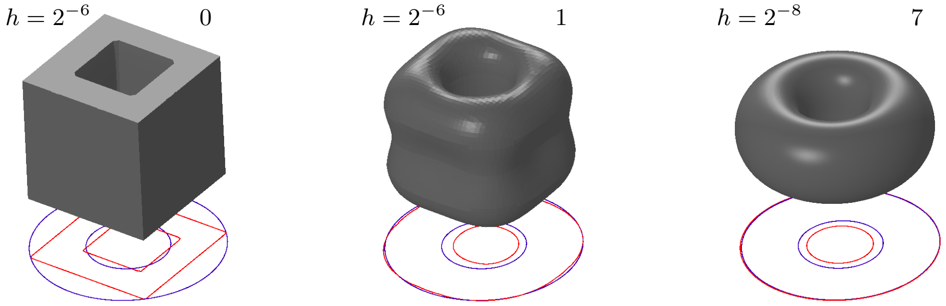
In dieser Arbeit leiten wir das zugehörige Phasenfeld-Modell her, indem wir mit der parametrischen Version starten und sie schrittweise in den Phasenfeld-Kontext übertragen. Anschließend beweisen wir die Existenz einer Lösung.

Ein lineares Modellproblem ist gegeben durch den biharmonischen Wärmefluss. Für diesen Fall präsentieren wir bewiesene Fehlerschranken und verifizieren sie numerisch. Anschließend testen wir unser Modell und vergleichen es mit einem semi-impliziten Modell von Du et al. [40, 42], wobei wir feststellen, dass unser Modell stabiler ist.



Evolution von zwei Kreisen unter Willmore-Fluss (obere Reihe), sowie die zugehörige diskrete Krümmung, farbkodiert auf dem Intervall $[-10, 10]$ *durch* .

Als Anwendung unseres geschachtelten, zeitdisketen Willmore-Fluss-Modells beschäftigen wir uns mit der Rekonstruktion von Hyperflächen anhand von gegebenen niederdimensionalen Huffman-Labelings. Dabei zählt das Huffman-Labeling die Anzahl der Schnittpunkte zwischen der Hyperfläche und dem Projektionsstrahl. Zur Rekonstruktion der Fläche minimieren wie eine Regularisierungsenergie bestehend aus der skalierten Flächen-, sowie der skalierten Willmore-Energie unter der Nebenbedingung, dass das Huffman-Labeling der minimierenden Fläche fast überall gleich dem gegebenen Huffman-Labeling ist. Wir lösen dieses Problem im Phasenfeld-Kontext durch einen alternierenden Algorithmus, der abwechselnd die Regularisierungsenergie und eine Mismatch-Energie minimiert und übertragen unser Problem von einem groben Startgitter auf immer feinere Gitter.



Rekonstruktion eines Torus anhand eines gegebenen Huffman-Labelings, dessen 1*-Niveaulinie in blau dargestellt ist. Die rote Linie zeigt die* 1*-Niveaulinie des Huffman-Labelings des* 3*-dimensionalen Objektes. Die aktuelle Gitterweite sowie der aktuelle Schritt auf dem jeweiligen Gitter werden oberhalb des Objektes angezeigt.*

In weiten Teilen der Arbeit lösen wir unser geschachteltes Minimierungsproblem zur Berechnung des Willmore-Flusses, indem wir die zugehörige Lagrange-Gleichung aufstellen und das daraus resultierende Sattelpunktproblem lösen. Alternativ beschäftigen wir uns im letzten Teil dieser Arbeit damit, wie wir sowohl das lineare Modellproblem, als auch unser geschachteltes, zeitdiskretes Problem mit Hilfe eines Augmented-Lagrange-Verfahrens lösen können.

Acknowledgements

First I would like to thank my advisor Martin Rumpf for the possibility of being part of his group for several years and for his support and guidance during this time.

Furthermore I am indepted to Benedikt Wirth for many fruitful discussions and successful cooperation on joint projects within this thesis.

In particular I want to thank Martin Rumpf and Benedikt Wirth for investigating on error estimates for the fully discrete biharmonic heat flow, presented in Chapter 4.

Moreover I am grateful to my colleagues at the Institute for Numerical Simulation, at the University of Bonn, for a pleasant working atmosphere over several years and for numerous interesting and helpful discussions.

Finally my thank goes to my family and my husband for their support over the years.

Martina Franken

Contents

1 Introduction **11**
1.1 Willmore flow in literature 13
1.2 Phase field based Willmore flow and mean curvature motion in literature 14

2 Foundations **17**
2.1 Phase Field approximations 17
2.2 Finite Element discretization 23
2.3 Survey of solving constrained minimization problems 25

3 Nested time discretization of Willmore flow **29**
3.1 Derivation of the model 29
  3.1.1 Natural time discretization of mean curvature motion 29
  3.1.2 Natural time discretization of Willmore flow 37
  3.1.3 The constrained optimization perspective 41
3.2 A fully practical numerical scheme 43
  3.2.1 Finite Element discretization in space 43
  3.2.2 Newton method for the constrained optimization problem . . . 45

4 The biharmonic heat equation - a linear model problem **49**
4.1 A nested scheme for the biharmonic heat equation 49
4.2 Numerical validation 53
  4.2.1 Implementation 53
  4.2.2 Validation 54

5 Numerical results and applications for Willmore flow **57**
5.1 Simulation based on the nested time discretization 57
5.2 Comparison with a semi-implicit scheme 63
  5.2.1 A semi-implicit phase field scheme for Willmore flow 64
  5.2.2 Newton method for the semi-implicit scheme 65
  5.2.3 Comparison 65
5.3 Application to an image restoration 66

6 Shape from apparent contour **71**
6.1 The model problem 71
6.2 Huffman labeling in literature 73
6.3 The variational phase field model 76
  6.3.1 Finite Element discretization 78

6.4 Numerical results 79
  6.4.1 Two-dimensional simulations 79
  6.4.2 Three-dimensional simulations 83
6.5 Further discussion 88

7 Augmented Lagrange method for computing Willmore and biharmonic heat flow 91
7.1 Augmented Lagrange method for the biharmonic heat equation 91
  7.1.1 Derivation of the model 92
  7.1.2 A fully practical numerical scheme 94
  7.1.3 Numerical results 95
7.2 Augmented Lagrange method for our Willmore flow model 98
  7.2.1 Derivation of the model 98
  7.2.2 A fully practical numerical scheme 100
  7.2.3 Numerical results 103
7.3 Future work 110

Bibliography 111

List of Figures 119

List of Algorithms 121

Nomenclature 123

Index 129

Introduction

In this work a phase field based nested variational time discretization for Willmore flow is presented. Willmore flow is the L^2-gradient flow minimizing the Willmore energy

$$w[x] = \frac{1}{2} \int_\Gamma \mathbf{h}^2(x) \, \mathrm{d}\mathcal{H}^{d-1} \tag{1.1}$$

on a hypersurface Γ parameterized over itself by the identity mapping x. Here $\mathbf{h}$ is the mean curvature of the hypersurface Γ and $\mathcal{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure. Willmore flow is also known as elastic flow and from a physical point of view Willmore energy approximates the stored energy in a thin elastic shell. Willmore flow and Willmore energy can be used for modeling edge sets in imaging [77, 75, 102, 22] or to applications in surface modeling [97, 13, 12, 85, 101]. In Section 1.1 we give a short survey on publications dealing with Willmore flow in general, while Section 1.2 presents some work focusing on phase field based Willmore and mean curvature flow.

The basic idea of our model is to approximate the mean curvature $\mathbf{h}$ by a time-discrete, approximate speed of the mean curvature motion. This speed is computed by a fully implicit time step of mean curvature motion, which forms the inner problem of our model. It is set up as a minimization problem taking into account the concept of natural time discretization. Therefore it is closely related to the variational time discretization of mean curvature motion as it is presented in the paper of Almgren, Taylor and Wang [1] and Luckhaus and Sturzenhecker [70]. Since Chambolle developed a corresponding level set algorithm for this type of time stepping [20] the inner variational problem is also closely related to this work.
The outer problem is a variational problem balancing between the L^2-distance of the surface at two consecutive time steps and the decay of the Willmore energy. This is a typical ansatz in case of natural time discretization as it is used in the inner problem. Within the Willmore energy the mean curvature is approximated as mentioned above. Consequently our model is a nested variational and leads to a PDE constraint optimization problem to compute a single time step.
A corresponding parametric version of this model based on finite elements on a triangulation of the evolving geometry was investigated by Olischläger and Rumpf in [81, 82]. Picking up the intention of these two works we aim for a stable time discretization allowing large time steps.

To set up the nested variational time discretization of Willmore flow based on phase field approximation we introduce some basic knowledge on phase field and finite element discretizations in Chapter 2. With respect to the two presented methods for solving the PDE constraint optimization problem we give a short survey of solving constraint minimization problems in the last Section of Chapter 2.

In Chapter 3 the derivation of the model starts with the natural time discretization of mean curvature motion. Based on this in Subsection 3.1.2 we continue with the natural time discretization of Willmore flow, the central model of this work. In this Subsection we also prove the existence of a time-discrete phase field solution of our model. To set up a fully practical numerical scheme consisting of finite element discretization in space as well as a Newton method for solving the constrained optimization problem (cf. Section 3.2) we first deal with the constrained optimization perspective of this model.

Since biharmonic heat flow is a linear model problem for our nested time discretization of Willmore flow we transfer our model to the linear case in Chapter 4. Moreover we present a result proven in [52] concerning error estimates for the fully discrete biharmonic heat flow and validate it numerically. Based on the analogy of the linear and the nonlinear model we expect an analog error estimate for our nested time discretized Willmore model.

This expectation will be supported by the results of a numerical test of our Willmore flow model presented in Chapter 5. This chapter focuses on numerical results and applications of the nested time discretized Willmore model and starts with the computation of the experimental error of our Willmore flow scheme for radially symmetric Willmore flow. The error order correlates with the one stated by Theorem 4.4 in the linear case. In addition we compare this result with the corresponding result computed with the semi-implicit phase field scheme for Willmore flow introduced by Du et al. [40, 42]. Their scheme is presented in Section 5.2 and the comparison based on our experiments reveals that the nested variational method presented in Chapter 3 is significantly more robust. Finally the last section of this chapter deals with an exemplary application of our model, which is an elastic model for edge restoration in images by Nitzberg et al. [77].

After concentrating on the nested time discretized Willmore model including numerical tests and a simple application in the previous chapters we focus on a further, more complex application in Chapter 6. The problem we are dealing with consists in reconstructing a hypersurface $\Gamma \subset \mathbb{R}^d$ corresponding to a given $(d-1)$-dimensional apparent contour or Huffman labeling. In Section 6.1 we derive the model problem and present a definition for the apparent contour and the Huffman labeling as it is used in this work. After this we give a short survey on Huffman labeling in literature. We start with the labeling Huffman introduced 1971 [58], continue with the labeling Williams worked with in [98] until we come to the Huffman labeling used by Bellettini et al.[9, 10]. Our work concerning reconstructing shapes from a given apparent contour is strongly related to the two works by Bellettini et al.. This strong relatedness is mirrored by the fact that apart from one property their kind of Huffman labeling is the same as our version. At the same time our version is completely different from the original labeling presented by Huffman in [58]. After this survey on Huffman labeling in literature we set up the phase field version of the apparent contour inverse problem and present some two and three-dimensional results. Although these results are convincing we have to discuss some shortcomings of this model in the last section of Chapter 6.1. This discussion motivates and leads over to the last chapter of this work.

In Chapter 7 we present an alternative method for solving the nested variational time discretized Willmore flow model introduced in Chapter 3. This alternative method consists in an Augmented Lagrange method which we introduce for the linear model problem first until we deal with the Augmented Lagrange method for our Willmore flow model. The main advantage of the Augmented Lagrange method in comparison to the Lagrange method used in Chapter 3 is that we have to solve a minimization problem in each iteration of the Augmented Lagrange method instead of finding a saddle point as in the Lagrange method used before. In the two-dimensional case this leads to convincing results, whereas the three-dimensional tests reveal some future work on more efficient solvers within the Trust-Region method.

1.1 Willmore flow in literature

Willmore flow as well as Willmore surfaces, which are defined as minimizer of the Willmore energy [99] are topics frequently treated in literature. In 1996 Polden [66, 84] analytically treated the Willmore flow of curves and surfaces. For sufficiently smooth initial surfaces Simonett presented a proof of the existence of a unique and locally smooth solution of Willmore flow [92]. In case of initial surfaces close to a sphere he proved exponential convergence to a sphere. In the same and the following year Kuwert and Schätzle [63, 64, 65] published a paper dealing with long time existence and regularity of solutions. Recently, Rivière [87] extended results of Kuwert and Schätzle [65] for co-dimension 1 to arbitrary co-dimension.

Also in the area of numerical treatment of Willmore flow several papers have been published. In [43] Dzuik, Kuwert and Schätzle focused on a theoretical and numerical treatment of Willmore flow of curves. In comparison to this Rusu [89] dealt with the numerical approximation of parametric Willmore flow of surfaces and presented a mixed scheme based on the surface parametrization x and the mean curvature vector $\mathbf{h}n$ (with n being the surface normal) as independent variables. An application of this ansatz in case of surface restoration can be found in the paper of Clarenz et al. [25]. A related level set formulation, which deals with a slightly different splitting based on the level set function ϕ and the curvature density function $\mathbf{h}|\nabla\phi|$ was presented by Droske and Rumpf in 2004 [38]. Deckelnick and Dzuik [31] focused on error analysis for spatially discretized, time-continuous Willmore flow of graphs. Within this work they used an analogous splitting in context of piecewise linear finite elements and proved $L^\infty(L^2)$- as well as $L^2(L^2)$-error bounds of order $O(h^2(\log h)^2)$ for the discretized graph solution. Later Deckelnick and Schieweck [33] focused on axially symmetric surfaces and proved convergence of a conforming finite element approximation. They proved an $L^2(L^2)$-error bound of order $O(h^4)$ and $L^\infty(H^i)$-error bounds for $i = 0, 1, 2$ of order $O(h^{4-i})$. In the same year Dziuk and Deckelnick [34] published a paper dealing with error analysis of elastic flow of curves. They proved $L^\infty(H^1)$- as well as $L^2(L^2)$-error bounds of order $O(h)$. For parametric Willmore flow Barrett, Garcke and Nürnberg [7] and Dziuk [44] presented an alternative finite element algorithm. In comparison to all these papers the one by Bobenko and Schröder [15] deals with a different topic. They present a discretized Willmore flow scheme which takes into account a circle pattern on the surface. The temporal evolution of this circle pattern reflects the invariance of Willmore flow with respect to Möbius transformations.

1.2 Phase field based Willmore flow and mean curvature motion in literature

Since phase field based mean curvature motion as well as phase field based Willmore flow are main topics of this work we give a short survey on papers dealing with these topics while we concentrate on a content based survey on phase field discretizations in Section 2.1.

Already 1977 Modica and Mortola proved in [74] for a properly chosen double well function Ψ the Γ-convergence of the nowadays called Modica-Mortola energy

$$a^\varepsilon[u] = \frac{1}{2}\int_\Omega \varepsilon|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x$$

to the area functional. Based on this Modica presented a phase field model for mean curvature motion in [73]. In 1994 Nochetto, Paolini and Verdi [80, 79] proved an error estimate of order $O(\varepsilon^2)$ between the exact evolution of an interface under mean curvature motion and the corresponding evolution of a diffuse interface computed with a phase field mean curvature motion model. More recently, Evans, Soner and Souganidis proved in [50] that a scaled Allen–Cahn equation leads to a generalized motion by mean curvature.

In [30] it is presented the so-called de Giorgi conjecture. It states that the functional

$$w^\varepsilon[u] = \frac{1}{2\varepsilon}\int_\Omega \left(-\varepsilon\Delta u + \frac{1}{2\varepsilon}\Psi'(u)\right)^2 \mathrm{d}x$$

Γ-converges to the Willmore functional. This functional is foundation for further analytical treatment by Loretti and March in [69] and Bellettini and Mugnai in [8]. Du et al. [39] focused on the Euler–Lagrange equation of the phase field formulation and proved by formal asymptotics that it converges to the Euler–Lagrange equation of the Willmore energy (1.1). A corresponding Γ-convergence result for a modified functional was finally presented by Röger and Schätzle in [88]. Recently Dondl, Mugnai and Röger used a phase field model for minimizing Euler's elastica energy of non-overlapping curves in a bounded domain [36].

In the area of papers dealing with numerically discretized phase field models Chen et al. [23] proved that the zero level set of the solution of the Allen-Cahn equation converges to the mean curvature motion as ε goes to zero if h, $\sqrt{\tau} = O(\varepsilon^p)$ for $p > 1$ for grid size h and the time step size τ. In 2005 Deckelnick, Dziuk and Elliott presented in [32] a survey of different discretization of mean curvature flow. They compare the parametric, graph, level set and phase field based ansatz. With respect to the simulation of fourth order PDEs based on phase field models, Elliott et al. [45] proved (among other estimates) for a splitting scheme for the Cahn–Hilliard evolution an $O(h^2)$ estimate for the $L^\infty(L^2)$-error. In 1992 Elliott and Larsson analyzed a corresponding time-discretized backward Euler scheme and verified an $O(h^2 + \tau)$ error estimate. The last paper we mention here are written by Du et al. [40, 42]. They presented a discrete semi-implicit scheme for Willmore flow based on the above phase field energy $w^\varepsilon[\cdot]$. Their model is important with respect to our work since we compare it with the one presented here in Section 5.2.

Chapter 2

Foundations

This chapter starts with a short survey of three mathematical topics that are fundamental for this work. Since we work with phase field approximations we shortly introduce two widespread kind of phase field functions and focus on the type we will work with. After this we present some basic knowledge on Finite Element discretizations and particularly present the notation used in this work. Finally this chapter is finished with a more detailed survey on solving constrained minimization problems.

2.1 Phase Field approximations

Phase field functions are appropriate to deal with discontinuity sets that often arise in computer vision or physics. In general there are several kinds of phase field functions. Two well known types are presented in this section: Ambrosio-Tortorelli type phase field functions, also called single well phase field functions and Modica-Mortola or double well phase field functions.

Single well phase field functions are often used for free discontinuity problems like finding the cartoon of a given image $y^0 : \Omega \to \mathbb{R}$ with image domain $\Omega \subset \mathbb{R}^d$. A cartoon is a pair (y, K) consisting of a piecewise smooth function $y : \Omega \setminus K \to \mathbb{R}$ and a discontinuity set $K \subset \Omega$. The cartoon of a given image y^0 can be found by minimizing the Mumford-Shah functional (cf. [76, 3])

$$e^l_{\mathrm{MS}}[y, K] = \int_{\Omega\setminus K} |\nabla y|^2 + \alpha |y - y^0|^2 \,\mathrm{d}x + \nu \mathcal{H}^{d-1}(K)\,. \tag{2.1}$$

The first term acts as a smoother, preserving edges, the second term measures how well y approximates the given image y^0 and the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$ controls the length of the discontinuity set K. As discontinuity sets are difficult to handle, one introduces a phase field function, in this case a single well phase field function u. It approximates $1 - \chi_K$, where χ_K denotes the characteristic function of the set K. Using this kind of phase field function we can approximate $\mathcal{H}^{d-1}(K)$ by the Ambrosio-Tortorelli approximation

$$e^{\varepsilon}_{\mathrm{AT}}[u] = \frac{1}{2}\int_\Omega \varepsilon |\nabla u|^2 + \frac{1}{\varepsilon}(1-u)^2 \,\mathrm{d}x,$$

while the first term of the Mumford-Shah functional (2.1) is approximated by

$$\int_{\Omega\setminus K} |\nabla y|^2 \,\mathrm{d}x \approx \int_\Omega u^2 |\nabla y|^2 \,\mathrm{d}x\ .$$

The so-called optimal profile [17] of such a phase field function is $u(x) = 1 - e^{-\frac{\operatorname{dist}(x,K)}{\varepsilon}}$, where $\operatorname{dist}(x, K)$ denotes the distance of a point x to the set K. In one dimension, following Berkels [11], we set $K = \{0\}$, $\Omega = [0, \infty)$ and get

Lemma 2.1 (Optimal profile of an Ambrosio-Tortorelli phase field function in $1D$)**.** *The function* $u^\varepsilon : [0, \infty) \to \mathbb{R}$, $x \mapsto 1 - e^{-\frac{x}{\varepsilon}}$ *minimizes*

$$\frac{1}{2} \int_0^\infty \varepsilon |u'(x)|^2 + \frac{1}{\varepsilon} (1 - u(x))^2 \,\mathrm{d}x \tag{2.2}$$

under the boundary conditions $u(0) = 0$ *and* $\lim_{x \to \infty} u(x) = 1$.

Proof. The second derivative of the function $u^\varepsilon(x) = 1 - e^{-\frac{x}{\varepsilon}}$ is

$$(u^\varepsilon)''(x) = -\frac{1}{\varepsilon^2} e^{-\frac{x}{\varepsilon}}\,.$$

Thus $u^\varepsilon(x)$ solves the Euler–Lagrange equation of (2.2)

$$-\varepsilon u''(x) - \frac{1}{\varepsilon}(1 - u(x)) = 0$$

and as $u^\varepsilon(0) = 0$ and $\lim_{x \to \infty} u^\varepsilon(x) = 1$ it fulfils both boundary conditions. Combined with the convexity of the functional and the fact that it is finite if it is evaluated in u^ε the proof is completed. □

Double well phase field functions or Modica-Mortola type phase field functions arise in a wide range of problems. One example of a phase transition problem is the study of the behavior of metal alloys consisting of two components [18, 46]. Moreover Modica-Mortola type phase field functions arise in the study of two-phase incompressible flows [91, 68, 54] as well as in the study of reaction-diffusion systems [53]. Another example is image segmentation where an image $y^0 : \Omega \to \mathbb{R}$ is given and has to be decomposed into two regions described by a piecewise constant function $w : \Omega \to \{-1, 1\}$. As presented in [100], a model for partitioning a given image y^0 into two regions depending on the deviation of the image from two given values $c_1, c_2 \in \mathbb{R}$ is

$$\min_w \mathcal{H}^{d-1}(\Omega \cap \partial\{x \in \Omega : w(x) = 1\}) + \int_\Omega (w + 1)|y^0 - c_1|^2 + (w - 1)|y^0 - c_2|^2 \,\mathrm{d}x$$

on the set $\{w : \Omega \to \{-1, 1\}\}$. In this case the $(d-1)$-dimensional Hausdorff measure of the discontinuity set acts as regularizer. To deal with the discontinuity set we use a Modica-Mortola type phase field function $u \in H^1(\Omega)$ which approximates the piecewise constant function $w : \Omega \to \{-1, 1\}$. Hence, the $(d-1)$-dimensional Hausdorff measure of the discontinuity set is approximated by the Modica-Mortola energy [74]

$$a^\varepsilon[u] = \frac{1}{2} \int_\Omega \varepsilon |\nabla u|^2 + \frac{1}{\varepsilon} \Psi(u) \,\mathrm{d}x, \tag{2.3}$$

where $\Psi : \mathbb{R} \to \mathbb{R}$ denotes a double well potential with global minima $\Psi(-1) = \Psi(1) = 0$. If we set $\Psi(u) = (1 - u^2)^2$, the optimal profile of the Modica-Mortola type phase field function is given by $u(x) = \tanh\left(\frac{\operatorname{sgndist}(x,K)}{\varepsilon}\right)$. In this formula the discontinuity set is again denoted by K and $\operatorname{sgndist}(x, K)$ is the signed distance of a point $x \in \Omega$ to the set $K \subset \Omega$ with negative sign in the interior of K.

Lemma 2.2 (Optimal profile of a Modica-Mortola type phase field function in $1D$)**.** *The function* $u^\varepsilon : \mathbb{R} \to \mathbb{R}$, $x \mapsto \tanh\left(\frac{x}{\varepsilon}\right)$ *minimizes*

$$\frac{1}{2}\int_{\mathbb{R}} \varepsilon |u'|^2 + \frac{1}{\varepsilon}\left(1-u^2\right)^2 \mathrm{d}x \tag{2.4}$$

under the boundary conditions $\lim\limits_{x\to-\infty} u(x) = -1$ *and* $\lim\limits_{x\to\infty} u(x) = 1$.

Proof. By the direct method in the calculus of variations one proves the existence of a minimizer. Following Braides [17] one estimates by applying Young's inequality

$$\frac{1}{2}\int_{\mathbb{R}} \varepsilon|u'|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x \geq \left|\int_{u(-\infty)}^{u(\infty)} \sqrt{\Psi(u)}\,\mathrm{d}u\right|$$

for all u. Since

$$\begin{aligned}\left(u^\varepsilon\right)'(x) &= \frac{1}{\varepsilon}\left(1-\tanh^2\left(\frac{x}{\varepsilon}\right)\right),\\ u^\varepsilon(0) &= 0\end{aligned}$$

u^ε solves the ordinary differential equation

$$\begin{aligned}u'(x) &= \frac{1}{\varepsilon}\sqrt{\Psi(u(x))},\\ u(0) &= 0\end{aligned}$$

which is equivalent to

$$\begin{aligned}u'(x) &= \frac{1}{\varepsilon}\left(1-u^2(x)\right),\\ u(0) &= 0\end{aligned}$$

with $\Psi(u) = \left(1-u^2\right)^2$. Moreover $\lim\limits_{x\to-\infty} u^\varepsilon(x) = -1$ and $\lim\limits_{x\to\infty} u^\varepsilon(x) = 1$.
Now let $\tilde{u}$ be a minimizer of (2.4) with $\lim\limits_{x\to-\infty} \tilde{u}(x) = -1$ and $\lim\limits_{x\to\infty} \tilde{u}(x) = 1$. Then

$$\begin{aligned}\frac{1}{2}\int_{\mathbb{R}} \varepsilon|\tilde{u}'|^2 + \frac{1}{\varepsilon}\Psi(\tilde{u})\,\mathrm{d}x &\leq \frac{1}{2}\int_{\mathbb{R}} \varepsilon|(u^\varepsilon)'|^2 + \frac{1}{\varepsilon}\Psi(u^\varepsilon)\,\mathrm{d}x\\ &= \int_{u^\varepsilon(-\infty)}^{u^\varepsilon(\infty)} \sqrt{\Psi(u^\varepsilon)}\,\mathrm{d}u^\varepsilon .\end{aligned}$$

Consequently

$$\left|\int_{\tilde{u}(-\infty)}^{\tilde{u}(\infty)} \sqrt{\Psi(\tilde{u})}\,\mathrm{d}\tilde{u}\right| \leq \frac{1}{2}\int_{\mathbb{R}} \varepsilon|\tilde{u}'|^2 + \frac{1}{\varepsilon}\Psi(\tilde{u})\,\mathrm{d}x \leq \int_{u^\varepsilon(-\infty)}^{u^\varepsilon(\infty)} \sqrt{\Psi(u^\varepsilon)}\,\mathrm{d}u^\varepsilon .$$

From this it follows that u^ε is a minimizer of (2.4). □

In this work we deal with Modica-Mortola type phase field functions. Thus we concentrate on this type and regard the Modica-Mortola energy (2.3) in more detail. The second term of the integrand in (2.3) consists of the double well potential, which in this work is $\Psi(u) = \left(1 - u^2\right)^2$. This term penalizes values of u being different from -1 or 1. However it does not prevent u from oscillating between these two values. Therefore we need the first term, containing the gradient of u. The scaling of the higher order term is explained by Braides in [17] for the one dimensional case: consider

$$\int_y^{y+\delta} \varepsilon |u'(x)|^2 \,\mathrm{d}x + \int_y^{y+\delta} \frac{1}{\varepsilon} \Psi(u(x)) \,\mathrm{d}x \tag{2.5}$$

on the interval $(y, y+\delta)$, where $u(y)$ is close to -1 and $u(y+\delta)$ close to 1. Thus δ is the interface transition width. As $\Psi(u(x))$ is bounded on $(y, y+\delta)$, the contribution of the second term is of order $\frac{\delta}{\varepsilon}$. With respect to the special choice of the interval $(y, y+\delta)$ the contribution of the first term is of order $\frac{\varepsilon}{\delta}$. Thus (2.5) is of order $\frac{\varepsilon}{\delta} + \frac{\delta}{\varepsilon}$ which is minimal and positive for $\varepsilon = \delta$. It follows: if (2.5) is bounded the number of such intervals is bounded and therefore u does not oscillate between -1 and 1.
Now we introduce the so-called Γ-convergence which is fundamental for the next theorem.

Definition 2.3 (Γ-convergence)**.** *Let X be a topological space and $f_i : X \to \mathbb{R} \cup \{\infty\}$, $i \in \mathbb{N}$ a sequence of functionals. Then we say f_i Γ-converges to $f : X \to \mathbb{R} \cup \{\infty\}$, if for every $x \in X$ the following two conditions hold:*

(i) For every sequence $x_i \in X$, $i \in \mathbb{N}$ with $x_i \xrightarrow{i \to \infty} x$

$$f(x) \le \liminf_{i \to \infty} f_i(x_i)\,.$$

(ii) There exists a sequence $x_i \in X$, $i \in \mathbb{N}$ with $x_i \xrightarrow{i \to \infty} x$, such that

$$f(x) \ge \limsup_{i \to \infty} f_i(x_i)\,.$$

In Chapter 3 we will use that $a^\varepsilon[\cdot]$ Γ-converges to $\frac{4}{3}\mathcal{H}^{d-1}[\cdot]$ for $\varepsilon \to 0$ and double well potential $\Psi(u) = \left(1 - u^2\right)^2$. Here the factor $\frac{4}{3}$ is induced by the fact that

$$\int_{-1}^{1} \sqrt{\Psi(s)} \,\mathrm{d}s = \frac{4}{3}$$

for $\Psi(u) = \left(1 - u^2\right)^2$. The more general version of this statement deals with an arbitrary $\Psi \in C^1(\mathbb{R})$ with global minima $\Psi(-1) = \Psi(1) = 0$.

Theorem 2.4. *Let*

$$a^\varepsilon[u] := \begin{cases} \frac{1}{2} \int_\Omega \varepsilon |\nabla u|^2 + \frac{1}{\varepsilon} \Psi(u) \,\mathrm{d}x & \text{if } u \in H^1(\Omega) \\ \infty & \text{else,} \end{cases}$$

then

$$\Gamma - \lim_{\varepsilon \to 0} a^\varepsilon = \int_{-1}^{1} \sqrt{\Psi(s)} \,\mathrm{d}s \ \mathrm{Per}(\cdot)\,,$$

where $\mathrm{Per} : L^1(\Omega) \to [0,\infty]$ *is the perimeter functional*

$$\mathrm{Per}(u) := \begin{cases} \mathcal{H}^{d-1}(\Omega \cap \partial\{x \in \Omega \,:\, u(x) = 1\}) & \text{if } u : \Omega \to \{-1, 1\} \\ \infty & \text{else} \end{cases}$$

and the Γ*-limit is taken with respect to the* $L^1(\Omega)$*-topology.*

For the proof we refer to [17]. In this work we present a two-dimensional computation which illustrates the impact of this Theorem. Regard the zero level set of u, denoted by $[u = 0]$ as curve with arc length parametrization $\gamma(s)$, where s denotes the arc length parameter. Then each point $x \in U_\delta([u = 0])$ in a δ-neighborhood of the zero level set of u can be written as $x(s,t) = \gamma(s) + tn(s)$ with $t \in \mathbb{R}$ and $n(s)$ being the outer unit normal vector to the curve γ in $\gamma(s)$. We choose δ small, but big enough such that $u(x)$ approximately equals 1 or -1 for all $x \in \Omega \setminus U_\delta([u = 0])$, consequently $\Psi(u) = |\nabla u| = 0$ on $\Omega \setminus U_\delta([u = 0])$. Thus we do a change of coordinates and approximate the Modica-Mortola energy

$$\begin{aligned} a^\varepsilon[u] &= \frac{1}{2}\int_\Omega \varepsilon|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x \\ &\approx \frac{1}{2}\int_{\gamma^{-1}([u=0])}\int_{-\delta}^{\delta}\left(\varepsilon|\nabla u^\varepsilon(s,t)|^2 + \frac{1}{\varepsilon}\Psi(u^\varepsilon(s,t))\right)|\det Dx(s,t)|\,\mathrm{d}t\,\mathrm{d}s\ . \end{aligned}$$

As before u^ε denotes the optimal profile (cf. Lemma 2.1 and 2.2).
At the point $\gamma(s)$, we denote the tangent vector to the curve γ with $t(s)$. Moreover let $\mathbf{h}$ be the mean curvature of γ and $(t(s)\,,\, n(s))$ the matrix consisting of the two columns $t(s)$ and $n(s)$. Then we get with respect to the Frenet formula $n'(s) = \mathbf{h}t(s)$

$$\begin{aligned} |\det(Dx(s,t))| &= |\det\left(\gamma'(s) + tn'(s)\,,\, n(s)\right)| \\ &= |(1 + t\mathbf{h})\det(t(s)\,,\, n(s))| \\ &= |1 + t\mathbf{h}|\,. \end{aligned} \tag{2.6}$$

In a δ-neighborhood of $[u = 0]$ and with δ small enough we can approximate

$$|\det(Dx(s,t))| \approx 1\,.$$

For the following corollary we need the definition

Definition 2.5. *We write*

$$f = g + O(h) \qquad \text{as } x \to x_0$$

provided there exists a constant C *such that*

$$|f(x) - g(x)| \le C|h(x)|$$

for all x *sufficiently close to* x_0*.*

In other words: $f = g + O(h)$ as $x \to x_0$ if and only if $f - g = O(h)$ as $x \to x_0$. Based on this notation we state:

Corollary 2.6. *Let* $u^\varepsilon : \mathbb{R} \to \mathbb{R}$, $x \mapsto \tanh\left(\frac{x}{\varepsilon}\right)$ *and* $a \in \mathbb{R}^{>0}$ *then*

$$\varepsilon \int_{-a}^{a} \left((u^\varepsilon)'(t)\right)^2 \mathrm{d}t = \frac{4}{3} + O(\sigma(a,\varepsilon))$$

with $\sigma(a,\varepsilon) := 2\left(1 - \tanh\left(\frac{a}{\varepsilon}\right)\right)^2 \left(1 + \tanh\left(\frac{a}{\varepsilon}\right)\right)$.
Moreover, for $a \geq -\frac{\varepsilon}{2}\left(\ln\frac{1}{2} + \frac{3}{2}\ln\delta\right)$ *with* $\delta \in \mathbb{R}^{>0}$

$$\varepsilon \int_{-a}^{a} \left((u^\varepsilon)'(t)\right)^2 \mathrm{d}t = \frac{4}{3} + O(\delta^3)\,.$$

Proof. From the proof of Lemma 2.2 we know $(u^\varepsilon)'(x) = \frac{1}{\varepsilon}\sqrt{\Psi(u^\varepsilon(x))}$. Thus, using a simple coordinate transformation we get

$$\begin{aligned}
\varepsilon \int_{-a}^{a} \left((u^\varepsilon)'(t)\right)^2 \mathrm{d}t &= \int_{u^\varepsilon(-a)}^{u^\varepsilon(a)} \sqrt{\Psi(u^\varepsilon)}\,\mathrm{d}u^\varepsilon \\
&= \int_{-1}^{1} \left(1 - (u^\varepsilon)^2\right) \mathrm{d}u^\varepsilon + \int_{[-1,1]\setminus[u^\varepsilon(-a),u^\varepsilon(a)]} \left((u^\varepsilon)^2 - 1\right) \mathrm{d}u^\varepsilon \\
&= \frac{4}{3} + O(\sigma(a,\varepsilon))\,.
\end{aligned}$$

In the last line we have used

$$\begin{aligned}
\left| \int_{[-1,1]\setminus[u^\varepsilon(-a),u^\varepsilon(a)]} \left((u^\varepsilon)^2 - 1\right) \mathrm{d}u^\varepsilon \right| &\leq \int_{[-1,1]\setminus[u^\varepsilon(-a),u^\varepsilon(a)]} |\,(u^\varepsilon)^2 - 1|\,\mathrm{d}u^\varepsilon \\
&\leq \int_{[-1,1]\setminus[u^\varepsilon(-a),u^\varepsilon(a)]} \left(1 - \tanh^2\left(\frac{a}{\varepsilon}\right)\right) \mathrm{d}u^\varepsilon \\
&= 2\left(1 - \tanh\left(\frac{a}{\varepsilon}\right)\right)\left(1 - \tanh^2\left(\frac{a}{\varepsilon}\right)\right) \\
&= \sigma(a,\varepsilon)\,.
\end{aligned}$$

As $1 - \tanh\left(\frac{a}{\varepsilon}\right) \leq 2e^{-2\frac{a}{\varepsilon}}$ we estimate

$$\sigma(a,\varepsilon) \leq 4\left(2e^{-2\frac{a}{\varepsilon}}\right)^2\,.$$

Moreover $a \geq -\frac{\varepsilon}{2}\left(\ln\frac{1}{2} + \frac{3}{2}\ln\delta\right)$ is equivalent to $2e^{-2\frac{a}{\varepsilon}} \leq \delta^{\frac{3}{2}}$. ☐

As $\Psi(u) = \left(1 - u^2\right)^2$ is the double well potential used before, $u^\varepsilon(s,t)$ is the optimal profile from Lemma 2.2, only depending on t and with Corollary 2.6 we get

$$\begin{aligned}
a^\varepsilon[u] &\approx \frac{1}{2}\int_{\gamma^{-1}([u=0])} \int_{-\delta}^{\delta} \varepsilon|\nabla u^\varepsilon(t)|^2 + \frac{1}{\varepsilon}\Psi(u^\varepsilon(t))\,\mathrm{d}t\,\mathrm{d}s \\
&= \frac{1}{2}\int_{\gamma^{-1}([u=0])} \int_{-\delta}^{\delta} \frac{2}{\varepsilon}\Psi(u^\varepsilon(t))\,\mathrm{d}t\,\mathrm{d}s \\
&\approx \frac{4}{3}\int_{\gamma^{-1}([u=0])} \mathrm{d}s \\
&\approx \frac{4}{3}\mathcal{H}^{d-1}([u=0])\,.
\end{aligned}$$

Thus this approximate computation leads to the same result as presented in Theorem 2.4.

2.2 Finite Element discretization

In most cases we use the linear Finite Element method for the spatial discretization of our models. Only in Chapter 6 we use the multilinear Finite Element method for some special terms. Thus we recall both versions in this section and introduce the notation. For further information we refer to [16].
In case of the linear Finite Element method our computational domain Ω, which is $[0,1]^d$ with $d = 2,3$ is covered by a regular and uniform simplicial mesh $\mathcal{T}$. The elements of this mesh are triangles in $2D$ and tetrahedra in $3D$, both denoted by $T \in \mathcal{T}$. Moreover I denotes the vertex index set of $\mathcal{T}$. With $\mathcal{P}_1$ denoting the space of affine functions the Finite Element space of continuous, piecewise affine functions is

$$\mathcal{V} := \left\{ \Phi \in C^0(\Omega) \, : \, \Phi|_T \in \mathcal{P}_1 \, \forall T \in \mathcal{T} \right\} .$$

This Finite Element space is spanned by the usual nodal basis $\{\Phi_i\}_{i \in I}$. Here, the basis function Φ_i is uniquely defined as the continuous, piecewise linear function with the nodal property

$$\Phi_i(x_j) = \begin{cases} 1 & \text{for } i = j \, , \\ 0 & \text{else,} \end{cases} \tag{2.7}$$

where x_i denotes a vertex of the grid $\mathcal{T}$. Now the Finite Element discretization of a continuous function $f : \Omega \to \mathbb{R}$ is $F = \sum_{i \in I} F_i \Phi_i$, where $F_i = f(x_i)$. Note, within this work we denote continuous functions with lower case letters and their Finite Element discretizations with upper case letters. The corresponding nodal coordinate vector is $\overline{F} = (F_i)_{i \in I}$.
In order to write down the matrix formulation of our discretized models we introduce the weighted lumped mass matrix

$$\mathbf{M}[\lambda] = \left(\int_\Omega \mathcal{I}_h \left(\lambda \Phi_i \Phi_j \right) \mathrm{d}x \right)_{i,j} ,$$

where $\lambda : \Omega \to \mathbb{R}$ is an arbitrary continuous weight function and $\mathcal{I}_h$ denotes the Lagrange interpolation. The canonical lumped mass matrix is denoted by $\mathbf{M} = \mathbf{M}[1]$ and the stiffness matrix is given by

$$\mathbf{L} = \left(\int_\Omega \nabla \Phi_i \cdot \nabla \Phi_j \, \mathrm{d}x \right)_{i,j} .$$

$\mathbf{M}[\lambda]$ and $\mathbf{M}$ are lumped mass matrices and consequently diagonal by definition. The stiffness matrix $\mathbf{L}$ is also a sparse matrix, but here all entries $L_{i,j}$ are different from zero for which the nodes x_i and x_j are either equal or adjacent in the mesh $\mathcal{T}$. To compute the nonzero entries we use midpoint quadrature, which is exact for linear finite elements.

For $f : \Omega \to \mathbb{R}^d$ being a vector valued function its Finite Element discretization is

$$F = \sum_{\substack{i \in I \\ k=1,\dots,d}} F_i^k \Phi_i^k ,$$

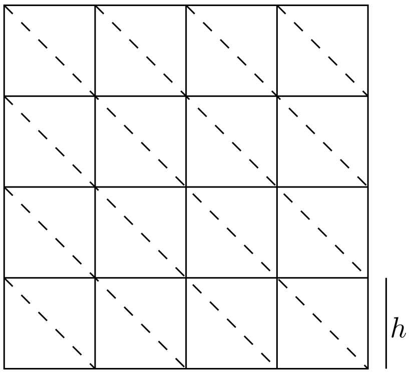


Figure 2.1: A uniform rectangular grid $\mathcal{C}$ (solid lines) and the corresponding uniform simplicial grid $\mathcal{T}$ (solid and broken lines). The vertices of both meshes are identical and the mesh resolution h, too.

where $\Phi_i^k = (0\dots0\,\Phi_i\,0\dots0)^T = \Phi_i e_k$ and Φ_i denotes the ith basis vector of $\mathcal{V}$. In this work we often deal with the problem of solving a system of equations of the form $\nabla_{(U,V,P)}\mathcal{L}[U,V,P] = 0$. We regard the three unknowns as one vector valued function with three components. As we use a Newton method to solve this problem we need to calculate a Hessian $D^2\mathcal{L}$ whose entries are matrices in the discretized version. Thus we have to deal with so-called block matrices, for which we introduce the following notation. In case of a $d \times d$ block matrix A we write

$$A = \left(A_{i,j}^{k,l}\right)_{i,j\in I,\, k,l=0,\dots d-1}\,,$$

where the indices k and l denote the position of the block entry and i, j are the indices within each block iterating over all entries of the vertex index set I.

In case of multilinear Finite Elements our computational domain Ω is covered by a uniform rectangular grid $\mathcal{C}$. Here, the elements are squares in $2D$ and cubes in $3D$. In both cases we denote them by $C \in \mathcal{C}$. As we can see in Figure 2.1 the vertex index set I is independent of the choice between linear and multilinear Finite Element discretization and the mesh resolution h is identical in both cases, too. Let $\mathcal{P}_1^m$ be the set of multilinear functions. Then the Finite Element space of continuous, piecewise multilinear functions is

$$\mathcal{V}_m := \left\{\Phi \in C^0(\Omega)\,:\,\Phi|_C \in \mathcal{P}_1^m\,\forall C \in \mathcal{C}\right\}\,.$$

Although the basis functions of $\mathcal{V}_m$ fulfill the nodal property (2.7), they are not identical to the basis functions of $\mathcal{V}$. Nevertheless, by a slight misuse of notation, we denote them by Φ_i, too. Within this work we specify in advance which terms of a functional are discretized by linear or multilinear Finite Elements, but afterwards we use the same notation, that means, lower case letters for continuous functions, upper case letters for discretized functions and upper case letters with bar on top for the corresponding nodal vector.

As we will see in Chapter 6 we use both discretizations within one functional. This is uncomplicated, because as we have seen the vertex index set of both kinds of discretizations is identical and thus the nodal vector we are working with is identical, too. We only have to be careful in being consistent in all derivatives of the functional. Moreover, in Section 6.3 we will have to evaluate functions of type

$$f(x_1, \dots, x_{d-1}) = \int_0^1 \tilde{f}(x)\,\mathrm{d}x_d$$

with $x \in \mathbb{R}^d$ and $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$ and integrals

$$\int_{[0,1]^{d-1}} f(x_1, \dots, x_{d-1})\,\mathrm{d}x_{1,\dots,d-1}\,. \tag{2.8}$$

Thus it is convenient to use multilinear Finite Element discretization with tensor product Gaussian quadrature [28]. This allows us to write an integral of type (2.8) as a single integral over the whole d-dimensional cube $[0,1]^d$. Within this work we use in case of multilinear Finite Elements Gaussian quadrature with two quadrature points in the direction of each coordinate axis at $(1 - \frac{1}{\sqrt{3}})/2$ and $(1 + \frac{1}{\sqrt{3}})/2$ [93].

2.3 Survey of solving constrained minimization problems

Solving a constrained minimization problem is a central issue in this work. Thus we give a short survey of this topic. We orient ourselves mainly on the book of Nocedal and Wright [78] and adapt it to our problem setup by treating equality constraints only. In the literature, the problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad c_i(x) = 0 \quad \forall i \in \mathcal{E}\,, \tag{2.9}$$

with f, $c_i : \mathbb{R}^n \to \mathbb{R}$ and finite index set $\mathcal{E}$ is often regarded. We will see in the following chapters that we always have to minimize functionals over function spaces. But at the end we regard the discretized case and consequently we have again a problem of type (2.9).

