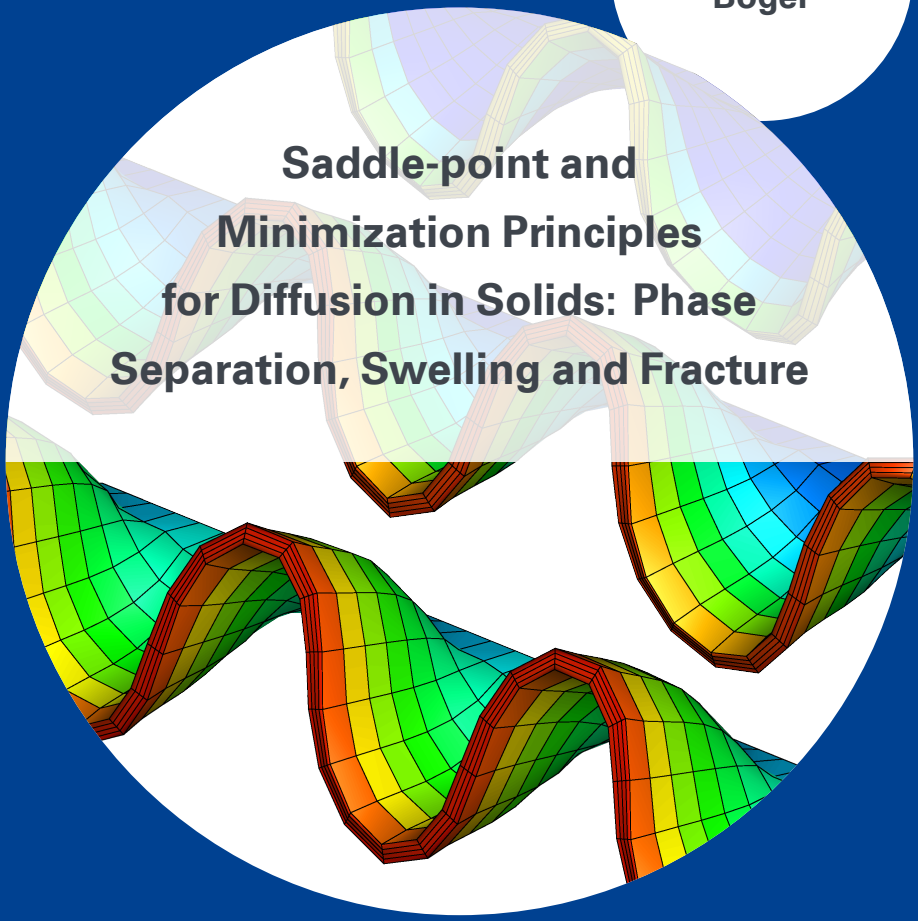
University of Stuttgart
Germany

Lukas
Böger

Saddle-point and Minimization Principles for Diffusion in Solids: Phase Separation, Swelling and Fracture



3

Publication series of the
Institute of Applied Mechanics (IAM)

Saddle-point and Minimization Principles for Diffusion in Solids: Phase Separation, Swelling and Fracture

Von der Fakultät Bau- und Umweltingenieurwissenschaften
der Universität Stuttgart zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

Vorgelegt von

Dipl.-Ing. Lukas Böger

aus Kiel

Hauptberichter:	Prof. Dr.-Ing. Holger Steeb
Mitberichterin:	Prof. Dr. Laura De Lorenzis

Tag der mündlichen Prüfung: 22. Februar 2019

Institut für Mechanik (Bauwesen) der Universität Stuttgart
Lehrstuhl für Kontinuumsmechanik, Prof. Dr.-Ing. H. Steeb

2020

Publication series of the Institute of Applied Mechanics (IAM), Volume 3
Institute of Applied Mechanics
University of Stuttgart, Germany, 2020

ISBN 978-3-937399-51-5
(D 93 – Dissertation, Universität Stuttgart)

Acknowledgments

This thesis is a result of my research under the supervision of Prof. Christian Miehe, my supporting, demanding, enthusiastic academic mentor who passed away too early. His death marked the disappearance of one of the most important people of my everyday life during the last years. The devotion to continuum mechanics he exemplified shaped his inimitable omnipresence at the institute, and I can barely remember a week without him entering the office, eager to hear about progress or challenges, willing to assist no matter how many hours that was about to take. He could be frenetic when discussing our work, questioned my competence and dedication (his criticism was always partly correct) and regularly left me in shaky self-doubt and anger. I did not exactly like him for these moments, and whether my uninspired periods stemmed from his way of overdoing supervision or my own mediocrity will never really be clear to me. About two points I am sure, however: First, Christian Miehe had a warm heart, and I appreciated that more often than I suffered from his inaccessibility. And second, it was impossible to not learn a lot from him. So far, I have not met a person with a more fundamental understanding of continuum mechanics, and he was always willing to share these insights with his students. This thesis is strongly influenced by him, and I am grateful for what I have learned during the last years!

Next, many thanks to Prof. Holger Steeb for his interest in my thesis and for taking over the examination of this work. It has been a strong and important support that he was willing to help me out from the moment on I entered his office. My gratitude goes to Prof. Laura De Lorenzis as well, again for her interest in this work and for being the co-examiner.

Many colleagues deserve to be mentioned. To start with, many thanks to Steffen Mauthe for the close cooperation and his unbreakable impetus to get things done, that saved me more than once. My gratitude goes to Felix Hildebrand, who was a great fellow for sharing things that stay in the office. He has always been capable of finding some nifty aspects of our daily work, even during dark days, when we agreed upon science being a large pot of shit everyone stirs from time to time. Thanks to Aref Nateghi for the collaboration and insights on Iranian politics, to Stephan Teichtmeister for his understanding of finite element spaces, to Fadi Aldakheel, Heike Ulmer, Gautam Ethiraj. To Nikolai Arnaudov, Felix Göküzüm and Alexander Müller, to Daniel Vallicotti for collective regret of our missing academic excellence, to Daniel Kienle, Dominic Zäh, Lisa Schänzel, Ashish Sridhar, Khiem Nguyen, Matthias Rambausek and Omkar Nadgir. There has never been room for lone fighters, envy or distrust in this research group. Thanks to Nadine Steinecke for calming everyone down in front of the lion's den, to Marc-André Keip for leading this group so impressively well when it was his turn.

The funding for my work has been supplied by the *German Research Foundation* (*DFG*) and the *Studienstiftung des dt. Volkes*. The latter institution assisted me in claiming more independence in working on my dissertation, which turned out to be invaluable. Additional thanks go to Prof. Klaas Enno Stephan from ETH Zürich, who helped this thesis by supporting that Lilian worked remotely from Stuttgart most of the time.

Last, thanks to Lilian that there is no need for cheesy words here, and for our even split of family and research obligations – it is good to know that I will be able to catch up for this.

Lukas Böger, May 2017

Contents

Zusammenfassung

Diffusion von Flüssigkeiten in Festkörpern ist ein zentraler Prozess in zahlreichen biomedizinischen, geotechnischen und auch alltäglichen Anwendungen. Wechselwirkungen mit makroskopischen Effekten sind vielfältig, auch die Formation von Mikrostrukturen kann als diffusive Phasenseparation erfolgen. Hierdurch werden Diffusionsphänomene für die moderne Kontinuumsmechanik interessant, da verschiedene räumliche Skalen wie auch multiple physikalische Effekte in Mehrfeld-Formulierungen eingebettet werden können. Diese Arbeit soll einen Beitrag leisten zu Theorie und computergestützter Simulation von spinodaler Entmischung und von Elastostatik in Wechselwirkung mit Fluidtransport und Materialversagen.

Zunächst werden die Kinematik großer Verzerrungen sowie Bilanzgleichungen der Newton'schen Mechanik für Kontinua eingeführt. Mit der Festkörpermaterie koexistiert eine einzelne flüssige Komponente. Deren Impuls wird vernachlässigt, Ansätze zur Beschreibung mehrerer Phasen in porösen Medien werden somit nicht angewandt. Ein Augenmerk gilt relevanten thermostatischen Prinzipien, da nur partiell offene Systeme den Aus- oder Eintritt von Flüssigkeit zulassen. Thermodynamische Hypothesen wie die lokale Längenskala für Gleichgewichtszustände von global heterogenen Zustandsgrößen werden aufgezeigt, und Elemente der konstitutiven Materialtheorie auf die zu lösenden Differentialgleichungen übertragen.

Diese Grundlagen werden auf drei spezifische Anwendungen zugeschnitten und jeweils in Variationsformulierungen für standard-dissipative und teils gradientenerweiterte Materialien integriert. Minimierungs- und/oder Sattelpunktformulierungen werden vorgestellt und mithilfe der Methode der finiten Elemente approximativ gelöst, anhand verschiedener Randwertprobleme untersucht und selektiv im Vergleich mit experimentellen Daten validiert.

Zunächst wird die Cahn-Hilliard-Gleichung auf einem starren Gebiet als Stationaritätsbedingung eines inkrementellen Potentials hergeleitet, dessen Existenz und Spezifikation ein signifikantes Novum darstellen. Die symmetrische Formulierung ist äquivalent zu bekannten Splitting-Verfahren, somit erhält man zwei Euler-Gleichungen zweiter Ordnung, die in drei Raumdimensionen mit feiner Vernetzung auf mehreren Prozessoren gelöst werden.

Anschließend wird die um Entmischung reduzierte Theorie auf elastisch deformierbare Hydrogele angewandt. Diese stark hydrophilen Polymere weisen ein breites Spektrum elastischer wie sensorischer Eigenschaften auf und zeigen mitunter extreme Wechselwirkungen zwischen Flüssigkeitsgehalt und Volumenänderung. Das zugehörige Variationsprinzip stellt der bekannten Sattelpunkt- ei-

ne neue Minimierungsvariante gegenüber, deren FE-Interpolation innerhalb der konformen Funktionenräume im Kontrast zur Sattelpunktformulierung kein Gegenstand einer Stabilitätsbedingung ist. Dieser Vorzug wird von komplizierteren Ansatzfunktionen und Randbedingungen begleitet, beide Varianten werden daher untersucht und gegeneinander abgewogen. Neben Studien zu geeignetem Elementdesign werden charakteristische Struktur-Instabilitäten simuliert und die wellenartigen Deformationsmuster mit experimentellen Ergebnissen verglichen.

Zuletzt wird das Phasenfeld-Verfahren der Bruchmechanik vorgestellt und auf Materialversagen von Hydrogelen angewendet. Das um einen Ordnungsparameter erweiterte, gekoppelte Problem kann ebenfalls als Sattelpunkt- und Minimierungsprinzip formuliert werden, und beide Varianten werden genutzt, um mechanisch und diffusions-induzierte spröde Versagensmechanismen mithilfe regularisierter Irreversibilität und gestaffelter Gleichungslöser zu simulieren.

Abstract

Liquid diffusion in solids plays a major role in countless biomedical, geotechnical and everyday applications. Interplays with various macroscopic effects are observable, and the formation of certain microstructures results from a diffusive phase separation, too. Diffusion phenomena are hence subject of modern continuum mechanics: different length scales and multiple physical effects can be embedded in multi-field formulations. This work will contribute to the theory and computational simulation of spinodal decomposition as well as elastostatics coupled to fluid transport and material failure.

To begin with, large strain kinematics and Newtonian balance equations for continua are introduced. Solid material and a single liquid with neglected momentum coexist, a theory of porous media for arbitrary constituents is hence not applied. Attention is paid to relevant thermostatic principles, as only partially open systems allow for the exchange of fluid mass. Hypotheses of thermodynamics like the local length scale of equilibrium postulates for globally heterogeneous state variables are illustrated, and elements of constitutive theory are transferred to the differential equations to be solved.

These foundations are employed for three specific problems, all of which are integrated into separate variational formulations for standard dissipative and possibly gradient-extended continua. Minimization and/or saddle point principles are presented and numerically solved with the finite element method, analyzed by means of suitable boundary value problems and partly validated by a comparison with experimental data.

First, the Cahn-Hilliard equation on a rigid domain is shown to be the Euler equation of an incremental potential – existence and specification of this potential is a key novelty of this part. The symmetric formulation is equivalent to known splitting methods, two second order differential equations are obtained and solved in three dimensions using a fine spatial discretization with parallel computing.

Second, the theory is ported to elastically deformable hydrogels, leaving out phase separation effects. These strongly hydrophilic polymers can have a broad range of elastic and sensory properties and are subject to a distinct, possibly extreme interplay between volume change and fluid content. The corresponding variational principle opposes the classical saddle point formulation with a new minimization formulation. Its finite element interpolation within the conforming function spaces is unconstrained, in contrast to the saddle point principle. This major advantage ships with additional complexities of the shape functions required for the sake of conformity and the application of

boundary conditions. Both formulations are analyzed and compared to each other. Studies of appropriate element design are performed, along with the simulation of characteristic structural deformation patterns which are matched against experimental findings.

Third, the phase-field method for brittle fracture is introduced and applied to crack initiation and growth in hydrogels. The coupled problem, equipped with an additional damage order parameter, can be formulated as saddle point and minimization formulation, too. Both methods are used for the simulation of mechanically and diffusion driven failure by a staggered solution scheme, while a non-smooth irreversibility condition is regularized by use of an energetic history field.

Chapter 1: Introduction

In this thesis, methods for the numerical modeling of possibly failing continua undergoing large deformation and fluid diffusion are developed. The treatise belongs to the field of computational continuum mechanics, with theoretical aspects and implementation details on a par. Abstractions and formalisms are necessary for the description of large deformation of solids, physical balance equations, traits of material laws and the construction of potentials that lead to the central equations by identification of a stationarity point. These partial differential equations can then be solved approximately by the finite element method that ships with its own technicalities like stability conditions and interpolation techniques.