One well known possibility to solve such problems consists in solving the corresponding saddle point problem. We set up the Lagrange function

$$\ell(x,p) = f(x) - \sum_{i \in \mathcal{E}} p_i c_i(x) \tag{2.10}$$

consisting of the function f we want to minimize and the scalar product of the Lagrange multiplier vector $p = (p_i)_{i \in \mathcal{E}}$ and the vector of constraints $c(x) = (c_i(x))_{i \in \mathcal{E}}$. A saddle point (x^*, p^*) of this Lagrange function has to fulfill the Karush-Kuhn-Tucker (KKT) conditions

$$\begin{aligned} \nabla_x \ell(x^*, p^*) = \nabla_x f(x^*) - \sum_{i \in \mathcal{E}} p_i^* \nabla_x c_i(x^*) &\overset{!}{=} 0\,, \\ \nabla_p \ell(x^*, p^*) = c(x^*) &\overset{!}{=} 0\,. \end{aligned}$$

This method is used in most parts of this work, but a disadvantage is that it requires solving a saddle point problem. Thus simple step size controls based on line search methods which aim to minimize an energy or function, do not work in general.

Another way to solve problem (2.9) requires setting up the quadratic penalty function

$$\mathfrak{q}(x;\lambda) = f(x) + \frac{\lambda}{2}\sum_{i\in\mathcal{E}} c_i^2(x)$$

with penalty parameter $\lambda > 0$. With this method one computes an approximate solution of problem (2.9) by minimizing the unconstrained function $\mathfrak{q}(x;\lambda)$ for a series of growing λ. A disadvantage of this method consists in the fact that, depending on the choice of λ the quadratic penalty function can be unbounded below although the original constrained problem has an exact solution (cf. [78]). Moreover the approximate minimizer x_k of the quadratic penalty function $\mathfrak{q}(x;\lambda_k)$ does not fulfill the constraint exactly. In [78] it is proven that

$$c_i(x_k) \approx -\frac{p_i^*}{\lambda_k}, \quad \forall i \in \mathcal{E}\,.$$

This will be prevented by using a different method, which is based on both methods previously presented in this section. In contrast to the quadratic penalty method, which is not used in this work, this other method will be used in Chapter 7.

This previously mentioned method is the Augmented Lagrange method or method of multipliers. It obviously bases on the two presented methods, as the structure of the Lagrange function $\ell(x,p)$ appears in the Augmented Lagrange function

$$\ell_A(x,p;\lambda) = f(x) - \sum_{i\in\mathcal{E}} p_i c_i(x) + \frac{\lambda}{2}\sum_{i\in\mathcal{E}} c_i^2(x)$$

as well as the structure of the quadratic penalty function $\mathfrak{q}(x;\lambda)$. Within this method one chooses a fixed estimate of the Lagrange multiplier vector p^k and a fixed $\lambda_k > 0$. For these parameters one minimizes $\ell_A(\cdot,p^k;\lambda_k)$ approximately until p^k and λ_k are updated to p^{k+1} and λ_{k+1} and $\ell_A(\cdot,p^{k+1};\lambda_{k+1})$ is approximately minimized. A suitable update formula for the Lagrange multiplier bases on the comparison of the optimality condition $\nabla_x \ell_A(x_k,p^k;\lambda_k) = 0$ for unconstrained minimization problem with the first Karush-Kuhn-Tucker condition $\nabla_x \ell(x^*,p^*)$ for a saddle point (x^*,p^*) of the Lagrange function (2.10). This update formula is

$$p_i^{k+1} = p_i^k - \lambda_k c_i(x_k) \quad \forall i \in \mathcal{E}\,.$$

In general convergence of this method can be assured even if λ is not increased indefinitely. Thus the problem of being ill conditioned is less critical than in the quadratic penalty method. Moreover Bertsekas proved in [14] that an approximate solution x^k is close to the exact solution x^* if p is the exact Lagrange multiplier or λ is big enough. However the main advantage of the Augmented Lagrange method is that we have to solve an unconstrained minimization problem in each Augmented Lagrange iteration and do not have to find a saddle point any more.
In most parts the concrete Algorithm for the Augmented Lagrange method is equal to an algorithm implemented in the LANCELOT software package [26]. The original

version handles equality and inequality constrains, but we implemented the equality constrained version, only. This algorithm can also be found in [78].
In this algorithm η^* and w^* denote final convergence tolerances: η^* guaranties that the L^2-norm of the constraint is lower than allowed and w^* is the tolerance which has to be fulfilled by the L^2-norm of the gradient of the Augmented Lagrange function. In comparison to this η_k and w_k are the corresponding temporal convergence tolerances which are adpated in each iteration.
The constants 100, 0.1 and 0.9 in the update formula for λ_{k+1} and η_{k+1} are unchanged in comparison to the original version of this algorithm presented in [78]. In general it is possible to slightly modify these constants without compromising the theoretical convergence properties of Algorithm 2.1. This special choice of constants guarantees: if the L^2-norm of the constraint is less or equal η_k than the Lagrange mutliplier is updated, λ_k is unchanged and the convergence tolerances η_k and w_k are reduced. Here the constant 0.9 in the update formula for η_{k+1} guarantees a stronger reduction of w_k than of η_k. If the L^2-norm of the constraint is greater than η_k the Lagrange multiplier is unchanged, while λ_k grows by a factor 100. In this case the convergence tolerances w_k and η_k also grow. Now the constant 0.1 guarantees a lower value for w_{k+1} than for η_{k+1}.

Algorithm 2.1: Augmented Lagrange method

```
given: initial point $x_0$ and initial Lagrange multipliers $p^0$;
given: convergence tolerances $\eta^*$ and $w^*$;
given: maximal number of Augmented Lagrange iterations $K_{max}$;
given: penalty parameter $\lambda_0$;
Set $w_0 = \frac{1}{\lambda_0}$ and $\eta_0 = \frac{1}{\lambda_0^{0.1}}$;
for $k = 0$ to $K_{max}$ do
    find an approximate minimizer $x_k$ of $\ell_A(\cdot, p^k; \lambda_k)$ such that
        $\|\nabla_x \ell_A(x_k, p^k; \lambda_k)\|_{L^2} \leq w_k$;
    if $\|c(x_k)\|_{L^2} \leq \eta_k$ then
        if $\|c(x_k)\|_{L^2} \leq \eta^*$ and $\|\nabla_x \ell_A(x_k, p^k; \lambda_k)\|_{L^2} \leq w^*$ then
            stop;
        end
        $p^{k+1} = p^k - \lambda_k c(x_k)$;
        $\lambda_{k+1} = \lambda_k$;
        $\eta_{k+1} = \frac{\eta_k}{\lambda_{k+1}^{0.9}}$;
        $w_{k+1} = \frac{w_k}{\lambda_{k+1}}$;
    else
        $p^{k+1} = p^k$;
        $\lambda_{k+1} = 100\lambda_k$;
        $\eta_{k+1} = \frac{1}{\lambda_{k+1}^{0.1}}$;
        $w_{k+1} = \frac{1}{\lambda_{k+1}}$;
    end
end
```

Nested time discretization of Willmore flow

3.1 Derivation of the model

Before we deal with the derivation of the whole model let us have a look at the natural time discretization of mean curvature motion, which is part of our model. Originally Luckhaus and Sturzenhecker [70] presented this ansatz to develop an implicit time discretization for mean curvature motion. Later Chambolle [21] transferred it to the level set version, and recently Olischläger and Rumpf [81, 82] used this ansatz to develop a two step time discretization for Willmore flow in the parametric case.

3.1.1 Natural time discretization of mean curvature motion

We consider the general case of a $(d-1)$-dimensional hypersurface Γ evolving under a gradient flow. Let the hypersurface Γ be contained in a computational domain $\Omega \subset \mathbb{R}^d$ and $x : \Gamma \to \Gamma$ be the identity on Γ. The Riemannian manifold consisting of all admissible hypersurfaces Γ is denoted by $\mathcal{M}$ and provided with a metric $g_\Gamma : T_\Gamma\mathcal{M} \times T_\Gamma\mathcal{M} \to \mathbb{R}$. Here $T_\Gamma\mathcal{M}$ is the tangent space to $\mathcal{M}$ in Γ and in this work the expression metric is used in the same sense as the expression scalar product. The energy which will be minimized by a gradient flow is $e : \mathcal{M} \to \mathbb{R}$. With this notation the gradient flow equation minimizing the energy $e[\cdot]$ with initial data $x^0 \in \mathcal{M}$ is

$$\partial_t\, x(t) = -\operatorname{grad}_{g_\Gamma} e[x(t)] \tag{3.1a}$$

$$x(0) = x^0, \tag{3.1b}$$

where the gradient is defined by

Definition 3.1. *Let $\mathcal{M}$ be a manifold, $T_\Gamma\mathcal{M}$ the tangent space to $\mathcal{M}$ in $\Gamma \in \mathcal{M}$, $x : \Gamma \to \Gamma$ the identity on Γ, $e : \mathcal{M} \to \mathbb{R}$ a scalar mapping and $g_\Gamma : T_\Gamma\mathcal{M} \times T_\Gamma\mathcal{M} \to \mathbb{R}$ a metric on $\mathcal{M}$. Then the gradient of e is defined by*

$$\operatorname{grad}_{g_\Gamma} e[x] = v \in T_\Gamma\mathcal{M} :\Leftrightarrow \forall w \in T_\Gamma\mathcal{M},\ g_\Gamma(v,w) = \langle e'[x], w\rangle,$$

where $\langle e'[x], w\rangle = \frac{\mathrm{d}}{\mathrm{d}\epsilon} e[c_w(\epsilon)]\big|_{\epsilon=0}$ with $c \in C^1([0,1],\mathcal{M})$ and $c(0) = x$, $\dot{c}_w(0) = w$.

To solve the gradient flow equation one can approximate the time derivative or use a natural time discretization, based on a variational ansatz, leading to a time discrete solution.

Definition 3.2 (Natural time discretization)**.** *Given a manifold $\mathcal{M}$ with metric $g_\Gamma : T_\Gamma\mathcal{M} \times T_\Gamma\mathcal{M} \to \mathbb{R}$ and a parametrization x^k of a hypersurface $\Gamma^k \in \mathcal{M}$ at time $k\tau$. Define the parametrization x^{k+1} of the hypersurface $\Gamma^{k+1} \in \mathcal{M}$ at time $(k+1)\tau$ by*

$$x^{k+1} := \operatorname*{argmin}_x \left(\operatorname{dist}(x, x^k)^2 + 2\tau e[x]\right), \tag{3.2}$$

where

$$\operatorname{dist}(x, x^k) := \inf_{c\in\mathcal{S}} \int_0^1 \sqrt{g_{c(s)}(\dot c(s), \dot c(s))}\,\mathrm{d}s$$

denotes the shortest path length on the manifold, with respect to the metric $g_{c(s)}$ and

$$\mathcal{S} := \{c \in C^1([0,1], \mathcal{M}) \,:\, c(0) = x^k, c(1) = x\}.$$

In [83], Otto proved that the Euler–Lagrange equation of the natural time discretization (3.2) coincides with the implicit Euler method for the gradient flow equation (3.1a). Additionally, by the definition of the natural time discretization we directly get an energy estimate, which guarantees the energy decay in each time step.

Corollary 3.3 (Energy estimate)**.** *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of parametrizations of a sequence of hypersurfaces $\left(\Gamma^k\right)_{k\in\mathbb{N}}$ solving the natural time discretization of a gradient flow minimizing $e[\cdot]$. Then,*

$$e[x^{k+1}] \leq e[x^k].$$

Proof. As x^{k+1} is the minimizing argument of $\operatorname{dist}(x, x^k)^2 + 2\tau e[x]$ and by the positiveness of $\operatorname{dist}(x, x^k)^2$ we get

$$e[x^{k+1}] \leq e[x^{k+1}] + \frac{1}{2\tau}\operatorname{dist}(x^{k+1}, x^k)^2 \leq e[x^k] + \frac{1}{2\tau}\operatorname{dist}(x^k, x^k)^2 = e[x^k].$$

□

In the special case of mean curvature motion we choose as metric $g_\Gamma : T_\Gamma\mathcal{M} \times T_\Gamma\mathcal{M} \to \mathbb{R}$ on the Riemannian manifold $\mathcal{M}$ the L^2-metric on the hypersurface Γ and denote it by $(\cdot,\cdot)_{L^2(\Gamma)}$. The parametrization of the initial hypersurface $\Gamma(0)$ is $x^0 : \Gamma(0) \to \Gamma(0)$, while each other hypersurface $\Gamma^k = \Gamma(k\tau)$ at time $k\tau$ is parametrized over the previous hypersurface Γ^{k-1} by $x^k : \Gamma^{k-1} \to \mathbb{R}^d$. Then the natural time discretization of mean curvature motion is

Definition 3.4 (Natural time discretization of mean curvature motion)**.** *Given a parametrization x^k of a hypersurface $\Gamma^k \in \mathcal{M}$ at time $k\tau$, define the parametrization x^{k+1} of the hypersurface $\Gamma^{k+1} \in \mathcal{M}$ at time $(k+1)\tau$ by*

$$x^{k+1} := \operatorname*{argmin}_{x:\Gamma^k\to\mathbb{R}} \|x - x^k\|^2_{L^2(\Gamma^k)} + 2\tau\mathcal{H}^{d-1}[x(\Gamma^k)]\,. \tag{3.3}$$

In a next step, we introduce the phase field version of the natural time discretization of mean curvature motion. We denote the phase field function, representing $x(\Gamma^k)$ by u and use Theorem 2.4 to replace $\mathcal{H}^{d-1}[x(\Gamma^k)]$ by $\frac{3}{4}a^\varepsilon[u]$. To replace the first term of (3.3) we introduce the following lemma.

Lemma 3.5. *Let* $u^\varepsilon : \mathbb{R} \to \mathbb{R}$, $t \mapsto \tanh\left(\frac{t}{\varepsilon}\right)$ *and* $\delta, a \in \mathbb{R}^{>0}$ *with* $\delta \ll 1$ *and* $a \geq -\frac{\varepsilon}{2}\left(\ln\frac{1}{2} + \frac{3}{2}\ln\delta\right)$. *Then*

$$\varepsilon \int_{-a}^{a} (u^\varepsilon(t) - u^\varepsilon(t-\delta)))^2 \,\mathrm{d}t = \frac{4}{3}\delta^2 \left(1 + O(\Theta(\delta,\varepsilon))\right) ,$$

where $\Theta(\delta,\varepsilon) := \frac{\sqrt{\delta}}{\varepsilon} + \frac{\delta^2}{\varepsilon^2} + \delta^3$.

Proof. The Taylor expansion of $u^\varepsilon(t-\delta)$ is

$$\begin{aligned} u^\varepsilon(t-\delta) &= u^\varepsilon(t) - (u^\varepsilon)'(t)\delta + \int_t^{t-\delta} (t-\delta-\tilde{s})\,(u^\varepsilon)''(\tilde{s})\,\mathrm{d}\tilde{s} \\ &= u^\varepsilon(t) - (u^\varepsilon)'(t)\delta + \int_0^1 (t-\delta-t+s\delta)\,(u^\varepsilon)''(t-s\delta)(-\delta)\,\mathrm{d}s \\ &= u^\varepsilon(t) - (u^\varepsilon)'(t)\delta + \delta^2\int_0^1 (1-s)\,(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s \end{aligned}$$

in which we have used a coordinate transformation $\tilde{s} := t - s\delta$ in the Lagrange Remainder. Comparing with Corollary 2.6 we get

$$\begin{aligned} \varepsilon \int_{-a}^{a} (u^\varepsilon(t) - u^\varepsilon(t-\delta))^2 \,\mathrm{d}t &= \varepsilon \int_{-a}^{a} \left((u^\varepsilon)'(t)\delta + \delta^2 \int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t \\ &= \varepsilon \int_{-a}^{a} \delta^2 \left((u^\varepsilon)'(t)\right)^2 \mathrm{d}t \\ &\quad + 2\varepsilon \int_{-a}^{a} (u^\varepsilon)'(t)\delta^3 \int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\,\mathrm{d}t \\ &\quad + \varepsilon \int_{-a}^{a} \delta^4 \left(\int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t \\ &= \delta^2\left(\frac{4}{3} + O(\delta^3)\right) \\ &\quad + 2\varepsilon \int_{-a}^{a} (u^\varepsilon)'(t)\delta^3 \int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\,\mathrm{d}t \\ &\quad + \varepsilon \int_{-a}^{a} \delta^4 \left(\int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t \\ &= \frac{4}{3}\delta^2\left(1 + O(\delta^3) + \vartheta(\delta,\varepsilon)\right) \end{aligned}$$

with

$$\begin{aligned} \vartheta(\delta,\varepsilon) := &\frac{3}{2}\varepsilon \int_{-a}^{a} (u^\varepsilon)'(t)\delta \int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\,\mathrm{d}t \\ &+ \frac{3}{4}\varepsilon \int_{-a}^{a} \delta^2 \left(\int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t \end{aligned}$$

Finally we have to show that $\vartheta(\delta,\varepsilon) = O\left(\frac{\sqrt{\delta}}{\varepsilon} + \frac{\delta^2}{\varepsilon^2}\right)$.

$$\begin{aligned}
|\vartheta(\delta,\varepsilon)| &\le \frac{3}{2}\varepsilon \int_{-a}^{a} |(u^\varepsilon)'(t)\delta \int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\,|\,\mathrm{d}t \\
&\quad + \frac{3}{4}\varepsilon\delta^2 \int_{-a}^{a} \left(\int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t \\
&\le \frac{3}{2}\varepsilon\delta \left(\int_{-a}^{a} \left((u^\varepsilon)'(t)\right)^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_{-a}^{a} \left(\int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t\right)^{\frac{1}{2}} \\
&\quad + \frac{3}{4}\varepsilon\delta^2 \int_{-a}^{a} \left(\int_0^1 (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t \\
&\le \frac{3}{2}\varepsilon\delta \left(\frac{C}{\varepsilon}\right)^{\frac{1}{2}} \left(\frac{C}{\delta\varepsilon^3}\right)^{\frac{1}{2}} + \frac{3}{4}\varepsilon\delta^2\frac{C}{\varepsilon^3} \\
&\le C\left(\frac{\sqrt{\delta}}{\varepsilon} + \frac{\delta^2}{\varepsilon^2}\right)
\end{aligned}$$

Here we have used the inequality $\int_{-a}^{a} (((u^\varepsilon)'(t))^2\,\mathrm{d}t \le C\varepsilon^{-1}$ from Corollary 2.6 and estimated

$$\begin{aligned}
\int_{-a}^{a} \left(\int_{-1}^{1} (s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}x\right)^2 \mathrm{d}t &\le \int_{-a}^{a}\int_{-1}^{1} (s-1)^2 \left((u^\varepsilon)''(t-s\delta)\right)^2 \mathrm{d}x\,\mathrm{d}t \\
&\le \int_{-a}^{a}\int_{-1}^{1} \left((u^\varepsilon)''(t-s\delta)\right)^2 \mathrm{d}x\,\mathrm{d}t \\
&= \int_{-a}^{a}\int_{t}^{t-\delta} \left((u^\varepsilon)''(\tilde{s})\right)^2 \left(-\frac{1}{\delta}\right) \mathrm{d}\tilde{s}\,\mathrm{d}t \\
&= \frac{1}{\delta}\int_{-a}^{a}\int_{t-\delta}^{t} \left((u^\varepsilon)''(\tilde{s})\right)^2 \mathrm{d}\tilde{s}\,\mathrm{d}t\ .
\end{aligned}$$

Using $(u^\varepsilon)'(x) = \frac{1}{\varepsilon}\sqrt{\Psi(u^\varepsilon(x))}$ from Lemma 2.2 we get

$$(u^\varepsilon)''(x) = \frac{\Psi'(u^\varepsilon(x))}{2\varepsilon^2}\,.$$

Consequently

$$\begin{aligned}\frac{1}{\delta}\int_{-a}^{a}\int_{t-\delta}^{t}\left((u^\varepsilon)''(\tilde{s})\right)^2\mathrm{d}\tilde{s}\,\mathrm{d}t &= \frac{1}{\delta}\int_{-a}^{a}\int_{u^\varepsilon(t-\delta)}^{u^\varepsilon(t)}\frac{1}{4\varepsilon^4}(-4u^\varepsilon)^2\left(1-(u^\varepsilon)^2\right)^2\frac{\varepsilon}{1-(u^\varepsilon)^2}\mathrm{d}u^\varepsilon\,\mathrm{d}t\\ &= \frac{4}{\delta\varepsilon^3}\int_{-a}^{a}\int_{u^\varepsilon(t-\delta)}^{u^\varepsilon(t)}(u^\varepsilon)^2\left(1-(u^\varepsilon)^2\right)^2\mathrm{d}u^\varepsilon\,\mathrm{d}t\\ &\le \frac{4}{\delta\varepsilon^3}\int_{-a}^{a}\int_{u^\varepsilon(t-\delta)}^{u^\varepsilon(t)}1\mathrm{d}u^\varepsilon\,\mathrm{d}t\\ &= \frac{4}{\delta\varepsilon^3}\int_{-a}^{a}\left(u^\varepsilon(t)-u^\varepsilon(t-\delta)\right)\mathrm{d}t\\ &\le \frac{4}{\delta\varepsilon^3}\int_{-a}^{a}2\,\mathrm{d}t\\ &\le \frac{16a}{\delta\varepsilon^3}\,.\end{aligned}$$

Since we use a fixed a we can find a constant C such that

$$\int_{-a}^{a}\left(\int_{-1}^{1}(s-1)(u^\varepsilon)''(t-s\delta)\,\mathrm{d}x\right)^2\mathrm{d}t \le \frac{C}{\delta\varepsilon^3}\,.$$

□

Lemma 3.5 is essential to prove the following theorem which tells us how to replace the first term of (3.3) in the phase field model of the natural time discretization of mean curvature motion.

Theorem 3.6. *For Γ being a $(d-1)$-dimensional regular hypersurface let $\tilde{u}^\varepsilon : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^d$, $d = 1, 2$ be a phase field function taking the optimal profile normal to this hypersurface Γ. Moreover let $\delta(x) = \varepsilon^{\beta(x)}$ with $\beta(x) > 2$ be some function on Γ and n the unit normal on Γ, where δ and n are assumed to be extended constantly in normal direction to Γ. Then*

$$\varepsilon\int_\Omega\left(u_\varepsilon(x)-u_\varepsilon(x-\delta(x)n(x))\right)^2\mathrm{d}x = \frac{4}{3}\int_\Gamma\delta^2(x)\mathrm{d}\mathcal{H}^{d-1}\left(1+O\left(\Sigma\left(\|\delta\|_\infty,\varepsilon\right)\right)\right) \quad (3.4)$$

with $\Sigma(\|\delta\|_\infty,\varepsilon) \to 0$ for $\varepsilon \to 0$.

Proof. We start to prove the case $d = 2$. Therefore we use a method presented in Section 2.1. For $d = 2$ the hypersurface Γ is a curve with arc length parametrization $\gamma(s)$, where s denotes the arc length parameter. Then we can write each point $x \in U_a(\Gamma)$ in an a-neighborhood of the curve Γ as $x(s,t) = \gamma(s) + tn(s)$ with $t \in \mathbb{R}$, $|t| < a$ and $n(s)$ being the unit normal vector to the curve γ in $\gamma(s)$. Now we proceed analogously as in Section 2.1 and use a coordinate transformation to rewrite the integral at the left hand side of (3.4). We choose

$$a = \max\left(-\frac{1}{2}\ln\left(\frac{1}{2}\right), \frac{3}{4}\max_{x\in\Omega}\beta(x)\right)(\varepsilon-\varepsilon\ln(\varepsilon))$$

such that $a \geq -\frac{\varepsilon}{2}\left(\ln\left(\frac{1}{2}\right) + \frac{3}{2}\ln(\delta(x))\right) = -\frac{\varepsilon}{2}\left(\ln\left(\frac{1}{2}\right) + \frac{3\beta(x)}{2}\ln(\varepsilon)\right)$ for all $x \in \Gamma$. Then

$$\begin{aligned}
&\varepsilon \int_{\Omega} \left(\tilde{u}^{\varepsilon}(x) - \tilde{u}^{\varepsilon}(x - \delta(x)n(x))\right)^2 \mathrm{d}x \\
&\overset{(2.6)}{=} \varepsilon \int_{\gamma^{-1}(\Gamma)} \int_{-a}^{a} \left(u^{\varepsilon}(t) - u^{\varepsilon}(t - \delta(x(s)))\right)^2 |1 + t\mathbf{h}| \,\mathrm{d}t \,\mathrm{d}s \\
&\quad + \varepsilon \int_{\Omega \setminus U_a(\Gamma)} \left(\tilde{u}^{\varepsilon}(x) - \tilde{u}^{\varepsilon}(x - \delta(x)n(x))\right)^2 \mathrm{d}x \; .
\end{aligned}$$

For the treatment of the second summand on the right hand side we use estimates known from the proof of Corollary 2.6 and get

$$\begin{aligned}
\varepsilon \int_{\Omega \setminus U_a(\Gamma)} \left(u_{\varepsilon}(x) - u_{\varepsilon}(x - \delta(x)n(x))\right)^2 \mathrm{d}x &\leq \varepsilon \int_{\Omega \setminus U_a(\Gamma)} \left(1 - \tanh\left(\frac{a}{\varepsilon}\right)\right)^2 \mathrm{d}x \\
&\leq \varepsilon \int_{\Omega \setminus U_a(\Gamma)} \left(2e^{-2\frac{a}{\varepsilon}}\right)^2 \mathrm{d}x \\
&\leq \varepsilon \int_{\Omega \setminus U_a(\Gamma)} \delta^3(x) \,\mathrm{d}x \\
&\leq \varepsilon |\Omega \setminus U_a(\Gamma)| \|\delta\|_{\infty}^3 \\
&\leq C\varepsilon \|\delta\|_{\infty}^3
\end{aligned}$$

and for the first one

$$\begin{aligned}
&\varepsilon \int_{\gamma^{-1}(\Gamma)} \int_{-a}^{a} \left(\tilde{u}_{\varepsilon}(t) - \tilde{u}_{\varepsilon}(t - \delta(x(s)))\right)^2 |1 + t\mathbf{h}| \,\mathrm{d}t \,\mathrm{d}s \\
&\leq \int_{\gamma^{-1}(\Gamma)} \varepsilon \int_{-a}^{a} \left(\tilde{u}_{\varepsilon}(t) - \tilde{u}_{\varepsilon}(t - \delta(x(s)))\right)^2 \mathrm{d}t \,\mathrm{d}s \; |1 + C(\varepsilon - \varepsilon\ln(\varepsilon))| \\
&\overset{\text{La}\,3.5}{\leq} \int_{\gamma^{-1}(\Gamma)} \frac{4}{3}\delta^2(x(s)) \left(1 + O(\delta^3(x(s))) + \vartheta(\delta(x(s)), \varepsilon)\right) \mathrm{d}s \, |1 + C(\varepsilon - \varepsilon\ln(\varepsilon))| \\
&\leq \int_{\Gamma} \frac{4}{3}\delta^2(x) \mathrm{d}\mathcal{H}^{d-1} \left(1 + O(\|\delta\|_{\infty}^3) + \vartheta(\|\delta\|_{\infty}, \varepsilon)\right) |1 + C(\varepsilon - \varepsilon\ln(\varepsilon))| \, .
\end{aligned}$$

Now we have to deal with $\left(1 + O(\|\delta\|_{\infty}^3) + \vartheta(\|\delta\|_{\infty}, \varepsilon)\right) |1 + C(\varepsilon - \varepsilon\ln(\varepsilon))|$. Therefore we define $\delta := \|\delta\|_{\infty}$. Moreover we use the convention that C denotes always a constant of the right value.

$$\begin{aligned}
&\left(1 + O(\delta^3) + \vartheta(\delta, \varepsilon)\right) |1 + C(\varepsilon - \varepsilon\ln(\varepsilon))| \\
&\overset{\text{proof of La}\,3.5}{\leq} \left(1 + C\delta^3 + C\left(\frac{\sqrt{\delta}}{\varepsilon} + \frac{\delta^2}{\varepsilon^2}\right)\right) \left(1 + C\left(\varepsilon - \varepsilon\ln(\varepsilon)\right)\right) \\
&\leq 1 + C\left(\frac{\sqrt{\delta}}{\varepsilon} + \frac{\delta^2}{\varepsilon^2} + \delta^3\right) + C\left(\varepsilon + \sqrt{\delta} + \frac{\delta^2}{\varepsilon} + \delta^3\varepsilon\right) - C\left(\varepsilon + \sqrt{\delta} + \frac{\delta^2}{\varepsilon} + \delta^3\varepsilon\right)\ln(\varepsilon) \\
&\leq 1 + C\left(\frac{\sqrt{\delta}}{\varepsilon} + \frac{\delta^2}{\varepsilon^2} + \delta^3 + \varepsilon\right) - C\left(\varepsilon + \sqrt{\delta} + \frac{\delta^2}{\varepsilon}\right)\ln(\varepsilon)
\end{aligned}$$

Thus we get

$$\varepsilon \int_{U_a(\Gamma)} \left(u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x))\right)^2 \mathrm{d}x = \frac{4}{3}\int_\Gamma \delta^2(x)\mathrm{d}\mathcal{H}^{d-1}\left(1 + O(\tilde\Sigma(\|\delta\|_\infty,\varepsilon))\right)$$

with $\tilde\Sigma(\|\delta\|_\infty,\varepsilon) = \left(\frac{\sqrt{\|\delta\|_\infty}}{\varepsilon} + \frac{\|\delta\|_\infty^2}{\varepsilon^2} + \|\delta\|_\infty^3 + \varepsilon\right) - \left(\varepsilon + \sqrt{\|\delta\|_\infty} + \frac{\|\delta\|_\infty^2}{\varepsilon}\right)\ln(\varepsilon)$.
As $\delta(x) = \varepsilon^{\beta(x)}$ with $\beta(x) > 2$, there exists one $\beta > 2$ such that $\|\delta\|_\infty = \varepsilon^\beta$ and therefore

$$\begin{aligned}\tilde\Sigma(\|\delta\|_\infty,\varepsilon) &= \varepsilon^{\frac{\beta}{2}-1} + \varepsilon^{2\beta-2} + \varepsilon^{3\beta} + \varepsilon - \left(\varepsilon + \varepsilon^{\frac{\beta}{2}} + \varepsilon^{2\beta-1}\right)\ln(\varepsilon)\\ &\le \begin{cases} \varepsilon - \varepsilon\ln(\varepsilon) & \text{for } \beta \ge 4\\ \varepsilon^{\frac{\beta}{2}-1} - \varepsilon\ln(\varepsilon) & \text{for } 2 < \beta < 4\end{cases} .\end{aligned}$$

In combination with

$$\frac{4}{3}\int_\Gamma \delta^2(x)\mathrm{d}\mathcal{H}^{d-1} = O(\varepsilon^{2\beta})$$

and

$$\varepsilon \int_{\Omega\setminus U_a(\Gamma)} \left(u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x))\right)^2 \mathrm{d}x = O(\varepsilon^{3\beta+1}) = O(\varepsilon^{2\beta})O(\varepsilon^{\beta+1})$$

we conclude there exists a function $\Sigma(\|\delta\|_\infty,\varepsilon)$ with $\Sigma(\|\delta\|_\infty,\varepsilon) \to 0$ for $\varepsilon \to 0$ such that

$$\varepsilon \int_\Omega \left(u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x))\right)^2 \mathrm{d}x = \frac{4}{3}\int_\Gamma \delta^2(x)\mathrm{d}\mathcal{H}^{d-1}\left(1 + O\left(\Sigma\left(\|\delta\|_\infty,\varepsilon\right)\right)\right) .$$

For $d = 3$ most parts of the proof can be done analogously to the case $d = 2$. Therefore we will focus on these parts which base on a different ansatz and do not detail those parts which can be done completely analogously to the two-dimensional case.
In contrast to the two-dimensional case, the hypersurface Γ is no curve, but a surface. Let this surface Γ be parametrized by $\gamma(s_1, s_1)$ and each parallel surface Γ_t in $U_a(\Gamma)$ will be parametrized by $\gamma_t(s_1, s_2) = \gamma(s_1, s_2) + tn(s_1, s_2)$, where $n(s_1, s_2)$ denotes a unit normal vector to the surface Γ in $\gamma(s_1, s_2)$. Now we split the integral at the left hand side of (3.4)

$$\begin{aligned}\varepsilon \int_\Omega \left(u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x))\right)^2 \mathrm{d}x &= \varepsilon \int_{-a}^{a}\int_{\Gamma_t} \left(u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x))\right)^2 \mathrm{d}\mathcal{H}^{d-1}\,\mathrm{d}t\\ &\quad + \varepsilon \int_{\Omega\setminus U_a(\Gamma)} \left(u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x))\right)^2 \mathrm{d}x .\end{aligned}$$

The integral over $\Omega \setminus U_a(\Gamma)$ will be treated as before, but the inner integral of the first summand will be rewritten as an integral over Γ. For a shorter notation we write

$s = (s_1, s_2)$.

$$\begin{aligned}
&\varepsilon \int_{-a}^{a} \int_{\Gamma_t} (u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x)))^2 \,\mathrm{d}\mathcal{H}^{d-1}\,\mathrm{d}t \\
&= \int_{-a}^{a} \int_{\gamma_t^{-1}(\Gamma_t)} (u_\varepsilon(x(s)) - u_\varepsilon(x(s) - \delta(x(s))n(x(s))))^2 \,\|\partial_{s_1}\gamma_t \times \partial_{s_2}\gamma_t\| \,\mathrm{d}s\,\mathrm{d}t \\
&= \int_{-a}^{a} \int_{\gamma^{-1}(\Gamma)} (u_\varepsilon(x(s)) - u_\varepsilon(x(s) - \delta(x(s))n(x(s))))^2 \,|1 - 2\mathbf{h}t + \mathbf{k}t^2| \|\partial_{s_1}\gamma \times \partial_{s_2}\gamma\| \,\mathrm{d}s\,\mathrm{d}t \\
&= \varepsilon \int_{-a}^{a} \int_{\Gamma} (u_\varepsilon(x) - u_\varepsilon(x - \delta(x)n(x)))^2 \,|1 - 2\mathbf{h}t + \mathbf{k}t^2| \mathrm{d}\mathcal{H}^{d-1}\,\mathrm{d}t
\end{aligned}$$

In the last two lines **k** denotes the Gaussian curvature and the step from line two to three can be found in [35]. The rest of the proof can be done analogously to the two-dimensional case.

□

Based on Theorem 2.4 and Theorem 3.6 we formulate the phase field model of the natural time discretization of mean curvature motion in Definition 3.4. Using Theorem 3.6 it is possible to write the phase field version of $\|x - x^k\|^2_{L^2(\Gamma^k)}$, the first term of (3.3) in the natural time discretization of mean curvature motion (MCM). Recall that the phase field function representing the interface $x(\Gamma^k)$ is denoted by u. Moreover u^k is the phase field function representing the interface Γ^k. We assume that both phase field functions take the optimal profile perpendicular to the interface, thus we can apply Theorem 3.6. Now u^k takes the role of $u_\varepsilon(x)$ and u the one of $u_\varepsilon(x - \delta(x)n(x))$. As x^k is the parametrization of Γ^k and x the parametrization of $x(\Gamma^k)$ the distance function $\delta = x - x^k$. Thus we get

$$\varepsilon \|u - u^k\|^2_{L^2(\Omega)} = \frac{4}{3} \|x - x^k\|^2_{L^2(\Gamma^k)} \left(1 + O(\|x - x^k\|_{L^\infty(\Gamma^k)}, \varepsilon)\right) .$$

As $O(\|x - x^k\|_{L^\infty(\Gamma^k)}, \varepsilon)$ converges to zero for $\varepsilon \to 0$ we can approximate the first term of (3.3) by $\frac{3}{4}\varepsilon\|u - u^k\|^2_{L^2(\Omega)}$. Using Theorem 2.4 both terms in the phase field version of (3.3) scale with $\frac{3}{4}$ and therefore we can neglect the scaling factor.

Definition 3.7 (Phase field version of the natural time discretization of MCM). *Given a phase field function u^k at time $k\tau$, define the phase field function u^{k+1} at time $(k+1)\tau$ by*

$$u^{k+1} := \underset{u \in H^1(\Omega)}{\operatorname{argmin}} \, \varepsilon \|u - u^k\|^2_{L^2(\Omega)} + 2\tau a^\varepsilon[u] \, .$$

Although we define u^{k+1} as minimizing argument of this special energy notice that it is not necessarily unique. We expect it to be unique only for small τ. Since this notation will be used frequently in this work we introduce the convention that a variable which is defined as minimizing argument of an energy does not have to be uniquely defined.

3.1.2 Natural time discretization of Willmore flow

After the derivation of the phase field version of the natural time discretization of mean curvature flow we deal with the nested time discretization of Willmore flow. We start with the parametric version and transfer it to the phase field set up.
As before, let the hypersurface Γ be contained in a computational domain $\Omega \subset \mathbb{R}^d$ and $x : \Gamma \to \Gamma$ be the identity on Γ. Then, the associated Willmore energy is $w[x] = \frac{1}{2}\int_\Gamma \mathbf{h}^2(x)\,\mathrm{d}\mathcal{H}^{d-1}$, where $\mathbf{h}$ denotes the mean curvature. The L^2-gradient flow, which minimizes the Willmore energy is called Willmore flow and the weak version of the gradient flow equation is

$$(\partial_t x, \vartheta)_{L^2(\Gamma)} = -\partial_x w[x](\vartheta)$$

for all test functions $\vartheta \in C^\infty(\Gamma, \mathbb{R}^d)$. In this equation $(\cdot,\cdot)_{L^2(\Gamma)}$ denotes the L^2-metric on the interface Γ and $\partial_x w[x](\vartheta)$ the variation of the Willmore energy in a direction ϑ. The hypersurface which one gets under Willmore flow at time t is denoted by $\Gamma(t)$ and $x(t)$ is the identity mapping on $\Gamma(t)$. Following the proceeding of Olischläger and Rumpf in [81, 82] one approximated the mean curvature by the time-discrete propagation speed of mean curvature motion. This ansatz is reasonable because the speed with which an interface evolves at some point under mean curvature motion equals the mean curvature of the interface at this point. Thus one introduces a variational problem in which one has to minimize the energy

$$e_{\mathrm{in}}[x,y] = \|y-x\|^2_{L^2(\Gamma)} + 2\tilde{\tau}\mathcal{H}^{d-1}[y(\Gamma)]\,, \tag{3.5}$$

which one knows from Definition 3.4. Here, the parametrization $x : \Gamma \to \Gamma$ of an interface Γ over itself is given and the parametrization $y : \Gamma \to \mathbb{R}^d$ of a hypersurface $y(\Gamma)$ denotes the parametrization which minimizes the inner energy (3.5). Thus it is the approximate solution of one step of mean curvature motion with time step size $\tilde{\tau}$ and therefore we can approximate the mean curvature $\mathbf{h}$ by $\frac{|y-x|}{\tilde{\tau}}$ and the Willmore energy $w[x] \approx \frac{1}{2}\int_\Gamma \frac{(y-x)^2}{\tilde{\tau}^2}\,\mathrm{d}\mathcal{H}^{d-1}$. Using Definition 3.2 of the natural time discretization we introduce the outer energy

$$e_{\mathrm{out}}[x^k, x, y] = \|x-x^k\|^2_{L^2(\Gamma^k)} + \frac{\tau}{\tilde{\tau}^2}\|y-x\|_{L^2(\Gamma)}\,. \tag{3.6}$$

Here, x^k is the identity on Γ^k and with a slight misuse of notation we write y instead of $y \circ x$. Starting with an approximation x^k to $x(k\tau)$ we compute the solution x^{k+1} at time $(k+1)\tau$ under Willmore flow by minimizing the outer energy (3.6) under the constraint that y is the minimizer of $e_{\mathrm{in}}[x,\cdot]$.

Definition 3.8 (Parametric nested variational time discretization of Willmore flow). *Given a parametrization x^k of a hypersurface Γ^k at time $k\tau$, define the parametrization x^{k+1} at time $(k+1)\tau$ by*

$$x^{k+1} = \underset{x \in \{\tilde{x}:\Gamma^k \to \mathbb{R}^d\}}{\operatorname{argmin}} e_{out}[x^k, x, y[x]],\ \text{where} \tag{3.7a}$$

$$y[x] = \underset{\tilde{y} \in \{\tilde{\tilde{y}}:x(\Gamma^k)\to\mathbb{R}^d\}}{\operatorname{argmin}} e_{in}[x,\tilde{y}]\,. \tag{3.7b}$$

Furthermore we denote $y[x^{k+1}]$ by y^{k+1}.

The transfer to the phase field version can be done straightforward. The phase field version of the inner energy has been treated in the previous subsection and the outer energy consists of two L^2- norms on interfaces Γ and Γ^k. We have seen how to deal with terms like this in the previous section, too.
Let u^k be the phase field function approximating the interface Γ^k, while u and v are the phase field approximations of $x(\Gamma^k)$ and $y(x(\Gamma^k))$. Assuming all phase field functions take the optimal profile perpendicular to the interface the outer and inner energy in the phase field version are

$$e^{\varepsilon}_{\mathrm{out}}[u^k, u, v] = \varepsilon \|u - u^k\|^2_{L^2(\Omega)} + \frac{\tau\varepsilon}{\tilde{\tau}^2} \|v - u\|^2_{L^2(\Omega)} \,, \tag{3.8a}$$

$$e^{\varepsilon}_{\mathrm{in}}[u, v] = \varepsilon \|v - u\|^2_{L^2(\Omega)} + 2\tilde{\tau} a^{\varepsilon}[v] \,. \tag{3.8b}$$

Here we used Theorem 3.6 to get

$$\begin{aligned} \varepsilon \|u - u^k\|^2_{L^2(\Omega)} &= \frac{4}{3} \|x - x^k\|_{L^2(\Gamma^k)} \left(1 + O(\|x - x^k\|_{L^\infty(\Gamma^k)}, \varepsilon)\right) \,, \\ \varepsilon \|v - u\|^2_{L^2(\Omega)} &= \frac{4}{3} \|y - x\|_{L^2(\Gamma)} \left(1 + O(\|y - x\|_{L^\infty(\Gamma)}, \varepsilon)\right) \end{aligned}$$

and Theorem 2.4 which states

$$\Gamma - \lim_{\varepsilon \to 0} a^{\varepsilon}[v] = \frac{4}{3} \mathcal{H}^{d-1} \left(y[\Gamma]\right) \,.$$

Consequently $e^{\varepsilon}_{\mathrm{out}}[u^k, u, v]$ is the phase field version of the outer energy $e_{\mathrm{out}}[x^k, x, y]$ (3.6) and $e^{\varepsilon}_{\mathrm{in}}[u, v]$ the phase field version of the inner energy $e_{\mathrm{in}}[x, y]$ (3.5). Thus for sufficiently small phase field parameter ε and assuming τ, $\tilde{\tau}$ to be small enough such that the local distance δ between the evolving hypersurfaces at two consecutive time steps of mean curvature motion and of Willmore flow is small enough in comparison to ε the nested variational time discretization of Willmore flow is

Definition 3.9 (Nested variational time discretization of Willmore flow)**.** *Given a phase field u^k at time $k\tau$ define the phase field u^{k+1} at time $(k+1)\tau$ by*

$$u^{k+1} = \operatorname*{argmin}_{u \in L^2(\Omega)} e^{\varepsilon}_{out}[u^k, u, v[u]] \,, \text{ where} \tag{3.9a}$$

$$v[u] = \operatorname*{argmin}_{\tilde{v} \in H^1(\Omega)} e^{\varepsilon}_{in}[u, \tilde{v}] \,. \tag{3.9b}$$

Furthermore, we denote $v[u^{k+1}]$ by v^{k+1}.

Notice, now we write v instead of $v[u]$ although v depends on u. An alternative way to formulate the constraint (3.9b) is to write it as Euler–Lagrange equation for (3.8b), that is, v has to solve

$$\int_\Omega \varepsilon(v - u)\vartheta + \frac{\tilde{\tau}}{2\varepsilon} \Psi'(v)\vartheta + \varepsilon\tilde{\tau} \nabla v \cdot \nabla\vartheta \,\mathrm{d}x = 0 \quad \forall \vartheta \in H^1(\Omega) \,.$$

The following theorem states an existence result for this time-discrete Willmore flow model:

Theorem 3.10 (Existence of a time-discrete phase field solution)**.** *Let Ω be a bounded set in $\mathbb{R}^d$ with Lipschitz boundary and suppose $u^0 \in L^2(\Omega)$. Then there exists a sequence $((u^k, v^k))_{k=1,\dots} \subset L^2(\Omega) \times H^1(\Omega)$ of solutions of* (3.9a) *and* (3.9b)*.*

For the proof of this theorem we need one theorem (cf. [61]):

Theorem 3.11. *Let V be a convex subset of a separable reflexive Banach space, $F : V \to \overline{\mathbb{R}}$ convex and lower semicontinuous. Then F is also lower semicontinuous with respect to weak convergence.*

and another theorem (cf. [49]):

Theorem 3.12 (Weak lower semicontinuity)**.** *Assume that L is smooth, bounded below, and in addition the mapping $p \mapsto L(Dw, w, x)$ is convex, for each $w \in \mathbb{R}$, $x \in U$. Moreover there exist constants $\alpha > 0$, $\beta \geq 0$ such that*

$$L(Dw, w, x) \geq \alpha |Dw|^q - \beta \quad \forall Dw \in \mathbb{R}^n,\ w \in \mathbb{R},\ x \in \Omega\,.$$

Then $I[\cdot]$ is weakly lower semicontinuous on $W^{1,q}(U)$.

$$I[w] := \int_U L(Dw(x), w(x), x)\,\mathrm{d}x$$

Now let us prove Theorem 3.10.

Proof. We start considering the variational problem (3.9b) for given $u^k \in L^2(\Omega)$ and regard a minimizing sequence $(v_j^k)_{j=1,\dots}$. Then we estimate $e_{\text{in}}^{\varepsilon}[u^k, v_j^k]$

$$\begin{aligned}
e_{\text{in}}^{\varepsilon}[u^k, v_j^k] &= \varepsilon \int_\Omega \left(v_j^k - u^k\right)^2 \mathrm{d}x + \tau \int_\Omega \varepsilon |\nabla v_j^k|^2 + \frac{1}{\varepsilon}\Psi(v_j^k(x))\,\mathrm{d}x \\
&\geq \varepsilon \int_\Omega ||v_j^k| - |u^k||^2\,\mathrm{d}x + \tau\varepsilon \int_\Omega |\nabla v_j^k|^2\,\mathrm{d}x \\
&\overset{\text{Young}}{\geq} \varepsilon \left(1 - \frac{1}{\delta}\right) \|v_j^k\|_{L^2(\Omega)}^2 + \varepsilon\,(1-\delta)\,\|u^k\|_{L^2(\Omega)}^2 + \tau\varepsilon \int_\Omega |\nabla v_j^k|^2\,\mathrm{d}x \\
&\geq \varepsilon \min((1 - \frac{1}{\delta}), \tau) \|v_j^k\|_{H^1(\Omega)}^2 + \varepsilon\,(1-\delta)\,\|u^k\|_{L^2(\Omega)}^2\,,
\end{aligned}$$

where we have used Young's inequality $2ab \leq \delta a^2 + \frac{1}{\delta} b^2$ for $a, b > 0$ and $\delta > 1$. From this it follows

$$\varepsilon \min((1 - \frac{1}{\delta}), \tau) \|v_j^k\|_{H^1(\Omega)}^2 \leq e_{\text{in}}^{\varepsilon}[u^k, v_j^k] + \varepsilon(\delta - 1)\|u^k\|_{L^2(\Omega)}^2\,.$$

We know $(v_j^k)_{j=1,\dots}$ is a minimizing sequence, thus there exists $N \in \mathbb{N}^{>0}$ such that $e_{\text{in}}^{\varepsilon}[u^k, v_j^k]$ is uniformly bounded (cf. [60]) for $j > N$. Moreover $u^k \in L^2(\Omega)$ and therefore we get the uniform boundedness of v_j^k in $H^1(\Omega)$. By the uniform boundedness of v_j^k in $H^1(\Omega)$ and by the fact that $H^1(\Omega)$ is a reflexive Banach space we extract a subsequence, again denoted by $(v_j^k)_{j=1,\dots}$ such that v_j^k converges weakly in $H^1(\Omega)$ to some v^k. It remains to show that $e_{\text{in}}^{\varepsilon}[u_j^k, \cdot]$ is weakly lower semicontinuous on $H^1(\Omega)$. This follows with Theorem 3.12. Thus we get by the direct method in the calculus of variations the existence of a minimizer $v^k = v^k[u^k] \in H^1(\Omega)$.

This allows us to consider for fixed $k > 0$ and given $u^{k-1} \in L^2(\Omega)$ a minimizing sequence $(u_j^k, v_j^k)_{j=1,\dots}$ with v_j^k being a minimizer of $e_{\text{in}}^{\varepsilon}[u_j^k, \cdot]$. The outer energy $e_{\text{out}}^{\varepsilon}$ can be estimated

$$\begin{aligned}
e_{\text{out}}^{\varepsilon}[u^{k-1}, u_j^k, v_j^k] &= \varepsilon \|u_j^k - u^{k-1}\|_{L^2(\Omega)}^2 + \frac{\tau\varepsilon}{\tilde{\tau}^2}\|v_j^k - u_j^k\|_{L^2(\Omega)}^2 \\
&\geq \varepsilon \int_\Omega |u_j^k - u^{k-1}|^2 \,\mathrm{d}x \\
&\geq \varepsilon \int_\Omega ||u_j^k| - |u^{k-1}||^2 \,\mathrm{d}x \\
&= \varepsilon \int_\Omega |u_j^k|^2 + |u^{k-1}|^2 - 2|u_j^k|\,|u^{k-1}| \,\mathrm{d}x \\
&\geq \varepsilon \|u_j^k\|_{L^2(\Omega)}^2 + \varepsilon \|u^{k-1}\|_{L^2(\Omega)}^2 - 2\varepsilon \int_\Omega \left(\frac{1}{2\delta}|u_j^k|^2 + \frac{\delta}{2}|u^{k-1}|^2 \right) \mathrm{d}x \\
&= \varepsilon \left(1 - \frac{1}{\delta}\right) \|u_j^k\|_{L^2(\Omega)}^2 + \varepsilon \left(1 - \delta\right) \|u^{k-1}\|_{L^2(\Omega)}^2 \,,
\end{aligned}$$

where we used Young's inequality with $\delta > 1$. From this it follows

$$\varepsilon(1 - \frac{1}{\delta})\|u_j^k\|_{L^2(\Omega)}^2 \leq e_{\text{out}}^{\varepsilon}[u^{k-1}, u_j^k, v_j^k] + \varepsilon(\delta - 1)\|u^{k-1}\|_{L^2(\Omega)}^2 \,.$$

We know that $(u_j^k, v_j^k)_{j=1,\dots}$ is a minimizing sequence, thus there exists $N \in \mathbb{N}^{>0}$ such that $e_{\text{out}}^{\varepsilon}[u^{k-1}, u_j^k, v_j^k]$ is uniformly bounded for $j > N$. Moreover $u^{k-1} \in L^2(\Omega)$ and therefore we get the uniform boundedness of u_j^k in $L^2(\Omega)$. In the same way as before we estimate $e_{\text{in}}^{\varepsilon}[u_j^k, v_j^k]$ and get the uniform boundedness of v_j^k in $H^1(\Omega)$, because u_j^k is uniformly bounded in $L^2(\Omega)$ and v_j^k is a minimizer of $e_{\text{in}}^{\varepsilon}[u_j^k, \cdot]$.
By the uniform boundedness of u_j^k and v_j^k in $L^2(\Omega)$ and $H^1(\Omega)$ and by the fact that $L^2(\Omega)$ and $H^1(\Omega)$ are reflexive Banach spaces we extract a subsequence, again denoted by $(u_j^k, v_j^k)_{j=1,\dots,}$, such that u_j^k converges weakly in $L^2(\Omega)$ to some u^k and v_j^k converges weakly in $H^1(\Omega)$ to some v^k. Furthermore we show that $e_{\text{out}}^{\varepsilon}[u^{k-1}, \cdot, \cdot]$ in (3.9a) is weakly lower semicontinuous, i. e. $\liminf_{j\to\infty} e_{\text{out}}^{\varepsilon}[u^{k-1}, u_j^k, v_j^k] \geq e_{\text{out}}^{\varepsilon}[u^{k-1}, u^k, v^k]$.
Therefore we define $\tilde{e}_{\text{out}}^{\varepsilon}[u^{k-1}, \tilde{u}, \tilde{v}] := \varepsilon\left(\tilde{u} - u^{k-1}\right)^2 + \frac{\varepsilon\tau}{\tilde{\tau}^2}\left(\tilde{v} - \tilde{u}\right)^2$ and use Fatou's Lemma

$$\begin{aligned}
\liminf_{j\to\infty} e_{\text{out}}^{\varepsilon}[u^{k-1}, \tilde{u}_j, \tilde{v}_j] &\geq \int_\Omega \liminf_{j\to\infty} \tilde{e}_{\text{out}}^{\varepsilon}[u^{k-1}, \tilde{u}_j, \tilde{v}_j] \,\mathrm{d}x \\
&= \int_\Omega \tilde{e}_{\text{out}}^{\varepsilon}[u^{k-1}, u, v] \,\mathrm{d}x
\end{aligned}$$

for $\tilde{v}_j \to v$ and $\tilde{u}_j \to u$ in $L^2(\Omega)$. As $e_{\text{out}}^{\varepsilon}[u^k, \cdot, \cdot]$ is convex we use Theorem 3.11 and get that $e_{\text{out}}^{\varepsilon}[u^{k-1}, \cdot, \cdot]$ is weakly lower semicontinuous.
It remains to prove that $e_{\text{in}}^{\varepsilon}[u^k, v^k] = \min_{v\in H^1(\Omega)} e_{\text{in}}^{\varepsilon}[u^k, v]$. Let us assume that there is a

v^* with $e^\varepsilon_{\text{in}}[u^k, v^k] - e^\varepsilon_{\text{in}}[u^k, v^*] > 0$, then we obtain

$$\begin{aligned}
0 < e^\varepsilon_{\text{in}}[u^k, v^k] - e^\varepsilon_{\text{in}}[u^k, v^*] &= \varepsilon \int_\Omega (u^k - v^k)^2 - (u^k - v^*)^2 \,\mathrm{d}x + 2\tilde{\tau}(a^\varepsilon[v^k] - a^\varepsilon[v^*]) \\
&= \varepsilon \int_\Omega 2u^k(v^* - v^k) + (v^k)^2 - (v^*)^2 \,\mathrm{d}x + 2\tilde{\tau}(a^\varepsilon[v^k] - a^\varepsilon[v^*]) \\
&\leq \liminf_{j\to\infty} \varepsilon \int_\Omega 2u^k_j(v^* - v^k_j) + (v^k_j)^2 - (v^*)^2 \,\mathrm{d}x + 2\tilde{\tau}(a^\varepsilon[v^k_j] - a^\varepsilon[v^*]) \\
&= \liminf_{j\to\infty} \left(e^\varepsilon_{\text{in}}[u^k_j, v^k_j] - e^\varepsilon_{\text{in}}[u^k_j, v^*] \right) .
\end{aligned}$$

In the third line we have used the lower semi-continuity of $a^\varepsilon[\cdot]$, the weak convergence of v^k_j to v^k in $H^1(\Omega)$ and the weak convergence of u^k_j to u^k in $L^2(\Omega)$, and that by Rellich's embedding theorem v^k_j already converges strongly in $L^2(\Omega)$, which implies $v^k_j u^k_j \rightharpoonup v^k u^k$ weakly in $L^1(\Omega)$. Hence, there exists an index j with $e^\varepsilon_{\text{in}}[u^k_j, v^k_j] - e^\varepsilon_{\text{in}}[u^k_j, v^*] > 0$, contradicting our assumption $e^\varepsilon_{\text{in}}[u^k_j, v^k_j] = \min_{v\in H^1(\Omega)} e^\varepsilon_{\text{in}}[u^k_j, v]$. Thus, $e^\varepsilon_{\text{in}}[u^k, v^k] = \min_{v\in H^1(\Omega)} e^\varepsilon_{\text{in}}[u^k, v]$ and (u^k, v^k) solves the variational problem (3.9a) and (3.9b). □

Now we have shown the existence of a solution of the variational problem given by (3.9a) and (3.9b). However remark that due to the non-convexity of Ψ the solutions of (3.9a) and (3.9b) are not necessarily unique.

3.1.3 The constrained optimization perspective

Computing one time step of Willmore flow with the nested variational time discrete Willmore model presented in Definition 3.9 involves the solution of a variational problem with a PDE constraint. This is equivalent to the problem of finding a saddle point of the corresponding Lagrangian ℓ (cf. [78]), which is defined as the sum of the outer energy functional (3.8a) and the variation of the inner energy functional (3.8b) in the direction of p. This p is a dual function and it takes the role of the Lagrange multiplier. Thus the Lagrangian is

$$\begin{aligned}
\ell[u^k, u, v, p] &= e^\varepsilon_{\text{out}}[u^k, u, v] + \partial_v e^\varepsilon_{\text{in}}[u, v](p) \\
&= \int_\Omega \varepsilon(u - u^k)^2 + \frac{\varepsilon\tau}{\tilde{\tau}^2}(v - u)^2 \,\mathrm{d}x + \int_\Omega 2\varepsilon(v - u)p + \frac{\tilde{\tau}}{\varepsilon}\Psi'(v)p + 2\varepsilon\tilde{\tau}\nabla v \cdot \nabla p \,\mathrm{d}x \; .
\end{aligned} \tag{3.10}$$

To find a saddle point $(u, v, p) \in L^2(\Omega) \times H^1(\Omega) \times H^1(\Omega)$ of this Lagrangian ℓ we have to solve the equation

$$\nabla_{(u,v,p)} \ell[u^k, u, v, p] = 0 \,. \tag{3.11}$$

The three components of $\nabla_{(u,v,p)}\ell[u^k,u,v,p]$ are

$$\begin{aligned}
\partial_u\ell[u^k,u,v,p](\vartheta) &= \int_\Omega 2\varepsilon(u-u^k)\vartheta - \frac{2\tau\varepsilon}{\tilde\tau^2}(v-u)\vartheta - 2\varepsilon\vartheta p \,\mathrm{d}x\,,\\
\partial_v\ell[u^k,u,v,p](\xi) &= \int_\Omega \frac{2\tau\varepsilon}{\tilde\tau^2}(v-u)\xi + 2\varepsilon p\xi + \frac{\tilde\tau}{\varepsilon}\Psi''(v)p\xi + 2\varepsilon\tilde\tau\nabla p\cdot\nabla\xi \,\mathrm{d}x\,,\\
\partial_p\ell[u^k,u,v,p](\varsigma) &= \int_\Omega 2\varepsilon\,(v-u)\,\varsigma + \frac{\tilde\tau}{\varepsilon}\Psi'(v)\varsigma + 2\varepsilon\tilde\tau\nabla v\cdot\nabla\varsigma \,\mathrm{d}x
\end{aligned}$$

with test functions $\vartheta\in L^2(\Omega)$, $\xi\in H^1(\Omega)$, and $\varsigma\in H^1(\Omega)$.
In order to understand the meaning of the three components of the equation (3.11) we present another ansatz to solve the variational problem presented in Definition 3.9. We have to keep in mind that the variable v in the outer energy in (3.9a) depends on u, thus $v=v[u]$. Hence, the variation of the outer energy $e^\varepsilon_{\mathrm{out}}$ in direction u with test function $\vartheta\in L^2(\Omega)$ reads

$$\partial_u\left(e^\varepsilon_{\mathrm{out}}[u^k,u,v[u]]\right)(\vartheta) = \left(\partial_u e^\varepsilon_{\mathrm{out}}\right)[u^k,u,v[u]](\vartheta) + \left(\partial_v e^\varepsilon_{\mathrm{out}}\right)[u^k,u,v[u]]\left(\partial_u v[u](\vartheta)\right)\,, \tag{3.12}$$

where the first summand is

$$\left(\partial_u e^\varepsilon_{\mathrm{out}}\right)[u^k,u,v[u]](\vartheta) = \int_\Omega 2\varepsilon(u-u^k)\vartheta - \frac{2\tau\varepsilon}{\tilde\tau^2}(v-u)\vartheta\,\mathrm{d}x\ .$$

To compute $\partial_u v[u]$, we start by noting that

$$\begin{aligned}
0 &= \partial_v e^\varepsilon_{\mathrm{in}}[v[u],u](\xi)\\
&= \int_\Omega 2\varepsilon\,(v[u]-u)\,\xi + \frac{\tilde\tau}{\varepsilon}\Psi'(v[u])\xi + 2\varepsilon\tilde\tau\nabla v[u]\cdot\nabla\xi\,\mathrm{d}x
\end{aligned}\tag{3.13}$$

for $\varsigma\in H^1(\Omega)$. The variation of the Euler–Lagrange equation in direction u is

$$\begin{aligned}
0 &= \partial_u\left(\partial_v e^\varepsilon_{\mathrm{in}}[v[u],u](\xi)\right)(\varrho)\\
&= \partial_v^2 e^\varepsilon_{\mathrm{in}}[v[u],u]\left(\xi,\partial_u v[u](\varrho)\right) + \partial_u\partial_v e^\varepsilon_{\mathrm{in}}[v[u],u](\xi,\varrho)
\end{aligned}\tag{3.14}$$

Following Hinze et al. [57] and Haslinger and Mäkinen [55] the Lagrangian dual problem consists of finding $p\in H^1(\Omega)$ such that

$$\partial_v^2 e^\varepsilon_{\mathrm{in}}[v[u],u](\xi,p) = -\partial_v e^\varepsilon_{\mathrm{out}}[u^k,u,v[u]](\xi) \tag{3.15}$$

$$\Leftrightarrow \int_\Omega 2\varepsilon p\xi + \frac{\tilde\tau}{\varepsilon}\Psi''(v[u])p\xi + 2\varepsilon\tilde\tau\nabla p\cdot\nabla\xi\,\mathrm{d}x = -\int_\Omega \frac{2\tau\varepsilon}{\tilde\tau^2}(v-u)\xi\,\mathrm{d}x$$

for $\xi\in H^1(\Omega)$. Using (3.15) and (3.14) we compute the second summand of (3.12)

$$\begin{aligned}
\left(\partial_v e^\varepsilon_{\mathrm{out}}\right)[u^k,u,v[u]](\partial_u v[u](\vartheta)) &\overset{(3.15)}{=} -\partial_v^2 e^\varepsilon_{\mathrm{in}}[v[u],u]\left(\partial_u v[u](\vartheta),p\right)\\
&= -\partial_v^2 e^\varepsilon_{\mathrm{in}}[v[u],u](p,\partial_u v[u](\vartheta))\\
&\overset{(3.14)}{=} \partial_u\partial_v e^\varepsilon_{\mathrm{in}}[v[u],u](p,\vartheta)\\
&= -2\varepsilon\int_\Omega \vartheta p\,\mathrm{d}x\ .
\end{aligned}$$

Consequently (3.12) is

$$\partial_u e^\varepsilon_{\text{out}}[u^k, u, v[u]](\vartheta) = \int_\Omega 2\varepsilon(u - u^k)\vartheta - \frac{2\tau\varepsilon}{\tilde{\tau}^2}(v - u)\vartheta - 2\varepsilon\vartheta p \,\mathrm{d}x \;. \tag{3.16}$$

Comparing this with the components of $\nabla_{(u,v,p)}\ell[u^k, u, v, p] = 0$ reveals $\partial_u \ell = 0$ is the actual Lagrangian multiplier formulation of the nested variational problem and corresponds to (3.16). The second component $\partial_v \ell = 0$ is the dual problem (3.15) defining a dual variable p and $\partial_p \ell = 0$ reflects the Euler–Lagrange equation of the inner variational problem (3.13).
To solve the Euler–Lagrange equation corresponding to ℓ in (3.10) we use a Newton scheme, because $\Psi(v) = (1 - v^2)^2$ and therefore the variation of the Lagrangian in v and p is nonlinear in v.

3.2 A fully practical numerical scheme

To formulate a fully practical numerical scheme we have to discretize our problem in space. After this we concentrate on the question of solving our nested minimization problem. Therefore, we introduce a Newton method for finding a saddle point of the discrete Lagrangian, corresponding to our problem.

3.2.1 Finite Element discretization in space

For the spatial discretization of our problem we regard a regular and uniform simplicial mesh $\mathcal{T}$ covering the computational domain Ω. As described in Section 2.2 we define the Finite Element space of continuous, piecewise affine functions

$$\mathcal{V} := \{\Phi \in C^0(\Omega) \,:\, \Phi|_T \in \mathcal{P}_1 \,\forall\, T \in \mathcal{T}\}$$

on this mesh. With the notation introduced in Section 2.2, i.e. $\mathbf{M}$ being the lumped mass matrix, the discrete version of the outer and inner energy $e^\varepsilon_{\text{out}}[\cdot]$ in (3.8a) and $e^\varepsilon_{\text{in}}[\cdot]$ in (3.8b) is

$$\mathcal{E}^\varepsilon_{\text{out}}[U^k, U, V] = \varepsilon\mathbf{M}\left(\overline{U} - \overline{U^k}\right) \cdot \left(\overline{U} - \overline{U^k}\right) + \frac{\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right) \;, \tag{3.17a}$$

$$\mathcal{E}^\varepsilon_{\text{in}}[U, V] = \varepsilon\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right) + 2\tilde{\tau}A^\varepsilon[V] \,. \tag{3.17b}$$

In the discrete inner energy $\mathcal{E}^\varepsilon_{\text{in}}$ the discrete phase field energy A^ε is

$$A^\varepsilon[V] = \frac{1}{2}\int_\Omega \varepsilon\nabla V \cdot \nabla V + \frac{1}{\varepsilon}\mathcal{I}_h\left(\Psi(V)\right)\mathrm{d}x = \frac{\varepsilon}{2}\mathbf{L}\overline{V} \cdot \overline{V} + \frac{1}{2\varepsilon}\mathbf{M}\overline{\Psi(V)} \cdot \overline{1} \,,$$

where $\mathcal{I}_h$ again denotes the Lagrange interpolation. This leads to the problem of solving the following fully discrete minimization problem in each time step.

Definition 3.13 (Fully discrete variational time discretization of Willmore flow)**.** *Given a discrete phase field $U^k \in \mathcal{V}$ at time $k\tau$ define the phase field $U^{k+1} \in \mathcal{V}$ at time $(k + 1)\tau$ by*

$$U^{k+1} = \operatorname*{argmin}_{U \in \mathcal{V}} \mathcal{E}^\varepsilon_{out}[U^k, U, V] \,, \text{ where} \tag{3.18a}$$

$$V = \operatorname*{argmin}_{\tilde{V} \in \mathcal{V}} \mathcal{E}^\varepsilon_{in}[\tilde{V}, U] \,, \tag{3.18b}$$

and denote by V^{k+1} *the solution* V *of* (3.18b) *for* $U = U^{k+1}$.

As $\mathcal{V}$ is a finite dimensional Euclidean space it is isomorphic to $\mathbb{R}^{\sharp I}$, where $\sharp I$ denotes the cardinality of the set I. Thus we identify each discrete function $U \in \mathcal{V}$ with its nodal coordinate vector $\overline{U}$ and get the following statement on the existence of a solution of the fully discretized Willmore flow model.

Theorem 3.14 (Existence of a fully discrete phase field solution)**.** *Let* Ω *be a bounded set in* $\mathbb{R}^d$*, covered with a regular and uniform simplicial mesh* $\mathcal{T}$ *and* I *being the vertex index set of* $\mathcal{T}$*. Suppose* $\overline{U}^0 \in \mathbb{R}^{\sharp I}$*. Then there exists a sequence* $((\overline{U}^k, \overline{V}^k))_{k=1,\ldots} \subset \mathbb{R}^{\sharp I} \times \mathbb{R}^{\sharp I}$ *of solutions of (3.18a) and (3.18b).*

Proof. We start considering the fully discrete variational problem (3.18b) for given $\overline{U}^k \in \mathbb{R}^{\sharp I}$ and regard a minimizing sequence $(\overline{V}^k_j)_{j=1,\ldots}$. Then we estimate $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^k_j]$

$$\begin{aligned}
\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^k_j] &= \varepsilon \mathbf{M} \left(\overline{V}^k_j - \overline{U}^k\right) \cdot \left(\overline{V}^k_j - \overline{U}^k\right) + 2\tilde{\tau} A^\varepsilon[\overline{V}^k_j] \\
&\geq \varepsilon \mathbf{M} \left(\overline{V}^k_j - \overline{U}^k\right) \cdot \left(\overline{V}^k_j - \overline{U}^k\right) \\
&\geq \varepsilon \lambda_{\min} \|\overline{V}^k_j - \overline{U}^k\|^2 \\
&\geq \varepsilon \lambda_{\min} |\|\overline{V}^k_j\| - \|\overline{U}^k\||^2 \\
&\geq \varepsilon \lambda_{\min} \left(\left(1 - \frac{1}{\delta}\right) \|\overline{V}^k_j\|^2 + (1-\delta) \|\overline{U}^k\|^2\right),
\end{aligned}$$

where $\lambda_{\min}$ denotes the smallest eigenvalue of $\mathbf{M}$ and we have used Young's inequality for $\delta > 1$. As $\|\overline{U}^k\| < \infty$ and $(\overline{V}^k_j)_{j=1,\ldots}$ is a minimizing sequence of $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \cdot]$ we get the uniform boundedness of $\overline{V}^k_j$ in $\mathbb{R}^{\sharp I}$. With Bolzano-Weierstraß we get the existence of a convergent subsequence. Thus it remains to show, that $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \cdot]$ is lower semicontinuous. As the first two summands of $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}]$ are quadratic in $\overline{V}$ and as $\Psi(V)$ is continuous in V, $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \cdot]$ is actually continuous. Hence, by the direct method in the calculus of variations, we get the existence of a minimizer $\overline{V}^k = \overline{V}^k[\overline{U}^k] \in \mathbb{R}^{\sharp I}$. This allows us to consider for fixed $k > 0$ and given $\overline{U}^{k-1} \in \mathbb{R}^{\sharp I}$ a minimizing sequence $(\overline{U}^k_j, \overline{V}^k_j)_{j=1,\ldots}$ of $\mathcal{E}^\varepsilon_{\text{out}}[\overline{U}^{k-1}, \cdot, \cdot]$ with $\overline{V}^k_j$ being the minimizer of $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k_j, \cdot]$. The outer energy $\mathcal{E}^\varepsilon_{\text{out}}$ can be estimated

$$\begin{aligned}
\mathcal{E}^\varepsilon_{\text{out}}[\overline{U}^{k-1}, \overline{U}^k_j, \overline{V}^k_j] &= \varepsilon \mathbf{M} \left(\overline{U}^k_j - \overline{U}^{k-1}\right) \cdot \left(\overline{U}^k_j - \overline{U}^{k-1}\right) + \frac{\tau\varepsilon}{\tilde{\tau}^2} \mathbf{M} \left(\overline{V}^k_j - \overline{U}^k_j\right) \cdot \left(\overline{V}^k_j - \overline{U}^k_j\right) \\
&\geq \varepsilon \lambda_{\min} \|\overline{U}^k_j - \overline{U}^{k-1}\|^2 \\
&\geq \varepsilon \lambda_{\min} \left(\left(1 - \frac{1}{\delta}\right) \|\overline{U}^k_j\|^2 + (1-\delta) \|\overline{U}^{k-1}\|^2\right),
\end{aligned}$$

where again $\lambda_{\min}$ denotes the smallest eigenvalue of $\mathbf{M}$ and $\delta > 1$. As $(\overline{U}^k_j, \overline{V}^k_j)_{j=1,\ldots}$ is a minimizing sequence of $\mathcal{E}^\varepsilon_{\text{out}}[\overline{U}^{k-1}, \cdot, \cdot]$ we get the uniform boundedness of $\overline{U}^k_j$ in $\mathbb{R}^{\sharp I}$. In the same way as before we estimate $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k_j, \overline{V}^k_j]$ and get the uniform boundedness of $\overline{V}^k_j$ in $\mathbb{R}^{\sharp I}$.

By the uniform boundedness of $\overline{U}_j^k$ and $\overline{V}_j^k$ in $\mathbb{R}^{\sharp I}$ we can extract a subsequence, again denoted by $(\overline{U}_j^k, \overline{V}_j^k)_{j=1,\dots}$, such that $\overline{U}_j^k$ converges to some $\overline{U}^k$ and $\overline{V}_j^k$ converges to some $\overline{V}^k$ both in $\mathbb{R}^{\sharp I}$. As $\mathcal{E}^\varepsilon_{\text{out}}[\overline{U}^{k-1}, \overline{U}, \overline{V}]$ is continuous in $(\overline{U}, \overline{V})$, we get

$$\mathcal{E}^\varepsilon_{\text{out}}[\overline{U}^{k-1}, \overline{U}^k, \overline{V}^k] = \lim_{j\to\infty} \mathcal{E}^\varepsilon_{\text{out}}[\overline{U}^{k-1}, \overline{U}_j^k, \overline{V}_j^k]\,.$$

It remains to prove that $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^k] = \min\limits_{\overline{V}\in\mathbb{R}^{\sharp I}} \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}]$. Therefore let us assume that there is a $\overline{V}^* \in \mathbb{R}^{\sharp I}$ with $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^k] - \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^*] > 0$, then we obtain by the continuity of $\mathcal{E}^\varepsilon_{\text{in}}[\cdot,\cdot]$ in both arguments

$$\begin{aligned} 0 &< \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^k] - \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^*] \\ &= \lim_{j\to\infty} \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}_j^k, \overline{V}_j^k] - \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}_j^k, \overline{V}^*]\,. \end{aligned}$$

Hence, there exists an index j with $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}_j^k, \overline{V}_j^k] - \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}_j^k, \overline{V}^*] > 0$. This contradicts our assumption that $\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}_j^k, \overline{V}_j^k] = \min_{\overline{V}\in\mathbb{R}^{\sharp I}} \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}_j^k, \overline{V}]$. Thus,

$$\mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}^k] = \min_{\overline{V}\in\mathbb{R}^{\sharp I}} \mathcal{E}^\varepsilon_{\text{in}}[\overline{U}^k, \overline{V}]$$

and $(\overline{U}^k, \overline{V}^k)$ solves the fully discrete variational problem (3.18a) and (3.18b). □

3.2.2 Newton method for the constrained optimization problem

As in Subsection 3.1.3 we replace the problem of solving the above fully discrete constrained optimization problem (3.18a) and (3.18b) by the equivalent problem of finding the saddle point of the corresponding Lagrangian. This discrete Lagrangian is

$$\begin{aligned} \mathcal{L}[U^k, U, V, P] &= \mathcal{E}^\varepsilon_{\text{out}}[U^k, U, V] + \partial_V \mathcal{E}^\varepsilon_{\text{in}}[U, V](P) \\ &= \varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k\right)\cdot\left(\overline{U} - \overline{U}^k\right) + \frac{\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\left(\overline{V} - \overline{U}\right) \\ &\quad + 2\varepsilon\mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\overline{P} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1 - V^2]\overline{V}\cdot\overline{P} + 2\varepsilon\tilde{\tau}\mathbf{L}\overline{V}\cdot\overline{P}\,. \end{aligned} \tag{3.19}$$

Now we have to solve $\nabla_{(U,V,P)}\mathcal{L}[U^k, U, V, P] = 0$ to find a saddle $(U, V, P) \in \mathcal{V}\times\mathcal{V}\times\mathcal{V}$ of the Lagrangian $\mathcal{L}$. Therefore we compute the three components

$$\begin{aligned} \partial_U\mathcal{L}[U,V,P](\Theta) &= 2\varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k\right)\cdot\overline{\Theta} - \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\overline{\Theta} - 2\varepsilon\mathbf{M}\overline{P}\cdot\overline{\Theta}\,, \\ \partial_V\mathcal{L}[U,V,P](\Xi) &= \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\overline{\Xi} + 2\varepsilon\mathbf{M}\overline{P}\cdot\overline{\Xi} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1 - 3V^2]\overline{P}\cdot\overline{\Xi} + 2\varepsilon\tilde{\tau}\mathbf{L}\overline{P}\cdot\overline{\Xi}\,, \\ \partial_P\mathcal{L}[U,V,P](\Sigma) &= 2\varepsilon\mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\overline{\Sigma} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1 - V^2]\overline{V}\cdot\overline{\Sigma} + 2\varepsilon\tilde{\tau}\mathbf{L}\overline{V}\cdot\overline{\Sigma} \end{aligned}$$

with Θ, Ξ, $\Sigma \in \mathcal{V}$. Due to the definition of the double well potential $\Psi(V) = \left(1 - V^2\right)^2$ the third summand of $\partial_V\mathcal{L}$ and the second summand of $\partial_P\mathcal{L}$ is nonlinear in V. Thus we set up a Newton method to find the saddle point of the Lagrangian

$\mathcal{L}$ and consequently we have to compute the Hessian of $\mathcal{L}$. The different components are

$$\begin{aligned}
\partial_U\partial_U\mathcal{L}[U,V,P](\Theta,\Lambda) &= 2\varepsilon\mathbf{M}\overline{\Lambda}\cdot\overline{\Theta}+\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\overline{\Lambda}\cdot\overline{\Theta}\\
\partial_V\partial_U\mathcal{L}[U,V,P](\Theta,\mathit{\Pi}) &= -\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\overline{\mathit{\Pi}}\cdot\overline{\Theta}\\
\partial_P\partial_U\mathcal{L}[U,V,P](\Theta,\Upsilon) &= -2\varepsilon\mathbf{M}\overline{\Upsilon}\cdot\overline{\Theta}\\
\partial_V\partial_V\mathcal{L}[U,V,P](\Xi,\mathit{\Pi}) &= \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\overline{\mathit{\Pi}}\cdot\overline{\Xi}+\frac{24\tilde{\tau}}{\varepsilon}\mathbf{M}[VP]\overline{\mathit{\Pi}}\cdot\overline{\Xi}\\
\partial_P\partial_V\mathcal{L}[U,V,P](\Xi,\Upsilon) &= 2\varepsilon\mathbf{M}\overline{\Upsilon}\cdot\overline{\Xi}-\frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2]\overline{\Upsilon}\cdot\overline{\Xi}+2\varepsilon\tilde{\tau}\mathbf{L}\overline{\Upsilon}\cdot\overline{\Xi}\\
\partial_P\partial_P\mathcal{L}[U,V,P](\mathit{\Sigma},\Upsilon) &= 0
\end{aligned}$$

and therefore the Hessian of the discretized Lagrangian is

$$D^2\mathcal{L}=\begin{pmatrix}
2\varepsilon\left(1+\frac{\tau}{\tilde{\tau}^2}\right)\mathbf{M} & -\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M} & -2\varepsilon\mathbf{M}\\
-\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M} & \frac{2\varepsilon\tau}{\tilde{\tau}^2}\mathbf{M}+\frac{24\tilde{\tau}}{\varepsilon}\mathbf{M}[VP] & 2\varepsilon\mathbf{M}-\frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2]+2\varepsilon\tilde{\tau}\mathbf{L}\\
-2\varepsilon\mathbf{M} & 2\varepsilon\mathbf{M}-\frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2]+2\varepsilon\tilde{\tau}\mathbf{L} & 0
\end{pmatrix}.$$

Now we define $\mathcal{F}[U,V,P]:=\nabla_{(U,V,P)}\mathcal{L}[U,V,P]$ and apply a Newton method to find a root of $\mathcal{F}$. Therefore we introduce some further notation: Let U_i, V_i, and P_i be given approximations for U^{k+1}, V^{k+1} and P^{k+1}, then we define $Z_i:=(U_i,V_i,P_i)$. In general in a Newton method we would have to iterate solving the following linear system of equations

$$D\mathcal{F}[Z_i](Z_{i+1}-Z_i)=-\mathcal{F}[Z_i]$$

for Z_{i+1} with $D\mathcal{F}=D^2\mathcal{L}$. To get a positive definite matrix on the left hand side we have to modify the linear system of equations, such that we solve

$$D\mathcal{F}[Z_i]^T D\mathcal{F}[Z_i](Z_{i+1}-Z_i)=-D\mathcal{F}[Z_i]^T\mathcal{F}[Z_i] \tag{3.20}$$

for Z_{i+1} within each Newton step. Notice, that in general $D\mathcal{F}[Z_i]^T D\mathcal{F}[Z_i]$ is only positive semi definite. However in our test cases $D\mathcal{F}[Z_i]$ has no eigenvalues being zero such that we can use the Cholesky solver from the CHOLMOD package [29, 24].
In Algorithm 3.1 it is documented that for a given phase field function U^0 we compute V^0 as a root of $\partial_P\mathcal{L}[U_0,\cdot,P]=0$ and P_0 by solving the linear system of equations $\partial_V\mathcal{L}[U_0,V_0,P]=0$ in P. At the beginning of each Newton step we initialize $Z_0=(U_0,V_0,P_0)$ with (U^k,V^k,P^k). Moreover it turned out that in all our applications presented in Chapter 5, a time step control in the Newton scheme is not required.

Algorithm 3.1: Newton method with Cholesky solver to solve the variational problem (3.18a) and (3.18b)

```
given: phase field function $U^0$;
given: maximal number of time steps $K_{max}$;
compute $V^0$ as root of $\partial_P \mathcal{L}[U^0, \cdot, P] = 0$;
compute $P^0$ as root of $\partial_V \mathcal{L}[U^0, V^0, \cdot] = 0$ ;
for $k = 0$ to $K_{max}$ do
    set counter of the Newton method $i = 0$;
    set $Z_i := (U_i, V_i, P_i) = (U^k, V^k, P^k)$;
    while $\|\mathcal{F}[Z_i]\| > \epsilon$ do
        solve $D\mathcal{F}[Z_i]^T D\mathcal{F}[Z_i](\tilde{Z}_i) = -D\mathcal{F}[Z_i]^T \mathcal{F}[Z_i]$ with Cholesky;
        $Z_{i+1} = Z_i + \tilde{Z}_i$;
        $i \leftarrow i + 1$;
    end
    $U^{k+1} = U_i$;
    $V^{k+1} = V_i$;
    $P^{k+1} = P_i$;
end
```

The biharmonic heat equation - a linear model problem

In this chapter we deal with a linear model problem for our nested time discretization of Willmore flow. This linear model problem is given by the L^2-gradient flow of the energy $w^l[u] = \frac{1}{2}\int_\Omega (\Delta u)^2 \, \mathrm{d}x$, where Ω is a polygonally bounded, convex domain. In [52] we presented some numerical analysis for this linear model problem and observed by numerical experiments the same convergence behavior for the nested time discretized Willmore model.

In addition to the notation introduced in the previous chapters, where we used upper case letters for spatially discretized quantities and lower case letters for spatially continuous quantities we distinguish between time-discrete quantities without tilde and time-continuous quantities with tilde. With this notation the L^2-gradient flow equation leading to a minimum of the energy $w^l[\tilde{u}]$ is

$$\partial_t \tilde{u} = -\Delta^2 \tilde{u} \quad \text{on } \Omega \tag{4.1}$$

with initial conditions $\tilde{u} = \tilde{u}^0$ at $t = 0$. For this partial differential equation we discuss three different types of boundary conditions. Natural boundary conditions as arose in Chapter 3 and as they will be used later on in this chapter, Dirichlet boundary conditions, as in the image restoration in Section 5.3, and periodic boundary conditions on a fundamental cell $\Omega = (0,1)^d$.

4.1 A nested scheme for the biharmonic heat equation

To develop an analog model for the biharmonic heat flow as our nested time discretized Willmore flow model we should notice that the energy functional $w^l[u]$ takes the role of the Willmore functional $w[u] = \frac{1}{2}\int_\Omega \mathbf{h}^2[u] \, \mathrm{d}x$. Thus, analog to the approximation of the mean curvature $\mathbf{h}$ we approximate the Laplacian of u by $\Delta u \approx \frac{v-u}{\tilde{\tau}}$, where v has to minimize the inner energy

$$e^l_{\text{in}}[u,v] = \|v-u\|^2_{L^2(\Omega)} + 2\tilde{\tau} a^l[v] \,. \tag{4.2}$$

Here a^l denotes the Dirichlet energy and is defined by $a^l[v] := \frac{1}{2}\int_\Omega |\nabla v|^2 \, \mathrm{d}x$. This inner energy follows the concept of the natural time discretization from Definition 3.2. The same concept and the discretization of Δu will be used to define the outer energy

$$e^l_{\text{out}}[u^k,u,v] = \|u-u^k\|^2_{L^2(\Omega)} + \frac{\tau}{\tilde{\tau}^2}\|v-u\|^2_{L^2(\Omega)} \,. \tag{4.3}$$

To formulate the time-discrete scheme for the biharmonic heat equation we have to specify the different ansatz spaces to realize the three kinds of boundary conditions. For Dirichlet boundary conditions $u = \Delta u = 0$ on $\partial\Omega$ we choose the ansatz space $\overline{H}^1(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ on } \partial\Omega\}$. Here $u = 0$ on $\partial\Omega$ is meant in a sense of trace. In case of natural boundary conditions $\partial_\nu u = \partial_\nu \Delta u = 0$ on $\partial\Omega$ with normal derivative ∂_ν on $\partial\Omega$ and outer normal ν we set $\overline{H}^1(\Omega) = H^1(\Omega) \cap L^2_0(\Omega)$ with $L^2_0(\Omega) = \{u \in L^2 : \int_\Omega u \,\mathrm{d}x = 0\}$ and finally for periodic boundary conditions on a fundamental cell $\Omega = (0,1)^d$ and with e_i being the ith unit normal vector the ansatz space is $\overline{H}^1(\Omega) = \{u \in H^1(\Omega) : u(x + e_i) = u(x) \text{ on } \partial\Omega\}$.
Thus, depending on the special choice of boundary conditions, respectively ansatz space $\overline{H}^1(\Omega)$ we get the following time-discrete scheme for the biharmonic heat equation.

Definition 4.1 (Time-discrete biharmonic heat flow). *For given $u^0 \in \overline{H}^1(\Omega)$, let the sequence $u^k \in \overline{H}^1(\Omega)$, $k \in \mathbb{N}$, be defined as the solution of*

$$u^{k+1} = \underset{u \in \overline{H}^1(\Omega)}{\operatorname{argmin}}\, e^l_{out}[u^k, u, v]\,, \text{ where} \tag{4.4a}$$

$$v = \underset{\tilde{v} \in \overline{H}^1(\Omega)}{\operatorname{argmin}}\, e^l_{in}[u, \tilde{v}]\,. \tag{4.4b}$$

In case of this linear model problem it is possible to write the operator version of the time-discrete biharmonic heat flow in Definition 4.1. Therefore we regard the Lagrangian corresponding to (4.4a) and (4.4b)

$$\begin{aligned}\ell^l[u^k, u, v, p] &= e^l_{\mathrm{out}}[u^k, u, v] + \partial_v e^l_{\mathrm{in}}[u, v](p) \\ &= \int_\Omega \left(u - u^k\right)^2 + \frac{\tau}{\tilde{\tau}^2}(v - u)^2 \,\mathrm{d}x + \int_\Omega 2(v - u)\,p + 2\tilde{\tau}\nabla v \cdot \nabla p \,\mathrm{d}x\ .\end{aligned}$$

From this we deduce the KKT conditions

$$0 = \int_\Omega 2\left(u - u^k\right)\vartheta + \frac{2\tau}{\tilde{\tau}^2}(u - v)\,\vartheta \,\mathrm{d}x - \int_\Omega 2p\vartheta \,\mathrm{d}x\,, \tag{4.5a}$$

$$0 = \int_\Omega \frac{2\tau}{\tilde{\tau}^2}(v - u)\,\xi \,\mathrm{d}x + \int_\Omega 2\xi p + 2\tilde{\tau}\nabla\xi \cdot \nabla p \,\mathrm{d}x\,, \tag{4.5b}$$

$$0 = \int_\Omega 2(v - u)\,\varrho + 2\tilde{\tau}\nabla v \cdot \nabla\varrho \,\mathrm{d}x \tag{4.5c}$$

with $\vartheta,\ \xi,\ \varrho \in \overline{H}^1(\Omega)$ which lead to the operator version of Definition 4.1. For deriving this operator version we follow from (4.5c)

$$v = (\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}\, u\,. \tag{4.6}$$

Here we denote by $\Delta_h u$ the solution $w \in L^2(\Omega)$ of

$$(w, \vartheta)_{L^2(\Omega)} = -(\nabla u, \nabla\vartheta)_{L^2(\Omega)} \quad \forall \vartheta \in \overline{H}^1(\Omega)\,.$$

Moreover from (4.5b) we get

$$p = \frac{\tau}{\tilde{\tau}^2}(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}(u - v)\ . \tag{4.7}$$

Combining (4.6) and (4.7) we deduce from (4.5a)

$$\begin{aligned} u - u^k &= \frac{\tau}{\tilde{\tau}^2}(v-u) + p \\ &\overset{(4.7)}{=} -\frac{\tau}{\tilde{\tau}^2}(u-v) + \frac{\tau}{\tilde{\tau}^2}\left(\mathrm{id} - \tilde{\tau}\Delta_h\right)^{-1}(u-v) \\ &= \frac{\tau}{\tilde{\tau}^2}\left(\left(\mathrm{id} - \tilde{\tau}\Delta_h\right)^{-1} - \mathrm{id}\right)(u-v) \\ &\overset{(4.6)}{=} -\frac{\tau}{\tilde{\tau}^2}\left(\left(\mathrm{id} - \tilde{\tau}\Delta_h\right)^{-1} - \mathrm{id}\right)^2 u\,. \end{aligned}$$

Therefore the operator version of Definition 4.1 is

Definition 4.2 (Operator version of the time-discrete biharmonic heat flow)**.** *For given $u^0 \in \overline{H}^1(\Omega)$, let the sequence $u^k \in \overline{H}^1(\Omega)$, $k \in \mathbb{N}$, be defined as the solution of*

$$u^{k+1} - u^k = -\frac{\tau}{\tilde{\tau}^2}\left(\left(\mathrm{id} - \tilde{\tau}\Delta_h\right)^{-1} - \mathrm{id}\right)^2 u^{k+1}\,.$$

To derive the fully discretized version of the time-discrete biharmonic heat flow in Definition 4.1 we proceed as follows: With $\mathcal{V}$ being the Finite Element space as defined in Section 2.2 we introduce the Finite Element space $\overline{\mathcal{V}} = \mathcal{V} \cap \overline{H}^1(\Omega)$, which is a Finite Element space with boundary conditions as in Definition 4.1. Then we write the discrete version of the linear outer and inner energy

$$\begin{aligned} \mathcal{E}^l_{\mathrm{out}}[U^k, U, V] &= \mathbf{M}\left(\overline{U} - \overline{U^k}\right)\cdot\left(\overline{U} - \overline{U^k}\right) + \frac{\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\left(\overline{V} - \overline{U}\right)\,, \\ \mathcal{E}^l_{\mathrm{in}}[V, U] &= \mathbf{M}\left(\overline{V} - \overline{U}\right)\cdot\left(\overline{V} - \overline{U}\right) + \tilde{\tau}\mathbf{L}\overline{V}\cdot\overline{V}\,. \end{aligned}$$

A comparison of these two discrete energies with the fully discrete outer and inner energy $\mathcal{E}^\varepsilon_{\mathrm{out}}$ and $\mathcal{E}^\varepsilon_{\mathrm{in}}$ in (3.17a) and (3.17b) demonstrates why this is the linear model problem of our nested time discretized Willmore flow model. With $\varepsilon = 1$ and without the nonlinear term $\frac{1}{2\varepsilon}\mathbf{M}\overline{\Psi(V)}\cdot\overline{1}$ in the inner energy (3.17b) the linear outer energy $\mathcal{E}^l_{\mathrm{out}}$ is equal to $\mathcal{E}^\varepsilon_{\mathrm{out}}$ and $\mathcal{E}^l_{\mathrm{in}}$ is equal to $\mathcal{E}^\varepsilon_{\mathrm{in}}$. Thus the fully discretized linear model problem of our nested time discretized Willmore flow model is

Definition 4.3 (Fully discretized biharmonic heat flow)**.** *For given $U^0 \in \overline{\mathcal{V}}$, the sequence $U^k \in \overline{\mathcal{V}}$, $k \in \mathbb{N}$, is defined as the solution of*

$$U^{k+1} = \operatorname*{argmin}_{U \in \overline{\mathcal{V}}} \mathcal{E}^l_{out}[U^k, U, V]\,, \text{ where} \tag{4.9a}$$

$$V = \operatorname*{argmin}_{\tilde{V} \in \overline{\mathcal{V}}} \mathcal{E}^l_{in}[U, \tilde{V}]\,. \tag{4.9b}$$

To avoid the necessity of analyzing an error caused by numerical quadrature we assume the fully discrete outer and inner energy $\mathcal{E}^l_{\mathrm{out}}$ and $\mathcal{E}^l_{\mathrm{in}}$ to be evaluated exactly such that $\mathcal{E}^l_{\mathrm{out}} = e^l_{\mathrm{out}}$ and $\mathcal{E}^l_{\mathrm{in}} = e^l_{\mathrm{in}}$. This is equivalent to assume at least second order quadrature. Moreover, in what follows, we consider quasiuniform triangulation with grid size h. The convergence behavior of the fully discrete model in Definition 4.3 will be presented in the following theorem. Therefore we need some further notation. As in [95] we define

$$\dot{H}^s := \left\{ v \in L^2(\Omega) \,:\, |v|_s = \left(\sum_{m=1}^{\infty} \lambda_m^s \left(v, \varphi_m\right)^2\right)^{\frac{1}{2}} < \infty \right\}\,,$$

the subspace of functions with s weak derivatives which is associated with the spectral norm belonging to Δ^s. Here $\{\lambda_m\}_{m=1}^{\infty}$ is a nondecreasing sequence of positive eigenvalues of the eigenvalue problem

$$-\Delta\varphi = \lambda\varphi \quad \text{in } \Omega, \qquad \text{with } \varphi = 0 \text{ on } \partial\Omega$$

and $\{\varphi_m\}_{m=1}^{\infty}$ is the corresponding sequence of eigenfunctions which form a basis of $L^2(\Omega)$. Furthermore we denote by $\Delta_h\Phi \in \overline{\mathcal{V}}$ for $\Phi \in \overline{\mathcal{V}}$ the discrete Laplacian, which is the unique solution of

$$(-\Delta_h\Phi, \Theta)_{L^2(\Omega)} = (\nabla\Phi, \nabla\Theta)_{L^2(\Omega)} \quad \forall\Theta \in \overline{\mathcal{V}}\,.$$

Finally, with $R_h u \in \overline{\mathcal{V}}$ for $u \in \overline{H}^1(\Omega)$ being the canonical Ritz projection, defined as the unique Finite Element solution of

$$(\nabla R_h u, \nabla\Theta)_{L^2(\Omega)} = (\nabla u, \nabla\Theta)_{L^2(\Omega)} \quad \forall\Theta \in \overline{\mathcal{V}}\,,$$

we introduce the projection $\mathring{\mathring{R}}_h u \in \overline{\mathcal{V}}$ for $u \in \dot{H}^5$, defined as $\mathring{\mathring{R}}_h u = \Delta_h^{-2} R_h \Delta^2 u$.