Most effects of interest fall into the category of Newtonian mechanics for solid continua, i.e., the study of the motion of bodies that are phenomenologically observable and not resolved on molecular levels. This abstraction is appropriate for many engineering tasks like the analysis of load carrying capacities or the study of failure phenomena. Nevertheless, two shortcomings of this approach frequently limit the precision of such models. First, the phenomenological length scale can be insufficient to correctly capture certain effects. Models must then be enriched by microstructural influences on the material behavior. Second, accounting only for motion and forces is too narrow for the modeling of processes during which electrical, thermal, chemical or diffusive phenomena come into play. Additional fields and balance laws then accompany the Newtonian equations. Modern continuum mechanics deals with these two extensions. This is often called a *multiscale*, *scale bridging* and/or *multiphysics* approach (terms that fit well into grant applications). Here, both additional length scales and complementary effects stem from diffusion of one fluid and crack growth in the solid.

First, fluid diffusion obeys a mass conservation law, which becomes manifest in an additional differential equation as well as fluid concentration and chemical potential as additional field quantities. This already captures standard diffusion along a negative concentration gradient (downhill diffusion). A widened diffusion model incorporates the separation of phases, which can be described by accounting for concentration gradients and a so-called microforce balance equation, leading to the Cahn-Hilliard equation. This encompasses the formation of material microstructures as a combination of up- and downhill diffusion and hence a smaller length scale. Note that throughout this work, the ensemble of fluid and solid is treated as one homogenized continuum. Solid porosity is hence

approximately taken into account by assumptions of material properties, not by the superposition of a skeleton and multiple components.

Second, the failure of solids is approached by a phase-field order parameter that represents a damage state. Its evolution is determined by a supplemental differential equation for the approximation of sharp crack discontinuities and is substantiated by ideas from fracture mechanics. Such a procedure must account for gradients of the damage variable, and even though crack growth is a macroscopic effect, the methodological treatment of its regularized description is identical to that of continuum theories tackling microstructural effects.

Large deformation, downhill diffusion, phase separation and crack growth hence form the recurrent theme of this work. These effects must however be narrowed down to specific applications to allow for implementation and analysis: phase separation is studied on a rigid domain, while downhill diffusion and crack propagation are analyzed in hydrogels. The latter are functional prototype polymers with a broad range of elastic properties and seminal perspective applications. They show strong coupling mechanisms between volumetric deformation and fluid content, which makes them an ideal choice for the study of fluid transport in elastic media. The polymer network is treated as an isotropic, elastic, approximately incompressible solid – properties that will be formulated as a material law. These constitutive laws are a central ingredient for continuum theories, because differential balance equations and associated initial and boundary conditions are not enough to enable a problem solution: material laws close this gap. Almost ten constitutive functions will be specified. The Neo-Hookean response function determines elastic stress states, the Flory-Rehner energy contribution specifies the chemical potential in crosslinked polymers, and a volumetric coupling term bonds volumetric deformation and fluid content. Fick's law of diffusion relates chemical potential gradients to fluid flux vectors. One material law specifies chemical microforces for given microtractions and chemical potential, another one defines chemical microtractions – both are ingredients of the Cahn-Hilliard theory. The criterion for a crack onset can be considered a further material law (the energetic Griffith criterion being the most prominent example) as well as the degradation function that reduces the elastic energy storage in damaged zones. Finally, a constitutive law will be specified for the exclusion of crack self-healing and a rate-dependent crack evolution.

These constitutive functions are either established choices or were proposed in recent years – they are no innovation per se. Solving the set of differential equations with the finite element method is neither a novelty. So where is the justification for filling more than a hundred pages with content? Three innovations are believed to come along with this treatise. They are sketched in the following.

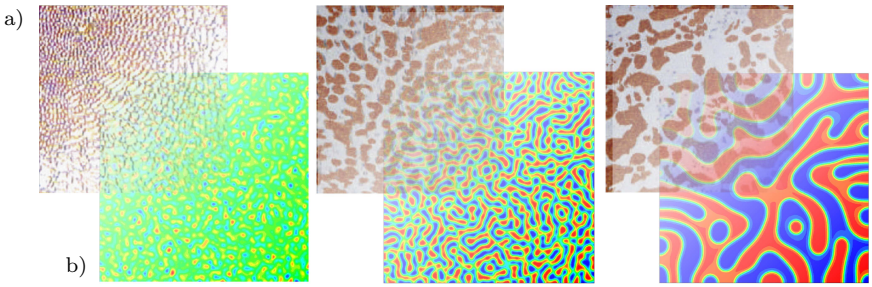


Figure 1.1: Microstructural phase separation: a) coarsening process in an Ag-Cu alloy, right after solidification and after 20 and 40 hours (left to right), taken from Anders et al. [2012], and b) exemplary finite element solution of the Cahn-Hilliard equation starting from random initial conditions.

1.1 A Variational Approach to the Cahn-Hilliard Equation

The Cahn-Hilliard equation was first formulated by Cahn & Hilliard [1958], and later subdivided into two balance equations and constitutive functions by Gurtin [1996]. This fourth order partial differential equation is an extended diffusion equation that accounts for interfacial energies and describes the continuous separation of two phases, e.g. during a decrease of temperature, see Figure 1.1. The process is often called spinodal decomposition. In contrast to reaction-diffusion equations, the Cahn-Hilliard equation does not consider creation or destruction due to chemical reactions, which renders the separation process purely diffusive. The order parameter governed by the Cahn-Hilliard equation is a normalized and conserved concentration that represents the phases.

Modeling of such decomposition phenomena is a key interest of non-equilibrium kinetics and important for the understanding of material microstructures. Accurate predictions of the process are even more important for the design of lithium-ion electrode materials that suffer from the large volumetric change associated with diffusive phase transformations during cycles of charging and discharging, see Anand [2012]; Di Leo et al. [2014] or Zhao et al. [2015]. Fortunately, the Cahn-Hilliard equation is well understood and has been subject to extensive studies. The existence of solutions was studied in dependence on material parameters by Elliott & Zheng [1986]. Approximate numerical solutions can be obtained by finite difference methods in the case of simple domain geometries, see e.g. Choo et al. [2004]. Finite element discretizations were proposed by Elliott & French [1987]; Feng & Prohl [2004]; Zhang et al. [2013] or Barrett et al. [2000]. These discretizations base on a split of the fourth order equation into two second order equations to circumvent the required C^1-continuity of interpolation functions. Alternative approaches can be based on the natural element method

(Rajagopal et al. [2010]), discontinuous Galerkin methods (Wells et al. [2006]) or isogeometric analysis (Gomez et al. [2008]; Anders & Weinberg [2011]).

While these contributions cover a broad spectrum of aspects that concern approximate solutions of the Cahn-Hilliard equation, they do not embed the equation into a variational formulation. In this work, continuous and incremental potentials for the description of phase separations are constructed that turn the problem into minimization and saddle point statements. Therewith, the phase decomposition can be integrated into theories for gradient-extended standard dissipative materials, see Miehe [2011]. This enforces the symmetry of the formulation, which offers possibilities to reduce the computational cost of a finite element solution, cf. Toh & Phoon [2008]. The potentials can also be interpreted to be a more instructive approach to the problem, though this point of view is not unquestionable. A spatial discretization of the time-discrete potentials leads to expressions for residual and tangent matrices that already comprise the split of the fourth order into two second order equations. Numerical benchmark simulations are provided to conclude the study of spinodal decomposition as the first diffusion aspect considered in this thesis.

1.2 Swelling in Hydrogels: a New Minimization Principle

To extend the study of fluid transport phenomena, downhill diffusion in hydrogels is considered in a next step. Hydrogels are functional materials that consist of a (possibly double) crosslinked polymer network. Their hydrophilic behavior can be accompanied by a high stiffness, which, together with the incompressibility of polymers, leads to strong swelling behavior when unsaturated gels are immersed into water. Chemical, thermal and electrical stimuli can furthermore control the macroscopic material behavior of hydrogels, which attributes a large variety of perspectives to this class of materials, see e.g. Hoffman [2012]; Kopecek [2009]; Murdan [2003]; Bruck [1973] or Calvert [2008]. Synthesis of hydrogels is e.g. described in Kong et al. [2003] or Chen & Park [2000], recent experiments on phase coexistence and separation in gels were contributed by Bai & Suzuki [2000], Bai & Suzuki [1999]. Swelling phenomena that lead to inhomogeneous deformation patterns or mechanical instabilities are the subject of Sultan & Boudaoud [2008]; Holmes et al. [2011] and Mora & Boudaoud [2006], see Figure 1.2.

In terms of modeling approaches, continuum theories were pioneered by Hong et al. [2008], with subsequent works by Chester & Anand [2010]; Duda et al. [2010] and Doi [2009], only to name a few. All of these approaches rely upon the constitutive energetic contribution of fluid content in crosslinked polymer networks formulated by Flory & Rehner Jr. [1943], coupled to elastic deformation and stresses. Thermal effects were studied by Chester & Anand [2011], chemical

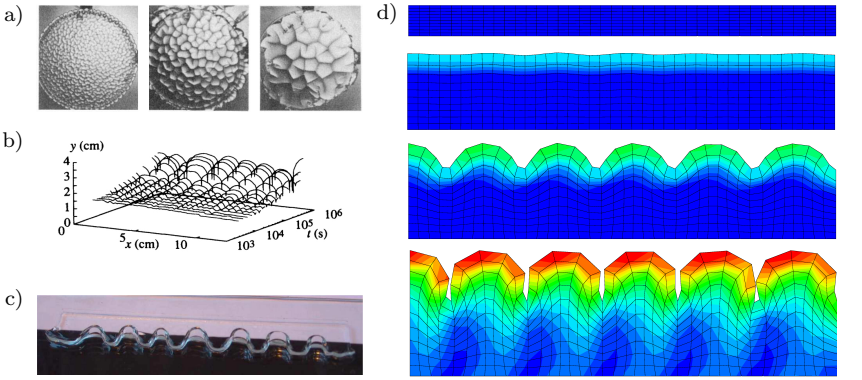


Figure 1.2: Formation of cusps as an instability pattern in hydrogel specimen caused by swelling and geometrical constraints: a) Top view of a petri dish, taken from Tanaka et al. [1987] b) swelling fold pattern, taken from Tanaka & Sigehuzi [1994] c) out of plane oscillations, taken from Sultan & Boudaoud [2008] d) showcase simulation of the mechanical instability, the contour depicts the chemical potential.

swelling by Dolbow et al. [2004] and Ballhause & Wallmersperger [2008], and electrical fields as external stimuli were analyzed by Wallmersperger & Ballhause [2008]. Finite element implementations with established or new element formulations were presented by Bouklas et al. [2015] or Krischok & Linder [2016], weakly imposed boundary conditions were studied by Ehlers & Acartürk [2008], and the formation of structural instability patterns was subject to the works of Zhang et al. [2009] and Liu et al. [2010].

An inherent characteristic of these approaches is the saddle point structure of the formulation. As a consequence, the finite element discretization is subject to the Ladyzhenskaya-Babuška-Brezzi (LBB) condition, see Babuška [1973] or Brezzi [1974]. Violations of this stability condition can lead to oscillatory patterns of primary variables, and, even worse, can be more subtle and stay unnoticed.

In this work, a new minimization principle is developed for the modeling of diffusion-deformation processes in hydrogels to overcome the shortcomings of the saddle point principle and to propose an alternative twin formulation. The formulation is similar to the one presented in Miehe et al. [2015a] for Biot's theory. Its main advantage is an independence of the LBB condition within the conforming element space. Also, the method seems to be slightly faster when comparing its computational cost with that of a classical saddle point principle. The new method comes along with the necessity to implement non-standard element formulations to fulfill the required conformity. Application of boundary conditions is furthermore complicated by the inversion of primary and secondary

variables. These challenges are addressed and shown to be resolvable by a set of boundary value problems.

1.3 Fracture Delay and Diffusion-Induced Failure in Hydrogels

The ability to predict material failure by models is an advantage for any engineering problem statement. As hydrogels have such a significance in biomedical applications (see the references in Section 1.2 or also Galaev & Mattiasson [1999] and Peppas et al. [2006]), the understanding of their rupture mechanisms is an important step towards trustworthy and reliable deployment of hydrogels as load bearing tissues, valves or similar components. To this end, the phase-field approach to fracture is taken as a promising method due to the simplicity of its implementation into existing finite element solvers and the ability to capture crack initiation or dendritic propagation.

Regularizing sharp crack interfaces to circumvent numerical difficulties associated with the handling of discontinuities has gained much attention during the last three decades. A reformulation of classical brittle fracture mechanics by Francfort & Marigo [1998] and Bourdin et al. [2000] led to Hakim & Karma's [2009] regularized interface description as a Ginzburg-Landau evolution equation. The idea was further established by Kuhn & Müller [2010]; Hesch & Weinberg [2014] and Miehe et al. [2010b], see Ambati et al. [2015] for a review of the different approaches. Besides the classical brittle fracture criterion, cohesive phase-field fracture was studied by Verhoosel & de Borst [2013], and recent advances in the modeling of ductile failure phenomena were contributed by e.g. Ambati et al. [2016] or Miehe et al. [2016]. Studies of fluid diffusion and phase-field fracture are often tailored to hydraulic fracturing, see e.g. Mikelic et al. [2013]; Wilson & Landis [2016] and Heider & Markert [2017]. Wu & Lorenzis [2016] presented an approach to model coupled diffusion and crack evolution in cement paste. Relevant experimental work besides measurements of the material strength deals with the failure delay of specimen, see Bonn et al. [1998], or the influence of the diffusing solute component, see Baumberger et al. [2006] or Baumberger et al. [2008].