Theorem 4.4 (Error estimates for fully discrete biharmonic heat flow)**.** *For given $\tilde{u}^0 \in \dot{H}^8$, let $\tilde{u}$ be the solution to the biharmonic heat equation* (4.1)*, and let U^k, $k \in \mathbb{N}$, be the solution of the fully discretized scheme* (4.9) *with $U^0 = \mathring{\mathring{R}}_h \tilde{u}^0$. On quasiuniform triangulations with grid size h and for uniform time steps τ and $\tilde{\tau}$ with $t_k = k\tau$*

$$\|U^k - \tilde{u}(t_k)\|_{L^2} \le C(t_k)(h^2 + \tau + \tilde{\tau})\,, \tag{4.10a}$$

$$\|U^k - \tilde{u}(t_k)\|_{H^1} \le C(t_k)\left(h + (1 + \tfrac{\sqrt{\tilde{\tau}}}{h})(\tau + \tilde{\tau})\right) \tag{4.10b}$$

holds for a continuous function $C : \mathbb{R}^+ \to \mathbb{R}^+$.

A proof of this theorem is presented in [52]. It bases on a splitting of the error $U^k - \tilde{u}(t_k)$ into $U^k - \tilde{U}(t_k)$, the difference between the fully discrete and the time continuous, but spatially discretized solution and $\tilde{U}(t_k) - \tilde{u}(t_k)$, the time continuous Finite Element error. This is a well known splitting to analyze condition and stability. As Theorem 4.4 states the scheme stably approximates the biharmonic heat flow for any combination of time steps τ and $\tilde{\tau}$. Nevertheless it does not make sense to increase $\tilde{\tau}$ beyond τ. The step size $\tilde{\tau}$ is related to the accuracy with which the Laplace operator is approximated and the outer step size τ defines the time scale at which the bi-Laplace is resolved. In the same way, the inner step size $\tilde{\tau}$ in the Willmore flow case is associated with the accuracy with which the mean curvature h is approximated and the outer step size defines the scale at which the Willmore flow is resolved. Thus in the nonlinear case it also does not make sense to increase $\tilde{\tau}$ beyond τ.

Apart from this similarity we will observe that the experimental convergence behavior of the phase field model for Willmore flow, presented in Chapter 3 fits to the convergence behavior of the linear model.

4.2 Numerical validation

In this section we numerically validate the convergence result presented in Theorem 4.4 by solving the Dirichlet boundary value problem

$$\partial_t \tilde{u} + \Delta^2 \tilde{u} = \tilde{f} \quad \text{ on } \Omega \times [0,T], \;\text{ with } \tilde{u} = \Delta \tilde{u} = 0 \text{ on } \partial\Omega \tag{4.11}$$

with $\Omega = [0,1]^2$ for a special choice of $\tilde{f}$.

4.2.1 Implementation

For the implementation of the validating test we consider the following lemma presenting a special right hand side and the corresponding analytical solution.

Lemma 4.5. *Let* $\Omega = [0,1]^2$ *and* $\tilde{f} : \Omega \times [0,T] \to \mathbb{R}$ *be*

$$\begin{aligned}\tilde{f}(t,x,y) :=\, & 20\, e^{20t}(x^4 - 2x^3 + x)\,(y^4 - 2y^3 + y) \\ & + 24\,(e^{20t} - 1)\,(y^4 - 2y^3 + y + (x^2 - x)\,(12y^2 - 12y) + x^4 - 2x^3 + x)\end{aligned}$$

then the solution of the Dirichlet boundary value problem (4.11) is

$$\tilde{u}(t,x,y) = (e^{20t} - 1)\,(x^4 - 2x^3 + x)\,(y^4 - 2y^3 + y)\,.$$

Proof. The property $\tilde{u}(t,x,y) = 0$ on $\partial\Omega$ for $\Omega = [0,1]^2$ can be seen directly. Moreover the gradient of $\tilde{u}$ in (x,y) is

$$\nabla \tilde{u}(t,x,y) = \begin{pmatrix} \left(e^{20t} - 1\right)\left(4x^3 - 6x^2 + 1\right)\left(y^4 - 2y^3 + y\right) \\ \left(e^{20t} - 1\right)\left(x^4 - 2x^3 + x\right)\left(4y^3 - 6y^2 + 1\right) \end{pmatrix}$$

and therefore

$$\Delta \tilde{u}(t,x,y) = \left(e^{20t} - 1\right)\left(\left(12x^2 - 12x\right)\left(y^4 - 2y^3 + y\right) + \left(x^4 - 2x^3 + x\right)\left(12y^2 - 12y\right)\right)\,.$$

Again it is obvious that $\Delta \tilde{u} = 0$ on $\partial\Omega$. Since

$$\partial_t \tilde{u}(t,x,y) = 20\, e^{20t}(x^4 - 2x^3 + x)\,(y^4 - 2y^3 + y)$$

the computation of

$$\Delta^2 \tilde{u}(t,x,y) = 24\,(e^{20t} - 1)\,(y^4 - 2y^3 + y + (x^2 - x)\,(12y^2 - 12y) + x^4 - 2x^3 + x)$$

finishes the proof. □

Since the right hand side $\tilde{f}$ of (4.11) is different from zero we have to replace the outer energy in (4.9) by

$$\mathcal{E}^{l,f}_{\text{out}}[U^k, U, V] := \mathcal{E}^{l}_{\text{out}}[U^k, U, V] - 2\tau \int_\Omega \mathcal{I}_h(\tilde{f}U)\,\mathrm{d}x\,,$$

while the constraint (4.9b) remains the same. As in Chapter 3 we set up the Lagrangian

$$\begin{aligned}\mathcal{L}^l[U^k, U, V, P] = \, & \mathbf{M}\left(\overline{U} - \overline{U^k}\right) \cdot \left(\overline{U} - \overline{U^k}\right) + \frac{\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right) \\ & - 2\tau \mathbf{M}\overline{\mathcal{I}_h(\tilde{f})} \cdot \overline{U} + 2\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{P} + 2\tilde{\tau}\mathbf{L}\overline{V} \cdot \overline{P}\,.\end{aligned}$$

The computation of the variation of the Lagrangian $\mathcal{L}$ in $\overline{U}$, $\overline{V}$ and $\overline{P}$

$$\begin{aligned}
\partial_U \mathcal{L}[U^k, U, V, P](\Theta) &= 2\mathbf{M}\left(\overline{U} - \overline{U}^k\right) \cdot \overline{\Theta} - \frac{2\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Theta} - 2\tau\mathbf{M}\overline{\mathcal{I}_h(\tilde{f})} \cdot \Theta \\
&\quad - 2\mathbf{M}\overline{P} \cdot \overline{\Theta}\,, \\
\partial_V \mathcal{L}[U^k, U, V, P](\Xi) &= \frac{2\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Xi} + 2\mathbf{M}\overline{P} \cdot \overline{\Xi} + 2\tilde{\tau}\mathbf{L}\overline{P} \cdot \overline{\Xi}\,, \\
\partial_P \mathcal{L}[U^k, U, V, P](\Sigma) &= 2\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Sigma} + 2\tilde{\tau}\mathbf{L}\overline{V} \cdot \overline{\Sigma}
\end{aligned}$$

reveals that the equation $\nabla_{(U,V,P)}\mathcal{L}[U^k, U, V, P] = 0$ is a linear equation. Consequently it is not necessary to use a Newton method for finding a saddle point. Nevertheless we use a Newton method to keep the model and the code as similar to the Willmore flow model in Chapter 3 as possible. Thus we compute the Hessian

$$D^2\mathcal{L} = \begin{pmatrix} 2\left(1 + \frac{\tau}{\tilde{\tau}^2}\right)\mathbf{M} & -\frac{2\tau}{\tilde{\tau}^2}\mathbf{M} & -2\mathbf{M} \\ -\frac{2\tau}{\tilde{\tau}^2}\mathbf{M} & \frac{2\tau}{\tilde{\tau}^2}\mathbf{M} & 2\mathbf{M} + 2\tilde{\tau}\mathbf{L} \\ -2\mathbf{M} & 2\mathbf{M} + 2\tilde{\tau}\mathbf{L} & 0 \end{pmatrix}.$$

To realize Dirichlet boundary conditions we introduce some notation. First, we define $\mathcal{F}[U^k, U, V, P] := \nabla_{(U,V,P)}\mathcal{L}[U^k, U, V, P]$ as before and denote the vertex index set of all Dirichlet nodes with I_D. Then we have to set

$$\mathcal{F}_i^k = 0 \quad \forall i \in I_D, k = 0, 1, 2$$

and for the Hessian of the Lagrangian

$$\begin{aligned}
D\mathcal{F}_{ij}^{kk} &= 0 \quad \forall i \in I_D,\, j \in I,\, i \neq j, k = 0, 1, 2\,, \\
D\mathcal{F}_{ij}^{kk} &= 0 \quad \forall i \in I,\, j \in I_D,\, j \neq i, k = 0, 1, 2\,, \\
D\mathcal{F}_{ii}^{kk} &= 1 \quad \forall i \in I_D,\, k = 0, 1, 2\,, \\
D\mathcal{F}_{ij}^{kl} &= 0 \quad \forall i \in I_D,\, j \in I,\, k, l = 0, 1, 2,\, k \neq l\,, \\
D\mathcal{F}_{ij}^{kl} &= 0 \quad \forall i \in I,\, j \in I_D,\, k, l = 0, 1, 2,\, k \neq l\,.
\end{aligned}$$

Different to the procedure in Chapter 3 we use a preconditioned biconjugate gradient method to solve the linear system of equations in each Newton step. It is faster than the direct solver in Chapter 3 and it is sufficient for this test. As preconditioner we use a block diagonal preconditioner [90], which works as follows. In case of a block matrix with $d \times d$ quadratic blocks of n rows it creates a $d \times d$ temporary matrix. Now, it iterates i from 0 to $n - 1$ and saves the ith diagonal entry of all blocks (j, k) at position (j, k) of the temporary matrix. For each i we compute the inverse of the temporary matrix, if it is invertible, or set the inverse to identity, if not. Finally all entries of the inverted matrix are sorted into the preconditioning block matrix the other way round they have been extracted from the block matrix which has to be preconditioned. Thus the preconditioning block matrix has diagonal matrices in each block.

4.2.2 Validation

As we have seen, for the special choice of $\tilde{f}$ the analytical solution $\tilde{u}$ is known. Hence we are able to compute the L^2- and the H^1-error of the evolution to validate the

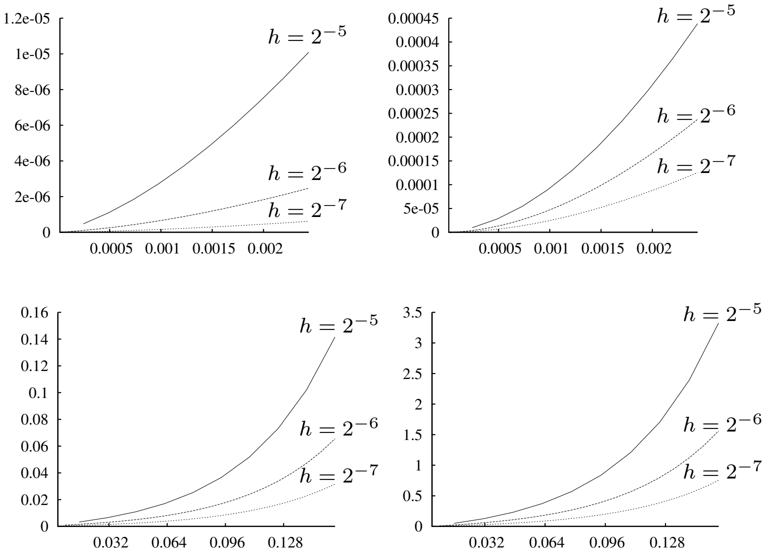


Figure 4.1: Temporal error evolution for the numerical solution of the biharmonic heat equation. Upper row: L^2-error for $\tau = h^2$ and $\tilde{\tau} = h^2$ (left) as well as $\tilde{\tau} = h$ (right). Lower row: H^1-error for $\tau = h$ and $\tilde{\tau} = h^2$ (left) as well as $\tilde{\tau} = h$ (right).

error order presented in Theorem 4.4. In Figure 4.1 there is presented the L^2-error for $\tau = \tilde{\tau} = h^2$ (upper left graph) and for $\tau = h^2$ and $\tilde{\tau} = h$ (upper right graph) as well as the H^1-error for $\tau = h$ and $\tilde{\tau} = h^2$ (lower left graph) and $\tau = \tilde{\tau} = h$ (lower right graph). In all four cases the error evolution is plotted on grids with grid sizes $h = 2^{-5}$, $h = 2^{-6}$ and $h = 2^{-7}$.

More precise than analyzing error plots is the computation of the error order. Therefore we take the numerical solution of the biharmonic heat equation at time $t = 0.15625$ for tests with step size $\tau = h$ and at time $t = 0.00244141$ for $\tau = h^2$. The error order for a transition from a grid with grid width $h = 2^{-k}$ to the next finer grid is called experimental order of convergence (eoc) and computed by

$$\text{eoc} = \frac{\ln\left(e_{k+1}(t)\right) - \ln\left(e_k(t)\right)}{-\ln 2}\,.$$

This order is presented in Table 4.1.
The above numerical analysis predicts an error decay by a factor $\frac{1}{4}$ for the L^2-error with $\tau = \tilde{\tau} = h^2$ and by a factor $\frac{1}{\sqrt{2}}$ for the H^1-error with $\tilde{\tau} = h$ and $\tau = h^2$ or $\tau = h$. In all other combinations presented in Figure 4.1 or Table 4.1 the numerical analysis predicts an error decay by a factor $\frac{1}{2}$. In case of the L^2-error the decay rates are confirmed by the numerical experiments. In contrast to this the decay rates in case of the H^1-error are better than expected for $\tilde{\tau} = h$ and $\tau = h^2$ or $\tau = h$. This suggests that the convergence estimates can be improved to get rid of the factor $\frac{\sqrt{\tilde{\tau}}}{h}$. Nevertheless $\tilde{\tau}$ may be chosen small. This compensates a factor $\frac{\sqrt{\tilde{\tau}}}{h}$ and does not influence the speed of the time stepping scheme.

	$\tau = h^2$, $\tilde{\tau} = h^2$		$\tau = h^2$, $\tilde{\tau} = h$		$\tau = h$, $\tilde{\tau} = h^2$		$\tau = h$, $\tilde{\tau} = h$	
k	L^2	H^1	L^2	H^1	L^2	H^1	L^2	H^1
6	2.032	1.020	0.886	0.959	1.417	1.113	1.085	1.093
7	2.009	1.005	0.919	0.969	1.259	1.048	1.045	1.050
8	2.002	1.001	0.953	0.979	1.148	1.021	1.023	1.026

Table 4.1: Experimental order of the L^2- and H^1-error for the numerical solution of the biharmonic heat equation with different time steps. The order $\frac{\ln(e_{k+1}(t))-\ln(e_k(t))}{-\ln 2}$ is shown for each transition from one grid with grid width $h = 2^{-k}$ to the next finer one. The error $e_k(t)$ was evaluated at time $t = 0.15625$ (cases with $\tau = h$) and $t = 0.00244141$ (cases with $\tau = h^2$), respectively.

	$\tau = h^2$, $\tilde{\tau} = h^2$		$\tau = h^2$, $\tilde{\tau} = h$		$\tau = h$, $\tilde{\tau} = h^2$		$\tau = h$, $\tilde{\tau} = h$	
	L^2	H^1	L^2	H^1	L^2	H^1	L^2	H^1
proven decay rate $\leq$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$

Table 4.2: Proven decay rate of the L^2- and H^1-error for the numerical solution of the biharmonic heat equation with different time steps.

Numerical results and applications for Willmore flow

In this chapter we present some numerical simulations demonstrating the reliability of the Willmore model presented in Definition 3.13. Moreover we compare it with a semi-implicit scheme by Du et al. [40, 42] and apply it to an image restoration problem in the last section.

5.1 Simulation based on the nested time discretization

We start to document the reliability of our Willmore model by observing the evolution of a circle in $\mathbb{R}^2$. This is an established test for Willmore flow models, because the evolution of the radius is known. If we start with a circle of initial radius r_0, Willmore flow leads to a family of concentric circles denoted by $\Gamma_{r(t)} \subset \mathbb{R}^2$, where $r(t)$ is the radius of the circle at time t (cp. [47, 37]). This radius solves the ordinary differential equation

$$\dot{r}(t) = \frac{1}{2}\frac{1}{r^3(t)} \qquad \text{with } r(0) = r_0\,.$$

By separation of variables, cp. [2], we get

$$r(t) = \sqrt[4]{2t + r_0^4}\,. \tag{5.1}$$

As we are working with phase field approximations the interface, in this case the circle, is given diffusely. Thus, to compute an L^2-error we have to generate an analytically "exact" phase field solution. As mentioned, the exact evolution of the radius is known. Moreover we expect our model to generate and preserve a phase field function with optimal profile perpendicular to the interface (cp. Lemma 2.2). Therefore we define a function

$$\tilde{u}_{k\tau} : \Omega \to \mathbb{R},\, x \mapsto \tanh\left(\frac{\operatorname{sgndist}\left(x, \Gamma_{r(k\tau)}\right)}{\varepsilon}\right)$$

which is the reference solution at time $k\tau$. With U^k being the numerical solution, computed by our Willmore model, the L^2-error is $\|e^k\|_{L^2} = \|\tilde{u}_{k\tau} - U^k\|_{L^2}$.

To benefit from a higher resolution and therefore get better results we start with a quarter circle of initial radius $r_0 = 0.4$ instead of a circle with initial radius $0.5 r_0$, centered around one corner of the computational domain $\Omega = [0, 1]^2$. In this test the

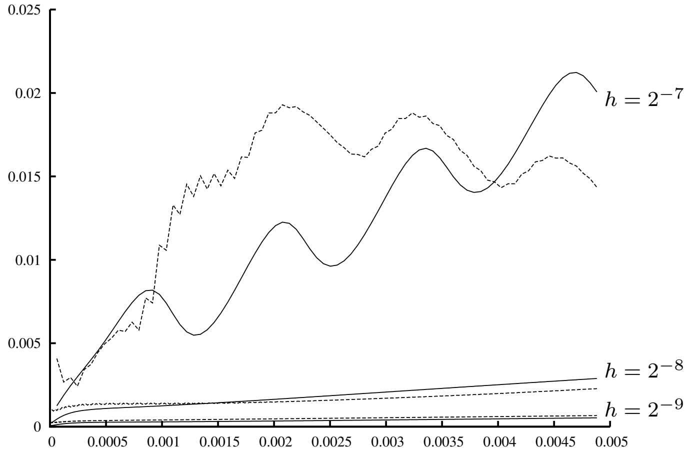


Figure 5.1: Evolution of the L^2-error for Willmore flow of a quarter circle in time. The initial radius is $r_0 = 0.4$, the interface transition thickness $\varepsilon = 2^{-6}$ and the time step sizes are $\tau = \tilde{\tau} = h^2$. Here, the grid sizes h vary from 2^{-7} to 2^{-9}. The solid lines show the error of our scheme, the dotted lines belong to the scheme proposed by Du et al. (cp. Section 5.2).

interface transition thickness is $\varepsilon = 2^{-6}$ and the time step sizes are $\tau = \tilde{\tau} = h^2$. The computation is performed on three different grids with grid sizes $h = 2^{-7}, 2^{-8}$ and 2^{-9}. In Figure 5.1 the results are visualized by solid lines. The oscillation of the error in the case $h = 2^{-7}$ indicates that the interface transition thickness $\varepsilon \sim 2h$ on this grid is still too small. For an adequate approximation of a fourth order partial differential equation we require $\varepsilon = 4h$. In Chapter 4, where we have treated the biharmonic heat equation as a linear model problem, we have proven for $\tau = \tilde{\tau} = h^2$ an error decay by a factor $\frac{1}{4}$ when halving the grid size. The results presented in Figure 5.1 indicate that it is reasonable to expect the same decay in case of our Willmore model.

As a further test simulation, let us consider Willmore flow of two exemplary geometries. In both cases we choose $\Omega = [0,1]^2$, $h = 2^{-9}$, $\varepsilon = 4h$, and $\tau = \tilde{\tau} = h^2$.
The first geometry is a rectangle of width 0.4 and height 0.1 and the corresponding phase field function is given by

$$u_{\text{rectangle}}(x) = \tanh\left(\frac{\max\left(|x_1 - 0.5| - 0.2,\ |x_2 - 0.5| - 0.05\right)}{\varepsilon}\right).$$

As presented in Figure 5.2 the evolution is initially concentrated at the corners. This leads to a locally concave object, which gets convex again within few time steps. Finally it evolves to a circle. Due to (5.1) this circle would continue to grow if we would compute more than 120 time steps.
The second geometry consists of two circles of radius $r_0 = 0.13$ centered at $(0.35, 0.5)$ and $(0.65, 0.5)$. The phase field function of this geometry is given by

$$u_{\text{circle}}(x) = \tanh\left(\frac{\min\left(f_{0.35}(x), f_{0.65}(x)\right)}{\varepsilon}\right)$$

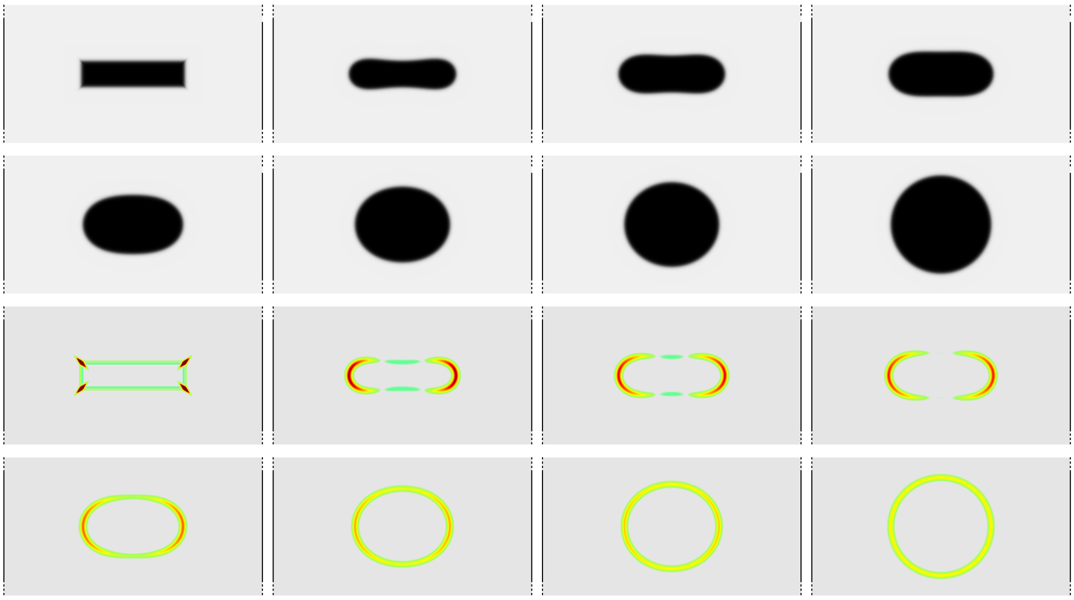


Figure 5.2: Computed evolution of a rectangle under Willmore flow for grid size $h = 2^{-9}$, $\varepsilon = 4h$, and $\tau = \tilde{\tau} = h^2$. Results are depicted at time steps $k = 0, 1, 6, 10, 20, 40, 60, 120$. The bottom rows show the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$, color-coding on the range $[-20, 20]$ as .

with

$$f_\alpha(x) = \sqrt{(x_1 - \alpha)^2 + (x_2 - 0.5)^2} - 0.13\,.$$

In this test simulation we observe a topological change (cp. Figure 5.3). Within the first time step both circles merge into one locally concave object. The concavity of this object is more pronounced, but similar to the concavity of the object at time step 1 in the evolution of the rectangle (cp. Figure 5.2). Therefore the further evolution of both objects is similar to each other. Different to the rectangular case it takes 450 time steps until the geometry is a circle.

Additionally to the evolution of the geometry itself the curvature approximation $(U^k - V^k)/\tilde{\tau}$ of the evolving geometries is displayed in both Figures 5.2 and 5.3. Notice, in case of the rectangle the curvature is color-coded on the range $[-20, 20]$ and in case of the two circles on the range $[-10, 10]$. This curvature information reveals a better understanding of the underlying geometry. In the fourth image of Figure 5.2 or the fifth image of Figure 5.3 it is not really possible to decide whether the geometry is convex. However, the curvature information displays this property since it is positive on the whole computational domain.

Finally we consider two three-dimensional simulations of Willmore flow, presented in Figure 5.4 and 5.5. In both cases the computation is performed on a grid with mesh size $h = 2^{-8}$, with interface transition thickness $\varepsilon = 4h$ and inner step size $\tilde{\tau} = h^2$. The first geometry is a cube (cp. Figure 5.4) with edge length 0.4. Consequently the

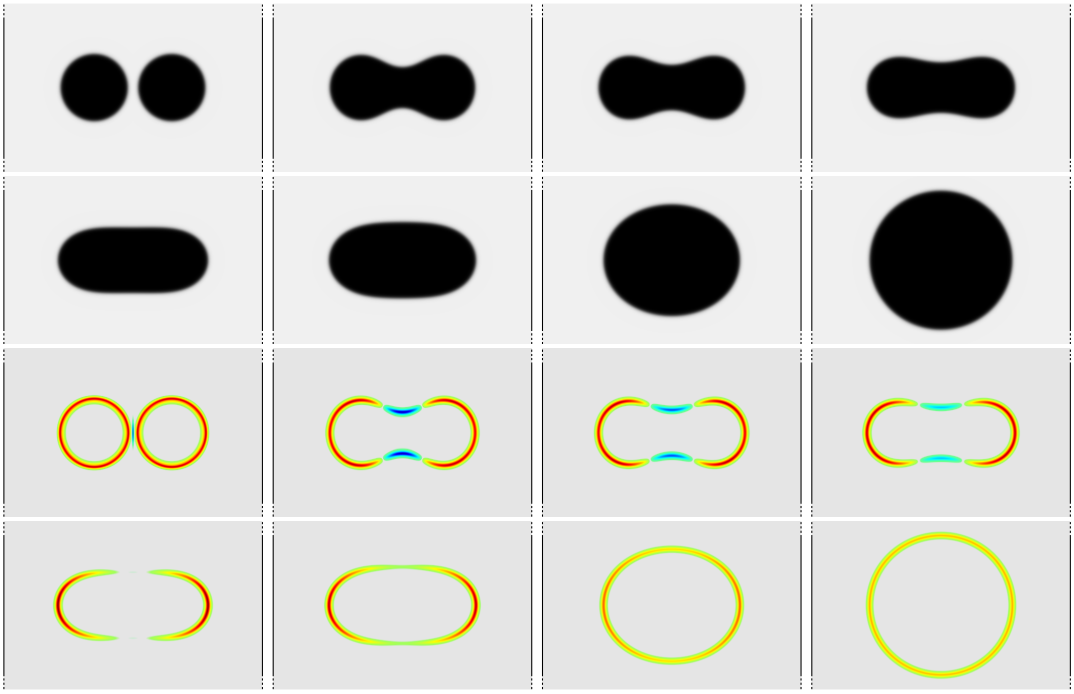


Figure 5.3: Computed evolution of two circles under Willmore flow on a grid of mesh size $h = 2^{-9}$, $\varepsilon = 4h$, and $\tau = \tilde{\tau} = h^2$. The visualized time steps are $k = 0$, 1, 2, 4, 20, 40, 160, 450. In the two lower lines the discrete curvature $(U^k - V^k)/\tilde{\tau}$ is color-coded on the range $[-10, 10]$ as .

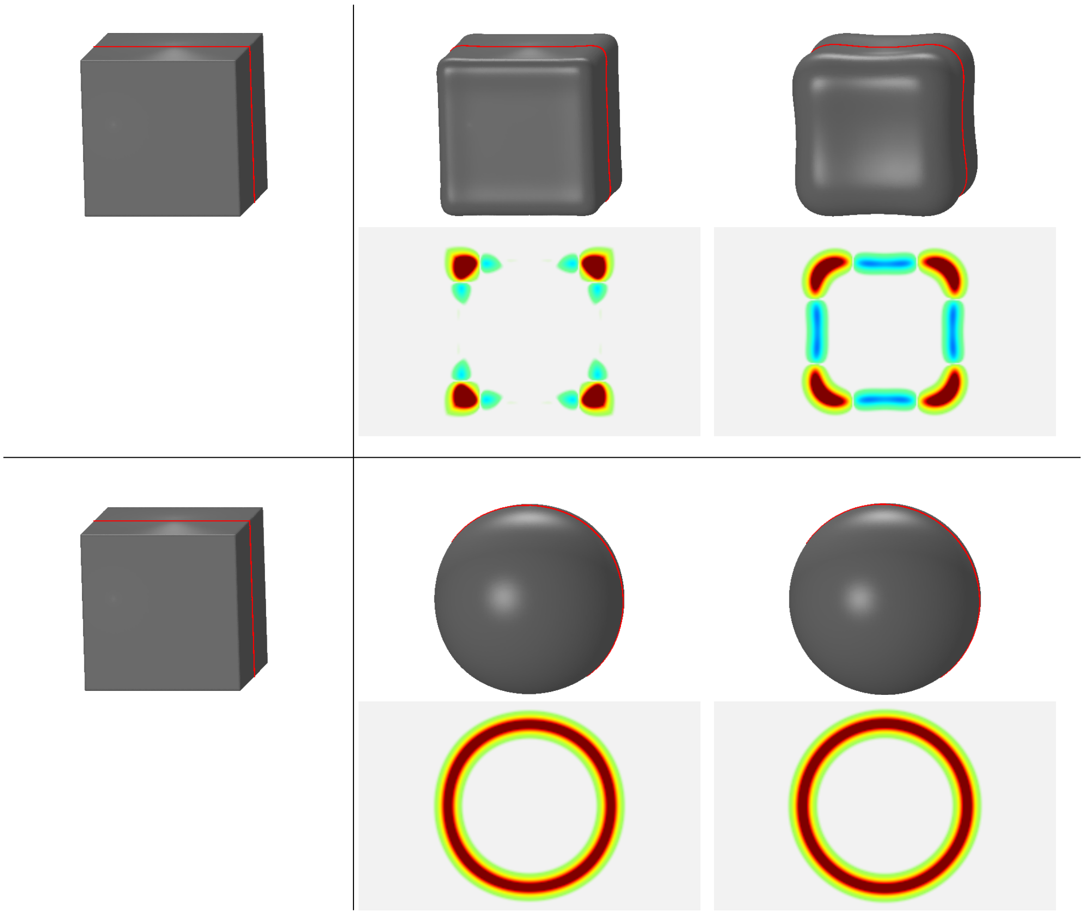


Figure 5.4: Different time steps of the discrete Willmore flow for a cube with edge length 0.4 as initial surface (left). The computational parameters are $h = 2^{-8}$, $\epsilon = 4h$, $\tilde{\tau} = h^2$. In the first line results are depicted at time steps $k = 1$, 20 with $\tau = h^3$ and in the second line at time steps $k = 1$, 10 with $\tau = 0.03h$. In addition to the evolving shapes we render the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$ on a planar slice indicated in red on the geometries and use a color-coding on the range $[-5, 5]$ as .

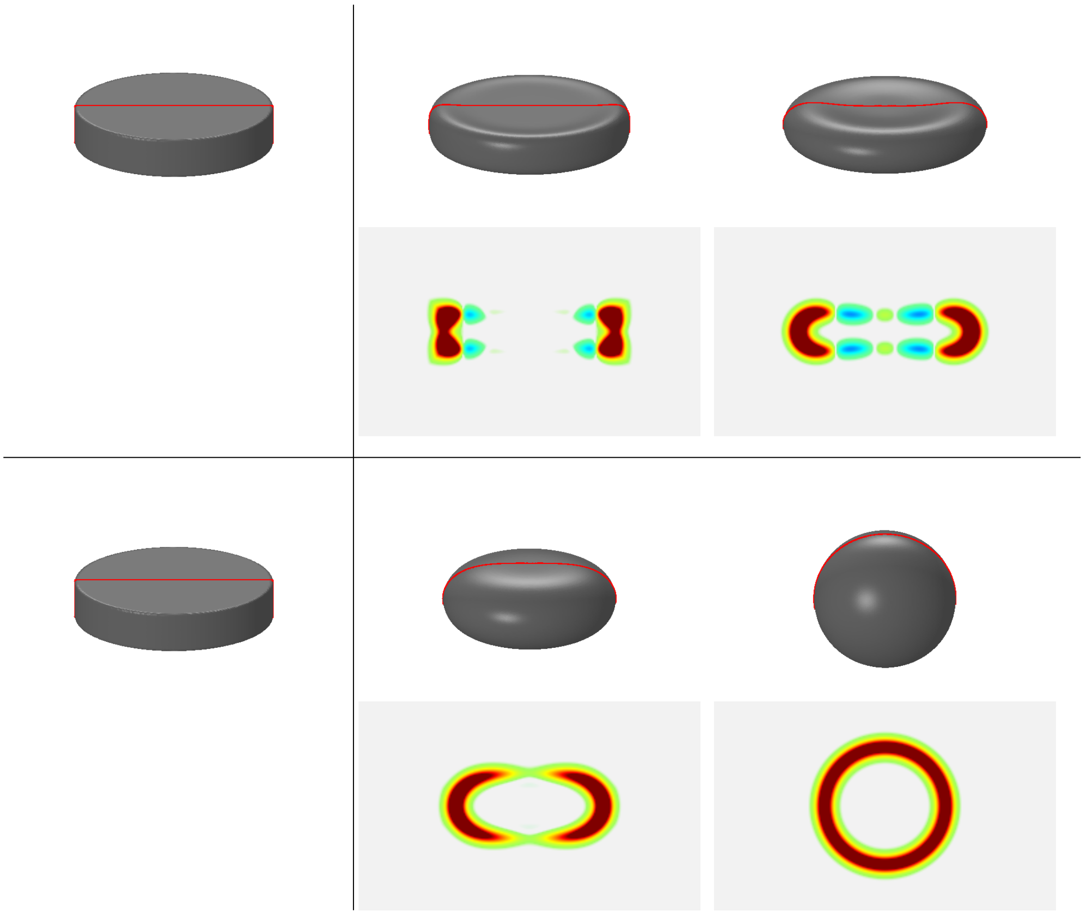


Figure 5.5: Different time steps of the discrete Willmore flow for a disk of diameter 0.5 and thickness 0.1 as initial surface (left). The computational parameters are $h = 2^{-8}$, $\epsilon = 4h$, $\tilde{\tau} = h^2$. In the first line results are depicted at time steps $k = 1, 20$ with $\tau = h^3$ and in the second line at time steps $k = 1, 10$ with $\tau = h^2$. In addition to the evolving shapes we render the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$ on a planar slice indicated in red on the geometries and use a color-coding on the range $[-5, 5]$ as .

corresponding phase field function is

$$u_{\text{cube}}(x) = \tanh\left(\frac{\max\left(\max\left(|x_1 - 0.5| - 0.2, |x_2 - 0.5| - 0.2\right), |x_3 - 0.5| - 0.2\right)}{\varepsilon}\right)$$

and the second geometry is a disc (cp. Figure 5.5) with diameter 0.5 and thickness 0.1, generated by

$$u_{\text{disc}}(x) = \tanh\left(\frac{\max\left(\sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2} - 0.25, |x_3 - 0.5| - 0.05\right)}{\varepsilon}\right).$$

In Figure 5.4 two examples for the evolution of a cube are presented and in Figure 5.5 two further examples for the evolution of a disc. In comparison to h, ε and $\tilde{\tau}$, which are the same in all four cases, the choice of τ varies. In the upper lines the evolution of the cube and the disc is presented after 1 and 20 time steps of size $\tau = h^3$. In the lower lines the evolution of both objects is visualized after 1 and 10 time steps. In case of the cube $\tau = 0.03h$ while $\tau = h^2$ for the disc.

Similar to the two-dimensional examples the discrete curvature $(U^k - V^k)/\tilde{\tau}$ is displayed on a cutting through the geometry, indicated by the red line on the three-dimensional shape. As in the two-dimensional case in Figure 5.2, the flow in both examples is most concentrated in regions with high curvature, leading temporarily to concave shapes. Apart from the example of the evolving cube with outer step size $\tau = 0.03h$ this behavior is displayed in all three examples.

5.2 Comparison with a semi-implicit scheme

After we have seen some results of numerical tests we now compare our variational time discretization of Willmore flow with a semi-implicit phase field scheme presented by Du et al. in [40] and [42]. They do not treat pure Willmore flow, but three-dimensional deformations of vesicle membranes and the corresponding elastic bending energy with constant surface area and volume. Canham, Evans and Helfrich treated this type of energy first [19, 48, 56]. For a membrane $\Gamma \subset \mathbb{R}^3$ the elastic bending energy is

$$e[\Gamma] = \int_\Gamma a_1 + a_2 \left(\mathbf{h} - c_0\right)^2 + a_3 \mathbf{k} \, \mathrm{d}\mathcal{H}^2 \tag{5.2}$$

and therefore much more complicated than the pure Willmore energy.

As Du et al. prescribed a constant surface area the surface tension a_1 is constant and can be neglected. Moreover they assume the bending rigidity a_2 to be constant and further neglect Gaussian curvature $\mathbf{k}$ entirely by setting the stretching rigidity a_3 to zero. Thus they treat the simplified bending energy

$$e_s[\Gamma] = \int_\Gamma \left(\mathbf{h} - c_0\right)^2 \mathrm{d}\mathcal{H}^2, \tag{5.3}$$

where $\mathbf{h}$ is the mean curvature of the membrane Γ and c_0 its spontaneous curvature. Consequently we only have to set the spontaneous curvature to zero and neglect the constraints on the surface area and the volume to get a semi-implicit phase field scheme for Willmore flow. This simplified version of the model presented by Du et al. is described and compared with our model in this section.

5.2.1 A semi-implicit phase field scheme for Willmore flow

Following [30], the Willmore energy in a phase field model can be approximated by

$$w^\varepsilon[u] = \frac{1}{2\varepsilon}\int_\Omega \left(-\varepsilon\Delta u + \frac{1}{2\varepsilon}\Psi'(u)\right)^2 \mathrm{d}x \; . \tag{5.4}$$

The first variation of (5.4), i.e.

$$\partial_u w^\varepsilon[u](\vartheta) = \int_\Omega \frac{1}{\varepsilon}\left(\varepsilon\Delta u + \frac{2}{\varepsilon}u\left(1-u^2\right)\right)\left(\varepsilon\Delta\vartheta + \frac{2}{\epsilon}\left(1-3u^2\right)\vartheta\right)\mathrm{d}x\,, \tag{5.5}$$

can be simplified by defining $f(u) := -\left(-\varepsilon\Delta u + \frac{1}{2\varepsilon}\Psi'(u)\right) = \varepsilon\Delta u + \frac{2}{\varepsilon}u(1-u^2)$ and $g(u) := \Delta f(u) + \frac{2}{\varepsilon^2}(1-3u^2)f(u)$. With this notation the variation is

$$\partial_u w^\varepsilon[u](\vartheta) = \int_\Omega g(u)\vartheta\,\mathrm{d}x \; . \tag{5.6}$$

To derive a semi-implicit time discretization ansatz Du, Liu and Wang [40] used the variable splitting

$$g(u,v) := \frac{1}{2}\Delta\left(f(u)+f(v)\right) + \frac{1}{\varepsilon^2}(1-u^2-uv-v^2)\left(f(u)+f(v)\right) \; .$$

This treatment of nonlinear terms has been used before, for example in [41], because it allows one to derive so-called discrete energy laws.
The semi-implicit time discretization ansatz $\varepsilon\frac{u^{k+1}-u^k}{\tau} = -g(u^{k+1},u^k)$, corresponding to this splitting leads to the following system of equations in u^{k+1} and f^{k+1}

$$\begin{aligned}
\int_\Omega \varepsilon\frac{u^{k+1}-u^k}{\tau}\vartheta\,\mathrm{d}x &= \int_\Omega \frac{1}{2}\nabla\left(f^{k+1}+f^k\right)\cdot\nabla\vartheta \\
&\quad - \frac{1}{\varepsilon^2}\left(1-\left(u^{k+1}\right)^2 - u^{k+1}u^k - \left(u^k\right)^2\right)\left(f^{k+1}+f^k\right)\vartheta\,\mathrm{d}x\,, \\
\int_\Omega f^{k+1}\vartheta\,\mathrm{d}x &= \int_\Omega -\varepsilon\nabla u^{k+1}\cdot\nabla\vartheta + \frac{2}{\varepsilon}u^{k+1}\left(1-(u^{k+1})^2\right)\vartheta\,\mathrm{d}x
\end{aligned}$$

for given u^k and $f^k \equiv f(u^k)$. In [40] Du et al. used Fourier spectral methods to discretize this time-discrete scheme in space, but for a better comparability with our model we prefer piecewise Finite Elements for the approximation of u^{k+1}, u^k, f^{k+1} and f^k. With the notation as before the discretized system of equations is

$$\begin{aligned}
\frac{\varepsilon}{\tau}\mathbf{M}\left(\overline{U}^{k+1}-\overline{U}^k\right) &= \frac{1}{2}\mathbf{L}\left(\overline{F}^{k+1}+\overline{F}^k\right) - \frac{1}{\varepsilon^2}\left(\mathbf{M}-\mathbf{M}[\left(U^{k+1}\right)^2]\right)\left(\overline{F}^{k+1}+\overline{F}^k\right) \\
&\quad + \frac{1}{\varepsilon^2}\left(\mathbf{M}[U^{k+1}U^k] + \mathbf{M}[\left(U^k\right)^2]\right)\left(\overline{F}^{k+1}+\overline{F}^k\right)\,, \\
\mathbf{M}\overline{F}^{k+1} &= -\varepsilon\mathbf{L}\overline{U}^{k+1} + \frac{2}{\varepsilon}\mathbf{M}\overline{U}^{k+1} - \frac{2}{\varepsilon}\mathbf{M}[\left(U^{k+1}\right)^2]\overline{U}^{k+1}\,.
\end{aligned}$$

5.2.2 Newton method for the semi-implicit scheme

To ensure the comparability with our model we solve this nonlinear system with the Newton method described in Subsection 3.2.2. We define a function

$$\mathcal{F}[U,F;U^k,F^k] = \left(\mathcal{F}_1[U,F;U^k,F^k],\, \mathcal{F}_2[U,F;U^k,F^k]\right)^T$$

with

$$\begin{aligned}
\mathcal{F}_1[U,F;U^k,F^k] &:= \frac{\varepsilon}{\tau}\mathbf{M}\left(\overline{U}-\overline{U}^k\right) - \frac{1}{2}\mathbf{L}\left(\overline{F}+\overline{F}^k\right) + \frac{1}{\varepsilon^2}\mathbf{M}\left(\overline{F}+\overline{F}^k\right) \\
&\quad - \frac{1}{\varepsilon^2}\left(\mathbf{M}[(U)^2] + \mathbf{M}[UU^k] + \mathbf{M}[\left(U^k\right)^2]\right)\left(\overline{F}+\overline{F}^k\right)\,, \\
\mathcal{F}_2[U,F;U^k,F^k] &:= \mathbf{M}\overline{F} + \varepsilon\mathbf{L}\overline{U} - \frac{2}{\varepsilon}\mathbf{M}\overline{U} + \frac{2}{\varepsilon}\mathbf{M}[(U)^2]\overline{U}\,.
\end{aligned}$$

To set up the Newton method we need to compute the first variation of $\mathcal{F}[U,F;U^k,F^k]$ with respect to U and F.

$$\begin{aligned}
\partial_U\mathcal{F}_1[U,F;U^k,F^k](\Theta) &= \frac{\varepsilon}{\tau}\mathbf{M}\overline{\Theta} - \frac{2}{\varepsilon^2}\mathbf{M}[U\left(F+F^k\right)]\overline{\Theta} - \frac{1}{\varepsilon^2}\mathbf{M}[U^k\left(F+F^k\right)]\overline{\Theta}\,, \\
\partial_F\mathcal{F}_1[U,F;U^k,F^k](\varSigma) &= -\frac{1}{2}\mathbf{L}\overline{\varSigma} + \frac{1}{\varepsilon^2}\left(\mathbf{M} - \mathbf{M}[U^2] - \mathbf{M}[UU^k] - \mathbf{M}[\left(U^k\right)^2]\right)\overline{\varSigma}\,, \\
\partial_U\mathcal{F}_2[U,F;U^k,F^k](\varUpsilon) &= \varepsilon\mathbf{L}\overline{\varUpsilon} - \frac{2}{\varepsilon}\mathbf{M}\overline{\varUpsilon} - \frac{6}{\varepsilon}\mathbf{M}[U^2]\overline{\varUpsilon}\,, \\
\partial_F\mathcal{F}_2[U,F;U^k,F^k](\Lambda) &= \mathbf{M}\overline{\Lambda}
\end{aligned}$$

We now define $Z_i := (U_i, F_i)$ and end up with solving the linear system of equations

$$D\mathcal{F}[Z_i]^T D\mathcal{F}[Z_i]\,(Z_{i+1}-Z_i) = -D\mathcal{F}[Z_i]^T\mathcal{F}[Z_i]$$

for Z_{i+1} with the Cholesky solver from the CHOLMOD package [29, 24]. A comparison with (3.20) reveals apart from the different definition of Z_i and $\mathcal{F}$ it is the same system of equations as in the nested variational model in Chapter 3.

5.2.3 Comparison

After presenting the semi-implicit scheme we compare it with our model. Therefore we choose the same test as presented at the beginning of this chapter, i.e. we start with a quarter circle with initial radius $r_0 = 0.4$, set $\varepsilon = 2^{-6}$, time step sizes $\tau = \tilde{\tau} = h^2$ and vary the grid size from 2^{-7} to 2^{-9}. The L^2-error of this evolution computed with the semi-implicit scheme is presented in Figure 5.1, where it is plotted in the same graph as the L^2-error computed with our method. The results are qualitatively equal.
For a better comparison of both schemes we listed the required Newton iterations and the resulting computing times of both methods in Table 5.1. These results are based on the simulation of 80 time steps on a single kernel of an Intel Xeon E5530 CPU. The comparison of the iteration numbers on a grid with grid width $h = 2^{-7}$ reveals that the Newton method needs about 4 iterations per time step for our nested time discretization, whereas the semi-implicit scheme requires up to 200 iterations. Consequently the computation times are significantly longer.

numerical scheme	$h = 2^{-7}$		$h = 2^{-8}$		$h = 2^{-9}$	
	iter	time [sec]	iter	time [sec]	iter	time [sec]
nested variational	288	1366	961	20651	2562	338602
semi-implicit [42]	6176	18746	3695	121531	5776	1051560

Table 5.1: Total number of Newton iterations and computation times for the simulations from Figure 5.1.

Furthermore we observed a stable convergence of the nonlinear solver in the nested variational scheme for step sizes up to $\tau = 0.03h$ whereas the nonlinear solver in the semi-implicit scheme revealed convergence problems for step sizes of this order.
Thus based on our experiments we conclude that the new method is significantly more robust.

5.3 Application to an image restoration

In this section we apply our nested variational Willmore model to an image restoration problem. This type of problem is widespread in computer vision [4, 5, 6, 67, 86, 77]. Since the result is more important than the temporal evolution image restoration is a problem in which high accuracy in time discretization is not as important as a robust numerical descent scheme. Thus this geometric variational problem is a nice application of our nested variational time discretization, which allows large time steps. To set up an adequate model we follow the edge restoration approach by Nitzberg et al. [77]. Therefore we slightly modify our model presented in Definition 3.9 by adding an area functional to the outer energy, leading to

$$e_{\text{out}}^{\varepsilon,\eta}[u^k, u, v] = e_{\text{out}}^{\varepsilon}[u^k, u, v] + 2\tau\eta a^{\varepsilon}[u]\,, \tag{5.7}$$

where $a^{\varepsilon}[u]$ is the phase field approximation of the surface area defined in (2.3).
By minimizing this weighted sum of Willmore energy and area functional we aim to find a continuation of a given image edge in a region which was destroyed in the original image. As before the image edge is given by a phase field function u^k. In contrast to Definition 3.9 we now have to deal with C^1 boundary conditions. This leads to the following model

Definition 5.1 (Image restoration iteration)**.** *Given a phase field u^k at time $k\tau$ we define the phase field u^{k+1} at time $(k+1)\tau$ by*

$$u^{k+1} = \operatorname*{argmin}_{u \in H^1(\Omega)} e_{out}^{\varepsilon,\eta}[u^k, u, v]\,, \text{ where} \tag{5.8a}$$

$$v = \operatorname*{argmin}_{\tilde{v} \in H^1(\Omega_\delta)} e_{in}^{\varepsilon}[u, \tilde{v}]\,, \text{ and} \tag{5.8b}$$

$$u^{k+1} = u^k \quad \text{on } \partial\Omega\,. \tag{5.8c}$$

Here Ω_δ denotes a δ-neighborhood of the reconstruction region Ω.

Solving this PDE constrained variational problem leads to the problem of finding a saddle $(u,v,p) \in H^1(\Omega) \times H^1(\Omega_\delta) \times H^1(\Omega_\delta)$ with $u = u^k$ on $\partial\Omega$ of the Lagrangian

$$\begin{aligned}
\ell^{\mathrm{ir}}[u^k,u,v,p] &= e_{\mathrm{out}}^{\varepsilon,\eta}[u^k,u,v] + \partial_v e_{\mathrm{in}}^{\varepsilon}[u,v](p) \\
&= \int_\Omega \varepsilon (u-u^k)^2 \,\mathrm{d}x + \tau\varepsilon \left(\frac{v-u}{\tilde{\tau}}\right)^2 \mathrm{d}x + \tau\eta \int_\Omega \varepsilon |\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x \\
&\quad + \int_\Omega 2\varepsilon(v-u)p + \frac{\tilde{\tau}}{\varepsilon}\Psi'(v)p + 2\varepsilon\tilde{\tau}\nabla v \cdot \nabla p \,\mathrm{d}x \,.
\end{aligned}$$

As we now have to deal with different boundary conditions as in Subsection 3.2.2 we want to treat the discretization in detail again, even if it is similar to the one we have already seen. Using our standard notation for the Finite Element discretization the fully discretized Lagrangian is

$$\begin{aligned}
\mathcal{L}^{\mathrm{ir}}[U^k,U,V,P] &= \mathcal{E}_{\mathrm{out}}^{\varepsilon,\eta}[U^k,U,V] + \partial_V \mathcal{E}_{\mathrm{in}}^{\varepsilon}[U,V](P) \\
&= \varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k\right) \cdot \left(\overline{U} - \overline{U}^k\right) + \frac{\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right) \\
&\quad + 2\tau\eta A^\varepsilon[U] + 2\varepsilon\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{P} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-V^2]\overline{V} \cdot \overline{P} \\
&\quad + 2\varepsilon\tilde{\tau}\mathbf{L}\overline{V} \cdot \overline{P} \,.
\end{aligned}$$

The first variation of this Lagrangian with respect to $\overline{U}$, $\overline{V}$ and $\overline{P}$ is

$$\begin{aligned}
\partial_U \mathcal{L}^{\mathrm{ir}}[U,V,P](\Theta) &= 2\varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k\right) \cdot \overline{\Theta} - \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Theta} + 2\varepsilon\eta\tau\mathbf{L}\overline{U} \cdot \overline{\Theta} \\
&\quad - \frac{4\tau\eta}{\varepsilon}\mathbf{M}[1-U^2]\overline{U} \cdot \overline{\Theta} - 2\varepsilon\mathbf{M}\overline{P} \cdot \overline{\Theta} \,, \\
\partial_V \mathcal{L}^{\mathrm{ir}}[U,V,P](\Xi) &= \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Xi} + 2\varepsilon\mathbf{M}\overline{P} \cdot \overline{\Xi} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2]\overline{P} \cdot \overline{\Xi} + 2\varepsilon\tilde{\tau}\mathbf{L}\overline{P} \cdot \overline{\Xi} \,, \\
\partial_P \mathcal{L}^{\mathrm{ir}}[U,V,P](\varSigma) &= 2\varepsilon\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\varSigma} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-V^2]\overline{V} \cdot \overline{\varSigma} + 2\varepsilon\tilde{\tau}\mathbf{L}\overline{V} \cdot \overline{\varSigma}
\end{aligned}$$

and the components of the Hessian of the Lagrange function are

$$\begin{aligned}
\partial_U\partial_U \mathcal{L}^{\mathrm{ir}}[U,V,P] &= 2\varepsilon\left(1 + \frac{\tau}{\tilde{\tau}^2}\right)\mathbf{M} + 2\varepsilon\tau\eta\mathbf{L} - \frac{4\tau\eta}{\varepsilon}\mathbf{M}[1-3U^2] \,, \\
\partial_V\partial_U \mathcal{L}^{\mathrm{ir}}[U,V,P] &= -\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M} \,, \\
\partial_P\partial_U \mathcal{L}^{\mathrm{ir}}[U,V,P] &= -2\varepsilon\mathbf{M} \,, \\
\partial_V\partial_V \mathcal{L}^{\mathrm{ir}}[U,V,P] &= \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M} + \frac{24\tilde{\tau}}{\varepsilon}\mathbf{M}[VP] \,, \\
\partial_P\partial_V \mathcal{L}^{\mathrm{ir}}[U,V,P] &= 2\varepsilon\mathbf{M} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2] + 2\tilde{\tau}\varepsilon\mathbf{L} \,, \\
\partial_P\partial_P \mathcal{L}^{\mathrm{ir}}[U,V,P] &= 0 \,.
\end{aligned}$$

Notice, $\mathcal{F}$ and $D\mathcal{F}$ which are used in the Newton method are not just defined by $\mathcal{F}[U,V,P] = \nabla_{(U,V,P)}\mathcal{L}^{\mathrm{ir}}[U,V,P]$ and $D\mathcal{F} = D^2\mathcal{L}^{\mathrm{ir}}$ as we have seen before. At this point we have to account for the different boundary conditions. In this case the uniform simplicial mesh $\mathcal{T}$ covers the computational domain Ω_δ and I denotes the

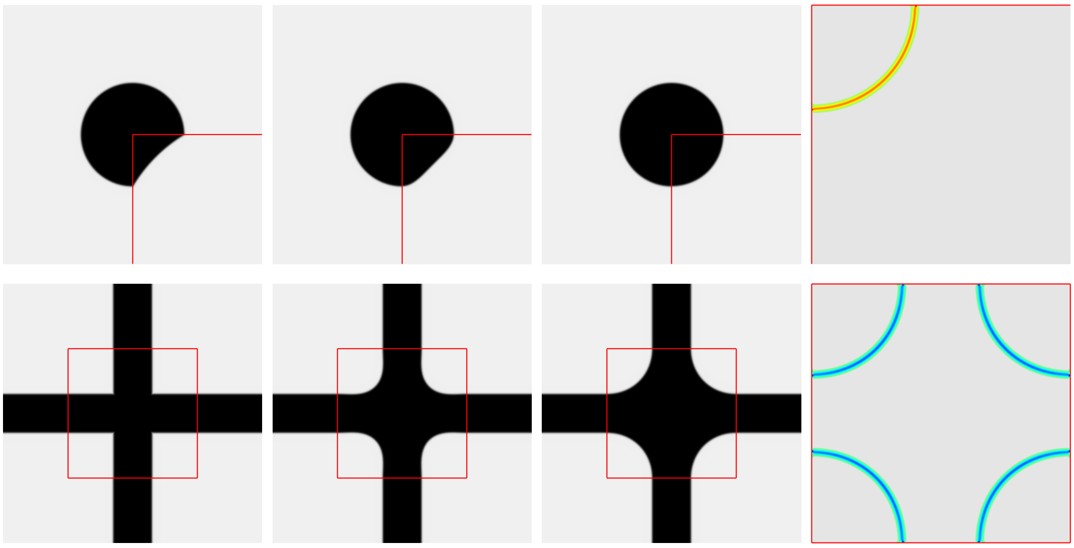


Figure 5.6: Time steps 0, 1, and 30 from two different edge restoration problems. The restoration region is a unit square, outlined in red and resolved by a $2^9 \times 2^9$ regular grid. Parameters are $\varepsilon = 4h$, $\eta = 1$, $\tau = 0.03h$, and $\tilde{\tau} = h^2$. In both rows the rightmost image shows a blowup of the restoration region at the last time step with the curvature color-coded as on the range $[-5, 5]$.

vertex index set of $\mathcal{T}$ as before. In addition to this the vertex index set of all Dirichlet nodes is denoted by $I_D \subset I$. Let $\overline{\mathcal{F}}$ being the nodal vector corresponding to the discrete vector valued function $\mathcal{F}$, than we have to set

$$\mathcal{F}_i^0 = 0 \quad \forall i \in I_D$$

and for the Hessian of the Lagrangian

$$\begin{aligned} D\mathcal{F}_{ij}^{00} &= 0 \quad \forall i \in I_D,\, j \in I,\, i \neq j\,, \\ D\mathcal{F}_{ij}^{00} &= 0 \quad \forall i \in I,\, j \in I_D,\, j \neq i\,, \\ D\mathcal{F}_{ii}^{00} &= 1 \quad \forall i \in I_D\,, \\ D\mathcal{F}_{ij}^{0k} &= 0 \quad \forall i \in I_D,\, j \in I,\, k = 1, 2\,, \\ D\mathcal{F}_{ij}^{k0} &= 0 \quad \forall i \in I,\, j \in I_D,\, k = 1, 2 \end{aligned}$$

to realize C^1 boundary conditions in U.

Results of this model are presented in Figure 5.6. In the first row we start with a damaged circle and in the second row we start with a cross as input data. In both cases the computational domain is outlined in red and resolved by a $2^9 \times 2^9$ regular grid. The parameters are $\varepsilon = 4h$, $\eta = 1$, $\tau = 0.03h$ and $\tilde{\tau} = h^2$ in both cases and the results are presented after 0, 1 and 30 iterations. As we can see this amount of iterations is enough to achieve satisfactory reconstruction results.

In a last test, presented in Figure 5.7, we vary the parameter η and keep all other parameters as before. As input data we choose a bar. These results illustrate the

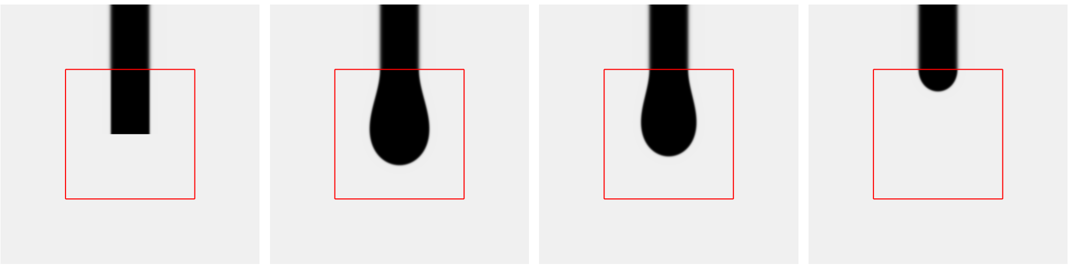


Figure 5.7: A variational C^1 edge continuation is computed via a gradient flow minimization of the sum of the Willmore energy and a weighted surface area starting from an initial phase field (left). The minimizing phase field is shown for the surface weight $\eta = 0.1$ (second), $\eta = 1$ (third), and $\eta = 10$ (fourth image).

influence of Willmore energy and surface area under different weighting. For $\eta = 0.1$ the Willmore energy is weighted higher than the surface area, thus it forms a big drop at the end of the bar. In contrast to this, if $\eta = 10$, the surface area is weighted higher than the Willmore energy such that the model shortens the length of the edge and we get a small cap at the end of the bar.

The presented results show that this model works fine for image restoration. However, due to the special properties of the Willmore energy and the absence of anisotropy it can be used for restoration of objects with C^1 boundary, only.

Chapter 6

Shape from apparent contour

In this chapter we deal with a further application of the variational time discrete Willmore model presented in Chapter 3. We infer the shape of a hypersurface from its so-called apparent contour. The corresponding model problem is presented in the first section and a survey of the basic terminology arising in literature in the second. After this we concentrate on the variational phase field model and present some numerical tests in two as well as in three dimensions.

6.1 The model problem

One application of our variational time discrete Willmore model is an inverse geometric problem. Given a smooth hypersurface $\Gamma \subset \mathbb{R}^d$ we define the corresponding apparent contour.

Definition 6.1 (Apparent contour)**.** *Let $\Gamma \subset \mathbb{R}^d$ be a $(d-1)$-dimensional hypersurface and $\pi : \mathbb{R}^d \to \mathbb{R}^{d-1}$ the parallel projection onto a $(d-1)$-dimensional viewing plane, without loss of generality assumed to be the $x_1, \ldots, x_{d-1}$-plane. Moreover let e_d be the unit normal vector to this plane. Then*

$$G(\Gamma) := \{\pi(x) \,:\, x \in \Gamma \wedge e_d \in T_x\Gamma\}$$

is the apparent contour of the hypersurface Γ.

The apparent contour separates the regions where the number of intersections between the hypersurface and the projection ray is constant. The labeling which specifies these intersection numbers is called Huffman labeling.

Definition 6.2 (Huffman labeling)**.** *Let $\Gamma \subset \Omega \subset \mathbb{R}^d$ with $\Omega = \omega \times [0,1]$ and $\omega \subset \mathbb{R}^{d-1}$ be a $(d-1)$-dimensional hypersurface and e_d the direction normal on the viewing plane. Then the Huffman labeling $\alpha = \alpha[\Gamma]$ of the hypersurface Γ is defined by*

$$\alpha[\Gamma](x) := \operatorname{card}\{t \in \mathbb{R} \,:\, x + te_d \in \Gamma\} \quad \text{for a.e. } x \in \omega\,.$$

In Figure 6.1 there is exemplarily visualized a two-dimensional torus in $\mathbb{R}^3$ and the corresponding two-dimensional Huffman labeling for projection direction e_3.
Now the aim of this chapter is to find a surface $\Gamma \in \mathbb{R}^d$ which belongs to a given $(d-1)$-dimensional Huffman labeling α^*. As we will see the solution of this problem is not unique. This can be illustrated in case of the Huffman labeling of a sphere and a cylinder with variable height, presented in Figure 6.2. The corresponding Huffman

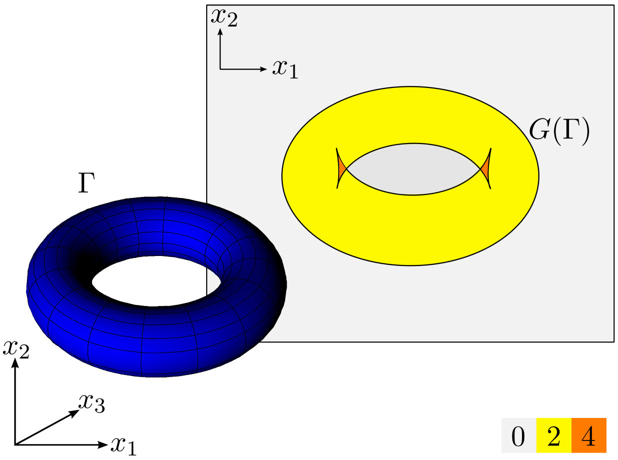

x_2
x_1
$G(\Gamma)$
Γ
x_2
x_3
x_1
0
2
4

Figure 6.1: A torus Γ in $\mathbb{R}^3$ and its apparent contour $G(\Gamma)$ as well as a color sketch of the corresponding two-dimensional Huffman labeling.

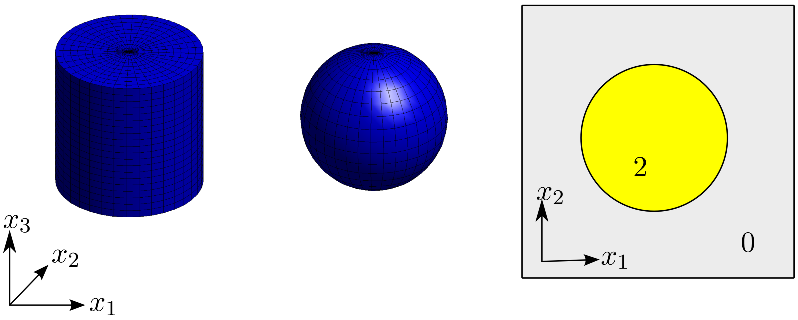

x_3
x_2
x_1
2
x_2
x_1
0

Figure 6.2: The Huffman labeling of a sphere and a cylinder can be the same.

labeling is almost everywhere the same. They distinguish from each other only on a null set, where the value is 1 in case of the sphere and ∞ in case of the cylinder. Neglecting such null sets, which is done in our variational model and in Definition 6.2, there are infinitely many surfaces which have the same Huffman labeling. To restrict the number of possible surfaces we introduce a suitable prior, in this case an energy consisting of surface plus Willmore energy, which has to be minimized by the surface whose Huffman labeling has to be equal to the given one α^* almost everywhere.

Problem 6.3 (Apparent contour inverse problem). *Let $\eta, \mu \in \mathbb{R}^{\geq 0}$ and $\alpha^* : \mathbb{R}^{d-1} \to \mathbb{R}$ be a given Huffman labeling. Find a hypersurface $\Gamma \subset \mathbb{R}^d$ minimizing the regularization energy*

$$e_{reg}[\Gamma] := \int_\Gamma \eta + \frac{\mu}{2}\boldsymbol{h}^2 \mathrm{d}\mathcal{H}^{d-1}$$

subject to $\alpha[\Gamma] = \alpha^*$ almost everywhere.

6.2 Huffman labeling in literature

The origin of the Huffman labeling presented in the previous section goes back to the paper “Impossible Objects as Nonsense Sentences” by Huffman in 1971 [58]. He dealt with the question how to differ impossible from possible two-dimensional images of three-dimensional objects. He mainly concentrated on polyhedra, but in the last part of his paper he treated smooth objects, too. The labeling he proposed discriminates between edges which have one associated surface (i.e. on the boundary of the polyhedral surface) and are marked with a single arrow “$>$” and folds having two associated surfaces and being marked with a double arrow “$\gg$”. The orientation of the arrows indicates on which side of the edge or fold the associated surface or surfaces lie. If we travel along a line in the direction the arrow points the associated surface or surfaces are to the right. Moreover each edge or fold has a depth-index counting the number of intersections until reaching the edge or fold and visible edges or folds are indicated by a solid line while broken lines stand for invisible edges or folds. In case of a two-dimensional torus in $\mathbb{R}^3$ this labeling is presented in part (a) of Figure 6.3.

Although there existed different kinds of labeling schemes [71, 96, 62] Williams dealt in “Topological Reconstruction of a Smooth Manifold-Solid from Its Occluding Contour” from 1995 [98] with the Huffman labeling, because this kind of labeling allows to treat smooth surfaces with and without boundary furthermore with visible and occluded contours. Given a labeled figure representing an occluding contour he aims to build a combinatorial model of the corresponding smooth manifold solid. Williams denotes by manifold solid an object with its interior. In comparison to this we regard the boundary of this manifold solid, only, and denote it as hypersurface. This procedure is called “paneling construction”. Thereby an occluded contour is defined as the image of points where the surface is tangent to the viewing direction. Thus the occluded contour is the same as our apparent contour defined in Definition 6.1 and it correlates to folds in Huffman’s labeling scheme. Edges as presented by Huffman do not arise in this work because Williams concentrates on smooth manifold-solids whose

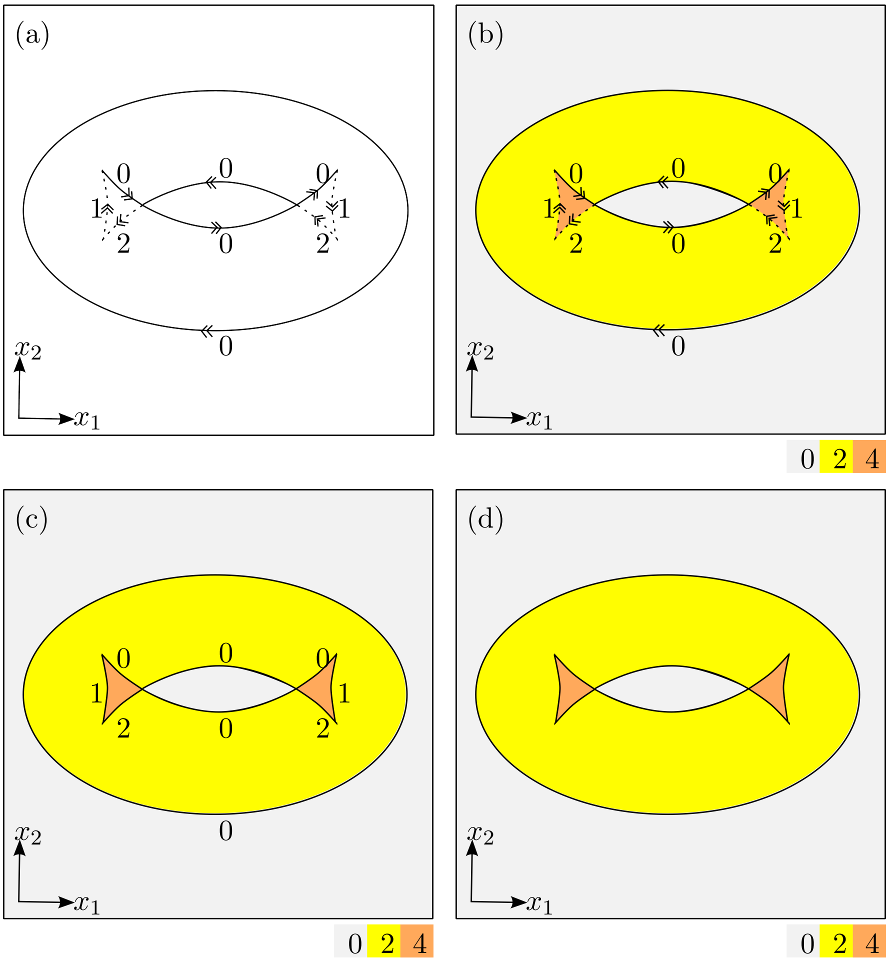

(a)
0
0
0
1
1
2
0
2
0
x_2
x_1
(b)
0
0
0
1
1
2
0
2
0
x_2
x_1
0 2 4
(c)
0
0
0
1
1
2
0
2
0
x_2
x_1
0 2 4
(d)
x_2
x_1
0 2 4

Figure 6.3: Four different versions of a two-dimensional Huffman labeling of a torus in $\mathbb{R}^3$: (a) Labeling as Huffman introduced it in [58]; (b) Huffman labeling as Williams used it in [98]; (c) Huffman labeling Bellettini et al. worked with in [9]; (d) The version of the Huffman labeling we are working with.

boundary is closed. For the paneling construction it is necessary to introduce a variable γ_R for each region R of the labeled figure. These variables have to solve a system of difference equations of the form $\gamma_A - \gamma_B = 2$ for every pair of neighboring regions A and B, when A lies to the right of B in the sense of Huffman's arrows. Remark that the solution of this system of difference equations is not unique. If $\{x_1, x_2, \dots, x_n\}$ is a solution, then $\{x_1 + c, x_2 + c, \dots, x_n + c\}$ is a solution, too. Nevertheless we can choose c such that these new variables correlate with what we are denoting by Huffman labeling. In Figure 6.3, part (b) this version of Huffman labeling is visualized for the same example of a torus, as in part (a).

In their paper "Topological and Variational Properties of a Model for the Reconstruction of Three-Dimensional Transparent Images with Self-Occlusions" from 2008 [9] Bellettini et al. worked with a further version of the Huffman labeling. In their work the Huffman labeling consists of two parts. The so-called depth ordering function defined on the arcs of the apparent contour, but not on crossings and cusps and the function $f_E(x)$ counting the number of layers of the surface in front of $x \in \Omega \subset \mathbb{R}^2$ on the projection plane. Thereby the depth ordering function is the same as the depth-index in Huffman's labeling scheme and the function $f_E(x)$ as the additional variable in the Huffman labeling Williams dealt with, thus the same as what we are denoting by Huffman labeling. The version of Huffman labeling used by Bellettini et al. is presented in part (c) of Figure 6.3. A comparison with Williams version of the Huffman labeling reveals that they only differ by the labeling of the contour with arrows. Bellettini et al. used this Huffman labeling for reconstructing a so called three-dimensional scene by a given gray level image. This three-dimensional scene consists of a three-dimensional solid and maybe layered shape. Within this problem they prove the existence of a smooth generic three-dimensional scene whose Huffman labeling equals the given one. Moreover they define a suitable equivalence relation for smooth generic three-dimensional scenes. In our notation this definition of the equivalence relation is:

Definition 6.4. *Let $E, F \subset (0,1) \times [0,1]^2$ be two finite perimeter sets. We say that E and F are equivalent, and write $E \sim F$, if there exists a homeomorphism $\Phi : (0,1) \times [0,1]^2 \to (0,1) \times [0,1]^2$ of the form $\Phi(t,x) = (\phi_x(t), x)$, such that ϕ_x is strictly increasing for any $x \in [0,1]^2$ and $\chi_E = \chi_F \circ \Phi$.*

They prove that with respect to this equivalence relation there exists a unique equivalence class whose Huffman labeling equals the given one. Moreover in a further work [10] the same authors used this kind of Huffman labeling. Within this work they deal with the completion of visible contours and the reconstruction of a three-dimensional scene by a given visible contour.

Finally we compare our version of the Huffman labeling which is visualized in part (d) of Figure 6.3 with the earlier versions. It turns out that our version is completely different from the version Huffman used originally, but it is strongly related to the version Bellettini et al. used, because it is only one of the two parts Bellettini's version of the Huffman labeling consists of.

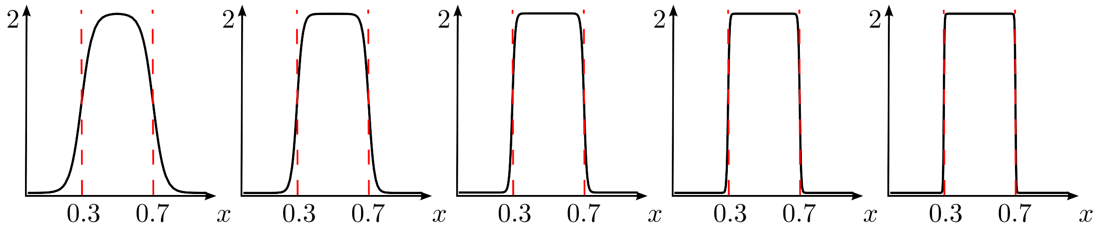


Figure 6.4: Phase field approximation $\alpha[u]$ of the one-dimensional Huffman labeling of a circle with radius $r = 0.2$ in $[0,1]^2$, centered in $(0.5, 0.5)$ and represented by the phase field function $u_c(x,y) = \tanh\left(\frac{\sqrt{(x-0.5)^2+(y-0.5)^2}-0.2}{\varepsilon}\right)$. From left to right with $\varepsilon = 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}$.

6.3 The variational phase field model

After the short excursion on different versions of Huffman labeling in the literature we derive the phase field model to the model problem presented in Section 6.1. As in the previous chapters let u be a Modica-Mortola type phase field function defined on $\Omega = [0,1]^d$ describing a surface. The Huffman labeling of this surface on the viewing plane normal to e_d can be approximated by

$$\alpha[u] := \frac{1}{2}\int_0^1 |\partial_{x_d} u| \, dx_d \,. \tag{6.1}$$

Figure 6.4 shows the function graph of the phase field approximation of the one dimensional Huffman labeling of a circle with radius $r = 0.2$ in $[0,1]^2$, centered in $(0.5, 0.5)$ and represented by the phase field function

$$u_c(x,y) = \tanh\left(\frac{\sqrt{(x-0.5)^2+(y-0.5)^2}-0.2}{\varepsilon}\right).$$

The sequence illustrates that this approximation of the Huffman labeling converges to the Huffman labeling in Definition 6.2 for $\varepsilon \to 0$.

To treat the constraint $\alpha[u] = \alpha^*$ we introduce a mismatch functional measuring the mismatch of the given Huffman labeling α^* and the Huffman labeling induced by the phase field function u.

$$j[u] := \gamma \int_{[0,1]^{d-1}} (\alpha^* - \alpha[u])^2 \, \mathrm{d}x_{1,\ldots,d-1} \tag{6.2}$$

The parameter $\gamma \in \mathbb{R}^{\geq 0}$ controls the influence of this mismatch functional which acts as penalty functional for large values of γ.

To formulate the phase field version of the apparent contour inverse Problem 6.3 we have to remember the surface energy scaled with η can be approximated by the Modica-Mortola energy (2.3) and the Willmore energy scaled with μ will be treated as in Chapter 3. That means we approximate the mean curvature by $\frac{v-u}{\tilde{\tau}}$ where v has to minimize the energy

$$e^{\varepsilon}_{\mathrm{in}}[u,v] = \varepsilon\|v-u\|^2_{L^2(\Omega)} + 2\tilde{\tau} a^{\varepsilon}[v] \,,$$

and $a^\varepsilon[v]$ (2.3) is defined by

$$a^\varepsilon[v] = \frac{1}{2}\int_\Omega \varepsilon|\nabla v|^2 + \frac{1}{\varepsilon}\Psi(v)\,\mathrm{d}x \;.$$

Then we get the phase field version of the apparent contour inverse Problem 6.3.

Problem 6.5 (Phase field apparent contour inverse problem). *Find a function $u \in H^1(\Omega)$ minimizing the penalized regularization energy*

$$e^\varepsilon_{reg,p}[u,v] = \frac{\eta}{2}\int_\Omega \frac{1}{\varepsilon}\Psi(u) + \varepsilon|\nabla u|^2\,\mathrm{d}x + \frac{\mu}{2}\int_\Omega \varepsilon\left(\frac{v-u}{\tilde{\tau}}\right)^2\mathrm{d}x \\ + \gamma\int_{[0,1]^{d-1}} \left(\alpha^* - \alpha[u]\right)^2\, dx_{1,\ldots,d-1}\,, \tag{6.3}$$

subject to

$$v = \operatorname*{argmin}_w e^\varepsilon_{in}[u,w] \\ = \operatorname*{argmin}_w \int_\Omega \varepsilon(w-u)^2 + \frac{\tilde{\tau}}{\varepsilon}\Psi(w) + \tilde{\tau}\varepsilon|\nabla w|^2\,\mathrm{d}x \;.$$

To solve this constrained variational problem we would like to apply a nested variational time discrete model as presented in Chapter 3 for Willmore flow. Therefore we would introduce an outer energy

$$e^{\varepsilon,rp}_{\mathrm{out}}[u^k,u,v] = \varepsilon\|u-u^k\|^2_{L^2(\Omega)} + 2\tau e^\varepsilon_{\mathrm{reg,p}}[u,v]\,,$$

set up the corresponding Lagrange function and consequently would have to compute the Hessian of all terms of the Lagrange function and therefore of the penalized regularization energy (6.3). Due to the special structure of the mismatch functional and the Huffman labeling $\alpha[u]$ the implementation of the Hessian of the mismatch energy requires to implement the expression

$$\int_{[0,1]^{d-1}} \frac{1}{2}\left(\int_0^1 \mathrm{sign}(\partial_{x_d}u)\partial_{x_d}\xi\,dx_d\right)\left(\int_0^1 \mathrm{sign}(\partial_{x_d}u)\partial_{x_d}\vartheta\,dx_d\right)\,dx_{1,\ldots,d-1}$$

with testfunctions ξ, $\vartheta \in H^1(\Omega)$. Thus we will circumvent the implementation of the Hessian of the mismatch functional and use an operator splitting strongly related to the Strang (also called Strang-Marchuk) operator splitting [51, 94, 72]. We split up the Problem 6.5 into the minimization of the regularization energy without mismatch functional

$$e^\varepsilon_{\mathrm{reg}}[u,v] = \frac{\eta}{2}\int_\Omega \frac{1}{\varepsilon}\Psi(u) + \varepsilon|\nabla u|^2\,\mathrm{d}x + \frac{\mu}{2}\int_\Omega \varepsilon\left(\frac{v-u}{\tilde{\tau}}\right)^2\mathrm{d}x$$

and the minimization of the mismatch functional (6.2). For the minimization of the regularization energy consisting of surface and Willmore energy we set up the Lagrange function

$$\ell^{\mathrm{ac}}[u^k,u,v,p] = \varepsilon\|u-u^k\|^2_{L^2(\Omega)} + 2\tau e^\varepsilon_{\mathrm{reg}}[u,v] + \partial_v e^\varepsilon_{\mathrm{in}}[u,v](p)\,. \tag{6.4}$$

Then we proceed analogously as in Chapter 3 and compute a saddle point $(u, v, p) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)$ of this Lagrangian by using a Newton method to solve the equation $\nabla_{(u,v,p)}\ell[u^k, u, v, p] = 0$.

The minimization of the mismatch functional $j[u]$ will be realized by an explicit gradient flow in u

$$\int_\Omega \varepsilon \left(u - u^k\right)^2 \vartheta \, \mathrm{d}x = -\gamma\tau\partial_u j[u^k](\vartheta)$$

with test function $\vartheta \in H^1(\Omega)$ and variation of the mismatch functional in u

$$\begin{aligned}
\partial_u j[u](\vartheta) &= \int_{[0,1]^{d-1}} 2\,(\alpha^* - \alpha[u])\, \partial_u \alpha[u](\vartheta)\, dx_{1,\dots,d-1} \\
&= \int_{[0,1]^{d-1}} 2\,(\alpha^* - \alpha[u]) \left(-\frac{1}{2}\int_0^1 \mathrm{sign}\,(\partial_{x_d} u)\, \partial_{x_d}\vartheta \, dx_d\right) dx_{1,\dots,d-1} \\
&= \int_\Omega (\alpha[u] \circ \pi - \alpha^* \circ \pi)\, \mathrm{sign}\,(\partial_{x_d} u)\, \partial_{x_d}\vartheta \, \mathrm{d}x \ .
\end{aligned}$$

To remember, π denotes the parallel projection from $\mathbb{R}^d$ to $\mathbb{R}^{d-1}$.

Finally we solve the phase field apparent contour inverse problem 6.5 by applying Algorithm 6.1. Thereby the initialization is analog to the initialization in Chapter 3.

Algorithm 6.1: Initialization and operator splitting for Problem 6.5

```
given: phase field function $u^0$;
given: time step size $\tau_0$;
given: maximal number of time steps $K_{max}$;
compute $v^0$ as root of $\partial_p \ell[u^0, \cdot, p] = 0$;
compute $p^0$ by solving $\partial_v \ell[u^0, v^0, p] = 0$ in $p$;
for $k = 0$ to $K_{max}$ do
    $\tau = 0.5\tau_0$;
    compute $(u^{k+\frac{1}{3}}, v^{k+\frac{1}{2}}, p^{k+\frac{1}{2}}) = \mathrm{argmin}\, \ell[u^k, u, v, p]$;
    $\tau = \tau_0$;
    compute $u^{k+\frac{2}{3}}$ by solving
        $\int_\Omega \varepsilon \left(u - u^{k+\frac{1}{3}}\right)^2 \vartheta \, \mathrm{d}x = -\gamma\tau\partial_u j[u^{k+\frac{1}{3}}](\vartheta)\,;$
    $\tau = 0.5\tau_0$;
    compute $(u^{k+1}, v^{k+1}, p^{k+1}) = \mathrm{argmin}\, \ell[u^{k+\frac{2}{3}}, u, v, p]$;
end
```

6.3.1 Finite Element discretization

The Finite Element discretization of the saddle point problem is done completely analog to the Finite Element discretization in Subsection 3.2.1.

In comparison to this we use a multilinear Finite Element approach on the underlying rectangular grid as presented in Section 2.2 to discretize the gradient flow in the mismatch functional (6.2). This is useful to simplify the numerical quadrature. For a rectangular grid with tensor product Gaussian quadrature it is easy to split up the integration over $[0,1]^d$ into one integral over$[0,1]$ and a second integral over $[0,1]^{d-1}$ or to simplify an integral like

$$\int_{[0,1]^{d-1}} f(x_1,\dots,x_{d-1})\,\mathrm{d}x_{1,\dots,d-1} \tag{6.5}$$

with

$$f(x_1,\dots,x_{d-1}) = \int_0^1 \tilde{f}(x)\,\mathrm{d}x_d\,,$$

$x \in \mathbb{R}^d$ and $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$ as mentioned in Subsection 3.2.1.
Hence, the discretized variation $\partial_u j[\cdot]$ coincides with the variation of the discretized functional $j[\cdot]$.

6.4 Numerical results

To test our shape reconstruction model we started with the sharp version of a given Huffman labeling as presented in Definition 6.2. But it turned out that we get better results if we use the approximated version (6.1) computed on the grid we are working with. Thus all tests in this section are set up as follows. In a first step we compute the phase field function representing an object we are planning to reconstruct and compute the corresponding Huffman labeling α^* by (6.1). This Huffman labeling is taken as given to reconstruct the corresponding higher dimensional object.
Moreover we use in each test a multilevel ansatz. We start with a given Huffman labeling on a coarse grid, apply Algorithm 6.1 until the computed shape and its corresponding Huffman labeling fit to the given Huffman labeling quite well and refine the grid as well as all included data, as for example the given Huffman labeling α^*. This procedure will be iterated over several grids of recursively finer grid width.