In this work, the homogenized continuum model for polymeric hydrogels is bonded to a damage phase-field, aiming at a first implementation of crack propagation in these materials. Continuous and time-discrete minimization and saddle point principles are constructed by analogy with the variational formulation for a damage-free continuum. Both formulations are implemented by means of the finite element method, whereby the minimization principle again profits from the independence of the LBB condition. Following Bonn et al. [1998], fracture delay is studied as a mechanically driven problem. Two further

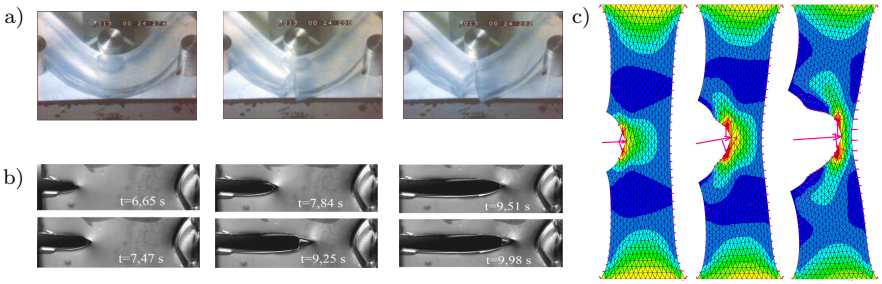


Figure 1.3: Tension induced, time dependent crack growth in simple specimen: a) Delayed fracture of gels in a three-point bending experiment, taken from Bonn et al. [1998] b) stepwise crack propagation, taken from Pizzocolo et al. [2013] and c) simulation of a creep test with viscous failure effects (an approach using configurational forces without fluid diffusion is shown as in Göküzüm [2013]).

benchmark simulations are presented, where the crack growth is caused by pure fluid transport.

1.4 Structure, Forecast, Characteristics of this Work

The structure of this work is as follows. Chapter 2 paves the way for a continuum field theory of Newtonian mechanics and fluid diffusion by a description of large deformations of continuous matter. This topic uses notions from differential geometry and is kept as short as possible to not reveal awkward misunderstandings of the author. It is followed by the introduction of two primary variables besides the deformation: concentration field and chemical potential. Both are integrated into ideas from non-equilibrium thermodynamics, before transport and stress quantities are introduced. Relevant balance laws are formulated, followed by the first and second axioms of thermostatics. Isothermal and quasi-static assumptions are postulated together with a restriction to elastic material behavior.

The necessity of constitutive laws is motivated at the beginning of Chapter 3. Established principles for the construction of material laws are formulated and applied to the constitutive functions in question. This includes an evaluation of the entropy inequality to impose constraints onto material laws. Anisotropic material behavior is furthermore excluded.

These first two chapters are a summary of state-of-the-art methods and conventions commonly employed for the formulation of problems in continuum mechanics. They also include several restrictions that narrow down the scope of the treatise. The most dominant one is the exclusion of viscoelastic, plastic or

viscoplastic solid material behavior. This simplification allocates the entropy production to fluid transport and the creation of fracture surfaces. An additional exclusion applies to mixture theories or the theory of porous media. While the study of diffusion in solids automatically brings up the question of how the porosity is taken into account, breaking it down to one scalar diffusivity parameter and a constitutive energetic contribution goes hand in hand with the assumption of an elastic solid response: it leaves room for a more sophisticated study of swelling, mechanical instabilities and failure phenomena. Finite element techniques and solution procedures are demanding enough to justify the above simplifications.

Spinodal decomposition is presented as the first application of the theory outline in Chapter 4. This part has been published in similar form as Miehe et al. [2014a]. It has a clear top-down characteristic: the innovation lies in the construction of incremental potentials for the Cahn-Hilliard equation. Finite element implementation and analysis of numerical benchmark problems serve as a proof of concept.

Free and constrained swelling in hydrogels, mechanical instabilities, singularities due to fluid-pressure interaction and a model validation against experimental data is presented in Chapter 5. This part ships with a novelty associated with variational statements, too, however, much more equilibrated with the effort necessary to implement a robust finite element solution: besides singular tangent matrices in boundary value problems with mechanical instabilities, the minimization principle demands for non-standard element formulation and complicates the application of boundary conditions. The work in this chapter has been published in similar form as Böger et al. [2017b] succeeding Böger et al. [2016].

An interlude is placed before the last application. Chapter 6 introduces the phase-field approach to brittle fracture mechanics. References to the relevant literature is collected, and notation and definitions are outlined. The phase-field evolution equation is developed as the Euler equation of a minimization principle in terms of the damage order parameter. The model follows Miehe et al. [2010b] and is a gradient damage theory with links to fracture mechanics where necessary. This model is said to be thermodynamically consistent because violations of the entropy inequality are excluded a priori and without further technicalities. The evolution equation however cannot be derived in Gurtin's elegant microforce approach, and a non-smooth irreversibility constraint requires special attention. To this end, an energetic history field is introduced following Miehe et al. [2010a] that goes hand in hand with a decoupled finite element solution of phase-field and other differential equations[1].

Chapter 7 combines this brittle fracture approach with the hydrogel model. The supplemental order parameter and its governing equation fit into both

[1]This approach is often called an operator split algorithm, which elegantly paraphrases the decoupling crudity.

minimization and saddle point formulation, hence again an emphasis is put on the variational formulation. This work has been published in similar form as Böger et al. [2017a] and is characterized by the junction of two methods: the link to the novel minimization formulation is collateral, computational models of the failure of hydrogels, however, are simply missing in the literature. The contribution hence fills this gap by proposing a method to solve the coupled problem.

Chapter 2:
Kinematics, Balance Laws and Non-Equilibrium Kinetics

In the following chapter, terms relevant for the description of an elastic solid with one or possibly two phases of fluid content undergoing large deformation at constant temperature are introduced. An overview of geometrical and physical quantities relevant for the present work is given together with balance laws stemming from mass conservation, Newtonian mechanics or different concepts. Furthermore, a link to thermodynamic restrictions and equilibrium states is made to sharpen the understanding of diffusive phenomena.

Throughout, the material is considered a mixture of one or two fluid constituents and solid (rendering the material *two-* or *three-phase*), simultaneously existing at the same location and at every material point, even if the number of fluid particles can be zero. Such an approach does neither take a porosity of the solid into account as in Biot's [1941] theory, nor does it include separate conservation laws for each component as it is done in the theory of porous media. This represents the most elementary view of diffusion in deforming solids, yet capable of capturing this process while simple enough to leave room for additional fields or phenomena of interest.

This chapter is influenced by the lectures C. Miehe taught at the University of Stuttgart and enriched by the notion of fluid components with associated balances and thermodynamic rationales. For embedding this outline into more complete studies of continuum mechanics, the reader is referred to e.g. Irgens [2008], Haupt [2002], Başar & Weichert [2000], Bonet & Wood [1997] or Wilmanski [1998]. On the side of thermodynamics, fluid transport is subject to non-equilibrium theories, often called *kinetics*. Here, the monographs of de Groot & Mazur [1984], Gyarmati [1967], Maugin [1999] and Lebon et al. [2008] should be mentioned as self-contained guides on the topic along with the very complete work of Balluffi et al. [2005] [2].

[2]Combining finite elasticity with diffusion and phase separation requires the conjunction of methodologies originally from continuum mechanics and those attributed to thermodynamics, which will exhibit some incompleteness here and there. This seems to be a common issue. Kestin [1992] states that ≪This marriage between thermodynamics and continuum mechanics was neither simple nor straightforward≫, and Truesdell & Toupin [1960, Sec. 245, p. 615 - 617] address similar challenges in their rant on thermodynamics.

2.1 Notation, Cartesian Base and Differential Operators

The notation in use will be as follows: uppercase symbols generally represent fields of the reference configuration, while lowercase symbols denote those of the current configuration (an explanation of these configurations is given below). Boldface symbols indicate tensor fields of an arbitrary order greater than zero (thus, vectors are included), lightface denote scalar fields. The dot product of two vectors is written as $(\cdot)\cdot(\cdot)$, the cross product as $(\cdot)\times(\cdot)$ and a dyadic product as $(\cdot)\otimes(\cdot)$. No operator glyph is used for the application of a tensor to another tensor. A colon refers to the contraction of two second-order tensors to a scalar.

Every tensorial object can be represented by a symbolic notation without any reference to basis vectors – this notation is preferred. Pure index notation referring to tensor components via sub- and superscript is avoided as well as a complete summation of tensor components and basis vectors. If the latter is used, sum signs are left out in favor of the Einstein summation convention. When the existence of a Cartesian coordinate system is presumed, orthonormal bases are written as $\boldsymbol{E}_A$, $\boldsymbol{E}^A$, $\boldsymbol{e}_a$ and $\boldsymbol{e}^a$. Possibly curvilinear coordinate systems span basis vectors $\boldsymbol{G}_A$, $\boldsymbol{G}^A$, $\boldsymbol{g}_a$ and $\boldsymbol{g}^a$, these coincide with the orthonormal bases most of the time, though. There is no formal need to distinguish between orthonormal bases in different configurations and vector spaces, but the ability to do so is retained to separate geometrical objects by their characteristic embedding.

The inverse of a tensor of second order or higher is written as $(\cdot)^{-1}$, its transpose $(\cdot)^T$, its inverse transpose $(\cdot)^{-T}$, its trace $\mathrm{tr}[(\cdot)]$ and its determinant $\det[(\cdot)]$. The unit tensor of second order is denoted by $\boldsymbol{1}$ with components represented by the Kronecker symbol, its fourth-order non-symmetric counterpart reads $\mathbb{I}$.

Differential operators are defined with respect to material coordinates without any subscript, and with a small boldface $\boldsymbol{x}$ subscript when referring to spatial coordinates (again, see below for an explanation). $\nabla(\cdot)$ is the gradient operator, $\nabla_{\boldsymbol{x}}(\cdot)$ its spatial counterpart. The divergence of a tensor is denoted by $\mathrm{Div}[(\cdot)]$ and $\mathrm{Div}_{\boldsymbol{x}}[(\cdot)]$, where the latter can also be written as $\mathrm{div}[(\cdot)]$ or $\mathrm{div}_{\boldsymbol{x}}[(\cdot)]$. The Laplace operator is referred to by an uppercase Delta $\Delta(\cdot)$, a variant for spatial coordinates by $\Delta_{\boldsymbol{x}}(\cdot)$. Finally, partial derivatives with respect to $(\cdot)$ usually listed as $\partial/\partial(\cdot)$ can be abbreviated as $\partial_{(\cdot)}$, and a *variational* or *functional* derivative of a scalar quantity $(\diamond)$ with respect to $(\cdot)$ is defined as $\delta_{(\cdot)}\diamond = \partial_{(\cdot)}\diamond - \mathrm{Div}[\partial_{\nabla(\cdot)}\diamond]$, which is a convenient notation for later use within variational formulations.

It is clearly beyond the scope of this chapter to comment on details of the above mentioned tensor operations, only a quick summary on the notation has been presented to guard against confusion. For details on tensor algebra and analysis, see e.g. Sokolnikoff [1951], Schade & Neemann [2009] or Itskov [2007]. In addition, a concise summary of tensor operations encountered in continuum mechanics is given in the appendix of Wriggers [2008].

2.2 Continuum Description of Large Deformation

The description of moving continuous matter without an emphasis on the cause of motion – usually called *kinematics* – acts as a starting point. Studying *large* or *finite* deformation merely means not studying *small deformation*; no assumptions are made to simplify geometric quantities by linearization. As finite deformation kinematics are often built on top of differential geometry, the subject tends to become complex, and the literature background is huge. The following shall be understood as a short overview, see Steinmann [2015][3], Jog & Hutter [2003], Marsden & Hughes [1994][4] for more detailed studies of kinematics in the context of differential geometry and Bröcker & Jänich [1973] or Carroll [1997] for differential geometry itself.

2.2.1 Motion of a Material Body and its Configurations

Characterization of the Motion of a Body. A *material body* $\mathcal{B}$ is defined as the collection of *material points* P that represent matter and continuously fill space. It is the carrier of all physical processes considered in this work. Each one-to-one mapping of material points to three-dimensional coordinates in the Euclidean space $\mathcal{R}^3$ is called a *configuration* and can be written as

$$\chi : \begin{cases} \mathcal{B} \to \chi(P) \subset \mathcal{R}^3 \\ P \mapsto \chi(P) = (x_1, x_2, x_3). \end{cases} \tag{2.1}$$

The bijective nature of χ ensures that matter does not interpenetrate itself. To identify the motion of a body, a set of configurations at instances of time is introduced, where the configuration at time $t \in T \subset \mathcal{R}_+$ is denoted by $\mathcal{S}$,

$$\chi_t : \begin{cases} \mathcal{B} \times T \to \mathcal{S} \subset \mathcal{R}^3 \\ (P, t) \mapsto \chi_t(P) = (x_1(t), x_2(t), x_3(t)). \end{cases} \tag{2.2}$$

A composition of one configuration and the inverse of another one can directly map coordinate triples between the two configurations. This ability is employed to avoid detouring to the particle set P; a specific configuration is chosen as a reference for the motion. This choice is arbitrary, but often associated with an initial state of the body (e.g. stress-free in case of purely mechanical problems) and enables for identifying material points by their reference coordinates – the associated configuration is called the *reference configuration*.

Mapping between Reference and Current Configuration. Relating configurations χ_t to a reference configuration called $\mathcal{B}$ induces the terminology *current configuration* $\mathcal{S}$ for $\chi_t(P)$ at time t. Reference and current configuration are

[3] The reader should expect a certain amount of frustration when studying this book.

[4] Some abstractions in this monograph seem to prove downward-compatibility of mathematics to engineering disciplines rather than clarifying the subject.