6.4.1 Two-dimensional simulations

As first test we consider the given Huffman labeling of a disc with radius $r_h = 0.25$. The computational parameters are $\varepsilon = 4h$, $\mu = 0.05$, $\gamma = 10$, $\eta = 0.05$, $\tau = 0.03h$, $\tilde{\tau} = h^2$ and we initialize u with the phase field function of a disk of radius $r_0 = 0.13$. The corresponding sequence is presented in Figure 6.5. The sharp version of the given Huffman labeling is visualized by a colored bar at the bottom. Directly above we see the function graph of the given Huffman labeling α^* (solid line) and the Huffman labeling of the current object $\alpha[u]$ (dotted line). The phase field function u of the object itself is visualized on top. The red lines are visual guides to facilitate the comparison and alignment of the different plots. During this test we computed 250, 25, 15 and 10 minimization steps on grids of mesh size $h = 2^{-m}$ for $m = 6, 7, 8, 9$. At first we observe a strong stretching, which aims to minimize the mismatch energy. Then the influence of the Willmore and surface energy gets more important such that the object evolves to an ellipse. Note that we actually do not expect the object to evolve

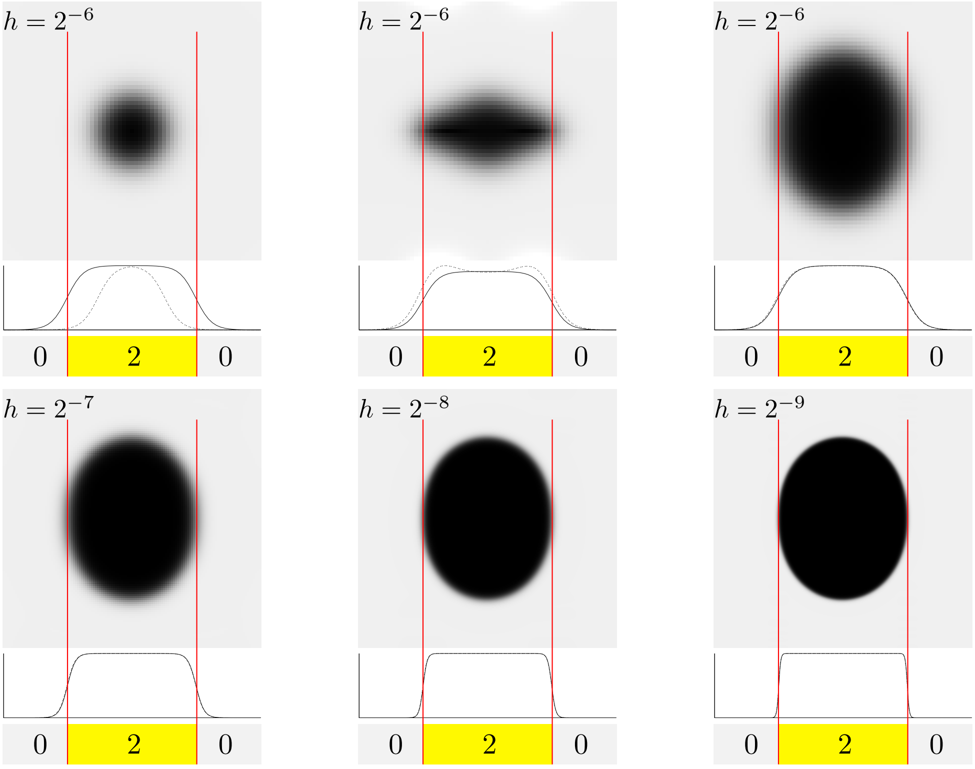


Figure 6.5: Reconstruction of an ellipse computed with a multilevel ansatz. The parameters are $\varepsilon = 4h$, $\mu = 0.05$, $\gamma = 10$ and $\eta = 0.05$ and for $\tau = 0.03h$ and $\tilde{\tau} = h^2$; the evolution is shown after 0, 1 and 250 steps on a grid with $h = 2^{-6}$ and after 25, 15 and 10 steps on grids with $h = 2^{-7}, 2^{-8}$ and 2^{-9}. The solid line represents the given Huffman labeling and the dotted line the Huffman labeling of the object in the image above. The colorbar shows the value of the sharp version of the given Huffman labeling.

to a circular disk. Due to the implicit restriction on a specified area by the Huffman labeling the object evolves qualitatively as presented in Figure 5.7 where we treated Dirichlet boundary conditions.

Within the next test we go into detail on the evolution of the single components of the energy. We define α^* as the Huffman labeling of an annulus, set $\varepsilon = 4h$, $\tau = 0.03h$, $\tilde{\tau} = h^2$, start with a grid of grid depth $h = 2^{-6}$ and end with a grid of grid depth $h = 2^{-9}$. The different parts of the energy are weighted as follows: the Willmore energy is scaled by $\mu = 0.0625$, the surface energy by $\eta = 0.125$ and the fitting energy by $\gamma = 3$. The computational results are shown in Figure 6.6, where various time steps of the descent scheme are depicted together with a plot of the decay of the different components. The visualization scheme of the Huffman labeling is the same as above. The first image shows the initial situation at the beginning of the computation on the grid of grid width $h = 2^{-6}$, while the second image shows the evolution after the second time step and the third one at the end of the computation on this grid after 40 time steps. The next two images visualize the situation at the beginning and at the end of the computation on the next finer grid with $h = 2^{-8}$, thus after 10 further time steps and the last image shows the final result on the finest grid of grid width $h = 2^{-9}$ after 60 time steps counted on all different grids. All six images are numbered and each of this numbers can be found in the graph at the top of the figure, where the single components of the energy are plotted over the number of time steps. Here one can see which time steps the single images visualize. Moreover the energy plot demonstrates what the single images indicate. During the very first time steps the fitting energy falls rapidly, which correlates with the facts that the object does not overstep the area marked by the red lines and that the Huffman labelings fit to each other quite well. Furthermore we observe that the Willmore energy being significantly bigger than all other energy parts although it is weighted with the smallest of the three weights and a jump in the energy by approximately one half for each transition to the next finer grid. The jumps are induced by a higher accuracy of the still diffuse interface on a finer grid. With respect to the additional hole in the middle the evolution fits to the evolution in the first example.

To generate two final two-dimensional tests, which are slightly more ambitious we regard a disc with two circular holes. In the one case the x_2 coordinate of the center of the two holes is the same and only the x_1 coordinate varies and in the other case the x_1 coordinate of both centers are close to each other while the x_2 coordinate differs significantly. In both cases the radii of the two holes are different from each other. In the second case the centers and the radii of the two holes are chosen in such a way that the corresponding Huffman labeling of the object takes values up to 6. A sketch of both geometries is presented in Figure 6.7. In comparison to the two previous tests we have to start the computation in these cases on a slightly finer grid with grid width $h = 2^{-7}$. In both cases we weight the Willmore energy with $\mu = 0.0625$, the surface energy with $\eta = 0.125$ and the fitting energy with $\gamma = 10$. To prevent that too big time steps τ destroy the structure of the objects we set $\tau = \tilde{\tau} = h^2$. The transition width is $\varepsilon = 4h$, as in all tests in this context.

In the first case the results in Figure 6.8 are based on 100, 80, and 30 descent steps on successively refined grids with grid size 2^{-7}, 2^{-8}, and 2^{-9}, respectively. The initial situation and the final one on the coarsest grid is visualized as well as the final object on the finest grid. The second image demonstrates why it is necessary to start on a finer grid than one with $h = 2^{-6}$ and step size $\tau = h^2$. On a coarser grid or with bigger time step size it would not be possible to dissolve this fine structure and we would get an object with one hole only. Even with this choice of grid and time step width the separation of the two holes is only weak at the end of the evolution on the grid with $h = 2^{-7}$. Nevertheless the Huffman labeling of the final object fits to the given one and apart from the effect we have just seen in the previous tests the object itself is recovered quite well.

This is completely different in the second case, whose results are presented in Figure 6.9. Here we have to remember that the mapping from object to Huffman labeling is not injective. Thus there exist Huffman labelings belonging to several different objects like for example the previously described initial one and the final object in Figure 6.9. In this case the algorithm “recovers” an object whose Huffman labeling is the given Huffman labeling, although it is not the one we started with to generate

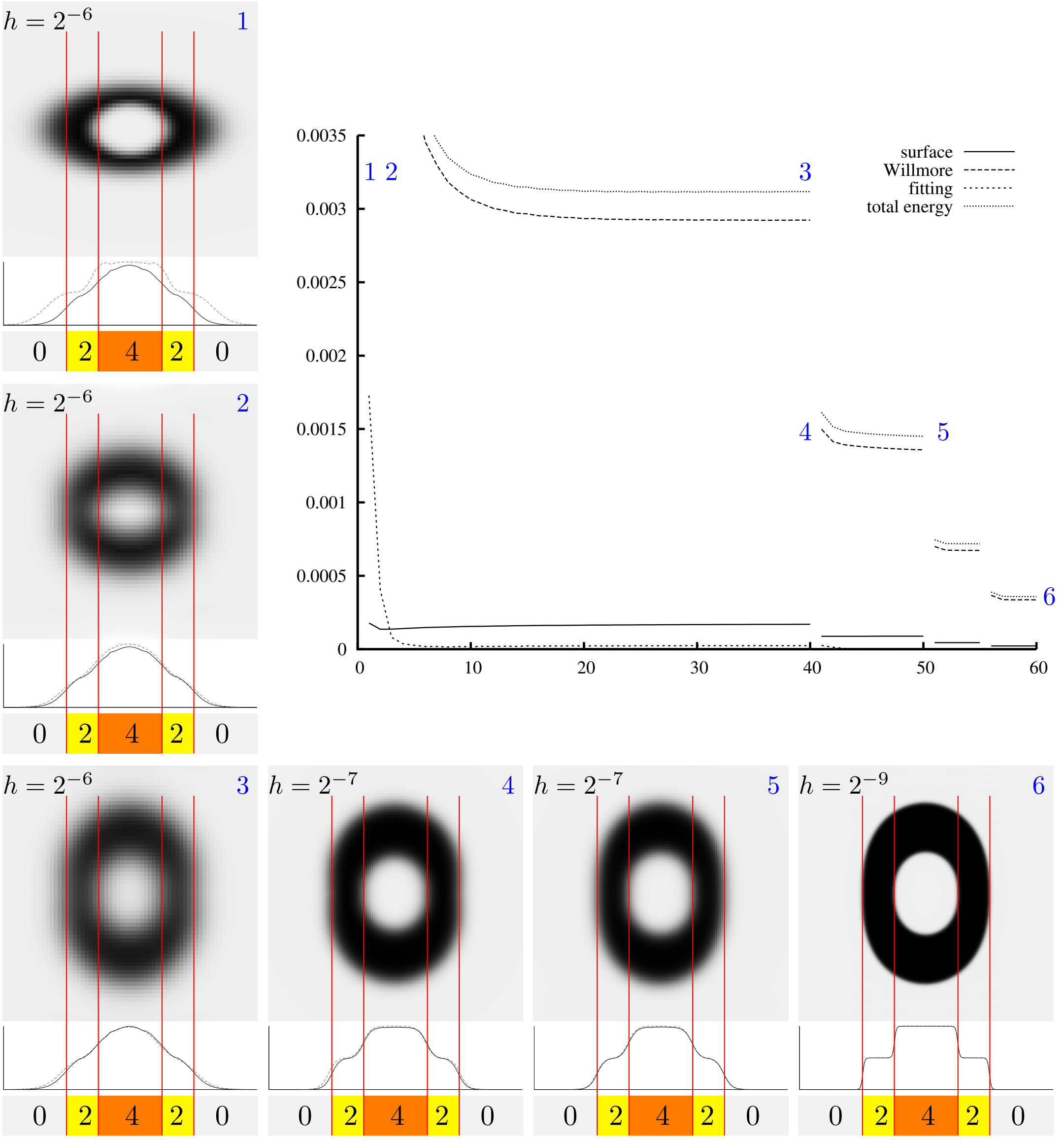


Figure 6.6: Reconstruction of an “annulus” with a multilevel ansatz, starting on a grid with $h = 2^{-6}$ and ending with $h = 2^{-9}$. Parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 3$, $\eta = 0.125$, $\tau = 0.03h$ and $\tilde{\tau} = h^2$. Different time steps and an energy decay plot, where the energy is plotted over the number of time steps, are shown.

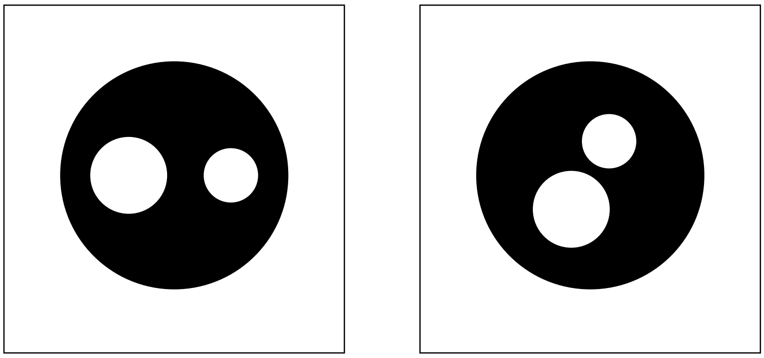


Figure 6.7: Sketch of the geometries used to generate the given Huffman labelings for the two numerical tests presented in Figure 6.8 (left) and 6.9 (right).

the “given” Huffman labeling. During this simulation we computed 100, 50 and 50 steps on grids with grid width 2^{-7}, 2^{-8} and 2^{-9} respectively.

6.4.2 Three-dimensional simulations

The two-dimensional tests are meant to test the minimization algorithm, but originally we are interested in the reconstruction of three-dimensional objects by a given two-dimensional Huffman labeling. In this higher dimensional case it is more complicated to find an adequate visualization of the given and the current Huffman labeling. In order to be able to compare both it would be nice to plot them over each other, but at the same time we want to be able to see both of them. Thus we decided to plot the jump set of the Huffman labeling or in case of the discretized Huffman labeling the 1-levelset. As we have seen at the beginning of this chapter this jump set or 1-levelset of the Huffman labeling is the apparent contour. Depending on the values the Huffman labeling takes we will use a different color coding, but this will be explained later.

The simplest three-dimensional test is the reconstruction of a sphere. Therefore we take a sphere with radius 0.3, compute the corresponding Huffman labeling and set it as given. As initial three-dimensional object we take an ellipsoid with semi-principal axes 0.15, 0.23 and 0.23. The computational parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 6$, $\eta = 0$ and $\tau = \tilde{\tau} = h^2$ and the results are presented in Figure 6.10. We computed 99 time steps on the coarsest grid with grid size $h = 2^{-6}$, 49 time steps on the next finer one and finally 9 time steps on the finest grid with $h = 2^{-9}$. The given apparent contour is visualized by a blue line and the current one by a red line. As in the two-dimensional case we observe that the minimization algorithm tries to decrease the fitting energy first by forming an object whose Huffman labeling fits to the given one. After some steps concentrating on the minimization of the fitting energy we observe the influence of the Willmore energy, which effects an evolution to a sphere. Finally we get a nice sphere whose Huffman labeling fits to the given one.

In a further step we reconstruct a torus with radii 0.25 and 0.1. The three-dimensional initial object is a cube with side length 0.5 and a hole with side length 0.235. Thereby

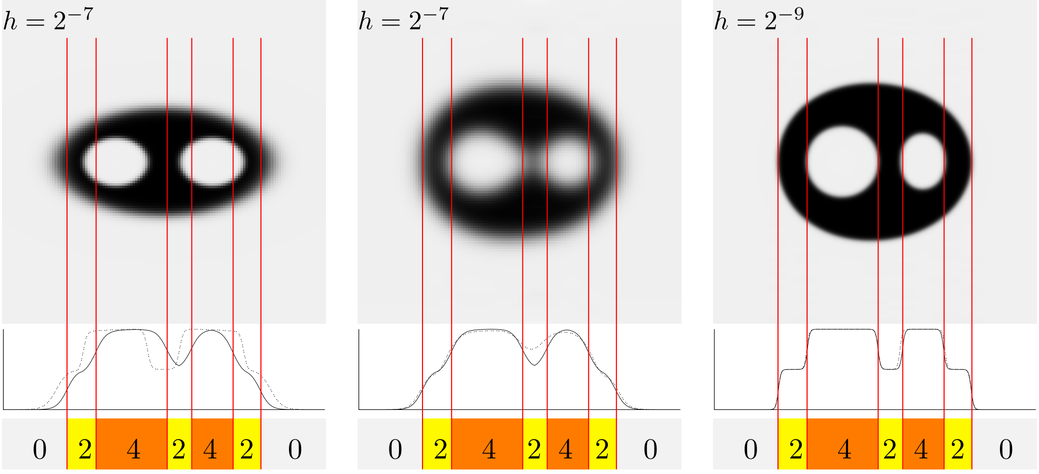


Figure 6.8: Reconstruction of a shape with two holes. The computational parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 10$, $\eta = 0.125$ and $\tau = \tilde{\tau} = h^2$. The computation bases on 100, 80 and 30 descent steps on successively refined grids with grid size 2^{-7}, 2^{-8} and 2^{-9}. The first image shows the initial situation, the second one the final evolution on the coarsest grid and the last one is the final result.

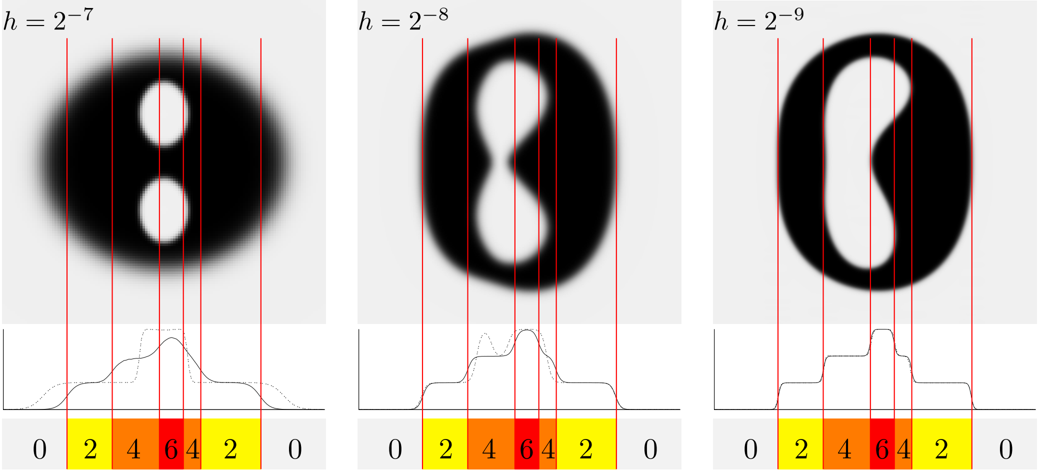


Figure 6.9: Reconstruction of a shape with a single non convex hole for a given Huffman labeling generated from a shape with two holes overlapping in the viewing direction. Parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 10$, $\eta = 0.125$ and $\tau = \tilde{\tau} = h^2$. The first two images show the initial situation on the particular grid while the last image shows the final situation after 100, 50 and 50 steps on grids with grid width 2^{-7}, 2^{-8} and 2^{-9} respectively.

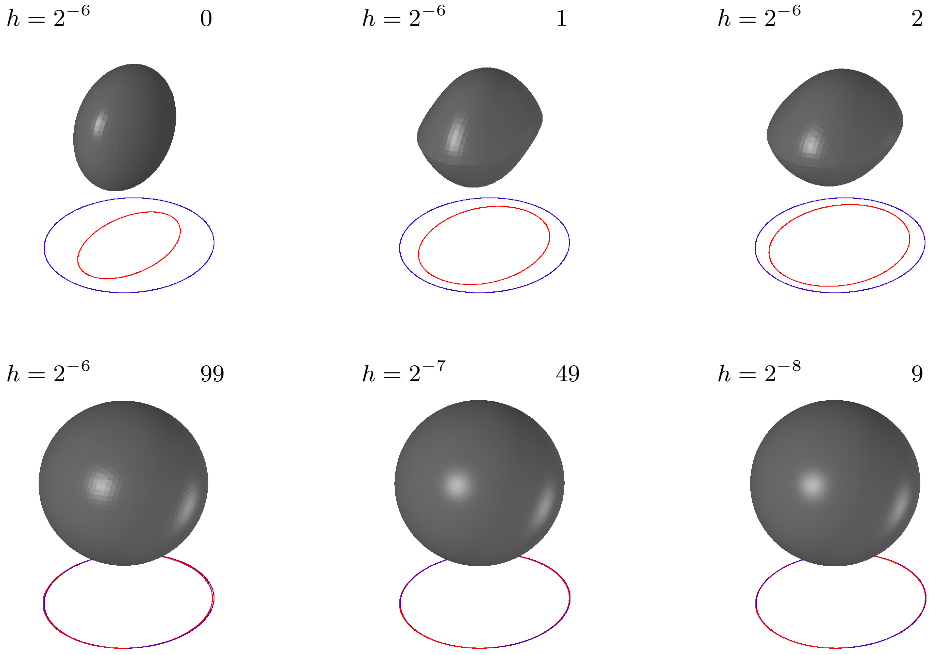


Figure 6.10: Reconstruction of a sphere for a given Huffman labeling whose 1-levelset is plotted in blue. The red line shows the 1-levelset of the Huffman labeling of the object above. The parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 6$, $\eta = 0$ and $\tau = \tilde{\tau} = h^2$. The current grid size and the current step on this grid is printed above each image.

the object is created this way that the hole points in the same direction as the hole of the torus we want to reconstruct. The weights of the single energy parts are not changed in comparison to the previous test and $\varepsilon = 4h$ and $\tau = \tilde{\tau} = h^2$ are unchanged, too. The minimization algorithm computes 99, 49 and 7 time steps on grids of successively finer grid size 2^{-6}, 2^{-7} and 2^{-8}. In Figure 6.11 we present the results. As before the given apparent contour is printed in blue and the current one in red. In case of this initial object with edges we observe the influence of the fitting and the Willmore energy during the first time step. The apparent contour of the object after the first time step indicates that the fitting energy decreases strongly during the first step. However if we compare the evolution of this object with the evolution of a rectangle under Willmore flow, cp. Figure 5.2, we clearly observe the influence of the Willmore energy which effects a strong evolution of the object around the edges and corners of the initial object. The fact that the evolution is concentrated at the edges and corners leads to a kind of dents on four sides of the object. During the next steps these dents vanish and we get a torus stretched in the projection direction until it evolves to a unstretched torus. A comparison of the apparent contours reveals that the hole of the torus is still too small, but we expect this mismatch to vanish on finer grids. At this point we do not continue the computation on finer grids, because until now we always compute on the whole grid, thus it would require a lot of memory and time. For a further discussion of this problem cp. Section 6.5

Until now we have only regarded constellations in which the Huffman labeling takes values 0 and 2. To reconstruct a torus rotated by 90° in comparison to the reconstructed torus in the previous example, we need a given Huffman labeling α^* taking values 0, 2 and 4. Apart from this rotation the characteristic parameters of the torus which we will reconstruct are the same as in Figure 6.11 and also the computational parameters like step size, transition width and weights of the single energy parts are unchanged. As initial three-dimensional object we take the initial object from the previous evolution rotated by 90° in the same direction as the torus, which we will reconstruct. In contrast to the visualization of the apparent contour in the previous figures we use the same color coding for the given and the current apparent contour in Figure 6.12. In both cases the blue line indicates the special part of the apparent contour which is equal to the set where the Huffman labeling jumps from 0 to 2. In case of the discretized Huffman labeling this is equal to the 1-levelset. Additionally the red line visualizes the part of the apparent contour which equals the set where the Huffman labeling jumps from 2 to 4 or the 3-levelset of the discretized Huffman labeling.

Within one step on the coarsest grid with $h = 2^{-6}$ the initial object evolves to an object with four handles and a hole in the middle. Thus it would be possible to put one finger in each of the four holes such that all four fingers touch each other in the middle. This object seems to be far away from a torus, but the apparent contour evolves in the right direction. One step later we get an object with a single hole pointing in the same direction the hole of the final torus shall point. After 99 time steps and therefore at the end of the evolution on the coarsest grid the object is clearly a torus, whose hole in the middle is still too small and which is slightly too big. These kinds of mismatch get smaller after 49 time steps on the next finer grid and finally 5 time steps on the finest grid with $h = 2^{-8}$. However the hole of the torus is still too

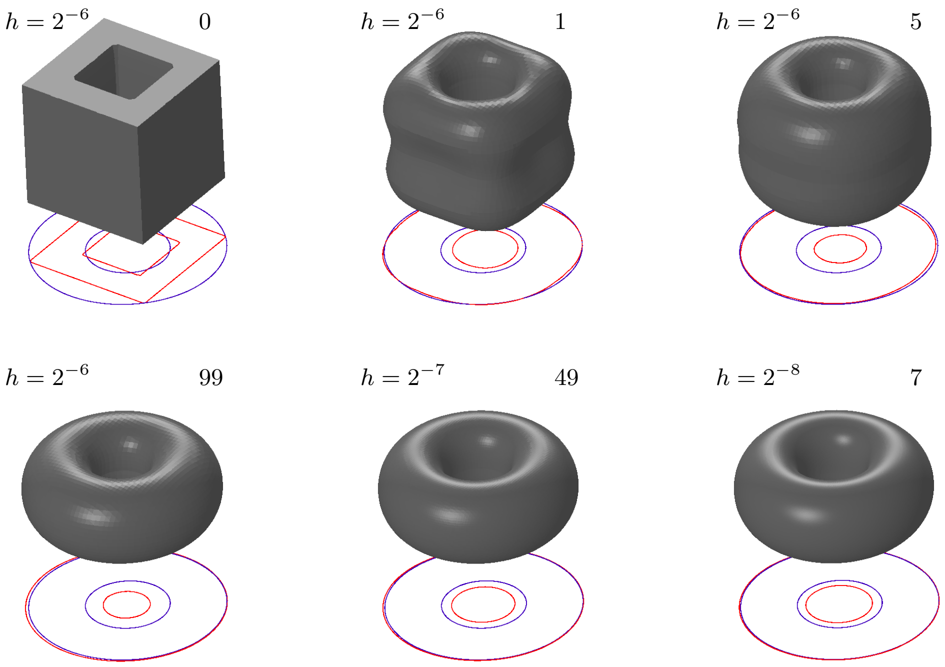


Figure 6.11: Reconstruction of a torus for a given Huffman labeling whose 1-levelset is plotted in blue. The red line shows the 1-levelset of the Huffman labeling of the object above. The parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 6$, $\eta = 0$ and $\tau = \tilde{\tau} = h^2$ and the current grid size is printed above each image as well as the current step on this grid.

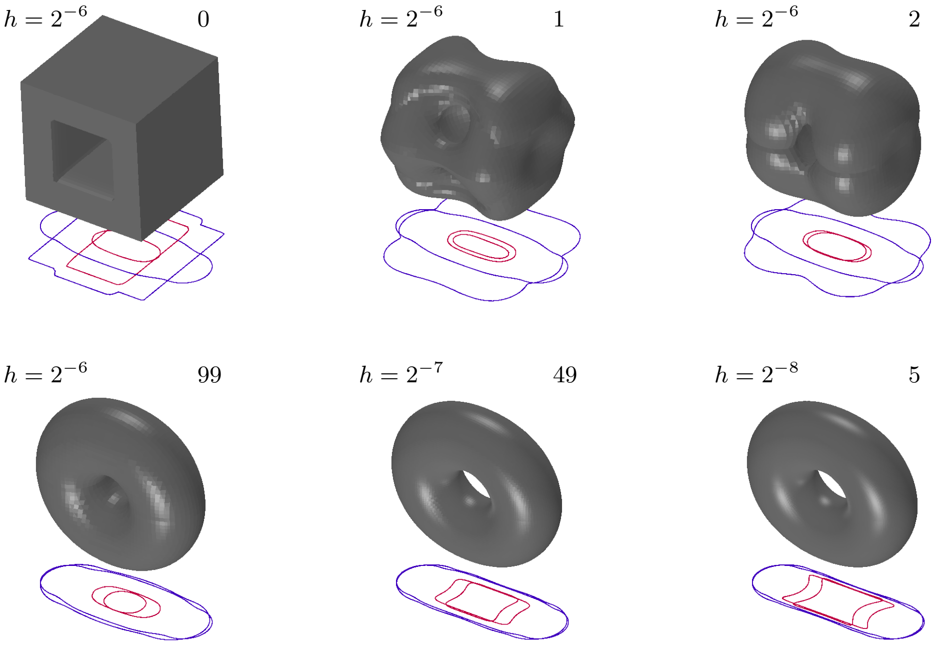


Figure 6.12: Reconstruction of a torus rotated by 90°. The blue line indicates where the Huffman labeling jumps from 0 to 2 and the blue line indicates where the Huffman labeling jumps from 2 to 4. The computational parameters are $\varepsilon = 4h$, $\mu = 0.0625$, $\gamma = 6$, $\eta = 0$ and $\tau = \tilde{\tau} = h^2$ and the evolution bases on 99, 49 and 7 time steps on successively refined grids of grid size 2^{-6}, 2^{-7} and 2^{-8} respectively.

small at the end of the evolution. Thus this result is qualitative equal to the result in case of the torus which is not rotated: the torus is reconstructed, but the hole is too small. In both cases this mismatch is expected to vanish on finer grids.

6.5 Further discussion

As we have seen in case of the torus and the rotated torus in the previous Subsection 6.4.2 the method presented in Section 6.3 works. However if we compare the given and the final Huffman labeling of the torus in Figure 6.11 and the rotated torus in Figure 6.12 there is still a mismatch.

On the one hand we could get better results by computing on finer grids. However as long as we compute on the whole grid this is not practicable. A solution would be to use a narrow band method. This ansatz is useful in case of phase field approximations because by definition there is only a small region around the interface where we observe changes. In most parts ob the computational domain the phase field function takes value 1 or -1 and stays unchanged. Thus a narrow band method would be efficient.

The use of adaptive grids would also be an alternative to realize computations on finer

grids. Since phase field approximations of an interface equal 1 or -1 in most parts of the computational domain it is reasonable to use coarse grids in these regions and grids of finer resolution around the interface.
Another problem is the fact that we decide in advance how many time steps to compute on which grid. Thus it is possible that we compute too many time steps on a special grid without further significant changes in the object. This effects a loss of time with respect to the whole computation. On the other hand it could be that it would be more effective to compute some more steps on a coarser grid instead of lots of on the finer one. This effects a loss of time, too.
Consequently we need a stopping criterion telling us the right moment to transfer the computation to the finer grid or to stop the computation completely. This stopping criterion would preferably be realized in combination with a step size control in the Newton method. Remembering Algorithm 6.1, which is an alternating one, we have to think about the kind of stopping criterion. Should it concentrate on a sufficient reduction of the fitting or the Willmore plus surface energy? We tried both versions, but a concentration on the reduction of the fitting energy involves a neglection of a sufficient reduction of the Willmore plus surface energy and vice versa. Another possibility we tried was a stopping criterion depending on the change in u from one time step to the next one. However apart from the fact that this change in u provides information on the step size, but not on the quality of the solution, it is not possible to combine it with a step size control.
A better solution would be to solve the phase field apparent contour inverse Problem 6.5 without an operator splitting. Solving this problem with the same ansatz as the problem without fitting energy would mean to introduce a Lagrange function

$$\ell^{\mathrm{ac}}[u^k, u, v, p] = \varepsilon \|u - u^k\|^2_{L^2(\Omega)} + 2\tau e^{\varepsilon}_{\mathrm{reg,p}}[u, v] + \partial_v e^{\varepsilon}_{\mathrm{in}}[u, v](p)$$

and to compute a saddle point $(u, v, p) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)$. This requires to compute the Hessian of the fitting energy. As discussed in Section 6.3 we prefer not to implement the Hessian and a way out of this is to approximate the Hessian of the fitting energy, as for example done in the BFGS method [78] and to add it to the exact Hessian of the rest of the Lagrange function. Although we have no alternating algorithm any more, we still have the difficulty, that the problem we want to solve is a minimization subject to an equality constraint, which leads in our case to a saddle point problem. For saddle point problems common step size controls based on a simple line search method do not work without modifications.

In case of unconstrained minimization problems in $\mathbb{R}^n$ we have an energy $e : \mathbb{R}^n \to \mathbb{R}$ to minimize and a line search method gets a starting point $x_k \in \mathbb{R}^n$ and a direction $p_k \in \mathbb{R}^n$ to compute a steps size τ which solves the problem

$$\min_{\tau > 0} e(x_k + \tau p_k)\,.$$

In case of constrained minimization problems

$$\min_{\mathbb{R}^n} e(x) \quad \text{subject to} \quad c(x) = 0$$

the step size control should not base on the energy alone, but also consider the constraint. Therefore we have to replace the energy e by a so-called merit function. This can be for example the l^1 merit function

$$\phi_{l^1}(x;\mu) = e(x) + \mu\|c(x)\|_{l^1} \,.$$

For further information cp. [78]. Thus in our case we have to rewrite our constraint

$$\begin{aligned} v &= \operatorname*{argmin}_w e^\varepsilon_{\mathrm{in}}[u,w] \\ &= \operatorname*{argmin}_w \int_\Omega \varepsilon(w-u)^2 + \frac{\tilde{\tau}}{\varepsilon}\Psi(w) + \tilde{\tau}\varepsilon|\nabla w|^2 \,\mathrm{d}x \end{aligned}$$

as equality constraint

$$c(u,v) := 2\varepsilon(v-u) + \frac{\tilde{\tau}}{\varepsilon}\Psi'(v) - 2\tilde{\tau}\varepsilon\Delta v = 0$$

and could work with the L^1 merit function

$$\phi_{L^1}[u^k,u,v;\mu] = \varepsilon\|u-u^k\|^2_{L^2(\Omega)} + 2\tau e^\varepsilon_{\mathrm{reg,p}}[u,v] + \mu\|c(u,v)\|_{L^1(\Omega)} \,.$$

This merit function could be used for a stopping criterion, too.

A third ansatz would be to replace the saddle point problem by a simple minimization problem. As we still have a constrained minimization problem this could be done by using an Augmented Lagrange method [78, 14]. Therefore one would introduce the Augmented Lagrange function

$$\ell^{\mathrm{ac}}_A[u^k,u,v,p;\lambda] = \varepsilon\|u-u^k\|^2_{L^2(\Omega)} + 2\tau e^\varepsilon_{\mathrm{reg,p}}[u,v] - (c[u,v],p)_{L^2(\Omega)} + \frac{\lambda}{2}\|c(u,v)\|_{L^2(\Omega)}$$

Due to the special property of the Augmented Lagrange method it remains to minimize this function in u and v.
We deal with this type of ansatz in the next chapter.

Chapter 7

Augmented Lagrange method for computing Willmore and biharmonic heat flow

In this chapter we deal with the question how to replace our saddle point problem presented in Subsection 3.2.2 by a minimization problem and induced by this how to solve the fully discrete constrained optimization problem (3.18a) and (3.18b) with the Augmented Lagrange method.
The motivation for this procedure was presented in Section 6.5, where we discussed the disadvantages of Algorithm 6.1 for solving the phase field apparent contour inverse problem 6.5. In this section we revealed the main advantage of the Augmented Lagrange method which consists in the fact that only an unconstrained minimization problem remains to solve in each Augmented Lagrange iteration. Thus we concluded that we want to solve Problem 6.5 with this method and set up the Augmented Lagrange function

$$\ell^{\mathrm{ac}}_A[u^k,u,v,p;\lambda] = \varepsilon\|u-u^k\|^2_{L^2(\Omega)} + 2\tau e^{\varepsilon}_{\mathrm{reg,p}}[u,v] - (c[u,v],p)_{L^2(\Omega)} + \frac{\lambda}{2}\|c(u,v)\|_{L^2(\Omega)}\,.$$

To keep complexity low we want to treat the Augmented Lagrange method not for the whole problem, but for the Willmore problem, as we have seen it in Chapter 3, only and as further simplification we start with the Augmented Lagrange method for the biharmonic heat equation from Chapter 4.

7.1 Augmented Lagrange method for the biharmonic heat equation

As we have seen in Chapter 4 the time-discrete biharmonic heat flow in Definition 4.1 belonging to the biharmonic heat equation

$$\partial_t u = -\Delta^2 u \quad \text{on } \Omega$$

with initial condition $u = u^0$ at $t = 0$ is the linear model problem of our nested time discretization of Willmore flow. Thus we will set up the Augmented Lagrange method for this problem, first. In contrast to the situation in Section 4.2 we regard Neumann boundary conditions instead of Dirichlet boundary conditions. This is reasonable because the theoretical results in Theorem 4.4 hold for both, thus we can test this new model and present the second kind of boundary condition at the same time. Due to the fact that we do no numerical analysis in this section we neglect the tilde-notation from Chapter 4 and write u instead of $\tilde{u}$.

7.1.1 Derivation of the model

To derive the Augmented Lagrange method for the time discrete biharmonic heat flow remember the outer and inner energy in the time-discrete biharmonic heat flow in Definition 4.1 which are given by

$$\begin{aligned} e^l_{\mathrm{out}}[u^k, u, v] &= \|u - u^k\|^2_{L^2(\Omega)} + \frac{\tau}{\tilde{\tau}^2}\|v - u\|^2_{L^2(\Omega)}\,, \\ e^l_{\mathrm{in}}[u, v] &= \|v - u\|^2_{L^2(\Omega)} + 2\tilde{\tau} a^l[v] \end{aligned}$$

with $a^l[v] := \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x$. Until now the constraint was given in the form

$$v = \operatorname*{argmin}_{\tilde{v}\in\overline{H}^1(\Omega)} e^l_{\mathrm{in}}[u, \tilde{v}]\,. \tag{7.1}$$

For the Augmented Lagrange method it is necessary to reformulate this constraint as equality constraint $c^l[u, v] = 0$ with $c^l[u, v] \in L^2(\Omega)$. Thus we compute the necessary condition for constraint (7.1) and reformulate it such that $\partial_v e^l_{\mathrm{in}}[u, v](\vartheta) = \left(c^l[u, v], \vartheta\right)_{L^2(\Omega)}$ for all test functions $\vartheta \in H^1(\Omega)$.

$$\begin{aligned} \partial_v e^l_{\mathrm{in}}[u, v](\vartheta) &= \int_\Omega 2\,(v - u)\,\vartheta \,\mathrm{d}x + 2\tilde{\tau}\int_\Omega \nabla v \cdot \nabla\vartheta \,\mathrm{d}x \\ &= \int_\Omega (2\,(v - u) - 2\tilde{\tau}\Delta_h v)\,\vartheta \,\mathrm{d}x \\ &=: 2\left(c^l[u, v], \vartheta\right)_{L^2(\Omega)} \end{aligned}$$

In the second line we denote by $\Delta_h v$ the solution $w \in L^2(\Omega)$ of

$$(w, \vartheta)_{L^2(\Omega)} = -\,(\nabla v, \nabla\vartheta)_{L^2(\Omega)} \quad \forall\vartheta \in \overline{H}^1(\Omega)$$

because, as presented in Chapter 4, u, $v \in \overline{H}^1(\Omega)$ and thus Δv is not defined in general. Moreover we eliminated the factor 2 in the equality constraint, such that we get the following version of the time-discrete biharmonic heat flow:

Definition 7.1 (Time-discrete biharmonic heat flow with equality constraint)**.** *For given $u^0 \in \overline{H}^1(\Omega)$, let the sequence $u^k \in \overline{H}^1(\Omega)$, $k \in \mathbb{N}$ be defined as the solution of*

$$\begin{aligned} u^{k+1} &= \operatorname*{argmin}_{u\in\overline{H}^1(\Omega)} e^l_{out}[u^k, u, v], \qquad \text{where} \\ c^l[u, v] &= 0\,. \end{aligned}$$

The constraint is

$$c^l[u, v] = v - u - \tilde{\tau}\Delta_h v\,. \tag{7.2}$$

As in Section 2.3 we set up the Augmented Lagrange function

$$\begin{aligned}\ell^l_A[u^k,u,v,p;\lambda] &= e^l_{\mathrm{out}}[u^k,u,v] - \left(c^l[u,v],p\right)_{L^2(\Omega)} + \frac{\lambda}{2}\|c^l[u,v]\|^2_{L^2(\Omega)}\\ &= \int_\Omega \left(u-u^k\right)^2 \mathrm{d}x + \frac{\tau}{\tilde\tau^2}\int_\Omega (v-u)^2\,\mathrm{d}x - \int_\Omega (v-u)\,p - \tilde\tau\Delta_h v\,p\,\mathrm{d}x\\ &\quad + \frac{\lambda}{2}\int_\Omega ((v-u)-\tilde\tau\Delta_h v)^2\,\mathrm{d}x\;.\end{aligned}$$

Here it is important to use a scalar product and its corresponding norm. In Section 2.3, where we presented a short survey on solving constrained minimization problems we introduced the Augmented Lagrange function with the Euclidean scalar product and the Euclidean norm, but here we have to minimize a functional over a function space, thus we have to adapt scalar product and norm. For further information cp. [59], where Ito and Kunish applied the Augmented Lagrange method to variational problems.

This Augmented Lagrange function has to be minimized in each Augmented Lagrange iteration in u and v. Therefore we have to solve the equation $\nabla_{(u,v)}\ell^l_A[u^k,u,v,p;\lambda]=0$ with

$$\begin{aligned}\partial_u\ell^l_A[u^k,u,v,p;\lambda](\vartheta) &= 2\int_\Omega \left(u-u^k\right)\vartheta\,\mathrm{d}x - \frac{2\tau}{\tilde\tau^2}\int_\Omega (v-u)\,\vartheta\,\mathrm{d}x + \int_\Omega p\vartheta\,\mathrm{d}x\\ &\quad + \lambda\int_\Omega ((v-u)-\tilde\tau\Delta_h v)\,(-\vartheta)\,\mathrm{d}x\,,\\ \partial_v\ell^l_A[u^k,u,v,p;\lambda](\xi) &= \frac{2\tau}{\tilde\tau^2}\int_\Omega (v-u)\,\xi\,\mathrm{d}x - \int_\Omega \xi p - \tilde\tau\Delta_h\xi p\,\mathrm{d}x\\ &\quad + \quad \lambda\int_\Omega ((v-u)-\tilde\tau\Delta_h v)\,(\xi-\tilde\tau\Delta_h\xi)\,\mathrm{d}x\;.\end{aligned}$$

Moreover we have to compute the Hessian with components

$$\begin{aligned}\partial_u\partial_u\ell^l_A[u^k,u,v,p;\lambda](\vartheta,\rho) &= 2\int_\Omega \rho\vartheta\,\mathrm{d}x + \frac{2\tau}{\tilde\tau^2}\int_\Omega \rho\vartheta\,\mathrm{d}x + \lambda\int_\Omega \rho\vartheta\,\mathrm{d}x\,,\\ \partial_v\partial_u\ell^l_A[u^k,u,v,p;\lambda](\vartheta,\sigma) &= -\frac{2\tau}{\tilde\tau^2}\int_\Omega \sigma\vartheta\,\mathrm{d}x - \lambda\int_\Omega (\sigma-\tilde\tau\Delta_h\sigma)\,\vartheta\,\mathrm{d}x\\ &= \partial_u\partial_v\ell^l_A[u^k,u,v,p;\lambda](\sigma,\vartheta)\,,\\ \partial_v\partial_v\ell^l_A[u^k,u,v,p;\lambda](\xi,\sigma) &= \frac{2\tau}{\tilde\tau^2}\int_\Omega \sigma\xi\,\mathrm{d}x + \lambda\int_\Omega (\sigma-\tilde\tau\Delta_h\sigma)\,(\xi-\tilde\tau\Delta_h\xi)\,\mathrm{d}x\end{aligned}$$

to apply a Trust-Region method [27]. This is the first time we use a Trust-Region method during this work. In general there exist Trust-Region methods for minimization problems as well as for saddle point problems, but the version which is already implemented in our QUOCMESH library [27, Algorithms 6.1.1 and 7.3.1-4] is meant for solving minimization problems. As we had to solve saddle point problems until now, we always decided to use a Newton method. The main conceptual difference between a Newton method with step size control and a Trust-Region method consists in the fact that one mainly uses line search techniques for step size control in Newton methods. A line search method chooses a special direction and looks for the optimal step size in this direction. In comparison to this a Trust-Region method chooses a step size and looks for the optimal direction. In [78] a short overview over the conceptional differences between both methods is given and in [27] different versions of Trust-Region methods are presented in detail. The motivation for using a Trust-Region method is to keep the linear model case as similar to the nonlinear case as possible and in the nonlinear case a Newton method with iterative solver inside does not work for all tests any more. As presented in the following sections it turns out that a Trust-Region method is the most robust method for solving the minimization problem within the Augmented Lagrange setting, especially in case of the nonlinear model.

For being able to compute biharmonic heat flow with this ansatz we need to talk about the spatial discretization, the right choice of initial values and the concrete algorithm. These questions will be treated in the next subsection.

7.1.2 A fully practical numerical scheme

To reveal a fully practical numerical scheme we have to specify the spatial discretization. As presented in Section 2.2 and used in most parts of this work we regard a regular and uniform simplicial mesh $\mathcal{T}$ covering the computational domain Ω and analogously to Section 4.1 $\overline{\mathcal{V}} := \mathcal{V} \cap \overline{H}^1(\Omega)$ denotes the Finite Element space with Neumann boundary conditions. Moreover all other notations are known from previous chapters.

Thus we can write the spatial discretized version of the outer energy and the constraint

$$\begin{aligned}\mathcal{E}^l_{\text{out}}[U^k, U, V] &= \mathbf{M}\left(\overline{U} - \overline{U^k}\right) \cdot \left(\overline{U} - \overline{U^k}\right) + \frac{\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right) , \\ C^l[U, V] &= \overline{V} - \overline{U} + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V} .\end{aligned}$$

Consequently the spatial discretized version of the Augmented Lagrange function is

$$\begin{aligned}\mathcal{L}^l_A[U^k, U, V, P; \lambda] = {} & \mathbf{M}\left(\overline{U} - \overline{U^k}\right) \cdot \left(\overline{U} - \overline{U^k}\right) + \frac{\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right) \\ & - \mathbf{M}\left(\left(\overline{V} - \overline{U}\right) + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}\right) \cdot \overline{P} \\ & + \frac{\lambda}{2}\mathbf{M}\left(\left(\overline{V} - \overline{U}\right) + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}\right) \cdot \left(\left(\overline{V} - \overline{U}\right) + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}\right) .\end{aligned}$$

Knowing this the components of the gradient of $\mathcal{L}^l_A[U^k, U, V, P; \lambda]$ can be computed straightforward

$$\begin{aligned}\partial_U \mathcal{L}^l_A[U^k, U, V, P; \lambda](\Theta) = {} & 2\mathbf{M}\left(\overline{U} - \overline{U^k}\right) \cdot \overline{\Theta} - \frac{2\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Theta} + \mathbf{M}\overline{P} \cdot \overline{\Theta} \\ & - \lambda\mathbf{M}\left(\left(\overline{V} - \overline{U}\right) + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}\right) \cdot \overline{\Theta} , \\ \partial_V \mathcal{L}^l_A[U^k, U, V, P; \lambda](\Xi) = {} & \frac{2\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{\Xi} - \mathbf{M}\overline{P} \cdot \overline{\Xi} - \tilde{\tau}\mathbf{L}\overline{P} \cdot \overline{\Xi} \\ & + \lambda\left(\mathbb{1} + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\right)^T \mathbf{M}\left(\left(\overline{V} - \overline{U}\right) + \tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}\right) \cdot \overline{\Xi} .\end{aligned}$$

For being able to apply the Trust-Region method we compute the components of the Hessian $D^2\mathcal{L}^l_A$

$$\begin{aligned}
\partial_U\partial_U\mathcal{L}^l_A[U^k,U,V,P;\lambda](\Theta,\Upsilon) &= \left(2+\frac{2\tau}{\tilde{\tau}^2}+\lambda\right)\mathbf{M}\overline{\Theta}\cdot\overline{\Upsilon}\,,\\
\partial_V\partial_U\mathcal{L}^l_A[U^k,U,V,P;\lambda](\Theta,\Sigma) &= \left(-\frac{2\tau}{\tilde{\tau}^2}-\lambda\right)\mathbf{M}\overline{\Theta}\cdot\overline{\Sigma}-\tilde{\tau}\lambda\mathbf{L}\overline{\Theta}\cdot\overline{\Sigma}\\
&= \partial_U\partial_V\mathcal{L}^l_A[U^k,U,V,P;\lambda](\Sigma,\Theta)\,,\\
\partial_V\partial_V\mathcal{L}^l_A[U^k,U,V,P;\lambda](\Xi,\varSigma) &= \left(\frac{2\tau}{\tilde{\tau}^2}+\lambda\right)\mathbf{M}\overline{\Xi}\cdot\overline{\varSigma}+2\tilde{\tau}\lambda\mathbf{L}\overline{\Xi}\cdot\overline{\varSigma}+\tilde{\tau}^2\lambda\mathbf{L}\mathbf{M}^{-1}\mathbf{L}\overline{\Xi}\cdot\overline{\varSigma}\,.
\end{aligned}$$

The algorithm for computing the time-discrete biharmonic heat flow with Augmented Lagrange method starts with an initial function $\overline{U}^0 \in \overline{\mathcal{V}}$ and the computation of $\overline{V}^0$ by solving

$$\mathbf{M}\left(\overline{V}-\overline{U}^0\right)\cdot\overline{\Theta}+\tilde{\tau}\mathbf{L}\overline{V}\cdot\overline{\Theta}=0\,.$$

All entries of $\overline{P}^0$ are set to 0.1, i.e. $\overline{P}^0 = \frac{1}{10}\overline{1}$ where $\overline{1}$ denotes the vector with 1 in each entry. The initial value of the penalty parameter is $\lambda^0_0 = 10$. Here the upper index denotes the time step and the lower index the Augmented Lagrange iteration. Finally the convergence tolerances η^* and w^* are set, before Algorithm 2.1 is applied. Regard that x in Algorithm 2.1 is $\left(\overline{U},\overline{V}\right)$ in this case. To compute the next time step we start with the results for $\overline{U}^1$, $\overline{V}^1$ and $\overline{P}^1$ and set $\lambda^1_0 = \frac{\lambda^0_k}{100^2}$, which is the value of the last but one λ in the previous call of the Augmented Lagrange Algorithm 2.1.

7.1.3 Numerical results

To test the Augmented Lagrange method for the biharmonic heat flow we have to find a good initial choice of u_0 and an analytical solution for computing the order of convergence. Therefore we look for a solution of the boundary value problem

$$\partial_t u+\Delta^2 u=0 \quad \text{on } \Omega\times[0,T] \text{ with } \partial_\nu u=\partial_\nu\Delta u=0 \text{ on } \partial\Omega\,. \tag{7.3}$$

Lemma 7.2. *Let* $\Omega=[0,1]^2$, *then*

$$u(t,x,y)=e^{-\pi^4 t}\left(\cos(\pi x)+\cos(\pi y)\right) \tag{7.4}$$

solves the Neumann boundary value problem (7.3).

Proof. First we have to compute the gradient of u in (x,y):

$$\nabla u(t,x,y)=e^{-\pi^4 t}\begin{pmatrix}-\pi\sin(\pi x)\\ -\pi\sin(\pi y)\end{pmatrix}.$$

Thus the Laplacian is given by

$$\Delta u(t,x,y)=e^{-\pi^4 t}\left(-\pi^2\cos(\pi x)-\pi^2\cos(\pi y)\right)$$

Algorithm 7.1: Augmented Lagrange method for the biharmonic heat equation problem

```
given: initial point $\overline{U}^0$;
given: convergence tolerances $\eta^*$ and $w^*$;
given: maximal number of time steps $Iter_{max}$ and Augmented Lagrange iterations $K_{max}$;
compute $\overline{V}^0$ by solving
    $$\mathbf{M}\left(\overline{V} - \overline{U}^0\right) \cdot \overline{\Theta} + \tilde{\tau}\mathbf{L}\overline{V} \cdot \overline{\Theta} = 0;$$
set initial penalty parameter $\lambda_0^0 = 10$;
set $\overline{P}_0^0 = \frac{1}{10}\overline{1}$;
for $i = 0$ to $Iter_{max}$ do
    apply Algorithm 2.1;
    $\overline{U}_0^{i+1} = \overline{U}_k^i$;
    $\overline{V}_0^{i+1} = \overline{V}_k^i$;
    $\overline{P}_0^{i+1} = \overline{P}_k^i$;
    $\lambda_0^{i+1} = \max\left(\frac{\lambda_k^i}{100^2}, 10\right)$;
end
```

and consequently

$$\nabla\Delta u(t,x,y) = e^{-\pi^4 t}\begin{pmatrix} \pi^3\sin(\pi x) \\ \pi^3\sin(\pi y) \end{pmatrix}$$

and

$$\Delta^2 u(t,x,y) = e^{-\pi^4 t}\pi^4\left(\cos(\pi x) + \cos(\pi y)\right) .$$

Thus it is easy to see that (7.4) solves the equation $\partial_t u + \Delta^2 u = 0$ on $\Omega \times [0,T]$. To proof the boundary conditions we regard the four parts of $\partial\Omega$ separately and start with the boundary from $(0,0)$ to $(1,0)$, called B_1. Here it holds

$$\partial_\nu u(t,x,0) = -\pi e^{-\pi^4 t}\begin{pmatrix} \sin(\pi x) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0\,,$$
$$\partial_\nu \Delta u(t,x,0) = \pi^3 e^{-\pi^4 t}\begin{pmatrix} \sin(\pi x) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0\,.$$

On B_2, the straight line from $(1,0)$ to $(1,1)$, there is

$$\partial_\nu u(t,1,y) = -\pi e^{-\pi^4 t}\begin{pmatrix} 0 \\ \sin(\pi y) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0\,,$$
$$\partial_\nu \Delta u(t,1,y) = \pi^3 e^{-\pi^4 t}\begin{pmatrix} 0 \\ \sin(\pi y) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0\,.$$

As $\sin(0) = \sin(\pi)$ and the normal ν only changes its sign on the opposite boundary the remaining two parts of the boundary $\partial\Omega$ can be neglected. ☐

	$\tau = h^2$, $\tilde{\tau} = h^2$		$\tau = h^2$, $\tilde{\tau} = h$		$\tau = h$, $\tilde{\tau} = h^2$		$\tau = h$, $\tilde{\tau} = h$	
k	L^2	H^1	L^2	H^1	L^2	H^1	L^2	H^1
5	1.952	1.685	0.749	0.825	0.878	0.879	0.875	0.875
6	1.988	1.406	0.870	0.902	0.966	0.966	0.996	0.996
7	2.054	1.159	0.934	0.949	1.004	1.004	1.039	1.039

Table 7.1: Experimental order of the L^2- and H^1-error for the numerical solution of the biharmonic heat equation with different time steps, computed with the Augmented Lagrange method and accuracy $w^* = \eta^* = 10^{-8}$. The order $\frac{\ln(e_{k+1}(t)) - \ln(e_k(t))}{-\ln 2}$ is shown for each transition from one grid with grid width $h = 2^{-k}$ to the next finer one. The error $e_k(t)$ was evaluated at time $t = 0.03125$ (cases with $\tau = h$) and $t = 0.00976562$ (cases with $\tau = h^2$), respectively.

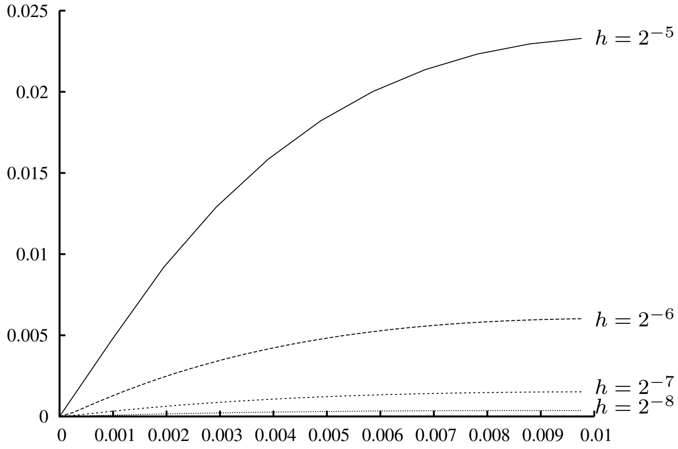


Figure 7.1: Temporal L^2-error evolution for the numerical solution of the biharmonic heat equation, computed with Augmented Lagrange method, accuracy $w^* = \eta^* = 10^{-8}$ and step sizes $\tau = \tilde{\tau} = h^2$.

Thus $u_0(x, y) = \cos(\pi x) + \cos(\pi y)$ is a good choice as an initial value. Based on this the error order of our method can be computed. We take the numerical solution of the biharmonic heat equation at time $t = 0.00976562$ for tests with step size $\tau = h^2$ and at time $t = 0.03125$ for $\tau = h$. Then we compute the error order for a transition from a grid with grid width $h = 2^{-k}$ to the next fine one by

$$\mathrm{eoc} = \frac{\ln(e_{k+1}(t)) - \ln(e_k(t))}{-\ln 2}$$

where $e_k(t)$ denotes the L^2-, respectively H^1-error, between the numerical and the analytical solution at time t. This error order is presented in Table 7.1. A comparison with Theorem 4.4 or the analytical decay rate presented in Table 4.2 reveals that the error orders from biharmonic heat flow computed with the Augmented Lagrange method agree with the analytical predictions. Only the error order for the H^1-error in the cases with $\tau = \tilde{\tau} = h$ and $\tau = h^2$, $\tilde{\tau} = h$ is better than expected.

Finally we regard the temporal L^2-error evolution for the numerical solution, computed with the Augmented Lagrange method. For $\tau = \tilde{\tau} = h^2$, $w^* = \eta^* = 10^{-8}$ and grid widths $h = 2^{-5}$, 2^{-6}, 2^{-7} and 2^{-8} it is presented in Figure 7.1. In this figure we observe the error order presented in Table 7.1 over most parts of the time. Consequently these tests confirm our new method.

7.2 Augmented Lagrange method for our Willmore flow model

As a further step we reveal the Augmented Lagrange method for our Willmore flow model. From Chapter 3 and 4 we know that both models are very similar to each other. Consequently most parts can be adapted straightforward. Nevertheless some aspects require a careful treatment.

7.2.1 Derivation of the model

In Section 4.1 we have just seen that the Willmore flow model differs from the biharmonic heat flow model by a scaling with ε and a nonlinear term in the inner energy. Since a scaling of an energy with ε does not influence the minimum we can use the same outer energy as in the linear model problem in the previous section. To keep the structure of the inner energy similar to the structure of the inner energy of the linear model problem we scale the inner energy (3.8b) with ε^{-1}, too, such that we get the following version of the nested variational time discretization of Willmore flow.

Definition 7.3 (Scaled nested variational time discretization of Willmore flow)**.** *Given a phase field u^k at time $k\tau$ define the phase field u^{k+1} at time $(k+1)\tau$ by*

$$u^{k+1} = \operatorname*{argmin}_{u \in L^2(\Omega)} e^w_{out}[u^k, u, v]\,, \text{ where} \tag{7.5a}$$

$$v = \operatorname*{argmin}_{\tilde{v} \in H^1(\Omega)} e^w_{in}[u, \tilde{v}]\,. \tag{7.5b}$$

Furthermore, we denote the solution v of the inner variational problem (7.5b) *for $u = u^{k+1}$ by v^{k+1}.*

Here the outer and inner energy are

$$e^w_{\text{out}}[u^k, u, v] = \|u - u^k\|^2_{L^2(\Omega)} + \frac{\tau}{\tilde{\tau}^2}\|v - u\|^2_{L^2(\Omega)}\,, \tag{7.6a}$$

$$e^w_{\text{in}}[u, v] = \|v - u\|^2_{L^2(\Omega)} + \frac{2\tilde{\tau}}{\varepsilon} a^\varepsilon[v] \tag{7.6b}$$

with area term $a^\varepsilon[v]$ known from (2.3).
Following the proceeding in the linear problem we have to formulate our constraint

as equality constraint $c^w[u,v] = 0$ with $c^w[u,v] \in L^2(\Omega)$.

$$\begin{aligned}\partial_v e^w_{\text{in}}[u,v](\vartheta) &= \int_\Omega 2\,(v-u)\,\vartheta \,\mathrm{d}x + \tilde{\tau}\int_\Omega 2\nabla v\cdot\nabla\vartheta + \frac{1}{\varepsilon^2}\Psi'(v)\vartheta\,\mathrm{d}x \\ &= \int_\Omega 2\left((v-u) - \tilde{\tau}\Delta_h v + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\right)\vartheta\,\mathrm{d}x \\ &=: 2\,(c^w[u,v],\vartheta)_{L^2(\Omega)}\end{aligned}$$

Consequently we get the following definition of a scaled nested variational time discretization of Willmore flow with equality constraint.

Definition 7.4 (Scaled nested variational time discretization of Willmore flow with equality constraint)**.** *Given a phase field u^k at time $k\tau$ define the phase field u^{k+1} at time $(k+1)\tau$ by*

$$u^{k+1} = \operatorname*{argmin}_{u\in L^2(\Omega)} e^w_{out}[u^k,u,v]\,,\ \text{where } v\in H^1(\Omega) \text{ solves} \tag{7.7a}$$

$$c^w[u,v] = 0\,. \tag{7.7b}$$

The constraint is

$$c^w[u,v] = v - u - \tilde{\tau}\Delta_h v + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\,. \tag{7.8}$$

Furthermore, we denote the solution v of the constraint (7.7b) *for $u = u^{k+1}$ by v^{k+1}.*

Comparing $c^w[u,v]$ in (7.8) with $c^l[u,v]$ in (7.2) demonstrates

$$c^w[u,v] = c^l[u,v] + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\,.$$

Thus we only added a nonlinear term and therefore the setup of the Augmented Lagrange method should be completely analog to the one in the linear case presented in Subsection 7.1.1. Indeed some numerical experiments revealed that it is better to weight the outer energy with ε^{-1}. This contradicts the scaling pesented at the beginning of this subsection, but as we will see in Subsection 7.2.3 it forces the Augmented Lagrange method to reach a higher precision in u. The new version of the Augmented Lagrange function is

$$\begin{aligned}\ell^w_A[u^k,u,v,p;\lambda] &= \frac{1}{\varepsilon}e^w_{\text{out}}[u^k,u,v] - (c^w[u,v],p)_{L^2(\Omega)} + \frac{\lambda}{2}\|c^w[u,v]\|^2_{L^2(\Omega)} \\ &= \int_\Omega \frac{1}{\varepsilon}\left(u-u^k\right)^2\mathrm{d}x + \frac{\tau}{\varepsilon\tilde{\tau}^2}\int_\Omega (v-u)^2\,\mathrm{d}x \\ &\quad - \int_\Omega (v-u)\,p - \tilde{\tau}\Delta_h v\,p + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)p\,\mathrm{d}x \\ &\quad + \frac{\lambda}{2}\int_\Omega\left((v-u) - \tilde{\tau}\Delta_h v + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\right)^2\mathrm{d}x\ .\end{aligned} \tag{7.9}$$

A further discussion of this scaling will be presented in Subsection 7.2.3.
The additional nonlinear term makes the computation and especially the implementation of the gradient and the Hessian of $\ell^w_A[u^k,u,v,p;\lambda]$ slightly more complicated.

The implementation will be treated in the next subsection and the spatial continuous versions of the single components of the gradient are

$$\begin{aligned}
\partial_u \ell_A^w[u^k,u,v,p;\lambda](\vartheta) &= \int_\Omega \frac{2}{\varepsilon}\left(u-u^k\right)\vartheta - \frac{2\tau}{\varepsilon\tilde{\tau}^2}\,(v-u)\,\vartheta\,\mathrm{d}x + \int_\Omega p\vartheta\,\mathrm{d}x \\
&\quad - \lambda\int_\Omega (v-u)\,\vartheta - \tilde{\tau}\Delta_h v\vartheta + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\vartheta\,\mathrm{d}x\,, \\
\partial_v \ell_A^w[u^k,u,v,p;\lambda](\xi) &= \frac{2\tau}{\varepsilon\tilde{\tau}^2}\int_\Omega (v-u)\,\xi\,\mathrm{d}x - \int_\Omega p\xi - \tilde{\tau}\Delta_h\xi p + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi''(v)p\xi\,\mathrm{d}x \\
&\quad + \lambda\int_\Omega \left((v-u) - \tilde{\tau}\Delta_h v + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\right) \\
&\qquad\qquad \left(\xi - \tilde{\tau}\Delta_h\xi + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi''(v)\xi\right)\mathrm{d}x\;.
\end{aligned}$$

Consequently the components of the Hessian are

$$\begin{aligned}
\partial_u\partial_u \ell_A^w[u^k,u,v,p;\lambda](\vartheta,\rho) &= \int_\Omega \frac{2}{\varepsilon}\vartheta\rho + \frac{2\tau}{\varepsilon\tilde{\tau}^2}\vartheta\rho\,\mathrm{d}x + \lambda\int_\Omega \vartheta\rho\,\mathrm{d}x\,, \\
\partial_v\partial_u \ell_A^w[u^k,u,v,p;\lambda](\vartheta,\varsigma) &= -\frac{2\tau}{\varepsilon\tilde{\tau}^2}\int_\Omega \vartheta\varsigma\,\mathrm{d}x - \lambda\int_\Omega \vartheta\varsigma\,\mathrm{d}x + \tilde{\tau}\lambda\int_\Omega \Delta_h\varsigma\vartheta\,\mathrm{d}x \\
&\quad - \lambda\frac{\tilde{\tau}}{2\varepsilon^2}\int_\Omega \Psi''(v)\vartheta\varsigma\,\mathrm{d}x \\
&= \partial_u\partial_v \ell_A^w[u^k,u,v,p;\lambda](\varsigma,\vartheta)\,, \\
\partial_v\partial_v \ell_A^w[u^k,u,v,p;\lambda](\xi,\varsigma) &= \frac{2\tau}{\varepsilon\tilde{\tau}^2}\int_\Omega \xi\varsigma\,\mathrm{d}x - \int_\Omega \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'''(v)p\xi\varsigma\,\mathrm{d}x \\
&\quad + \lambda\int_\Omega \left(\varsigma - \tilde{\tau}\Delta_h\varsigma + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi''(v)\varsigma\right) \\
&\qquad\qquad \left(\xi - \tilde{\tau}\Delta_h\xi + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi''(v)\xi\right)\mathrm{d}x \\
&\quad + \lambda\int_\Omega \left((v-u) - \tilde{\tau}\Delta_h v + \frac{\tilde{\tau}}{2\varepsilon^2}\Psi'(v)\right)\left(\frac{\tilde{\tau}}{2\varepsilon^2}\Psi'''(v)\xi\varsigma\right)\mathrm{d}x\;.
\end{aligned}$$

As presented in the linear case we will use a Trust-Region method for solving the minimization problem arising in each iteration of the Augmented Lagrange method.

7.2.2 A fully practical numerical scheme

For the spatial discretization we proceed as in Subsection 3.2.1 and regard a regular and uniform simplicial mesh $\mathcal{T}$ covering the computational domain Ω. Then the Finite Element space of continuous, piecewise affine functions on this mesh is denoted by $\mathcal{V} := \{\Phi \in C^0(\Omega) : \Phi|_T \in \mathcal{P}_1 \forall T \in \mathcal{T}\}$. Following the notation presented in Section 2.2 the discrete version of the Augmented Lagrange function $\ell_A^w[u^k,u,v,p;\lambda]$ (7.9) is

$$\begin{aligned}\mathcal{L}_A^w[U^k,U,V,P;\lambda] = {}& \frac{1}{\varepsilon}\mathbf{M}\left(\overline{U}-\overline{U^k}\right)\cdot\left(\overline{U}-\overline{U^k}\right) + \frac{\tau}{\varepsilon\tilde{\tau}^2}\mathbf{M}\left(\overline{V}-\overline{U}\right)\cdot\left(\overline{V}-\overline{U}\right)\\ & - \mathbf{M}\left(\left(\overline{V}-\overline{U}\right)+\tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}+\frac{\tilde{\tau}}{2\varepsilon^2}\overline{\Psi'(V)}\right)\cdot\overline{P}\\ & + \frac{\lambda}{2}\mathbf{M}\left(\left(\overline{V}-\overline{U}\right)+\tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}+\frac{\tilde{\tau}}{2\varepsilon^2}\overline{\Psi'(V)}\right)\\ & \cdot\left(\left(\overline{V}-\overline{U}\right)+\tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}+\frac{\tilde{\tau}}{2\varepsilon^2}\overline{\Psi'(V)}\right).\end{aligned}$$

Taking into account that the mass matrix $\mathbf{M}$ as well as the weighted mass matrix $\mathbf{M}[\lambda]$ with $\lambda : \Omega \to \mathbb{R}$ being an arbitrary continuous weight function are lumped matrices and therefore diagonal matrices and $\Psi'(v) = -4v\left(1-v^2\right)$ we get

$$\begin{aligned}\partial_V\left(\mathbf{M}\overline{\Psi'(V)}\cdot\overline{P}\right)(\Theta) &= -4\mathbf{M}\mathbf{D}[1-3V^2]\overline{\Theta}\cdot\overline{P}\\ &= -4\mathbf{M}[1-3V^2]\overline{\Theta}\cdot\overline{P}\\ &= -4\mathbf{M}[1-3V^2]\overline{P}\cdot\overline{\Theta}.\end{aligned}$$

Here

$$\mathbf{D}[\lambda] := \left(\delta_{ij}\lambda_i\right)_{i,j}$$

with $\lambda_i = \lambda(x_i)$ and therefore $\mathbf{M}\mathbf{D}[\lambda] = \mathbf{M}[\lambda]$. Consequently the components of the gradient of the Lagrange function $\mathcal{L}_A^w[U^k,U,V,P;\lambda]$ are

$$\begin{aligned}\partial_U\mathcal{L}_A^w[U^k,U,V,P;\lambda](\Theta) = {}& \frac{2}{\varepsilon}\mathbf{M}\left(\overline{U}-\overline{U^k}\right)\cdot\overline{\Theta} - \frac{2\tau}{\varepsilon\tilde{\tau}^2}\mathbf{M}\left(\overline{V}-\overline{U}\right)\cdot\overline{\Theta} + \mathbf{M}\overline{P}\cdot\overline{\Theta}\\ & - \lambda\mathbf{M}\left(\left(\overline{V}-\overline{U}\right)+\tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}+\frac{\tilde{\tau}}{2\varepsilon^2}\overline{\Psi'(V)}\right)\cdot\overline{\Theta},\\ \partial_V\mathcal{L}_A^w[U^k,U,V,P;\lambda](\Xi) = {}& \frac{2\tau}{\varepsilon\tilde{\tau}^2}\mathbf{M}\left(\overline{V}-\overline{U}\right)\cdot\overline{\Xi} - \mathbf{M}\overline{P}\cdot\overline{\Xi} - \tilde{\tau}\mathbf{L}\overline{P}\cdot\overline{\Xi}\\ & + \frac{2\tilde{\tau}}{\varepsilon^2}\mathbf{M}[1-3V^2]\overline{P}\cdot\overline{\Xi} + \lambda\left(\mathbb{1}+\tilde{\tau}\mathbf{L}\mathbf{M}^{-1}-\frac{2\tilde{\tau}}{\varepsilon^2}\mathbf{D}[1-3V^2]\right)\\ & \mathbf{M}\left(\left(\overline{V}-\overline{U}\right)+\tilde{\tau}\mathbf{M}^{-1}\mathbf{L}\overline{V}+\frac{\tilde{\tau}}{2\varepsilon^2}\overline{\Psi'(V)}\right)\cdot\overline{\Xi}.\end{aligned}$$

Now the computation of the discretized versions of $\partial_u\partial_u\ell_A^w[u^k,u,v,p;\lambda](\vartheta,\varsigma)$ and $\partial_v\partial_u\ell_A^w[u^k,u,v,p;\lambda](\vartheta,\upsilon)$ can be done straightforward and leads to

$$\begin{aligned}\partial_U\partial_U\mathcal{L}_A^w[U^k,U,V,P;\lambda](\Theta,\Sigma) &= \left(\frac{2}{\varepsilon}+\frac{2\tau}{\varepsilon\tilde{\tau}^2}+\lambda\right)\mathbf{M}\overline{\Theta}\cdot\overline{\Sigma},\\ \partial_V\partial_U\mathcal{L}_A^w[U^k,U,V,P;\lambda](\Theta,\Upsilon) &= \left(-\frac{2\tau}{\varepsilon\tilde{\tau}^2}-\lambda\right)\mathbf{M}\overline{\Theta}\cdot\overline{\Upsilon} - \tilde{\tau}\lambda\mathbf{L}\overline{\Theta}\cdot\overline{\Upsilon}\\ &\quad + \frac{2\lambda\tilde{\tau}}{\varepsilon^2}\mathbf{M}[1-3V^2]\overline{\Theta}\cdot\overline{\Upsilon}.\end{aligned}$$

The computation of the discretized version of $\partial_v\partial_v\ell^w_A[u^k,u,v,p;\lambda](\xi,\pi)$ is more complicated such that we do it step by step. To simplify it we rewrite the first variation of the Augmented Lagrange function in V in direction Ξ

$$
\begin{aligned}
&\partial_V\mathcal{L}^w_A[U^k,U,V,P;\lambda](\Xi)\\
&= \frac{2\tau}{\varepsilon\tilde\tau^2}\mathbf{M}\left(\overline{V}-\overline{U}\right)\cdot\overline{\Xi} - \mathbf{M}\overline{P}\cdot\overline{\Xi} - \tilde\tau\mathbf{L}\overline{P}\cdot\overline{\Xi} + \frac{2\tilde\tau}{\varepsilon^2}\mathbf{M}[1-3V^2]\overline{P}\cdot\overline{\Xi} + \lambda\mathbf{M}\left(\overline{V}-\overline{U}\right)\cdot\overline{\Xi}\\
&\quad + \lambda\tilde\tau\mathbf{L}\overline{V}\cdot\overline{\Xi} + \lambda\frac{\tilde\tau}{2\varepsilon^2}\mathbf{M}\overline{\Psi'(V)}\cdot\overline{\Xi} + \lambda\tilde\tau\mathbf{L}\left(\overline{V}-\overline{U}\right)\cdot\overline{\Xi} + \lambda\tilde\tau^2\mathbf{L}\mathbf{M}^{-1}\mathbf{L}\overline{V}\cdot\overline{\Xi}\\
&\quad + \frac{\lambda\tilde\tau^2}{2\varepsilon^2}\mathbf{L}\mathbf{M}^{-1}\mathbf{M}\overline{\Psi'(V)}\cdot\overline{\Xi} - \lambda\frac{2\tilde\tau}{\varepsilon^2}\mathbf{D}[1-3V^2]\mathbf{M}\left(\overline{V}-\overline{U}\right)\cdot\overline{\Xi}\\
&\quad - \lambda\frac{2\tilde\tau^2}{\varepsilon^2}\mathbf{D}[1-3V^2]\mathbf{M}\mathbf{M}^{-1}\mathbf{L}\overline{V}\cdot\overline{\Xi} - \lambda\frac{\tilde\tau^2}{\varepsilon^4}\mathbf{D}[1-3V^2]\mathbf{M}\overline{\Psi'(V)}\cdot\overline{\Xi}\,.
\end{aligned}
$$

Moreover it is helpful to remember that $\mathbf{M}[UV]\overline{P} = \mathbf{M}[UP]\overline{V}$, which can be easily verified

$$
\begin{aligned}
\mathbf{M}[UV]\overline{P} &= \mathbf{M}\mathbf{D}[UV]\overline{P}\\
&= \mathbf{M}\left(\sum_j \delta_{ij}U_iV_iP_j\right)_i\\
&= \mathbf{M}\left(U_iV_iP_i\right)_i\\
&= \mathbf{M}\left(\sum_j \delta_{ij}U_iP_iV_j\right)_i\\
&= \mathbf{M}\mathbf{D}[UP]\overline{V}\\
&= \mathbf{M}[UP]\overline{V}\,.
\end{aligned}
$$

Based on this we compute

$$
\begin{aligned}
&\partial_V\partial_V\mathcal{L}^w_A[U^k,U,V,P;\lambda](\overline{\Xi},\overline{\mathit{\Pi}})\\
&= \left(\frac{2\tau}{\varepsilon\tilde\tau^2}+\lambda\right)\mathbf{M}\overline{\Xi}\cdot\overline{\mathit{\Pi}} - \frac{12\tilde\tau}{\varepsilon^2}\mathbf{M}[VP]\overline{\Xi}\cdot\overline{\mathit{\Pi}} - \frac{4\tilde\tau\lambda}{\varepsilon^2}\mathbf{M}[1-3V^2]\overline{\Xi}\cdot\overline{\mathit{\Pi}} + 2\tilde\tau\lambda\mathbf{L}\overline{\Xi}\cdot\overline{\mathit{\Pi}}\\
&\quad + \frac{4\tilde\tau^2}{\varepsilon^4}\lambda\mathbf{M}[(1-3V^2)^2]\overline{\Xi}\cdot\overline{\mathit{\Pi}} - \frac{2\tilde\tau^2\lambda}{\varepsilon^2}\mathbf{M}[1-3V^2]\mathbf{M}^{-1}\mathbf{L}\overline{\Xi}\cdot\overline{\mathit{\Pi}}\\
&\quad - \frac{2\tilde\tau^2\lambda}{\varepsilon^2}\mathbf{L}\mathbf{M}^{-1}\mathbf{M}[1-3V^2]\overline{\Xi}\cdot\overline{\mathit{\Pi}} + \tilde\tau^2\lambda\mathbf{L}\mathbf{M}^{-1}\mathbf{L}\overline{\Xi}\cdot\overline{\mathit{\Pi}} + \frac{12\tilde\tau\lambda}{\varepsilon^2}\mathbf{M}[(V-U)V]\overline{\Xi}\cdot\overline{\mathit{\Pi}}\\
&\quad - \frac{24\tilde\tau^2}{\varepsilon^4}\lambda\mathbf{M}[(1-V^2)V^2]\overline{\Xi}\cdot\overline{\mathit{\Pi}} - \frac{12\tilde\tau^2\lambda}{\varepsilon^2}\mathbf{M}[V\Delta_hV]\overline{\Xi}\cdot\overline{\mathit{\Pi}}\,.
\end{aligned}
$$

Finally, to set up the Augmented Lagrange algorithm for our Willmore flow model we have to deal with the computation of the initial value $\overline{V}^0$. Apart from the different Augmented Lagrange function this is the only difference of Algorithm 7.1 to the new one. But the computation of the initial value $\overline{V}^0$ can be done analogously as in Algorithm 3.1.

Algorithm 7.2: Augmented Lagrange method for our Willmore flow model

```
given: initial point $\overline{U}^0$;
given: convergence tolerances $\eta^*$ and $w^*$;
given: maximal number of time steps $Iter_{max}$ and Augmented Lagrange iterations $K_{max}$;
compute $\overline{V}^0$ by solving
    $$\mathbf{M}\left(\overline{V}-\overline{U}^0\right)\cdot\overline{\Theta}-\tilde{\tau}\mathbf{L}\overline{V}\cdot\overline{\Theta}+\frac{\tilde{\tau}}{2\varepsilon^2}\mathbf{M}\overline{\Psi'(V)}=0;$$
set initial penalty parameter $\lambda_0^0 = 10$;
set $\overline{P}_0^0 = \frac{1}{10}\overline{1}$;
for $i = 0$ to $Iter_{max}$ do
    apply Algorithm 2.1;
    $\overline{U}_0^{i+1} = \overline{U}_k^i$;
    $\overline{V}_0^{i+1} = \overline{V}_k^i$;
    $\overline{P}_0^{i+1} = \overline{P}_k^i$;
    $\lambda_0^{i+1} = \max\left(\frac{\lambda_k^i}{100^2}, 10\right)$;
end
```

In Algorithm 2.1, which is applied in Algorithm 7.2, one has to find an approximate minimizer of the Augmented Lagrange function $\ell_A(\cdot, p^k, \lambda_k)$. The index k denotes the number of Augmented Lagrange iterations. Since k is the index denoting the time step in this chapter we will denote the index counting the Augmented Lagrange iteration with j. Thus in this special context the problem equals the problem of finding an approximate minimizer of the Augmented Lagrange function $\ell_A^w[u^k, \cdot, \cdot, p^j; \lambda_j]$. This can be done by a Newton method with an iterative or direct solver inside or with a Trust-Region method. In the following subsection we present some results computed with a Newton method with preconditioned biconjugate gradient method within the Augmented Lagrange method and some with Trust-Region method. This originates from the fact that we started with a Newton method combined with an iterative solver, which works fine in some cases, but is not stable enough for all arising test cases. Thus we tested a Newton method combined with direct solver revealing the same lack of stability. The method working stable in all cases is the Trust-Region method. It only leads to problems in three dimensions needing too much main memory. We will go into detail on this problem in the following subsection.

7.2.3 Numerical results

Following the procedure in Section 5.1 we start to test our new method by analyzing the evolution of a quarter circle under Willmore flow, computed with this method. For being able to compare the results with the results from Section 5.1 we choose the same setup. Thus we start with the double well function representing a quarter circle of initial radius $r_0 = 0.4$, set the interface transition thickness to $\varepsilon = 2^{-6} = 0.015625$

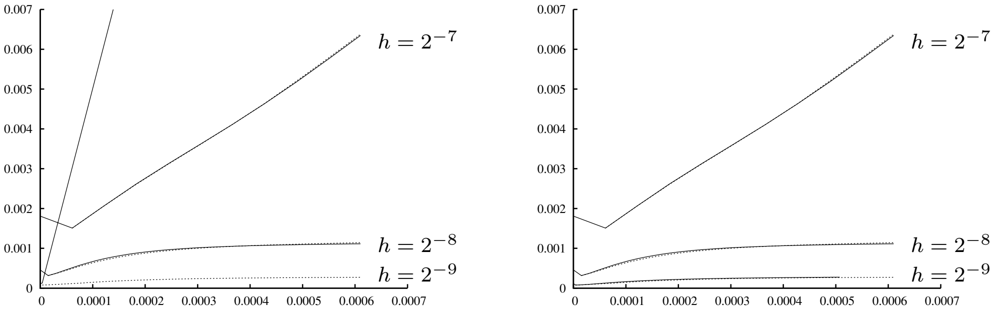


Figure 7.2: Temporal L^2-error evolution for the numerical solution of the Willmore flow, computed with Augmented Lagrange method. The parameters are $\varepsilon = 2^{-6}$, $\tau = \tilde{\tau} = h^2$ and for the left image as well as for $h = 2^{-7}$, 2^{-8} in the right image $w^* = 10^{-8}$, $\eta^* = 10^{-13}$. For $h = 2^{-9}$ in the right image we set $w^* = 10^{-10}$. The dotted line presents the L^2-error computed with the model presented in Chapter 3 and the solid one the L^2-error computed with the Augmented Lagrange method.

and the step sizes $\tau = \tilde{\tau} = h^2$. To guarantee a good approximation of the mean curvature we have to enforce a high precision in the constraint $c^w[u, v] = 0$ which we get by setting $\eta^* = 10^{-13}$. Moreover we start with the unscaled version of the Augmented Lagrange function

$$\ell_A^w[u^k, u, v, p; \lambda] = \cancel{\frac{1}{\varepsilon}} e_{\text{out}}^w[u^k, u, v] - (c^w[u, v], p)_{L^2(\Omega)} + \frac{\lambda}{2} \|c^w[u, v]\|_{L^2(\Omega)}^2$$

which is more intuitive when coming from the linear model problem. The second stopping criterion which measures the L^2-norm of the gradient of the Augmented Lagrange function $\|\nabla \ell_A^w\|_{L^2(\Omega)} \leq w^*$ is $w^* = 10^{-8}$. Additionally we still use a Newton method with preconditioned biconjugate gradient method as iterative solver inside. Later it turns out that this solver is not sufficient for all test cases, but for this first test it works. The result is presented in the left image of Figure 7.2. Thereby the solid line shows the L^2-error computed with the Augmented Lagrange method presented in this chapter, while the dotted line presents the L^2-error computed with the method presented in Chapter 3. On the two grids with grid widths $h = 2^{-7}$ and $h = 2^{-8}$ the L^2-errors computed with both methods are qualitatively the same, but on the finest grid with $h = 2^{-9}$ the L^2-error computed with the Augmented Lagrange method grows rapidly.

This problem on the finest grid can be solved by choosing a finer stopping criterion $w^* = 10^{-10}$. The result of the corresponding test is presented in the right image of Figure 7.2 and demonstrates a good accordance with the L^2-error computed with the old method from Chapter 3.

A comparison of the L^2-norm of the single components of the gradient of the Lagrange, respectively Augmented Lagrange function, after the first time step reveals the reason

$\Vert\partial_u \ell(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$	$= 1.339 \cdot 10^{-16}$	$\Vert\partial_u \ell_A^w(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$	$= 3.150 \cdot 10^{-16}$
$\Vert\partial_v \ell(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$	$= 8.288 \cdot 10^{-9}$	$\Vert\partial_v \ell_A^w(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$	$= 8.288 \cdot 10^{-9}$
$\Vert\partial_p \ell(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$	$= 1.121 \cdot 10^{-11}$	$\Vert c^w(u^1,v^1)\Vert_{L^2(\Omega)}$	$= 4.631 \cdot 10^{-8}$
$\Vert\partial_u \ell(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$	$= 3.719 \cdot 10^{-9}$	$\Vert\partial_u \ell_A^w(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$	$= 3.719 \cdot 10^{-9}$
$\Vert\partial_v \ell(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$	$= 8.354 \cdot 10^{-9}$	$\Vert\partial_v \ell_A^w(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$	$= 8.354 \cdot 10^{-9}$
$\Vert\partial_p \ell(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$	$= 4.096 \cdot 10^{-18}$	$\Vert c^w(u_A^1,v_A^1)\Vert_{L^2(\Omega)}$	$= 1.684 \cdot 10^{-9}$

Table 7.2: Single components of $\Vert\nabla_{(u,v,p)}\ell\Vert_{L^2(\Omega)}$ and $\Vert\nabla_{(u,v)}\ell_A^w\Vert_{L^2(\Omega)}$ evaluated after one time step in the evolution of a quarter circle with grid width $h = 2^{-6}$, $\tau = 0.03h$, $\tilde{\tau} = h^2$, $\varepsilon = 0.0625$, $w^* = 10^{-8}$ and $\eta* = 10^{-13}$. (u^1,v^1,p^1) denotes the solution computed with the Lagrange method from Chapter 3 and (u_A^1,v_A^1,p_A^1) the solution computed with the Augmented Lagrange method.

for the necessity of choosing a more precise stopping criterion w^*. We regard the evolution of a quartercircle with initial radius $r_0 = 0.4$ with step size $\tau = 0.03h$, $\tilde{\tau} = h^2$ and $\varepsilon = 0.0625$ after one time step. On the one hand computed with the Lagrange method presented in Chapter 3 with stopping criterion $\Vert\nabla_{(u,v,p)}\ell\Vert_{L^2(\Omega)} \leq 10^{-8}$ and on the other hand computed with the Augmented Lagrange method and stopping criteria $\Vert\nabla_{(u,v)}\ell_A\Vert_{L^2(\Omega)} \leq w^* = 10^{-8}$ and $\Vert c^w\Vert_{L^2(\Omega)} \leq \eta^* = 10^{-13}$. As we use a different ε scaling in the Augmented Lagrange method we adapt the Lagrange method such that the results are comparable. These results are presented in Table 7.2. Here ℓ denotes the rescaled version of the Lagrange function and ℓ_A^w the Augmented Lagrange function. Moreover (u^1,v^1,p^1) denotes the solution for u, v and p after one time step computed with the Lagrange method, while (u_A^1,v_A^1,p_A^1) denotes the solution for u, v and p after one time step computed with the Augmented Lagrange method.
Especially the comparison of $\Vert\partial_u\ell(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$, $\Vert\partial_u\ell_A^w(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$, $\Vert\partial_u\ell(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$ and $\Vert\partial_u\ell_A^w(u^0,u_A^1,v_A^1,p_A^1)\Vert_{L^2(\Omega)}$ reveals that the Lagrange method from Chapter 3 leads to a better solution for u, because $\Vert\partial_u\ell(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$ and $\Vert\partial_u\ell_A^w(u^0,u^1,v^1,p^1)\Vert_{L^2(\Omega)}$ are significantly smaller than the corresponding values for (u_A^1,v_A^1,p_A^1). Thus a scaling of the energy e_{out}^w with ε^{-1} emphasizes this part of the Augmented Lagrange function and leads to a higher precision in the u component of the gradient of the Lagrange function without choosing a finer stopping criterion w^*.

Based on the previous results we use the scaled version of the Augmented Lagrange function

$$\ell_A^w[u^k,u,v,p;\lambda] = \frac{1}{\varepsilon}e_{\text{out}}^w[u^k,u,v] - (c^w[u,v],p)_{L^2(\Omega)} + \frac{\lambda}{2}\Vert c^w[u,v]\Vert_{L^2(\Omega)}^2$$

for all following tests.
First we present the corresponding test to the one presented in Figure 7.2. We take again the phase field function of a quarter circle with initial radius $r_0 = 0.4$ on a grid with grid width $h = 2^{-9}$ and $\tau = \tilde{\tau} = h^2$, $\varepsilon = 2^{-6}$ and different to the previous tests $w^* = \eta^* = 10^{-8}$. The result is presented in Figure 7.3. The solution computed with the Augmented Lagrange method is presented by the solid line, while the dotted line shows the reference solution from Chapter 3 again. Here it is nice to see, that the L^2-errors are all qualitatively equal although we use coarser stopping criteria w^* and η^*.

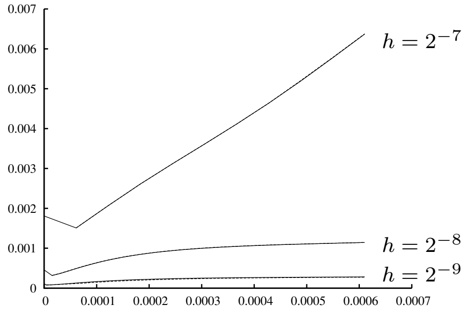


Figure 7.3: Temporal L^2-error evolution for the numerical solution of the Willmore flow, computed with Augmented Lagrange method and new scaling. The parameters are $\varepsilon = 2^{-6}$, $\tau = \tilde{\tau} = h^2$, $h = 2^{-7}$, 2^{-8}, 2^{-9} and $w^* = \eta^* = 10^{-8}$. The dotted line presents the L^2-error computed with the model presented in Chapter 3 and the solid one the L^2-error computed with the Augmented Lagrange method.

In a further test we take the same initial phase field function as in Figure 5.3. This is an image consisting of two circles with initial radius $r_0 = 0.13$ centered at $(0.35, 0.5)$ and $(0.65, 0.5)$. Different to the test presented in Figure 5.3 we take a grid with grid width $h = 2^{-8}$, $\varepsilon = 4h$ and step sizes $\tau = 0.03h$, $\tilde{\tau} = h^2$. As mentioned above the stopping criteria for the Augmented Lagrange method are $w^* = \eta^* = 10^{-8}$. The new results are presented in Figure 7.4. A comparison with Figure 5.3 reveals that the evolution is completely different. With the old method the two circles merge to one object within one time step, while this new method finds a second possible solution, in which both objects touch each other in one point. The plot of the curvature approximation demonstrates that the angle between both objects is $\pi/2$, such that there is mean curvature zero. Due to the typical evolution under Willmore flow the curved parts of both objects grow to minimize the curvature. This leads to boundary effects which influence the objects significantly such that they grow together and form one bar. Thus we get zero mean curvature in the whole computational domain $\Omega = [0, 1]^2$.
A change of grid width to $h = 2^{-9}$ and step size $\tau = \tilde{\tau} = h^2$ does not change the type of solution found by the Augmented Lagrange method. In Figure 7.5 we observe that even on this finer grid and with finer step size the two circles do not evolve into one object. They form noses touching each other, again.