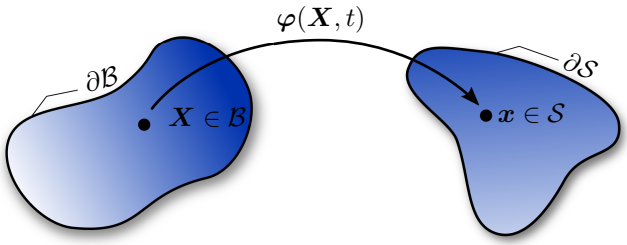


Figure 2.1: Thou shalt not write a PhD thesis in continuum mechanics without this drawing. The deformation map $\boldsymbol{\varphi}$ maps points $\boldsymbol{X}$ of the reference configuration $\mathcal{S}$ onto points $\boldsymbol{x}$ of the deformed configuration $\mathcal{S}$.

in addition named *material* and *spatial* or *Lagrangian* and *Eulerian* configuration. When material coordinate triples of the reference configuration are denoted $\boldsymbol{X} = (X_1, X_2, X_3) \in \mathcal{B}$, the mapping

$$\boldsymbol{\varphi} : \begin{cases} \mathcal{B} \times T \to \mathcal{S} \\ (\boldsymbol{X}, t) \mapsto \boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t) \end{cases} \tag{2.3}$$

can be introduced as the *deformation map*. It characterizes the motion of the material body with respect to the reference configuration and transforms material coordinates $\boldsymbol{X}$ into spatial coordinates $\boldsymbol{x}$ at time t. This bijective map is assumed to be continuously differentiable once with respect to material coordinates and twice with respect to the time. Fixing material coordinates $\boldsymbol{X}$ gives the time-dependent path of one particular material point, starting at the reference configuration. Fixing time t renders all material points of the body at a particular frozen time t.

Spatial and Material Quantities and their Derivatives. Whenever a quantity describes physical events in the body (e.g. temperature, fluid concentration, tractions), it can be allocated to a specific particle out of P. With the notion of reference and current configuration at hand, these quantities are understood as *fields* parameterized by current or reference coordinates and time. Consider the arbitrary field W as an example,

$$W(\boldsymbol{X}, t) = W(\boldsymbol{\varphi}^{-1}(\boldsymbol{x}, t), t) = w(\boldsymbol{x}, t). \tag{2.4}$$

It has been expressed as a material and as a spatial quantity and yields an identical value. The so-called *material time derivative* reads

$$\frac{d}{dt} W(\boldsymbol{X}, t) = \frac{\partial}{\partial t} W(\boldsymbol{X}, t) = \dot{W}(\boldsymbol{X}, t). \tag{2.5}$$

Here, total and partial time derivative coincide, because the material coordinates do not depend on time (this is not necessarily the case, see the mechanics

of configurational forces). In terms of spatial coordinates the material time derivative is

$$\frac{d}{dt}w(\boldsymbol{x},t) = \frac{\partial}{\partial t}w(\boldsymbol{x},t) + \nabla_{\boldsymbol{x}}w(\boldsymbol{x},t)\cdot\frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{\varphi}^{-1}(\boldsymbol{x},t),t) = \dot{w}(\boldsymbol{x},t). \tag{2.6}$$

The result has a local and a convective summand, but is independent of material coordinates. The latter is especially useful in fluid dynamics, where the reference configuration is unknown. Without such a restriction, material time derivatives are mostly applied to quantities parameterized by material coordinates in the following.

Velocity and Acceleration. With the above definitions at hand, the velocity of a material point is defined as the material time derivative of the deformation map,

$$\boldsymbol{V}(\boldsymbol{X},t) = \frac{\partial}{\partial t}\boldsymbol{\varphi}(\boldsymbol{X},t) = \frac{d}{dt}\boldsymbol{\varphi}(\boldsymbol{X},t), \tag{2.7}$$

while the second derivative yields the acceleration of a material point,

$$\boldsymbol{A}(\boldsymbol{X},t) = \frac{\partial^2}{\partial t^2}\boldsymbol{\varphi}(\boldsymbol{X},t) = \frac{d^2}{dt^2}\boldsymbol{\varphi}(\boldsymbol{X},t). \tag{2.8}$$

The velocity can be parameterized by spatial coordinates by inserting the inverse deformation map, which reads

$$\boldsymbol{v}(\boldsymbol{x},t) = \boldsymbol{V}(\boldsymbol{\varphi}^{-1}(\boldsymbol{x},t),t). \tag{2.9}$$

Material and spatial velocity describe an identical kinematical quantity with only the parametrization that differs. The same applies to the acceleration,

$$\boldsymbol{a}(\boldsymbol{x},t) = \boldsymbol{A}(\boldsymbol{\varphi}^{-1}(\boldsymbol{x},t),t). \tag{2.10}$$

2.2.2 The Most Fundamental Strain Measure

The deformation map and its time derivatives capture the characteristics of global motion. For the formulation of physical quantities and balance laws however, kinematical quantities that describe *local* phenomena of motion are needed. To fulfill this demand, numerous *strain measures* exist. Each has its own mapping characteristics and a different geometrical meaning. Here, the scope is limited to the most important ones, starting with the *deformation gradient*.

Consider a curve $\boldsymbol{W}(\beta)$ in the undeformed configuration, parameterized by the scalar β. The material points it exists on are mapped onto spatial points by (2.3), and the curve in the deformed configuration is denoted by $\boldsymbol{w}(\beta)$. The relation between the two curves reads

$$\boldsymbol{w}(\beta) = \boldsymbol{\varphi}(\boldsymbol{W}(\beta),t). \tag{2.11}$$

A vector tangent to the spatial curve is given by its differential, and using the chain rule yields

$$d\boldsymbol{w}(\beta) = \frac{d}{d\beta}\boldsymbol{w}(\beta) = \frac{d}{d\beta}\boldsymbol{\varphi}(\boldsymbol{W}(\beta), t) = \partial_{\boldsymbol{X}}\boldsymbol{\varphi}(\boldsymbol{X}, t) d\boldsymbol{W}(\beta). \tag{2.12}$$

This identifies a linear *tangent map*, which renders a spatial tangent vector by application to its material counterpart tangent. This map is defined as the deformation gradient

$$\boldsymbol{F}(\boldsymbol{X}, t) = \nabla\boldsymbol{\varphi}(\boldsymbol{X}, t) \tag{2.13}$$

and its determinant is denoted by

$$J = \det[\boldsymbol{F}]. \tag{2.14}$$

The deformation gradient describes the deformation of an infinitesimal line element and thus captures the deformation in the local neighborhood of $\boldsymbol{X}$.

There are two additional geometrical mappings related to $\boldsymbol{F}$, with the first being called the *area map*: consider a second material curve $\bar{\boldsymbol{W}}(\bar{\beta})$ with tangent vector $d\bar{\boldsymbol{W}}(\bar{\beta})$ (lowercase $\bar{\boldsymbol{w}}$ for the spatial counterpart). Its cross product with $d\boldsymbol{W}(\beta)$ spans an infinitesimal area element, and Nanson's formula gives

$$d\boldsymbol{w}(\beta) \times d\bar{\boldsymbol{w}}(\bar{\beta}) = J\boldsymbol{F}^{-T} d\boldsymbol{W}(\beta) \times d\bar{\boldsymbol{W}}(\bar{\beta}). \tag{2.15}$$

As cross products form normal vectors, this equation reveals the nature of the inverse transpose deformation gradient $\boldsymbol{F}^{-T}$, mapping normal vectors of the reference configuration onto normal vectors in the deformed configuration.

After infinitesimal line and area elements, an infinitesimal volume element is considered. This parallelepiped can be constructed as the triple product of $d\boldsymbol{W}(\beta)$, $d\bar{\boldsymbol{W}}(\bar{\beta})$ and the tangent vector $d\tilde{\boldsymbol{W}}(\tilde{\beta})$ of a third curve $\tilde{\boldsymbol{W}}(\tilde{\beta})$ (again, lowercase $\tilde{w}$ for the curve in the current configuration). Again leaving out the details, it can be summarized that

$$d\boldsymbol{w}(\beta) \cdot \left(d\bar{\boldsymbol{w}}(\bar{\beta}) \times d\tilde{\boldsymbol{w}}(\tilde{\beta})\right) = J\left[d\boldsymbol{W}(\beta) \cdot \left(d\bar{\boldsymbol{W}}(\bar{\beta}) \times d\tilde{\boldsymbol{W}}(\tilde{\beta})\right)\right]. \tag{2.16}$$

Thus, the determinant of the deformation gradient maps infinitesimal volume elements of the reference configuration onto their deformed volume elements in the current configuration. This *volume map* and the area map will come in handy when formulating balance principles, which first appear as integrals over domain or surface in the current configuration, but can then be traced back to material integrals over a domain or surface that does not change with time.

Another aspect that arises from the identification of the volume map concerns a constraint of $\boldsymbol{\varphi}(\boldsymbol{X}, t)$. As stated earlier, the deformation map is bijective,

which already ensures $\det[\boldsymbol{F}] \neq 0$. Additionally, the volume map dictates the assumption

$$\det[\boldsymbol{F}] > 0, \tag{2.17}$$

which strengthens the constraint on (2.3) by preventing non-physical deformation states. This demand is sometimes called the *permanence axiom of matter*.

2.2.3 Manifolds, Tangent and Cotangent Spaces

So far, neither a definition of spaces on which mappings and quantities are defined, nor a characterization of reference and current configurations has been proposed (besides being sets of points). Doing so is a non-trivial task: the established, canonical approach in continuum mechanics employs a fair amount of differential geometry. Thus, the somewhat arbitrary choice where to limit an introduction to terms of differential geometry comes into play – one might even feel tempted to skip the whole block. Some characteristics of the fields in use can however best be understood with a minimum exposure to the ideas of *tangent* and *cotangent spaces* on *differentiable manifolds*[5]. For an in-depth discussion of this topic, see Jänich [1993], Tu [2010] or the more comprehensive lecture notes by Carroll [1997]. Again, Marsden & Hughes [1994] cannot be omitted here.

Manifolds add a layer of abstraction to the description of bodies and surfaces, there is no necessity to describe the whole geometry in a Euclidean vector space. It is instead assumed that the body can be *locally* approximated by the Euclidean space. This local approximation is called a *chart* and a transformation that maps a subset of material or spatial coordinates one-to-one into the Euclidean space. A collection of non-overlapping or smoothly overlapping charts that cover the whole set (material or spatial points) is called an *atlas*. The set of points together with an atlas that covers the whole body is called the *manifold* – thus, for the following, there are two manifolds, one that represents the reference configuration and one for the current configuration.

Multiple charts can be composed into a coordinate transformation that is C^∞-continuous by assumption and possibly more complicated than linear transformations. The manifold is then said to be C^∞-continuous; its coordinate transformations between charts are arbitrarily often continuously differentiable. For the purpose of this work, a C^2-manifold suffices, but the generalization to C^∞-continuity comes without additional difficulty. Recall that (2.3) is a composition of mappings and was assumed to be differentiable at least once with respect to material coordinates and twice with respect to time. This demand is a priori satisfied with $\mathcal{B}$ and $\mathcal{S}$ being C^∞-manifolds.

[5]The unofficial reasoning for this subsection is that it is hard for PhD students to not write down the theory parts they have understood to some extent.

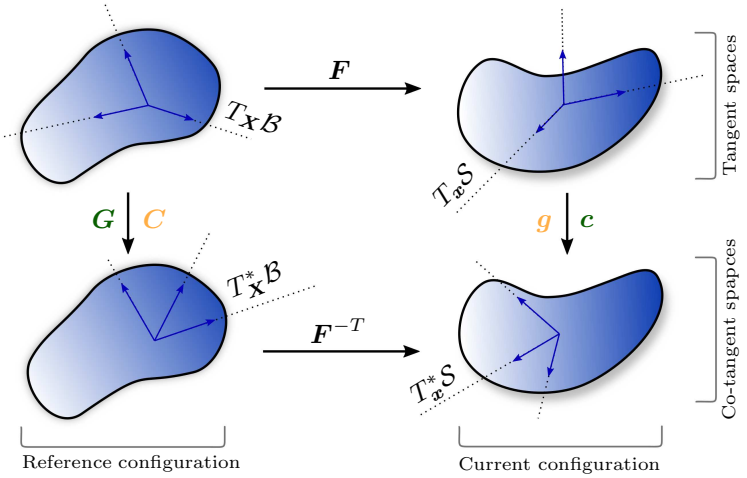


Figure 2.2: Overview of tangent and co-tangent spaces on reference and current configuration. The metric tensors $\boldsymbol{G}$ and $\boldsymbol{g}$ map objects from tangent to co-tangent spaces, as do the left and right Cauchy-Green tensors $\boldsymbol{C}$ and $\boldsymbol{c}$. Red and blue fonts indicate which metrics are related by push-forward and pull-back operations.

Tangent Spaces. In the formalism of differential geometry, manifolds are commonly defined intrinsically, i.e., without reference to an ambient Euclidean space. Therefore, the natural objects to integrate over a manifold are differential forms (and the balance laws introduced in Section 2.5 are all integral in nature). Vectors should thus be understood as directional derivative operators on curves and not pointing from some location to another. A vector space in which these operators are embedded is called the *tangent space*, denoted by $T_{\boldsymbol{X}}\mathcal{B}$ and $T_{\boldsymbol{x}}\mathcal{S}$ for reference and current configuration, respectively. Their coordinate bases are $\boldsymbol{G}_A$ and $\boldsymbol{g}_a$, and elements of the space are called *covariant* vectors.

Cotangent Spaces. An element that maps an arbitrary tensorial object on $T_{\boldsymbol{X}}\mathcal{B}$ or $T_{\boldsymbol{x}}\mathcal{S}$ onto tensorial objects on $\mathcal{R}$ is called a *one-form* and exists in a vector space *dual* to the tangent space. These two dual *cotangent spaces* are denoted $T^*_{\boldsymbol{X}}\mathcal{B}$ and $T^*_{\boldsymbol{x}}\mathcal{S}$, again for reference and current configuration, respectively. Suitable one-form examples are the gradient of a scalar function or normal vectors on surfaces (e.g. $d\boldsymbol{W} \times d\tilde{\boldsymbol{W}}$ from above). Elements of the cotangent spaces are *contravariant* quantities with coordinate bases $\boldsymbol{G}^A$ and $\boldsymbol{g}^a$.

Metric Tensors. The need to compute lengths and angles by scalar products on a possibly curved body gives rise to the *metric tensors* of second degree. Though there is no physical meaning attached to the mapping of a tangent to a one-form or the inverse, they are often depicted as such maps. Another effect of using these

maps is raising and lowering of indices. Here, $\boldsymbol{G}$ is denoted as the Lagrangian metric and $\boldsymbol{g}$ as the Eulerian. For orthonormal coordinate systems, they collapse to unity tensors.

Pull-Back and Push-Forward. *Pull-back* operations stem from the composition of a field defined on $\mathcal{S}$ (parameterized in terms of spatial coordinates $\boldsymbol{x}$) and the deformation map (2.3), rendering the field pulled back by a function of material coordinates $\boldsymbol{X}$. This concept cannot be trivially transferred to vectorial or tensorial objects, nor is there an intrinsic *push-forward* counterpart. Since (2.3) is however assumed to be invertible and smooth and the inverse of (2.3) again smooth, the push-forward of a field depending on material coordinates is straightforward. An additional outcome of this demand is the construction of pull-backs and push-forwards for arbitrary tensors by application of the chain rule, see e.g. Carroll [1997, p. 129 et sqq.], resulting in mappings that are formed by the deformation gradient, its transpose, its inverse or the inverse transpose. It is not misleading to think of these operations as generalized coordinate transformations, but crucial to keep in mind that co- and contravariant objects are pulled back and pushed forward differently.

2.2.4 Further Conventional Measures of Strain

First, with the notions of tangent and cotangent spaces at hand, the deformation gradient can be reformulated more precisely by its mapping characteristics. It maps tangent vectors in $T_{\boldsymbol{X}}\mathcal{B}$ onto tangent vectors in $T_{\boldsymbol{x}}\mathcal{S}$. Likewise, the normal map $\boldsymbol{F}^{-T}$ can now be understood as a mapping from $T^*_{\boldsymbol{X}}\mathcal{B}$ to $T^*_{\boldsymbol{x}}\mathcal{S}$. Both mappings are *two-point tensors*, as their mixed bases stem from tangent and co-tangent spaces. Because the deformation gradient is non-singular, it can multiplicatively be decomposed into a rotation and a positive definite tensor using the polar decomposition theorem. The *right* and *left polar decomposition* read

$$\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} = \boldsymbol{v}\boldsymbol{R}. \tag{2.18}$$

Due to $\boldsymbol{R}$ being a proper orthogonal tensor, its mapping of vectors can geometrically be interpreted as a rotation, whereas $\boldsymbol{U}$ and $\boldsymbol{v}$[6] represent pure stretches of a vector.

Next, consider a scalar product of the two tangent vectors $d\boldsymbol{W}$ and $d\bar{\boldsymbol{W}}$ from above, where the dependence on β and $\bar{\beta}$ has been dropped for a short notation. A comparison with the same contraction of deformed tangent quantities yields

$$d\boldsymbol{w} \cdot d\bar{\boldsymbol{w}} = \boldsymbol{F}d\boldsymbol{W} \cdot \boldsymbol{F}d\bar{\boldsymbol{W}} = d\boldsymbol{W} \cdot \boldsymbol{F}^T\boldsymbol{F}d\bar{\boldsymbol{W}}, \tag{2.19}$$

[6] Confusion between the stretch tensor and the spatial velocity due to identical boldface symbols can be prevented by brainy context-awareness.

where the symmetric, positive semi-definite *right Cauchy-Green tensor* can be identified as

$$\boldsymbol{C} = \boldsymbol{F}^T \boldsymbol{F}, \tag{2.20}$$

mapping vectors of the material tangent space $T_{\boldsymbol{X}}\mathcal{B}$ onto material one-forms in $T^*_{\boldsymbol{X}}\mathcal{B}$. Note that this strain tensor is often referred to as the Eulerian metric pulled back to the reference configuration. And indeed, a pull-back of $\boldsymbol{g}$ results in $\boldsymbol{F}^T\boldsymbol{g}\boldsymbol{F}$, which is identical to (2.20) as can be seen when performing the contraction of non-orthonormal basis vectors in both expressions.

By analogy, the *left Cauchy-Green tensor* can be found by comparing the scalar product of two spatial tangent vectors with that of the corresponding undeformed objects,

$$d\boldsymbol{W} \cdot d\bar{\boldsymbol{W}} = \boldsymbol{F}^{-1} d\boldsymbol{w} \cdot \boldsymbol{F}^{-1} d\bar{\boldsymbol{w}} = d\boldsymbol{w} \cdot \boldsymbol{F}^{-T}\boldsymbol{F}^{-1} d\bar{\boldsymbol{w}}, \tag{2.21}$$

where the symmetric, positive semi-definite tensor $\boldsymbol{c}$ is identified as

$$\boldsymbol{c} = \boldsymbol{F}^{-T}\boldsymbol{F}^{-1}. \tag{2.22}$$

Both left and right Cauchy-Green tensor are unaffected by rigid body motions such as rotations. This can be seen by their relation to the stretch tensors from a polar decomposition,

$$\boldsymbol{C} = \boldsymbol{U}^2 \quad \text{and} \quad \boldsymbol{c}^{-1} = \boldsymbol{v}^2, \tag{2.23}$$

and allows for a use as objective strain measures, which will be of value when facing objectivity demands later on.

Further strain measures include the spatial *Almansi* and the corresponding material *Green* strain tensor (the latter has special meaning as it boils down to the small strain tensor upon linearization) or the *Seth-Hill* strain tensors. The treatise could be continued to the study of strain rates and objective strain rates, but such an extent is clearly beyond the scope of the current introductory chapter.

2.3 Fluid Concentration and Chemical Potential

Diffusion is the transport of atoms or molecules (called particles in the following) caused by their random Brownian motion and collisions between these particles. Collisions happen more frequently when the number of particles in some volume is high and the so-called *mean free path* is shorter, which already gives a hint on the most basic macroscopically observable mechanism of diffusion: the transport of particles from regions with a higher number of particles to those with a lower number. This obviously resembles the logic behind heat transfer, an analogy which is as legitimate as it motivated Fick [1855] to formulate a law for species migration of salt in water analog to Fourier's description of heat transfer, see Crank [1975]. Fick's experiments, modeled after Graham's [1861] pioneering

experimental work, dealt with an aggregate of two substances called a *dilute solution*, in which less numerous *solute* particles move faster than the *solvent* particles. However spatial inhomogeneities of the number of particles not only force the diffusion of dilute solute particles in a medium. Self-diffusion by vacancy exchanges or in chemically homogeneous solutions with two constituents is driven by such gradients, as well as the transport of substitutional particles in metals, see Balluffi et al. [2005].

There are more circumstances besides spatial variations of the particle density which cause mass transport in a medium. Facing the above mentioned Brownian motion as the cause of diffusion, the most obvious ones are non-isothermal conditions that lead to different activity levels of particles and thus thermo-diffusion. Electric fields and charged solutions, stress fields or the evolution of dislocations due to plastic deformation should exemplary be mentioned as the most important effects.

The literature relevant for diffusive mass transport is vast, and only a few references can be given here. For treatments of diffusion as an irreversible thermodynamic process, see e.g. Bokstein et al. [2005], or the references in the introductory paragraphs of this chapter. Analytical solutions of the diffusion equation are given in Crank [1975], Barrer [1941] or Ghez [1988], while Mehrer [2007] and Balluffi et al. [2005] give a very complete overview of diffusion processes including mass transport in various media and phase separation phenomena. The latter is the unmixing process of two fluid constituents in a *binary solution* and results from a combination of particle transport along and against the gradient of the particle density (*uphill* and *downhill* diffusion). This phenomenon shows non-local diffusion effects and can describe a microstructure evolution, which renders its significance for microstructure-aware continuum theories.

Some assumptions and terms will be used throughout the following, and they deserve a clarification at the beginning. First, mass transport is limited to fluid mass transport (this fact is reflected in the mass balances, see Section 2.5.1). Though this is not necessary for the validity of diffusion theories, keeping terms and equations generalized for mass transport of other substances than fluids does not reveal insights relevant for this work. Second, fluid mass transport is limited to *one or two components*. In the latter case, the diffusive state can be described by one set of variables associated with a component. This is due to the assumption that available space in a solid skeleton can be occupied by either one or the other component, but voids are impossible. The two constituents are then called a *binary solution*, and the spatial distribution of one component is identified with the space not occupied by the other. An alternative interpretation of this single variable describing the spatial distribution of a binary solution is its role as an *order parameter* that indicates domains within a material. This will be clarified in Chapter 4. For now, the most important consequence concerns the notation, as the circumstances described in this paragraph allow a simple labeling for mass

transport variables: an arbitrary number of components (which is usually taken into account), each of which is identified by a subscript, is unnecessary here; one single set of variables without subscripts is sufficient.

Chemical reactions, phase transitions, electric fields, thermal diffusion, capillarity effects or kinds of diffusion other than fluid mass transport are not considered in this work. This reduces the complexity of energetic contributions to the body as well as terms relevant for fluid mass balances (chemical reactions could create a mass supply, for example). These assumptions furthermore narrow down the effects that cause fluid diffusion to the aforementioned spatial variations of the number of solvent particles and thus provide a good precondition for bringing Fick's [1855] findings to fluid diffusion in a solid.

The following list of terms summarizes the above outline by identification of terms with solid and fluid matter:

- A *solution* is composed of *solvent* and *solute particles*. An *ideal solution* obeys certain constitutive assumptions parallel to an ideal gas, a *dilute solution* is characterized by a small number of solute particles compared to the number of solvent particles as well as a low velocity of solvent compared to solute particles. The variety of molecules forming the solvent is not restricted.
- A *fluid* is an ensemble of solute particles, water is the simplest example.
- The *material body* is the solid formation of all solvent particles into a skeleton, e.g. a crystal or a polymer chain network. Possible complexities of the material structure below the macroscopic level are neglected by this viewpoint.
- A *binary solution* comprises two fluids without any restriction on the ratio of particle numbers. Thus, the identification of solvent and solute particles does not make sense here.

One should keep in mind that these illustrations are meant to include a loss of generality and point out how terms are used in the present work, not across-the-board.

2.3.1 Scalar Concentration Field per Unit Volume

With a little effort, the term *concentration* has been omitted in the paragraphs above to make the following definition appear less profane. Consider a number of solute particles of one kind moving through a domain of time-dependent shape, i.e., the current configuration of the material body. Again, the material skeleton consists of solvent particles that form solid matter and move by far less fast than solute particles.

Number of Particles per Volume. At every point $\boldsymbol{x} \in \mathcal{S}$, a positive number of solute particles, superimposed onto or co-existing with solid matter, is measured *per unit volume* by the *concentration* $c_s(\boldsymbol{x}, t)$. The pull-back of this scalar field

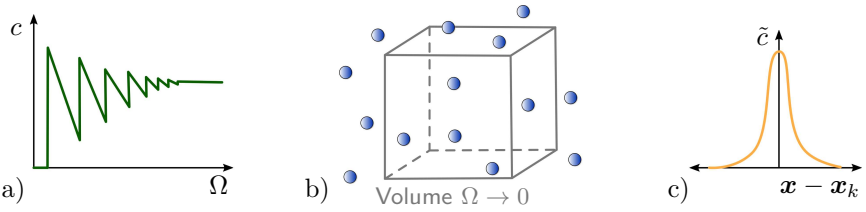


Figure 2.3: The challenge of defining a suitable concentration field stems from the limit case of infinitesimal volumes: In- or exclusion of one particle leads to a non-smooth concentration function as depicted in a) and b), while the convolution function $\tilde{c}(\boldsymbol{x} - \boldsymbol{x}_k)$ in c) is a weight of the particle k, which by (2.26) provides a smooth concentration function (after Balluffi et al. [2005, p. 8-9]).

to the reference configuration is denoted by $c(\boldsymbol{X}, t)$ (free of any subscripts) and a central field to the problem statements in this work. The mapping characteristics of c are identified as

$$c : \begin{cases} \mathcal{B} \times T \to \mathcal{R}_+ \\ (\boldsymbol{X}, t) \mapsto c(\boldsymbol{X}, t). \end{cases} \tag{2.24}$$

Material and spatial concentration can be related by comparing the number of species particles per deformed and undeformed volume element and a composition with the deformation map,

$$c = J c_s \circ \boldsymbol{\varphi}. \tag{2.25}$$

As intuitive as this first mention of a spatially varying concentration field has been, as important is a look ahead side note on its relation to the methodology of classical thermodynamics, where statements exclude inhomogeneous states. Stating a dependence of c and c_s on the position is thus an assumption that will be justified not before an equilibrium length scale comes into play in Section 2.6.

Convolution Function. One technical challenge that arises from this definition shall be faced here and in line with Balluffi et al. [2005]: counting the number of particles does not lead to a smooth function when the volume element becomes smaller and smaller. Instead, a particle excluded from the reduced volume element will cause the concentration to jump to a lower value. By further lessening the volume without exclusion of a particle, the concentration will continuously raise again, until the next particle drops out the volume element, see Figure 2.3. To overcome this issue, the spatial concentration is not defined by the number of particles per volume element, but by weights of all particles depending on the position $\boldsymbol{x}$. To this end, let N_Ω denote the number of particles

in a volume element $\Omega \subset \mathcal{S}$. With the *local convolution* $\tilde{c}(\boldsymbol{x})$, the concentration reads

$$c_{\mathcal{S}} = \frac{\sum_{k=1}^{N_\Omega} \tilde{c}\,(\boldsymbol{x} - \boldsymbol{x}_k)}{\int_\Omega \tilde{c}(\boldsymbol{x}) dv}. \tag{2.26}$$

In the limit case, $c_{\mathcal{S}}$ evaluates to approximately $N_\Omega/|\Omega|$, and for arbitrary volume elements, the concentration is continuously differentiable depending on the choice of $\tilde{c}$, which typically resembles some distribution function shaped by a *coarse-grained length* parameter, see Figure 2.3. This choice of $\tilde{c}$ shall, however, not be of importance, as from now on and again following Balluffi et al. [2005], the specific convolution function is expected to have negligible influence on the resulting concentration field.