In case of a phase field function representing a rectangle with side length 0.4 and 0.1 the solution equals the solution computed with the old method. This is demonstrated in the following test, where we take a grid of grid width $h = 2^{-9}$, $\varepsilon = 4h$, step sizes $\tau = \tilde{\tau} = h^2$ and accuracy $w^* = \eta^* = 10^{-8}$. In Figure 7.6 the evolution of the shape as well as the corresponding approximated curvature $(U^k - V^k)/\tilde{\tau}$ are presented. Com-

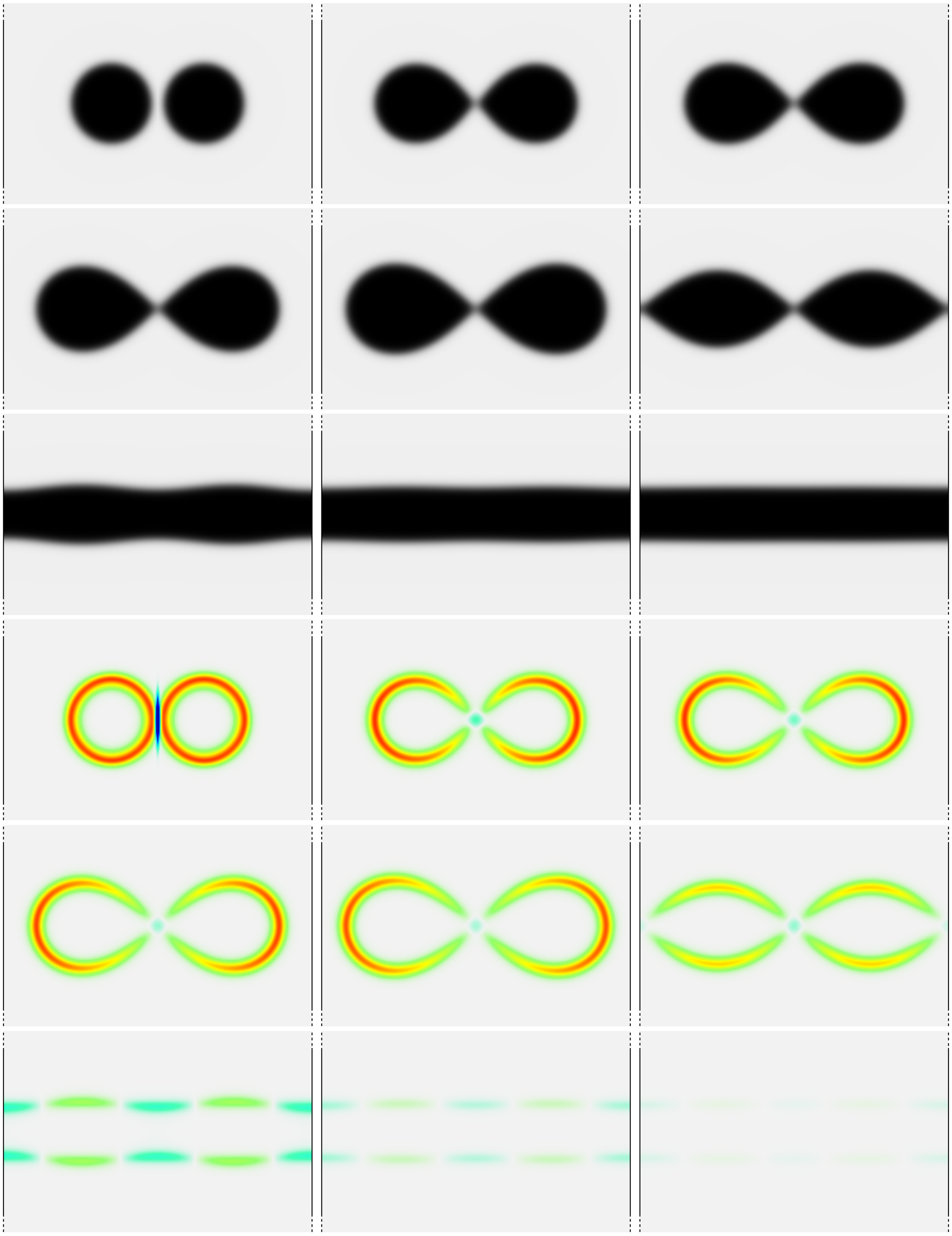


Figure 7.4: Evolution of two circles with initial radius $r_0 = 0.13$ and center points $(0.35, 0.5)$ and $(0.65, 0.5)$, step sizes $\tau = 0.03h$, $\tilde{\tau} = h^2$, $\varepsilon = 4h$ and $w^* = \eta^* = 10^{-8}$ on a grid with grid width $h = 2^{-8}$ after 0, 1, 2, 4, 6, 7, 8, 9 and 10 time steps. In the three lower rows the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$ is color-coded on the range $[-10, 10]$ by .

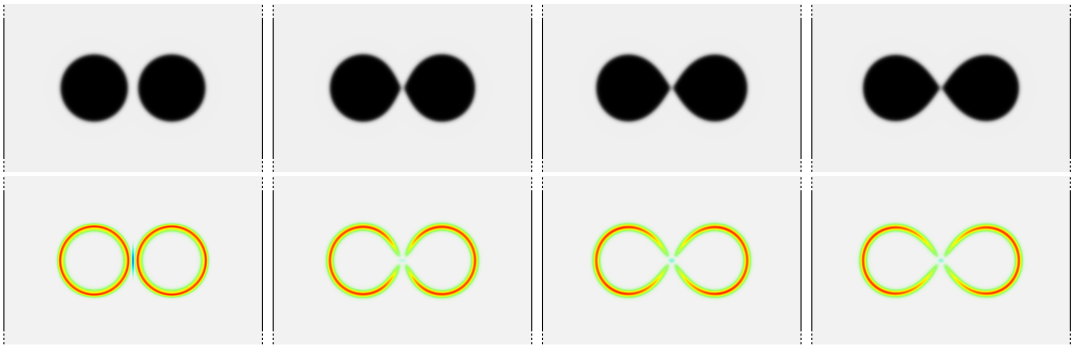


Figure 7.5: Evolution of two circles with initial radius $r_0 = 0.13$ and center points $(0.35, 0.5)$ and $(0.65, 0.5)$, step sizes $\tau = \tilde{\tau} = h^2$, $\varepsilon = 4h$ and $w^* = \eta^* = 10^{-8}$ on a grid with grid width $h = 2^{-9}$ after 0, 1, 10 and 18 time steps. In the two lower rows the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$ is color-coded on the range $[-10, 10]$ by .

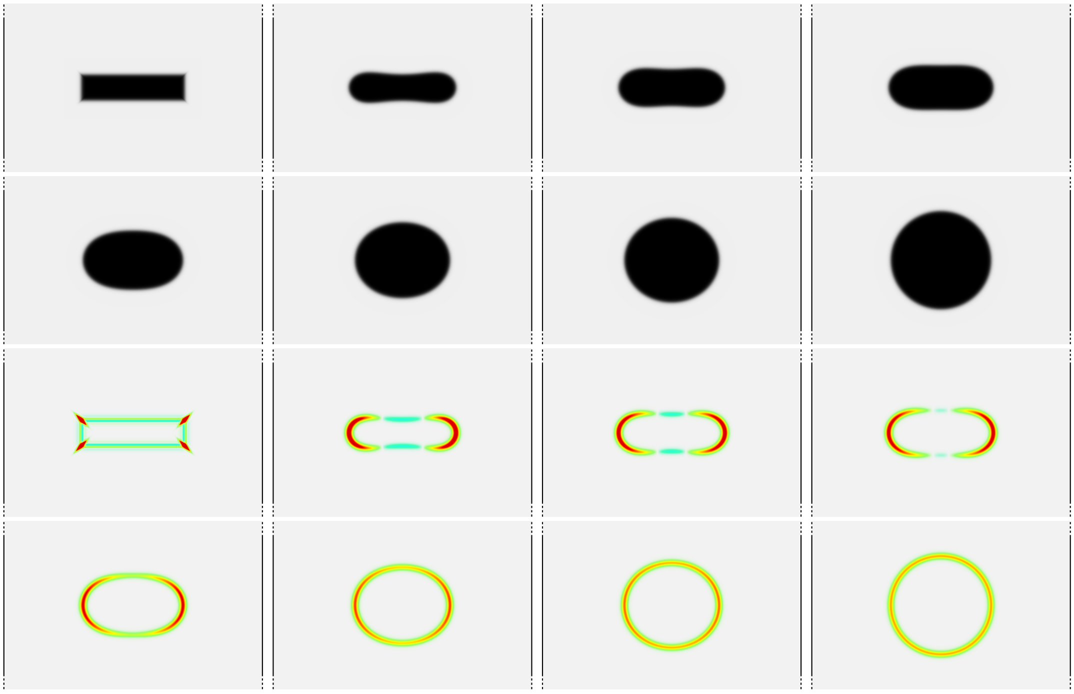


Figure 7.6: Evolution of a rectangle with step sizes $\tau = \tilde{\tau} = h^2$, $\varepsilon = 4h$ and $w^* = \eta^* = 10^{-8}$ on a grid with grid width $h = 2^{-9}$ after 0, 1, 6, 10, 20, 40, 60 and 120 time steps. In the two lower rows the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$ is color-coded on the range $[-20, 20]$ by .

grid width	min. mem	mem to store a CSM	mem needed by the TRM
2^{-3}	0.3	2	10
2^{-4}	2.0	12	58
2^{-5}	14.8	44	740
2^{-6}	113.1	198	14807
2^{-7}	884.4	1575	

Table 7.3: Minimal main memory (mem) in MB needed to save a 2×2 block matrix in the three-dimensional case, main memory needed to store a Cholmod sparse matrix (CSM) and main memory needed during the application of the Trust-Region method (TRM).

paring these results with the one presented in Figure 5.2 does not lead to differences.

The presented two-dimensional tests demonstrate that the Augmented Lagrange method works reliable for solving our Willmore model. Consequently it would be nice to test it in the three-dimensional case, too. Unfortunately in this case the Trust-Region method needs too much main memory on finer grids. In general a three-dimensional grid of grid depth (gd) consists of $(2^{\text{gd}} + 1)^3$ nodes. As our unknowns consist of two components u and v and we are working on a triangulated grid this leads to 2×27 matrix entries per node. If each entry needs a memory of 2^3byte we need at least $(2^{\text{gd}} + 1)^3 * 54 * 2^3$byte memory to store one block matrix. This is only the minimal memory which is needed. In real terms within the Trust-Region method there is applied a Cholmod factorization which saves a so-called Cholmod sparse matrix and this matrix needs more memory. Additionally the Trust-Region method saves two matrices of this type and needs some further main memory. The concrete numbers are presented in Table 7.3. Especially the last column reveals that the factor between the needed memory grows from one grid to the next finer one from 5.8 over 12.76 to 20.01. Thus a rough extrapolation reveals that we need 20×14GB $= 280$GB main memory on a grid of grid width $h = 2^{-7}$ and therefore 20×280GB $= 5600$GB on the next finer grid with $h = 2^{-8}$. Since our machine with the most main memory is a shared memory machine with 1024GB main memory a computation on a grid with grid width $h = 2^{-8}$ is not possible. Theoretically a computation on the next coarser grid should be possible, but the test to find out the needed memory turned out that this computation takes too much time for being practicable. Thus a computation on a grid of grid width $h = 2^{-6}$ is the best we can expect for three-dimensional computations.
Consequently we generated a phase field function representing a cube of side lenght 0.4 on a grid of grid width $h = 2^{-6}$ and started the computation of one step Willmore flow of time step size $\tau = \tilde{\tau} = h^2$ with accuracy $w^* = \eta^* = 10^{-8}$ and interface transition width $\varepsilon = 4h$. Twelve days later the computation of the first time step was not yet finished. Consequently even a computation on this grid is not practicable until now.

7.3 Future work

Recalling the previous sections we conclude that the Augmented Lagrange method works reliable for solving the linear model problem presented in Subsection 7.1.1 as well as our Willmore model in two dimensions. Therefore we can apply it for solving the phase field apparent contour inverse problem 6.5 in the two-dimensional case. This induces further work such as the implementation of a Hessian which is partwise approximated by a BFGS method and the application of a suitable step size control. For further information cp. Section 6.5.

To make the Augmented Lagrange method working stable in the three-dimensinal case we have to deal with the problem of finding a more efficient possibly iterative solver to be used within the Trust-Region method.
Another ansatz is to reduce complexity of the three-dimensional problem by using an efficient narrow band method. As already mentioned in Section 6.5 narrow band methods are suitable in combination with double well phase field approximations since the phase field function approximately equals 1 or -1 in most parts of the computational domain.

Bibliography

[1] Fred Almgren, Jean E. Taylor, and Lihe Wang. Curvature-driven flows: a variational approach. *SIAM J. Control Optim.*, 31(2):387–438, 1993.

[2] Herbert Amann. *Gewöhnliche Differentialgleichungen.* Walter de Gruyter, 1983.

[3] L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Bollettino dell'Unione Matematica Italiana, Sezione B*, 6(7):105–123, 1992.

[4] L. Bar, N. Sochen, and N. Kiryati. Semi-blind image restoration via Mumford-Shah regularization. *IEEE Transactions on Image Processing*, 15(2):483–493, 2006.

[5] L. Bar, N. Sochen, and N. Kiryati. Restoration of images with piecewise space-variant blur. In *Proceedings of the First International Conference on Scale Space Methods and Variational Methods in Computer Vision*, pages 533–544. Springer, 2007.

[6] Leah Bar, Benjamin Berkels, Martin Rumpf, and Guillermo Sapiro. A variational framework for simultaneous motion estimation and restoration of motion-blurred video. In *Eleventh IEEE International Conference on Computer Vision (ICCV 2007)*, 2007.

[7] J. W. Barrett, H. Garcke, and R. Nürnberg. A parametric finite element method for fourth order geometric evolution equations. *J. Comp. Phys.*, 222:441–467, 2007.

[8] G. Bellettini and L. Mugnai. On the approximation of the elastica functional in radial symmetry. *Calculus of Variations and Partial Differential Equations*, 24(1):1–20, 2005.

[9] Giovanni Bellettini, Valentina Beorchia, and Maurizio Paolini. Topological and variational properties of a model for the reconstruction of three-dimensional transparent images with self-occlusions. *J. Math. Imaging Vision*, 32(3):265–291, 2008.

[10] Giovanni Bellettini, Valentina Beorchia, and Maurizio Paolini. Completion of visible contours. *SIAM Journal on Imaging Sciences*, 2(3):777–799, 2009.

[11] Benjamin Berkels. *Joint methods in imaging based on diffuse image representations.* Dissertation, University of Bonn, 2010.

[12] M. Bertalmio, A.L. Bertozzi, and G. Sapiro. Navier-stokes, fluid dynamics, and image and video inpainting. In *IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 355–362, 2001.

[13] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proc. of SIGGRAPH 2000*, pages 417–424, New Orleans, USA, July 2000.

[14] D.P. Bertsekas. *Constrained optimization and Lagrange multiplier methods.* Optimization and neural computation series. Athena Scientific, 1996.

[15] I. A. Bobenko and P. Schröder. Discrete Willmore flow. In *SIGGRAPH (Courses)*, pages 101–110. ACM Press, 2005.

[16] Dietrich Braess. *Finite Elemente.* Springer, 2nd edition, 1997. Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie.

[17] Andrea Braides. Γ*-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications.* Oxford University Press, Oxford, 2002.

[18] John W. Cahn and John E. Hilliard. Free energy of a nonuniform system. i. interfacial free energy. *Journal of Chemical Physics*, 28(2):258–267, 1958.

[19] P.B. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the red blood cell. *J. Theor. Biol.*, 26:61–81, 1970.

[20] Antonin Chambolle. An algorithm for mean curvature motion. *Interfaces and free Boundaries*, 6:195–218, 2004.

[21] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, November 2004.

[22] T. F. Chan, S. H. Kang, and J. Shen. Euler's elastica and curvature-based inpainting. *SIAM Appl. Math.*, 63(2):564–592, 2002.

[23] Xinfu Chen, Charlie M. Elliott, Andy Gardiner, and Jennifer Jing Zhao. Convergence of numerical solutions to the Allen-Cahn equation. *Appl. Anal.*, 69(1-2):47–56, 1998.

[24] Y. Chen, T. A. Davis, W. W. Hager, and S. Rajamanickam. Algorithm 887: CHOLMOD, supernodal sparse Cholesky factorization and update/downdate. *ACM Transactions on Mathematical Software*, 35(3):22:1–22:14, 2009.

[25] U. Clarenz, U. Diewald, G. Dziuk, M. Rumpf, and R. Rusu. A finite element method for surface restoration with smooth boundary conditions. *Computer Aided Geometric Design*, 21(5):427–445, 2004.

[26] A. Conn, G.I.M. Gould, and P.L. Toint. *Lancelot: A Fortran Package for Large-Scale Nonlinear Optimization (Release A).* Springer Series in Computational Mathematics. Springer, 2010.

[27] A. R. Conn, N. I. M Gould, and P. L. Toint. *Trust-Region Methods.* SIAM, 2000.

[28] W. Dahmen and A. Reusken. *Numerik für Ingenieure und Naturwissenschaftler.* Springer-Lehrbuch. Springer, 2008.

[29] T. A. Davis and W. W. Hager. Dynamic supernodes in sparse Cholesky update/downdate and triangular solves. *ACM Transactions on Mathematical Software*, 35(4):27:1–27:23, 2009.

[30] Ennio De Giorgi. Some remarks on Γ-convergence and least squares method. In *Composite media and homogenization theory (Trieste, 1990)*, volume 5 of *Progr. Nonlinear Differential Equations Appl.Progr. Nonlinear Differential Equations Appl.*, pages 135–142. Birkhäuser Boston, Boston, MA, 1991.

[31] K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces and Free Boundaries*, 8:21–46, 2006.

[32] K. Deckelnick, G. Dziuk, and C. M. Elliott. Computation of geometric partial differential equations and mean curvature flow. *Acta Numerica*, 14:139–232, 2005.

[33] K. Deckelnick and F. Schieweck. Error analysis for the approximation of axisymmetric willmore flow by $c1$-elements. Technical Report 23, Universität Magdeburg, 2009. to appear in Interfaces and Free Boundaries.

[34] Klaus Deckelnick and Gerhard Dziuk. Error analysis for the elastic flow of parametrized curves. *Math. Comp.*, 78(266):645–671, 2009.

[35] Manfredo P. Do Carmo. *Differentialgeometrie von Kurven und Flächen.* Vieweg, 1976.

[36] Patrick W. Dondl, Luca Mugnai, and Matthias Röger. Confined elastic curves. arXiv:submit/0044093 [math.AP].

[37] M. Droske and M. Rumpf. A level set formulation for Willmore flow. *Interfaces and Free Boundaries*, 6(3):361–378, 2004.

[38] M. Droske and M. Rumpf. A variational approach to non-rigid morphological registration. *SIAM Journal on Applied Mathematics*, 64(2):668–687, 2004.

[39] Qiang Du, Chun Liu, Rolf Ryham, and Xiaoqiang Wang. A phase field formulation of the willmore problem. *Nonlinearity*, 18:1249–1267, 2005.

[40] Qiang Du, Chun Liu, and Xiaoqiang Wang. Simulating the deformation of vesicle membranes under elastic bending energy in three dimensions. *Journal of Computational Physics*, 212(2):757–777, 2006.

[41] Qiang Du and R. A. Nicolaides. Numerical analysis of a continuum model of phase transition. *SIAM Journal on Numerical Analysis*, 28:1310–1322, 1991.

[42] Qiang Du and Xiaoqiang Wang. Convergence of numerical approximations to a phase field bending elasticity model of membrane deformations. *International Journal of Numerical Analysis and Modeling*, 4(3–4):441–459, 2007.

[43] G. Dziuk, E. Kuwert, and R. Schätzle. Evolution of elastic curves in $\mathbb{R}^n$: existence and computation. *SIAM J. Math. Anal.*, 33, no. 5(5):1228–1245 (electronic), 2002.

[44] Gerhard Dziuk. Computational parametric Willmore flow. *Numer. Math.*, 111(1):55–80, 2008.

[45] C. M. Elliott, D. A. French, and F. A. Milner. A second order splitting method for the Cahn-Hilliard equation. *Numer. Math.*, 54:575–590, 1989.

[46] C.M. Elliott. The Cahn-Hilliard model for the kinetics of phase separation. *Num. Math.*, 1988.

[47] Selim Esedoglu, Steven J. Ruuth, and Richard Tsai. Threshold dynamics for high order geometric motions. *Interfaces and Free Boundaries*, 2008. to appear.

[48] E. A. Evans. Bending resistance and chemically induced moments in membrane bilayers. *Biophys. J.*, 14(12):923–931, 1974.

[49] L. C. Evans. *Partial Differential Equations*. Americal Math. Soc., 1998.

[50] L. C. Evans, H. M. Soner, and P. E. Souganidis. Phase transitions and generalized motion by mean curvature. *Communications on Pure and Applied Mathematics*, 45:1097–1123, 2006.

[51] I. Faragó. Operator splittings and numerical methods. In Ivan Lirkov, Svetozar Margenov, and Jerzy Wasniewski, editors, *Large-Scale Scientific Computing*, volume 3743 of *Lecture Notes in Computer Science*, pages 347–354. Springer Berlin / Heidelberg, 2006.

[52] Martina Franken, Martin Rumpf, and Benedikt Wirth. A phase field based pde constraint optimization approach to time discrete willmore flow. *International Journal of Numerical Analysis and Modeling*, 2011. accepted.

[53] Karl Glasner and Robert Almgren. Dual fronts in a phase field model. *Physica D*, 146:328–340, 2000.

[54] Maurizio Grasselli and Dalibor Prazak. Longtime behavior of a diffuse interface model for binary fluid mixtures with shear dependent viscosity. *Interfaces and Free Boundaries*, 13:507–530, 2011.

[55] J. Haslinger and R.A.E. Mäkinen. *Introduction to Shape Optimization: Theory, Approximation, and Computation.* Advances in Design and Control. Society for Industrial and Applied Mathematics, 2003.

[56] W. Helfrich. Elastic properties of lipid bilayers: Theory and possible experiments. *Zeitschrift für Naturforschung*, 28c:693–703, 1973.

[57] Michael Hinze, Rene Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical modelling*. Springer, Dordrecht, NL, 2009.

[58] D. A. Huffman. Impossible objects as nonsense sentences. *Machine Intelligence*, 6:295–323, 1971.

[59] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*. Advances in design and control. Society for Industrial and Applied Mathematics, 2008.

[60] Jürgen Jost. *Postmodern analysis*. Universitext. Springer-Verlag, Berlin, third edition, 2005.

[61] Jürgen Jost and Xianqing Li-Jost. *Calculus of variations*. Cambridge University Press, 1998.

[62] Takeo Kanade. Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence*, 17:409 – 460, 1981.

[63] E. Kuwert and R. Schätzle. The Willmore flow with small initial energy. *J. Differential Geom.*, 57(3):409–441, 2001.

[64] E. Kuwert and R. Schätzle. Gradient flow for the Willmore functional. *Comm. Anal. Geom.*, 10(5):1228–1245 (electronic), 2002.

[65] Ernst Kuwert and Reiner Schätzle. Removability of point singularities of Willmore surfaces. *Ann. of Math. (2)*, 160(1):315–357, 2004.

[66] Christophe Le Potier. Schéma volumes finis pour les opérateurs de diffusion fortement anisotropes sur des maillages non structurés. *Comptes Rendus Mathematique*, 340:921–926, 2005. C. R. Acad. Sci. Paris, Ser. I.

[67] Sang Hwa Lee, Nam Su Moon, and Choong Woong Lee. Recovery of blurred video signals using iterative image restoration combined with motion estimation. In *Proceedings of the 1997 IEEE International Conference on Image Processing (ICIP '97)*, volume 1, pages 755–758, 1997.

[68] Chun Liu and Jie Shen. A phase field model for the mixture of two incompressible fluids and its approximation by a fourier-spectral method. *Physica D: Nonlinear Phenomena*, 179(3-4):211 – 228, 2003.

[69] P. Loreti and R. March. Propagation of fronts in a nonlinear fourth order equation. *European Journal of Applied Mathematics*, 11(2):203–213, 2000.

[70] S. Luckhaus and Th. Sturzenhecker. Implicit time discretization for the mean curvature flow equation. *Calc. Var.*, 3:253–271, 1995.

[71] Jitendra Malik. Interpreting line drawings of curved objects. *International Journal of Computer Vision*, 1(1):73–103, 1987.

[72] Gurij Ivanovich Marchuk. Some application of splitting-up methods to the solution of mathematical physics problems. *Applications of Mathematics*, 13(2):103–132, 1968.

[73] Luciano Modica. The gradient theory of phase transitions and the minimal interface criterion. *Archive for Rational Mechanics and Analysis*, 98(2):123–142, 1987.

[74] Luciano Modica and Stefano Mortola. Un esempio di Γ^--convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.

[75] D. Mumford. Elastica and computer vision. In C. Bajaj, editor, *Algebraic Geometry and Its Applications*, pages 491–506. Springer, New York, 1994.

[76] David Mumford and Jayant Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.

[77] M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation and Depth (Lecture Notes in Computer Science Vol. 662).* Springer-Verlag Berlin Heidelberg, 1993.

[78] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer, New York / Berlin, 1999.

[79] R. H. Nochetto, M. Paolini, and C. Verdi. Optimal interface error estimates for the mean curvature flow. *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie IV*, 21(2):193–212, 1994.

[80] R. H. Nochetto, M. Paolini, and C. Verdi. Quadratic rate of convergence for curvature dependent smooth interfaces: a simple proof. *Applied Mathematics Letters. An International Journal of Rapid Publication*, 7(4):59–63, 1994.

[81] Nadine Olischläger and Martin Rumpf. Two step time discretization of Willmore flow. In *Proceedings of the 13th IMA International Conference on Mathematics of Surfaces XIII*, pages 278–292, Berlin, Heidelberg, 2009. Springer-Verlag.

[82] Nadine Olischläger and Martin Rumpf. A nested variational time discretization for parametric Willmore flow. *Interfaces and Free Boundaries*, 14(4):431–454, 2012.

[83] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[84] A. Polden. Curves and Surfaces of Least Total Curvature and Fourth-Order Flows. *Dissertation, Universität Tübingen*, page , 1996.

[85] S. D. Rane, J. Remus, and G. Sapiro. Wavelet-domain reconstruction of lost blocks in wireless image transmission and packet-switched networks. In *Image Processing. 2002. Proceedings. 2002 International Conference on 22-25 Sept. 2002, Vol.1*, 2002.

[86] A. Rav-Acha and S. Peleg. Restoration of multiple images with motion blur in differentdirections. In *Proceedings of the fifth IEEE Workshop on Applications of Computer Vision*, pages 22–28, 2000.

[87] Tristan Rivière. Analysis aspects of Willmore surfaces. *Invent. Math.*, 174(1):1–45, 2008.

[88] Matthias Röger and Reiner Schätzle. On a modified conjecture of De Giorgi. *Mathematische Zeitschrift*, 254(4):675–714, December 2006.

[89] R. Rusu. An algorithm for the elastic flow of surfaces. *Preprint Mathematische Fakultät Freiburg*, 01-35:, 2001.

[90] Lars Ole Schwen. *Composite Finite Elements for Trabecular Bone Microstructures.* Dissertation, University of Bonn, 2010.

[91] Jie Shen and Xiaofeng Yang. Energy stable schemes for cahn-hilliard phase-field model of two-phase incompressible flows. *Chin. Ann. Math., Ser. B*, 31(5):743–758, 2010.

[92] G. Simonett. The Willmore Flow near spheres. *Diff. and Integral Eq.*, 14(8):1005–1014, 2001.

[93] J. Stoer and R. Bulirsch. *Introduction to numerical analysis*, volume 12 of *Texts in Applied Mathematics.* Springer-Verlag, New York, third edition, 2002. Translated from the German by R. Bartels, W. Gautschi and C. Witzgall.

[94] Gilbert Strang. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):pp. 506–517, 1968.

[95] Vidar Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series in Computational Mathematics.* Springer, Berlin, 2nd edition, 2006.

[96] David Waltz. Understanding line drawings of scenes with shadows. In *The Psychology of Computer Vision.* McGraw-Hill, 1975.

[97] William Welch and Andrew Witkin. Variational surface modeling. *Computer Graphics*, 26(2):157–166, 1992.

[98] Lance R. Williams. Topological reconstruction of a smooth manifold-solid from its occluding contour. In *Proceedings of the Third European Conference on Computer Vision*, pages 36–47, 1995.

[99] T.J. Willmore. *Riemannian Geometry.* Claredon Press, Oxford, 1993.

[100] Benedikt Wirth. *Variational methods in shape space.* Dissertation, University Bonn, 2010.

[101] Guoliang Xu and Qing Pan. G^1 surface modelling using fourth order geometric flows. *Computer-Aided Design*, 38(4):392–403, 2006.

[102] Shin Yoshizawa and Alexander G. Belyaev. Fair triangle mesh generation with discrete elastica. In *Proceedings of the Geometric Modeling and Processing; Theory and Applications (GMP'02)*, pages 119–123, Washington, DC, USA, 2002. IEEE Computer Society.

List of Figures

2.1 Rectangular and corresponding uniform simplicial grid 24

4.1 Temporal error evolution for the numerical solution of the biharmonic heat equation 55

5.1 Evolution of the L^2-error for Willmore flow of a quarter circle in time 58
5.2 Evolution of a rectangle under Willmore flow 59
5.3 Evolution of two circles under Willmore flow 60
5.4 Evolution of a cube under Willmore flow 61
5.5 Evolution of a disc under Willmore flow 62
5.6 Two different edge restoration problems 68
5.7 A variational C^1 edge continuation 69

6.1 Huffman labeling of a torus 72
6.2 Huffman labeling of a sphere and a cylinder 72
6.3 Huffman labeling in literature 74
6.4 Refinement of Huffman labeling 76
6.5 Reconstruction of an ellipse 80
6.6 Reconstruction of an annulus 82
6.7 Sketch of two geometries underlying the given Huffman labelings in Figure 6.8 and 6.9 83
6.8 Reconstruction of an object with two holes beside each other 84
6.9 Reconstruction of an object with two holes over each other 84
6.10 Reconstruction of a sphere 85
6.11 Reconstruction of a torus 87
6.12 Reconstruction of a rotated torus 88

7.1 Temporal L^2-error evolution for the numerical solution, computed with Augmented Lagrange method, of the biharmonic heat equation 97
7.2 Temporal L^2-error evolution for the numerical solution, computed with Augmented Lagrange method, of the Willmore flow 104
7.3 Temporal L^2-error evolution for the numerical solution, computed with Augmented Lagrange method and new scaling, of the Willmore flow . 106
7.4 Evolution of two circles under Willmore flow computed with Augmented Lagrange 107
7.5 Evolution of two circles under Willmore flow on a finer grid computed with Augmented Lagrange 108

7.6 Evolution of a rectangle under Willmore flow computed with Augmented Lagrange 108

List of Algorithms

2.1 Augmented Lagrange method 28

3.1 Newton method with Cholesky solver to solve the variational problem (3.18a) and (3.18b) 47

6.1 Initialization and operator splitting for Problem 6.5 78

7.1 Augmented Lagrange method for the biharmonic heat equation problem 96
7.2 Augmented Lagrange method for our Willmore flow model 103

Nomenclature

A Set Notation

$\mathbb{R}$	real numbers	17
$\mathbb{R}^d$	d-dimensional real Euclidean space	17
Ω	computational domain	29
∂A	boundary of the set A	18
$\sharp I$	cardinality of the set I	44
$[u = 0]$	zero level set of u	21
$U_\delta(A)$	δ-neighborhood of the set A	21
$\overline{H}^1(\Omega)$	different ansatz spaces to realize three kinds of boundary conditions	50
$\dot{H}^s$	subspace of functions with s weak derivatives	51
$L_0^2(\Omega)$	$L_0^2(\Omega) = \{u \in L^2 \,:\, \int_\Omega u \,\mathrm{d}x = 0\}$	50

B Analytic Expressions

$(\cdot,\cdot)_{L^2(\Gamma)}$	L^2-metric on the hypersurface Γ	30
$\\|\cdot\\|$	Euclidean norm	44
$\\|\cdot\\|_{l^1}$	l^1-norm	90
$\\|\cdot\\|_{L^1(\Omega)}$	L^1-norm	90
χ_K	characteristic function of the set K	17
δ_{ij}	Kronecker delta	101
p	Lagrange multiplier (vector)	25
$\mathrm{dist}(\cdot,\cdot)$	shortest path length on a manifold	30
$\mathrm{dist}(x, K)$	distance of a point x to the set K	18

$\operatorname{sgndist}(x, K)$	signed distance of a point $x \in \Omega$ to the set $K \subset \Omega$	18
Per	perimeter functional	21
D^2	Hessian	46
$\Gamma - \lim$	Γ-limit	20
card	cardinality	71
$\mathcal{H}^{d-1}$	$(d-1)$-dimensional Hausdorff measure	17
R_h	canonical Ritz projection	52
$\hat{\hat{R}}_h$	projection from $\dot{H}^5$ to $\overline{\mathcal{V}}$	52

C Geometric Analysis

e_d	unit normal vector in x_d-direction	71
π	parallel projection from $\mathbb{R}^d$ onto a $(d-1)$-dimensional viewing plane	71
Γ	hypersurface	29
$\mathcal{M}$	Riemannian manifold	29
g_Γ	metric on a hypersurface Γ	29
$\operatorname{grad}_{g_\Gamma}$	gradient with respect to the metric g_Γ	29
$T_\Gamma \mathcal{M}$	tangent space to $\mathcal{M}$ in $\Gamma \in \mathcal{M}$	29
ν	outer normal	50
∂_ν	normal derivative	50
$\mathbf{h}$	mean curvature	11
$\mathbf{k}$	Gaussian curvature	36
$G(\Gamma)$	apparent contour of a hypersurface Γ	71
$\alpha[\Gamma](\cdot)$	Huffman labeling of a hypersurface Γ	71
α^*	given Huffman labeling	73

D Energies and Functionals

$w[\cdot]$	Willmore energy, parametric version	37
$w^\varepsilon[\cdot]$	phase field version of the Willmore energy	64

$a^\varepsilon[\cdot]$	Modica-Mortola energy	18
$A^\varepsilon[\cdot]$	discrete version of the Modica-Mortola energy	43
$\Psi : \mathbb{R} \to \mathbb{R}$	double well potential	18
$e^\varepsilon_{\mathrm{AT}}$	Ambrosio-Tortorelli functional	17
$\mathfrak{q}(\cdot;\lambda)$	quadratic penalty function	26
e^l_{MS}	Mumford-Shah functional	17
$e_{\mathrm{out}}[\cdot,\cdot,\cdot]$	parametric version of the outer energy for Willmore flow	37
$e_{\mathrm{in}}[\cdot,\cdot]$	parametric version of the inner energy for Willmore flow	37
$e^\varepsilon_{\mathrm{out}}[\cdot,\cdot,\cdot]$	phase field version of the outer energy for Willmore flow	38
$\mathcal{E}^\varepsilon_{\mathrm{out}}[\cdot,\cdot,\cdot]$	discrete version of the outer energy for Willmore flow	43
$e^\varepsilon_{\mathrm{in}}[\cdot,\cdot]$	phase field version of the inner energy for Willmore flow	38
$\mathcal{E}^\varepsilon_{\mathrm{in}}[\cdot,\cdot]$	discrete version of the inner energy for Willmore flow	43
$\ell(\cdot,p)$	Lagrange function	25
$\mathcal{L}[\cdot,\cdot,\cdot,\cdot]$	discrete Lagrangian	45
$e^l_{\mathrm{out}}[\cdot,\cdot]$	outer energy for the linear model problem (the biharmonic heat equation)	49
$\mathcal{E}^l_{\mathrm{out}}[\cdot,\cdot,\cdot]$	discrete version of the linear outer energy	51
$e^l_{\mathrm{in}}[\cdot,\cdot]$	inner energy for the linear model problem (the biharmonic heat equation)	49
$\mathcal{E}^l_{\mathrm{in}}[\cdot,\cdot,\cdot]$	discrete version of the linear inner energy	51
$a^l[\cdot]$	Dirichlet energy	49
$w^l[\cdot]$	$w^l[u] = \frac{1}{2}\int_\Omega (\Delta u)^2 \,\mathrm{d}x$	49
$\mathcal{E}^{l,f}_{\mathrm{out}}[\cdot,\cdot,\cdot]$	discrete version of the outer energy with respect to nonzero right hand side in the linear case	53
$\mathcal{F}[\cdot,\cdot,\cdot]$	gradient of the discrete Lagrangian for Willmore flow	46
$e^{\varepsilon,\eta}_{\mathrm{out}}[\cdot,\cdot,\cdot]$	outer energy for the image restoration problem	66
$\ell^{\mathrm{ir}}[\cdot,\cdot,\cdot,\cdot]$	Lagrangian in case of the image restoration problem	67
$\mathcal{L}^{\mathrm{ir}}[\cdot,\cdot,\cdot,\cdot]$	discrete Lagrangian in case of the image restoration problem	67
$e[\Gamma]$	elastic bending energy for a membrane $\Gamma \subset \mathbb{R}^3$	63

$e_{\mathrm{s}}[\Gamma]$	simplified version of the elastic bending energy treated by Du et al. in [40] and [42]	63
$e_{\mathrm{reg}}[\Gamma]$	regularization energy of a hypersurface Γ in case of shape reconstruction	73
$e^{\varepsilon}_{\mathrm{reg,p}}[\cdot,\cdot]$	penalized regularization energy in case of shape reconstruction	77
$e^{\varepsilon,rp}_{\mathrm{out}}[\cdot,\cdot,\cdot]$	outer energy in case of shape reconstruction	77
$j[\cdot]$	mismatch functional	76
$\ell^{\mathrm{ac}}[\cdot,\cdot,\cdot,\cdot]$	Lagrange function in case of shape reconstruction	77
$c(\cdot,\cdot)$	constraint written as equality constraint	90
$c^{l}[\cdot,\cdot]$	equality constraint in the linear model case	92
$C^{l}[\cdot,\cdot]$	discrete version of the equality constraint in the linear model case	94
$c^{w}[\cdot,\cdot]$	equality constraint in the scaled version of the Willmore model	99
$\ell_A(\cdot,p;\lambda)$	Augmented Lagrange function	26
$\ell^{l}_A[\cdot,\cdot,\cdot,\cdot;\cdot]$	Augmented Lagrange function in the linear model case	93
$\mathcal{L}^{l}_A[\cdot,\cdot,\cdot,\cdot;\cdot]$	discrete version of the Augmented Lagrange function in the linear model case	94
$\ell^{\mathrm{ac}}_A[\cdot,\cdot,\cdot,\cdot;\cdot]$	Augmented Lagrange function in case of shape reconstruction	90
$e^{w}_{\mathrm{out}}[\cdot,\cdot,\cdot]$	outer energy in the scaled version of the Willmore flow model	98
$e^{w}_{\mathrm{in}}[\cdot,\cdot]$	inner energy in the scaled version of the Willmore flow model	98
$\ell^{w}_A[\cdot,\cdot,\cdot\cdot;\cdot]$	Augmented Lagrange function in the scaled version of the Willmore flow model	99
$\mathcal{L}^{w}_A[\cdot,\cdot,\cdot,\cdot;\cdot]$	discrete version of the Augmented Lagrange function in the scaled version of the Willmore flow model	101

E Discretization

$\mathbb{1}$	identity matrix	94
τ	time step size of the outer problem	37
Δ_h	discrete Laplacian	52
$\tilde{\tau}$	time step size of the inner problem	37
$\alpha[u]$	approximated Huffman labeling	76

$\mathcal{V}$	Finite Element space of continuous, piecewise affine functions	23
$\overline{\mathcal{V}}$	Finite Element space with boundary conditions	51
$\mathcal{V}_m$	Finite Element space of continuous, piecewise multilinear functions	24
I	vertex index set	23
I_D	vertex index set of all Dirichlet nodes	54
$\mathcal{C}$	uniform rectangular grid	24
C	elements of a uniform rectangular grid	24
$\mathcal{T}$	uniform simplicial mesh	23
T	element of a uniform simplicial mesh $\mathcal{T}$	23
h	mesh resolution	24
$\{\Phi_i\}_{i\in I}$	nodal basis of a Finite Element space	23
$\mathcal{I}_h$	Lagrange interpolation	23
$\mathcal{P}_1$	space of affine functions	23
$\mathcal{P}_1^m$	set of multilinear functions	24
u	spatially continuous version of the function u	49
U	spatially discretized version of the function u	49
$\tilde{u}$	time-continuous version of the function u (only used in Chapter 4) in comparison to the time-discrete version without tilde	49
$\mathbf{M}$	canonical lumped mass matrix	23
$\mathbf{M}[\lambda]$	weighted lumped mass matrix with arbitrary continuous weight function λ	23
$\mathbf{L}$	stiffness matrix	23
$\mathbf{D}[\lambda]$	$\mathbf{D}[\lambda] := (\delta_{ij}\lambda_i)_{i,j}$	101

F Other Notation

η	scaling factor in front of the surface energy	73
γ	scaling factor in front of the mismatch energy	76
μ	scaling factor in front of the Willmore energy	73
$\phi_{l^1}(\cdot;\cdot)$	l^1 merit function	90
$\phi_{L^1}[\cdot,\cdot,\cdot;\cdot]$	L^1 merit function	90

Index

A

Ambrosio-Tortorelli functional 17
apparent contour 12, 71ff., 75, 83, 86
apparent contour inverse problem 73
    phase field version . . 76ff., 89, 91, 110
Augmented Lagrange
    function . . 26, 90f., 93f., 99f., 102–105
    method . . . 13, 26, 28, 90–93, 95–100, 103–106, 109

B

Banach space
    reflexive . . . . 39f.
bending rigidity 63
bi-Laplace 52
biconjugate gradient method . . . . 54, 103f.
biharmonic heat flow 12, 49, 52, 94f., 97f.
    error estimates 52
    fully discretized 12, 51
    time-discrete 50, 91, 95
    time-discrete, operator version 51
    time-discrete, with equality constraint 92
block matrix 24, 54, 109
boundary condition . . 18f., 49f., 67, 92, 96
    C^1 . . . . 66, 68
    Dirichlet 49f., 54, 80, 91
    natural . . . . 49f.
    Neumann 91, 94
    periodic . . . . 49f.

C

Cholesky solver 46f., 65
CHOLMOD 46, 65
Cholmod
    factorization 109
    sparse matrix 109
constraint . 37f., 53, 63, 76, 90, 92, 94, 99, 104
    equality 25, 89f., 92, 99
    minimization problem 12
    PDE 11f., 41
    variational problem 77
    vector of 25
convergence 13, 26, 66
    Γ- 20
    behavior . . . . 49, 51f.
    estimate 55
    experimental order of 55
    problem 66
    result 53
    tolerance 95f., 103
    weak . . . . 39ff.
convex . . . . 18, 39f., 49, 58f.
    non 41, 84
    subset 39

D

depth ordering function 75
depth-index 73, 75
direct method 19, 39, 44
double well potential 18ff., 22, 45
dual function 41
dual problem . . . . 42f.

E

edge 11, 17, 66, 73, 86
    continuation 69
    invisible 73
    length 59, 61, 69
    restoration 12, 66, 68
    visible 73
energy
    decay 30, 82

  estimate 30
error order 12, 55, 97f.
Euclidean
  norm 93
  scalar product 93
  space 44
Euler–Lagrange equation 14, 18, 30, 38, 42f.

F

Finite Element method
  linear 23
  multilinear 23
Finite Element space 23, 51, 94
  piecewise affine 23, 43, 100
  piecewise multilinear 24
fold 73
Frenet formula 21

G

Gaussian curvature 36, 63
gradient 20, 29, 53, 94f., 99ff., 104f.
gradient flow 11, 29f., 37, 49, 78f.
  equation 29f., 37
  minimization 69

H

Hessian 24, 46, 54, 67f., 77, 89, 93, 95, 99f., 110
Huffman labeling 12, 71, 79ff., 83–88
hypersurface 11f., 29f., 37f., 71, 73

I

image restoration 49, 57, 66, 69
intersection 71, 73
isomorphic 44

K

Karush-Kuhn-Tucker condition 25f., 50

L

Lagrange
  function 25f., 67, 77, 89, 101, 105
  interpolation 23, 43
  multiplier 26, 28, 41
  multiplier vector 25f.
Lagrangian 41, 43, 45f., 50, 53f., 67f., 78
  discrete 43
LANCELOT 26
Laplace operator 52
Lipschitz boundary 39
lumped mass matrix 43
  canonical 23
  weighted 23

M

mean curvature 11, 14, 21, 37, 49, 52, 63, 106
  approximation 11, 37, 76, 104
  flow 11, 14, 37
  motion 11f., 14, 29f., 33, 36ff.
  vector 13
mesh 23f., 43, 100
  regular and uniform simplicial 43
  resolution 24
  size 59f., 79
  uniform simplicial 23, 44, 67, 94, 100
method of multipliers 26
metric 29f., 37
minimizing sequence 39f., 44
mismatch 76, 86, 88
  energy 77, 79
  functional 76–79
Modica-Mortola
  energy 14, 18f., 21, 76
  phase field function 18f., 76
Mumford-Shah functional 17

N

natural time discretization 11f., 29f., 37, 49
  mean curvature motion 29f., 33, 36f.
  Willmore flow 37
Newton method 12, 24, 43, 45ff., 54, 65, 67, 78, 89, 93f., 103f.
nodal basis 23
nonlinear 43, 45
  case 52, 94
  model 12, 94
  solver 66
  system 65
  term 51, 64, 98f.

O

occluded contour 73

P
paneling construction 73, 75
penalty functional 76
perimeter functional 21
phase field function . . . . . . 17f., 30, 33, 36, 38, 46f., 57f., 63, 66, 76, 78f., 88, 105f., 109f.
  Ambrosio-Tortorelli . . . . 17f.
  double well 17
  Modica-Mortola 17
  single well 17
prior 73
profile
  optimal 18f., 21f., 33, 36, 38, 57

Q
quadratic penalty function 26

R
Rellich's embedding theorem 41
Riemannian manifold . . . . 29f.

S
saddle point13, 25f., 41, 43, 45, 54, 78, 89
  problem 25, 78, 89ff., 93
semicontinuity
  lower 39, 44
  weak lower . . . . 39f.
stiffness matrix 23
stretching rigidity 63
subsequence . . . . 39f., 44f.
surface tension 63

T
time step control 46

U
uniqueness 13, 23, 36, 41, 52, 71, 75

W
Willmore energy 11, 13f., 37, 63f., 66, 69, 73, 76f., 80f., 83, 86
Willmore flow . 11–15, 29, 37f., 41, 52, 54, 57–63, 77, 86, 98, 102ff., 106, 109
  fully discrete variational time discretization of . . . . 43f.
  natural time discretization of 12
  nested variational time discretization of 12f., 37f., 49, 51, 91, 98
  parametric nested variational time discretization of 37
  scaled nested variational time discretization of . . . . 98f.
  scaled nested variational time discretization, with equality constraint 99
  semi-implicit phase field scheme . . 64

Y
Young's inequality 39f., 44