Characteristics of the Concentration. Now the concentration is defined for arbitrarily small volume elements, a state without any solute particles is obviously rendered by $c = 0$, and there is no a priori upper limit for the concentration. Most often, mass transport takes place and the concentration field evolves, when ∇c does not vanish, though this is no obligatory condition.

Note that the concentration field may be related to an arbitrary reference number n_{ref}, such that the number of particles per unit volume is cn_{ref} and values of the concentration itself are in another range, e.g. $[0, 1]$ when n_{ref} is the maximum number of particles in a unit volume. This will be done in Chapter 4. Another possible modification is the multiplication of c with an associated volume per particle. This transforms the concentration into a *swelling field* and captures the volume of the material due to species migration. This approach is chosen in the Chapters 5 and 7.

2.3.2 Chemical Potential for Fluid Mass Transport

The chemical potential is defined per substance and measures its tendency ≪to change their location, composition or state of aggregation≫ (Job & Herrmann [2006]). A high chemical potential thus indicates a high tendency for a change in one or more of these characteristics. In general, it is a function of temperature, pressure and phase. Chemical potentials of solvents additionally depend on solute concentration and solvent properties (see Job & Herrmann [2006] again). The appropriate unit is that of an energy contribution, i.e., Joule, and when related to molar quantities, Joule per mole or even ≪Gibbs≫, see Wiberg [1972]. Denoting the quantity by μ seems to be the established convention.

Though this is not the approach within a continuum theory, values of the chemical potential can be obtained by looking up experimental data listed in large tables in chemistry books. There, they appear per substance and aggregation as the difference to the referential elements that the substance consists of. Substances with a negative chemical potential are therefore stable, because they have a lower

tendency to change aggregation than their referential elements (positive values of the chemical potential are possible, but appear less frequent). Values cannot be measured directly, but are instead computed e.g. via voltage measurements in electro-chemical cells.

For processes with mass transport, there is no point in referring to fix values recorded for some referential circumstances, because the concentration changes with time and position and with it the chemical potential. The chemical potential must thus be constituted by a material law. Without anticipating much, two exemplary constitutive relations for the chemical potential are shortly given in the following paragraph.

Constitutive Examples. With the *Boltzmann constant* k, the chemical *activity coefficient* γ and the solution-dependent atomic fraction Y, the constitutive chemical potential for a solution reads

$$\widehat{\mu}_{\text{mixture}} = \mu_0 + kT \log[\gamma Y], \tag{2.27}$$

where μ_0 is the *unitary* part and independent on the concentration of the solution, and the second the concentration-dependent *communal* or *mixing* part and originates from the entropy of mixing, see Sato [2004]. This law can be simplified for ideal solutions, in which case Raoult's [1887] law with $Y \approx 1$ applies, or for dilute solutions, where Henry's [1803] law with $Y \approx c$ and γ being a constant is valid. Another well-established expression for the chemical potential is that of an ideal gas,

$$\widehat{\mu}_{\text{ideal gas}} = -kT \log \left[\frac{V}{N} \left(\frac{2\pi m k T}{h_0^2} \right)^{3/2} \right]. \tag{2.28}$$

Here, h_0 is some constant, V the volume of the gas, N the number of molecules and m the mass of one molecule. A shared trait of these two expressions is that (2.28) and the communal part of (2.27) tend to negative infinity when the concentration approaches zero, which Job & Herrmann [2006] interpret as a basic law rather than a common outcome of constitutive choices for the chemical potential.

Relation to Concentration and Pressure. With these two examples at hand, it seems that concentration and temperature are the main reasons for a chemical potential to vary (for the mixture (2.27), the atomic fraction Y incorporates the concentration, while it is the quotient of volume and particle number for the ideal gas (2.28)). At constant temperature, an increase of c leads to an increase of μ, and a concentration gradient has already been identified as a reason for mass transport. Could the chemical potential thus be dropped for simplicity, leaving its responsibility for solute diffusion to a concentration gradient?

Adopting an exemplary experiment from Job & Herrmann [2006] clarifies the enhanced information carried by the chemical potential: when ether is given into a homogeneous mixture of iodine and water, the equilibrium is characterized by

two distinct phases with almost no iodine dissolved in the water – there is a concentration gradient of solute iodine particles, but the chemical potential is homogeneous throughout water and ether because $\mu_{0,\,\text{water}} \neq \mu_{0,\,\text{ether}}$.

The chemical potential thus offers a layer of abstraction to a concentration gradient: it can cause solute particles to travel along a concentration gradient, but additional effects can be included, too.

This point can be stretched to a similar statement for the pressure. As easily imaginable for gases, the pressure increases with increasing concentration, because more particles randomly move and collide in the same domain. Therefore, pressure and concentration can often be linked to each other, and indeed, the above mentioned laws for ideal or dilute solution are originally formulated in terms of the pressure, and spatial variations of the pressure can often be understood as a mechanism behind solute diffusion. This quantity is nevertheless not chosen as a state variable. Again, the pressure leaks generality and extensibility towards diffusion mechanisms caused by alternative circumstances. Additionally, the mechanical state of the solid cannot be described by a scalar and no additive split of the stresses is intended as it is done within Biot's [1941] theory, where a fluid pressure can be retrieved from decomposed total stresses. Finally, the pressure is a quantity related to mechanical energy, while the chemical potential is still necessary to formulate energetic contributions of mass transport (see Section 2.6.1) and therefore is indispensable.

Chemical Potential as a Field. For now, these characteristics of the chemical potential shall be sufficient. For additional information dedicated to an understanding of this quantity, the reader is referred to Baierlein [2001] or Cook & Dickerson [1995]. In order to generalize the notion of a chemical potential such that its behavior can be prescribed by a constitutive material law later on, the scalar quantity is introduced as an intensive field parameterized by material coordinates. It thus has the mapping properties

$$\mu : \begin{cases} \mathcal{B} \times T \to \mathcal{R} \\ (\boldsymbol{X}, t) \mapsto \mu(\boldsymbol{X}, t). \end{cases} \tag{2.29}$$

When understood as a global field, the chemical potential enters the problem as a variable that can be controlled in the environment or at the boundary of the solid. This makes sense after having identified its role as the driving quantity behind solute particle diffusion. A remaining conceptual challenge is the link to a constitutive law like (2.27) or (2.28) – why should μ be a global field for which a solution will be sought, when a material law dictates its value anyhow? In fact, the inverse relation is used: the concentration can be determined at every material or spatial point with given values of the chemical potential. Thus, control of the diffusion driving quantity is retained while allowing for arbitrarily complex constitutive functions for μ.

Similar to the comment after the introduction of the concentration field, the position-dependence of μ shall be understood as anticipated. A discussion on the circumstances of this dependence can be found in the Section 2.4.2.

2.4 Local Equilibrium States, Exchange and Interaction

Since deformation map and strain measures have been introduced and concentration as well as chemical potential fields are identified as descriptions of diffusive state and diffusion driving force, a closer look at transport and interaction phenomena from a thermodynamic viewpoint is due. Four aspects will be considered that arise for deforming solids with mass transport. First, the *thermodynamic system* is characterized by stating which quantities can be exchanged with the environment. Then, the notion of *equilibrium states* is linked with length scales, and non-mechanical transport and finally mechanical interaction quantities are embedded into the continuum theory.

2.4.1 Partially Open System and Reservoir

A system to which thermodynamical statements can be applied to is easily found. Here, $\partial\mathcal{S}$ is the surface that encloses the thermodynamic system, embedded in an environment – thermodynamic system and material body are thus exchangeable notions. Whether a quantity can be exchanged with its environment constitutes the thermodynamic system to be *closed* or *open* with respect to the particular quantity.

But what are the categories of quantities in question for an exchange? There are three of them; matter, *work* and *heat*. Matter has been introduced in Section 2.2, but limited to the notion of a solid material body. Additional constituents can come into play, and within the context of this work, there is exactly one such constituent, i.e., fluid particles. Heat is a form of energy transfer between objects caused by their unequal temperature and irrespective of the actual mechanism, see Schroeder [2000, p. 18]. Work is any other form of energy that can be supplied to or produced by a system (often, only mechanical work is considered). The difference between heat and mechanical work is qualitatively, yet concisely pointed out by Lebon et al. [2008, p. 7]: ≪Microscopically, mechanical work is related to coherent correlated motions of the particles while heat represents that part of motion, which is uncorrelated, say incoherent≫. The importance of these terms lies in the necessity to grasp physical effects by properly understanding their energetic contributions to the system[7]. This will be further specified in Section 2.6.1.

[7]This is because there is no concept more fundamental than energy.

Several terms have been established for systems and their exchange characteristics throughout the literature. The following list is presented in analogous, concise form in e.g. Kuhl [2004].

- An *isolated system* does not exchange any quantity with its environment. This is the most boring case.
- An *adiabatically closed system* is allowed to exchange work, but no heat or matter with its environment.
- Heat as well as work may be exchanged in *closed systems*, but matter is not allowed to do so.
- In an *open system*, work, heat and matter can be transported across the system boundary.

These characteristics can be associated with a membrane that encloses the system and that is, according to the system characteristics, deformable or rigid, permeable or impermeable, diathermal or adiabatic.

The exchange of solid mass between an open system and its environment imposes a conceptual modification of all balance laws (compared to systems closed for solid mass) and extensive quantities to be defined per unit mass rather than unit volume, see Haupt [2002]. This approach is applied in e.g. biomechanical theories of growth, but has no relevance for this work: no solid mass exchange is assumed across material borders, not even across borders of cut-out submanifolds. In contrast, fluid mass must obviously be allowed to cross borders and migrate into the system or out of it. The associated membrane is thus *semipermeable*, and the system shall be called *partially open* here, which is not a standard term, but appropriate for diffusion processes in deforming solids and in line with e.g. Kowalski [2003, p. 46].

Not very astonishingly, work may be exchanged between the environment and the solid – that is, tractions can act on the surface of the material body as the most common form of mechanical work. Non-mechanical work associated with an in- or outflux of fluid matter, on the macroscopic as well as on a microscopic scale, is permitted as a consequence of the openness with respect to the fluid.

Exchange of thermal energy is possible, this is however a trait without much influence, because isothermal conditions will be assumed later on. The actual consequence of this openness is a conceptual one: constant temperature must be assumed in the material body and throughout its environment. If adiabatic processes are assumed, there is no need for assumptions for the surrounding temperature.

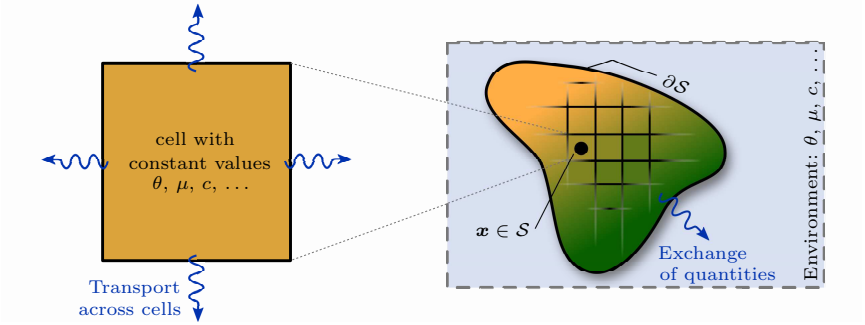


Figure 2.4: Local equilibrium hypothesis: a macroscopic gradient of a state variable is considered a mosaic of unit cells with constant values for which thermostatic equilibrium holds. Transport of quantities across unit cell borders is always allowed, while the exchange with the surrounding across $\partial\mathcal{S}$ depends on the assumptions made for the system.

2.4.2 Local Equilibrium Postulate for Unit Cells

The specification of the thermodynamic system shapes what is understood as an *equilibrium state*: a state associated with the absence of spatial variations of state variables – those variables that are chosen to determine the macroscopic state of the system. It is free of gradients of those variables, which contradicts some parts of the outline so far: a chemical potential gradient has been made responsible for diffusive transport of solute particles, but chemical equilibrium only holds when the chemical potential is constant throughout the material body and its environment. In general, state variables like chemical potential, concentration or temperature are ≪well defined and single-valued at equilibrium≫ (Tadmor et al. [2012]), such that they lack a proper definition in non-equilibrium states.

This last statement should make a reader tremble with fear when confronted with it for the first time. Why should basic fields such as temperature or energetic states be impugned just because they are non-uniform in the body? The answer can best be given in two steps. First, the values of intensive fields in question are based upon atomistic scales as statistical averages – temperature captures Brownian motion, chemical potential is associated with energetic states per particle, and so on. When these underlying microscopic states quickly change over time and in space, it becomes quite unclear, what the average abstraction

actually represents, cf. Kowalski [2003, p. 46]. Second, statements about an equilibrium state without a length scale are incomplete. Measuring a quantity like temperature is non-unique when spatial and temporal fluctuations obscure an observed value – thus, protection against uncertainty of such a quantity can be established by choosing a length scale in which these variations vanish, and that means: in which the quantity is in equilibrium.

Bridging the study of equilibrium states (sometimes called *thermostatics* to emphasize the superiority of theories for process rates) to non-equilibrium states is the challenge that comes along with spatially inhomogeneous states. The conceptual dodge that allows for using the equilibrium-state variable definitions within non-equilibrium processes is two-fold. First, the spatial scope or length scale for equilibrium statements is narrowed down to a microscopic equilibrium cell. Within this cell, temperature, chemical potential and other thermodynamic quantities are assumed to be constant and legitimately approximated by averages. This *local equilibrium cell* is thus in a thermodynamic system on its own at position $\boldsymbol{X}$ and in an equilibrium state, in contrast to the overall system which forms a system with spatial variations of quantities being constant in individual cells. Second, an irreversible rate process is interpreted as a *quasi-static* sequence of perturbed and neighboring non-equilibrium states in time, i.e., the duration of one perturbation compared to duration of the complete process towards equilibrium is small, or when relating the macroscopic and microscopic time scales, atomistic equilibrium is reached substantially faster than a change in macroscopic variables occurs (this second assumption can in fact be made for spatially constant state variables, too). The first assumptions is called the local equilibrium postulate (cf. Haase [1964], de Groot & Mazur [1984] or Alberty [1997]) and together with the second one, it enables for an analysis of irreversible processes by integration of the infinitesimal changes to independent state variables that appeared during the quasi-static sequence of states[8]. Global quantities can additionally be formulated in an integral form with their local, constant cell values as integrands, see Kestin [1992]. This methodology can be interpreted as an analogy to the generalization of particle-wise Newtonian ideas to continua (Kowalski [2003, p. 47]). It should finally be noted that these assumptions are valid for non-equilibrium states *near equilibrium*, though the meaning of *near* could not be cleared yet with complete rigor.

Any quantity that appeared within a statement of thermodynamic equilibrium can now be treated as a field depending on position and time. This has already been done for concentration and chemical potential in the previous section – in fact, position dependence of state variables are often assumed in the continuum mechanics literature without any references to the local

[8]The notion of *entropy* (see Section 2.6.2) is however still not definite between non-equilibrium states, see Lebon et al. [2008, p. 9]. The local equilibrium hypothesis is thus to be understood as an assumption with satisfactory foundation, not as a first order principle.

thermodynamic equilibrium postulate, though it is tacitly accepted by allowance of spatial variations of a quantity. The picture of undeformed and deformed solid ($\mathcal{B}$ and $\mathcal{S}$) in their environment must now be extended to a spatial level of refinement, see Figure 2.4. In this sense, they are composed by equilibrium cells with constant variables at time t. Note finally that assumptions on the openness and closeness of the macroscopic thermodynamic system equally applies to the equilibrium cells. That is, transport of fluid particles, heat or other energetic quantities takes place between cells, while solid matter is not exchanged.

2.4.3 Fields, Fluxes, Production and Supply

The way is now paved for a precise description of this exchange sketched in the above, i.e., between equilibrium cells and across the system surfaces $\partial\mathcal{S}$ or $\partial\mathcal{B}$. To this end, a *flux* is the transport of a quantity across a surface, straightforward in its physical meaning (e.g. solute particles) or abstract (e.g. an energetic flow from one region or cell to another), while sources supply or produce and sinks destroy or absorb the quantity being accounted. As a preliminary classification, stresses are not understood as fluxes of a mechanical force or traction quantity. These terms are rather taken as non-transport fields and treated as carriers of mechanical interaction later in this section.

With this distinction between mechanical and transport fields at hand, all fluxes are scalar quantities per unit area, *flux vectors* are vectorial objects in the tangent spaces of current and reference configuration, while *production* and *supply* terms are again scalar fields.

To begin with, a field $i_s(\boldsymbol{x},t)$ measures a scalar quantity per unit deformed configuration. Its material analog, measured per unit undeformed volume, is denoted without any subscript, $i(\boldsymbol{X},t)$, and in general, it has the mapping characteristics

$$i : \begin{cases} \mathcal{B}\times T \to \mathcal{R} \\ (\boldsymbol{X},t) \mapsto i(\boldsymbol{X},t). \end{cases} \tag{2.30}$$

One such field is the concentration and has already been introduced in (2.24) (with a mapping onto $\mathcal{R}_+$, which of course supersedes the definition (2.30)).

Spatial and material supply terms for the quantity i are again fields defined per unit deformed and undeformed volume, respectively, and identified by $r_i(\boldsymbol{x},t)$ and $R_i(\boldsymbol{X},t)$. Similarly, $\gamma_i(\boldsymbol{x},t)$ and $\Gamma_i(\boldsymbol{X},t)$ denote spatial and material scalar production terms. The production of a quantity is quite rare (it could e.g. originate from chemical reactions), but will in fact occur. Because their relation to the unit volume, supply and production terms as well as the original fields can be related by (2.16) and (2.3), which reads

$$i = Ji_s \circ \boldsymbol{\varphi}, \qquad \Gamma_i = J\gamma_i \circ \boldsymbol{\varphi} \quad \text{and} \quad R_i = Jr_i \circ \boldsymbol{\varphi}. \tag{2.31}$$

The flux *per unit area* of the deformed solid is denoted by h_i, its material counterpart is H_i. These scalars measure the amount of quantity i that flows out of a region through a unit deformed area: the sign convention accounts for the direction of a positive outward normal vector. By assuming a linear relationship between flux and unit outward normal, the spatial and material flux vectors $\boldsymbol{j}_i \in T_{\boldsymbol{x}}\mathcal{S}$ and $\boldsymbol{J}_i \in T_{\boldsymbol{X}}\mathcal{B}$ can be introduced by

$$h_i = \boldsymbol{j}_i \cdot \boldsymbol{n} \quad \text{and} \quad H_i = \boldsymbol{J}_i \cdot \boldsymbol{N}. \tag{2.32}$$

The relationship of spatial and material fluxes stems from the identity

$$h_i da = H_i dA \tag{2.33}$$

which, after insertion of (2.32) and the area map, reads (2.15)

$$\boldsymbol{J}_i = J\boldsymbol{F}^{-1}\boldsymbol{j}_i. \tag{2.34}$$

Both spatial and material flux vectors are equally valid for transport between equilibrium cells and across system borders $\partial\mathcal{S}$ or $\partial\mathcal{B}$.

Note that when specifying the transport of matter across a surface, spatial and material flux vectors can optionally be defined in terms of a barycentric particle velocity relative to the velocity (2.9) or (2.7), see e.g. de Groot & Mazur [1984]. This is common in theory of mixtures, but comes along with an additional point map of fluid particles and further balance laws – as those ingredients are neglected in this work, a fluid particle flux can be taken as is, without further substantiations.

2.4.4 Mechanical State Described by Stress Tensors

Mechanical work between $\mathcal{S}$ and its surrounding is exchanged via tractions. Similarly, tractions capture the mechanical interaction between hypothetical parts of $\mathcal{S}$. For the definition of a traction vector, the force per unit deformed area da is called a *traction vector* $\boldsymbol{t}(\boldsymbol{x}, t)$ and is naturally a quantity of $T_{\boldsymbol{x}}\mathcal{S}$. Newton's third law dictates that the traction vector exerts a force per unit area with identical norm and inverse orientation on the submanifold as well as the remaining surface.

Cauchy's theorem is developed by considering equilibrium conditions for a tetrahedron and yields

$$\boldsymbol{t} = \boldsymbol{\sigma}\boldsymbol{n}, \tag{2.35}$$

where $\boldsymbol{\sigma}$ is the second order *Cauchy stress tensor* field. It is a mapping of normal onto traction vectors and thus completely describes the stress state in the current configuration.

There are two objective reasons to construct alternative stress tensors. First, it will be advantageous to end up with balance laws formulated in terms of quantities (partially) defined on the reference configuration. And second, exploitation of stress tensor symmetries influence the computational efficiency of finite element implementations. The choice of stress measures mainly stays a matter of taste, though – here, only the three most commonly employed stress tensors besides the Cauchy stress are introduced.

The *Kirchhoff* stress has the same mapping characteristics as the Cauchy stress and an identical *work-conjugate* strain rate, but relates a contraction with it to a unit undeformed volume, which explains the scaling by J in its definition,

$$\boldsymbol{\tau} = J\boldsymbol{\sigma}. \tag{2.36}$$

Next, a pull-back of $\boldsymbol{\tau}$ can be used to define a dual stress measure for the reference configuration. This tensor is called the *second Piola-Kirchhoff stress tensor*,

$$\boldsymbol{S} = \boldsymbol{F}^{-1}\boldsymbol{\tau}\boldsymbol{F}^{-T} \tag{2.37}$$

and retains the symmetry properties of the two preceding. It does not yield a physically motivated mapping of a normal to a traction-like vector. Last, the *first Piola-Kirchhoff stress tensor* is a two-point (and thus unsymmetrical) tensor and, when additionally preceded by the current metric tensor, work-conjugate to the (rate of the) deformation gradient: its popularity is caused by exactly this simplicity. To construct an expression for the first Piola-Kirchhoff stress, a force vector per unit undeformed area with identical magnitude of the resulting force is taken as the result of a mapping of a reference normal vector. This operation is achieved by a tensor of second order,

$$\boldsymbol{P} = J\boldsymbol{\sigma}\boldsymbol{F}^{-T}, \tag{2.38}$$

which is identified as the first Piola-Kirchhoff stress tensor with the identity

$$\boldsymbol{P}\boldsymbol{N}dA = \boldsymbol{\sigma}\boldsymbol{n}da. \tag{2.39}$$

Note that the modified stress tensor $\boldsymbol{g}\boldsymbol{P}$ can be used as a more convenient stress measure, see Figure 2.5. In the case of orthonormal coordinate bases, it collapses to (2.38).

These stress tensors will be more than sufficient for the purpose of this work, see e.g. Başar & Weichert [2000] for a more detailed characterization of tensorial stress quantities.

2.5 Five Balances Principles for Extensive Quantities

In what follows, five balance laws will be constructed for notions of solid and fluid mass, momentum and microforces, covering both balances for conserved and non-conserved quantities. Each balance is formulated as an integral expression over a

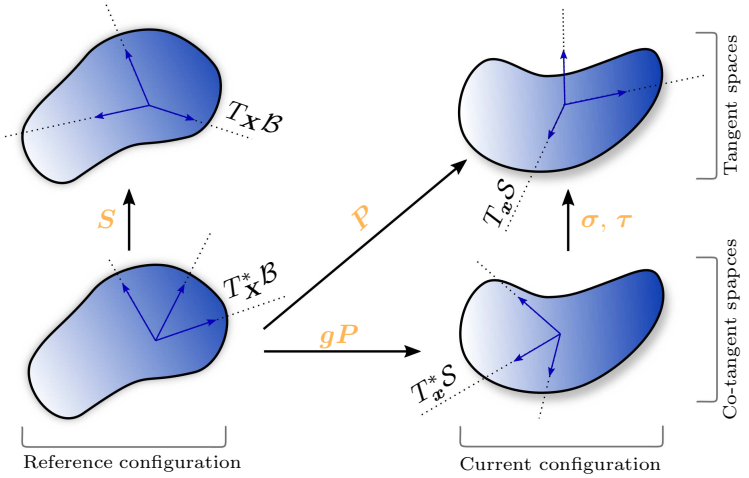


Figure 2.5: Visualization of the properties of common stress tensors that map between tangent and co-tangent spaces in reference and current configuration. $\boldsymbol{\sigma}$ denotes the Cauchy, $\boldsymbol{S}$ the second Piola-Kirchhoff, $\boldsymbol{\tau}$ the Kirchhoff and $\boldsymbol{P}$ the first Piola-Kirchhoff stress tensor.

submanifold $\mathcal{P}_\mathcal{S} \in \mathcal{S}$ and transferred to its referential counterpart domain $\mathcal{P}_\mathcal{B}$. For the subsequent elaboration, let $\partial\mathcal{P}_\mathcal{S}$ and $\partial\mathcal{P}_\mathcal{B}$ denote the surfaces of the spatial and the material submanifolds, da a deformed area element of $\mathcal{P}_\mathcal{S}$ and dA its referential counterpart. Spatial and material outward unit normal to $\mathcal{P}_\mathcal{S}$ and $\mathcal{P}_\mathcal{B}$ are denoted by $\boldsymbol{n}(\boldsymbol{x}, t)$ and $\boldsymbol{N}(\boldsymbol{X}, t)$, respectively.

The balance laws share a common structure, which can be referred to as a master balance law. This master balance does however not strengthen the understanding of the underlying physics, but rather satisfies the needs of individuals with an excessive desire for categories and is thus omitted here. Common characteristics of the balances are rather underlined by a notational convention for fluxes, flux vectors, production and supply terms of the quantity that is accounted for.

2.5.1 Solvent Solid and Solute Fluid Mass Balances

Conservation of Solid Mass. The material mass density of the solid material is the scalar field ρ with its spatial counterpart $\rho_\mathcal{S}$. Integrating the latter over the submanifold $\mathcal{P}_\mathcal{S}$ yields the mass of the body at time t, and conservation of solid

mass demands that the total time derivative of this expression equals zero in the absence of any mass production (e.g. due to chemical reactions),

$$\frac{d}{dt}\int_{\mathcal{P}_\mathcal{S}} \rho_s dv = 0. \tag{2.40}$$

By transferring the integral to the referential submanifold $\mathcal{P}_\mathcal{B}$, using the volume map (2.16) and the material solid mass density reveals two objections: first, the referential density field is constant over time, and second, a relation between material and spatial mass density is given by the volume map,

$$\frac{d}{dt}\rho = 0 \quad \text{and} \quad \rho = J\rho_s \circ \boldsymbol{\varphi}. \tag{2.41}$$

The second expression can be used within further balance equations to relate terms involving the spatial solid mass density to the referential domain of integration $\mathcal{P}_\mathcal{B}$ and is sometimes referred to as the *Lagrangian equation of continuity*.

Balance of Fluid Mass. The fluid mass is understood as a non-conserved quantity, though chemical reactions are not taken into account and hence no creation of fluid particles. There may nevertheless be sources inside of $\mathcal{P}_\mathcal{S}$ as pumps that inject (or absorb) species particles. It is straightforward to assume zero fluid supply or absorption later on, but without this loss of generality at this point, the global balance of fluid mass reads

$$\frac{d}{dt}\int_{\mathcal{P}_\mathcal{S}} c_s dv = \int_{\mathcal{P}_\mathcal{S}} r_c dv - \int_{\partial\mathcal{P}_\mathcal{S}} h_c da. \tag{2.42}$$

Using (2.31), (2.32) and (2.33), integration by parts and the standard localization argument, the local form of the balance of fluid mass is obtained as

$$\dot{c} = -\operatorname{Div}[\boldsymbol{J}_c] + R_c, \tag{2.43}$$

which, boiled down to one dimension, can be used to construct Fick's so-called second law by insertion of a linear constitutive relation of species flux and concentration gradient (in fact, Fick's first law)[9].

2.5.2 Newtonian Balances of Mechanical Momenta

Balance of Linear Momentum. Next, a generalization of Newton's first and second axiom to deformable continua is performed by associating Newtonian forces that act on a rigid particle with *body forces* per unit volume $\boldsymbol{r}_{\boldsymbol{\varphi}}$ integrated

[9]It is, however, not advised to give up the separation of balance laws and constitutive assumptions, thus only remembering Fick's first law is fine.

over $\mathcal{P}_\mathcal{S}$ and tractions integrated over $\partial\mathcal{P}_\mathcal{S}$, and similarly formulating the *impulse* of $\mathcal{P}_\mathcal{S}$ as the sum of all particle-wise impulses,

$$\frac{d}{dt}\int_{\mathcal{P}_\mathcal{S}} \rho_s \boldsymbol{v} dv = \int_{\mathcal{P}_\mathcal{S}} \boldsymbol{r}_{\boldsymbol{\varphi}} dv + \int_{\partial\mathcal{P}_\mathcal{S}} \boldsymbol{t} da. \tag{2.44}$$

Following the same steps as before together with (2.35) and (2.38), this global (sometimes called substantial) balance is localized and transferred to the reference configuration, which reads

$$\mathrm{Div}[\boldsymbol{P}] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \rho \boldsymbol{A}, \tag{2.45}$$

often called the *local impulse balance* or *equation of motion*. With a localization in terms of spatial quantities, a similar equation arises depending on the Cauchy stress (which will however not play a role in the following, it is thus not explicitly given).

Balance of Angular Momentum. The rotational counterpart of the linear momentum is the angular momentum, for which a similar balance law exists. The total angular momentum is constructed by integrating the cross-product of the local impulse and the position (relative to an arbitrary point of reference, here the coordinate system origin is chosen) over the submanifold $\mathcal{P}_\mathcal{S}$. Its time derivative is in balance with an external torque formed by volume forces and traction on the surface of $\mathcal{P}_\mathcal{S}$,

$$\frac{d}{dt}\int_{\mathcal{P}_\mathcal{S}} \boldsymbol{\varphi} \times \rho_s \boldsymbol{v} dv = \int_{\mathcal{P}_\mathcal{S}} \boldsymbol{\varphi} \times \boldsymbol{r}_{\boldsymbol{\varphi}} dv + \int_{\partial\mathcal{P}_\mathcal{S}} \boldsymbol{\varphi} \times \boldsymbol{t} da. \tag{2.46}$$

The steps to reduce this global equation to a local statement are technical details of educational value, but beyond the scope of this outline. The reader is referred to any of the textbooks mentioned at the beginning of this chapter. Summarizing, the transformation yields a well-known symmetry-like relation for the first Piola-Kirchhoff stress

$$\boldsymbol{F}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{F}^T, \tag{2.47}$$

which reflects the actual symmetry condition $\boldsymbol{\sigma}^T = \boldsymbol{\sigma}$.

2.5.3 Balance Law for Microstructural Arrangements

The next balance law is on the one hand easy to swallow due to its simplicity, as fewer terms and tensorial operations are involved. On the other hand, it accounts for *microstructural arrangements* like ordering of atoms (thus, not the simple existence of atoms or particles, but the microstructure they form), which is hardly a physical process one can easily associate with a force and a balance. The idea behind this balance law goes back to Gurtin [1996], who aimed at the disjunction of constitutive material equations and balance principles. While this separation of concerns works out well and is of methodological value, a motivation

of the balance law beyond formal merits seems to lag behind its structural virtue. As the microforce balance is, however, a central ingredient of Chapter 4, it is wise to postulate the equation for future use and regardless of the doubts that come along with it.

To this end, consider the aforementioned microscopic arrangement of particles on a length scale far below the one associated with mechanical tractions acting on a deforming surface[10]. The spatial body force that performs work along the particle arrangement parameterized by c_s is assumed to be a scalar called the *chemical microforce* and is defined *per unit volume* in $\mathcal{P}_s$. The scalar *microtraction* h_ξ acts on the surface $\partial\mathcal{P}_s$, again conjugate to c_s. As the notation will indicate, r_ξ, γ_ξ and h_ξ are supply term, production and flux, respectively. More specifically, r_ξ is an external body microforce that Gurtin [1996] additively equipped with an internal production variant, here γ_ξ, which is identified as the microforce itself. Then, postulating the balance leads to

$$0 = \int_{\mathcal{P}_S} r_\xi + \gamma_\xi dv + \int_{\partial\mathcal{P}_S} h_\xi da. \tag{2.48}$$

When formulated locally on the reference configuration by the steps analogously taken as for the balance laws above, it takes the form

$$0 = \mathrm{Div}[\boldsymbol{J}_\xi] + R_\xi + \Gamma_\xi, \tag{2.49}$$

which will be referred to as the local microforce balance. The referential scalar microforce is the material production term Γ_ξ.

2.6 Energy and Entropy Balances and the Second Law

Two fundamental aspects of deformation-diffusion processes have not been covered so far: a comparison of energetic contributions and their direction. While energetic contributions are tackled by the first law of thermodynamics, the process direction is constrained by the second law. Both postulates are phenomenological in nature and have been worded in many ways. The more interesting and controversial law is the second, giving rise to the fanciful quantity *entropy*. After its first versions attributed to Carnot, Clausius or Kelvin, who stated that certain processes are impossible to happen, the notion of entropy with a macroscopic definition for reversible and irreversible processes came into play and changed the shape of the second law towards a formula-based representation. In the second half of the 19th century, this understanding of entropy was superseded by Boltzmann's microscopic definition in terms of *multiplicity* as a number of possible microstates.

[10]The different length scales as a motivation for the microforce balance again go back to Gurtin [1996].

In the following, no attempt is made to compile a poor, minimal version of the history of thermodynamics as e.g. outlined in Müller & Weiss [2005, chapter 21] – instead, the way is paved for an understanding of energy and entropy contributions within the present continuum theory as well as for the exploitation of the second law for choosing the best thermodynamic state variables and restricting the scope of constitutive material laws. Finally, the notion of *dissipation* is introduced and related to the production of entropy.

2.6.1 Non-Conserving Balances for Energy and Entropy

Balance of Energy. In order to set up a balance law for the *energy* of the material body, a balance is first established for the *total specific energy* e_s in the current configuration,

$$\frac{d}{dt}\int_{\mathcal{P}_s} e_s dv = \int_{\mathcal{P}_s} r_e dv - \int_{\partial\mathcal{P}_s} h_e da. \tag{2.50}$$

This equation lacks in meaning. The actual content of the balance of energy, known as the *first law of thermodynamics*[11] (usually written down in terms of global, spatially constant contributions) is the association of total energy density, flux and source term with heat, mechanical and non-mechanical work and *internal energy*. Their appearance in such an energy balance postulates an equivalence of work and heat, i.e., they can be transformed into each other, but cannot be destroyed.

Work and heat enter the balance equation in their incremental forms (*power*). The internal energy is further specified by the detachment of kinetic energy: the total energy of the system is assumed to be a sum of internal (potential) and kinetic energy. The latter is due to non-zero linear and angular momenta: mechanical power exerted on the system not only leads to a change of internal energy, but becomes manifest in a change of particle movement as well (cf. Tadmor et al. [2012, p. 140]). Flux and source terms have multiple additive components, too: one due to the incremental work of mechanical body forces in $\mathcal{P}_s$ and tractions on $\partial\mathcal{P}_s$, one accounting for heat supply again in $\mathcal{P}_s$ and through $\partial\mathcal{P}_s$, a third for a macroscopic energy decrease (sometimes an increase) due to solvent particles injected by a pump or supplied over the surface (here, the chemical potential reveals its characteristics by measuring the change of energy of the system that results from the addition of one solvent particle), and finally a fourth and last one that induces energy changes resulting from microstructural arrangements.

Denoting the spatial internal energy by u_s leads to the global statement

$$\frac{d}{dt}\int_{\mathcal{P}_s} \left(u_s + \tfrac{1}{2}\,\rho_s \boldsymbol{v}\cdot\boldsymbol{v}\right) dv = \int_{\mathcal{P}_s} \left(\boldsymbol{r}_{\boldsymbol{\varphi}}\cdot\boldsymbol{v} + r_q + r_c\mu\circ\boldsymbol{\varphi}^{-1} + r_\xi\dot{c}_s\right) dv$$

[11] Historically, it was discovered after the second law, though (see Lebon et al. [2008])

$$+\int_{\partial\mathcal{P}_S}\left(\boldsymbol{t}\cdot\boldsymbol{v}-h_q-h_c\mu\circ\boldsymbol{\varphi}^{-1}+h_\xi\dot{c}_s\right)da. \tag{2.51}$$

When this form is expressed in terms of referential quantities and integrals together with the Cauchy theorem (2.35), (2.38) as well as (2.32), its Lagrangian, substantial form reads

$$\frac{d}{dt}\int_{\mathcal{P}_\mathcal{B}}\left(u+\tfrac{1}{2}\,\rho\boldsymbol{V}\cdot\boldsymbol{V}\right)dV=\int_{\mathcal{P}_\mathcal{B}}\left(\boldsymbol{R}_{\boldsymbol{\varphi}}\cdot\boldsymbol{V}+R_q+R_c\mu+R_\xi\dot{c}\right)dV$$

$$+\int_{\partial\mathcal{P}_\mathcal{B}}\left(\boldsymbol{P}\boldsymbol{N}\cdot\boldsymbol{V}-\boldsymbol{J}_q\cdot\boldsymbol{N}-\boldsymbol{J}_c\cdot\boldsymbol{N}\mu+\boldsymbol{J}_\xi\cdot\boldsymbol{N}\dot{c}\right)dA. \tag{2.52}$$

Here, energy contributions arise with variables that are explicitly non-uniform in $\mathcal{P}_s$, the energy is thus assumed to spatially vary – note again that this goes back to the local equilibrium postulate outlined in Section 2.4.2 and differs from versions of the first law that are usually found in undergraduate thermodynamics textbooks. Another detail that differs from these treatises is that the local internal energy is a density related to the unit undeformed volume and not to the unit mass. The advantage of this approach is revealed in Section 3.2.

Upon application of the divergence theorem and the chain rule, rearrangement of terms and insertion of local microforce, linear momentum and fluid mass balances, the local balance of energy collapses to

$$\dot{u}=\boldsymbol{P}:\dot{\boldsymbol{F}}+R_q-\mathrm{Div}[\boldsymbol{J}_q]+(\mu-\Gamma_\xi)\dot{c}-\boldsymbol{J}_c\cdot\nabla\mu+\boldsymbol{J}_\xi\cdot\nabla\dot{c}, \tag{2.53}$$

which shows that the internal energy is clearly a non-conserved quantity (sometimes, this fact is confused by interpreting the first law of thermodynamics as a statement of conservation of energy – whether this understanding is valid depends on the choice of the thermodynamic system or whether the environment of the system is part of the conservation statement). Besides, the aforementioned equivalence of heat and (mechanical and non-mechanical) work together with a definition of these work terms is again the main outcome of the balance of energy. Its local form (2.53) will in addition be used in Section 2.6.3.

Balance of Entropy. Any extensive thermodynamic state function can be accounted for in a balance law, thus stating the existence of an extensive *entropy* formally leads to a balance. Its meaning is not immediately visible, first of all because entropy fluxes or source terms do not automatically ship with an intuitive interpretation as it is the case for a solute flux. And second, neither the global nor the local statement directly lead to an insight (this will, however, change when addressing the interpretation of entropy in Section 2.6.2). The global balance of entropy is formulated as

$$\frac{d}{dt}\int_{\mathcal{P}_S}s_s dv=\int_{\mathcal{P}_S}r_s+\gamma_s dv-\int_{\partial\mathcal{P}_S}h_s da. \tag{2.54}$$

After application of the divergence theorem, equations (2.16), (2.32) and (2.33) as well as the localization theorem, a local balance can be established as before, which reads

$$\dot{s} = -\operatorname{Div}[\boldsymbol{J}_s] + R_s + \Gamma_s. \tag{2.55}$$

Similarly to the balance of energy, an identification of $\boldsymbol{J}_s$ and R_s with quantities introduced up to this point will be the important part related to this equation. The spatial dependence of s is again exclusively due to the concept of local equilibrium states, see Section 2.4.2.

2.6.2 Characterization of Entropy and Temperature

Statistical Entropy and Entropy of Mixing. Entropy is defined according to Boltzmann's findings, often called the statistical definition of entropy. This foils the continuum and thus macroscopic scope of the theory outlined so far, but yields the best understanding of the mechanism behind this term. The *multiplicity* of a system is the number of possible microstates for one macrostate, denoted by Ω: each possible arrangement of particles in the system is one microstate, while an interchange of two identical particles does not count as a separate one (particles are indistinguishable, this is a sensible assumption as it resolves the Gibb's paradox of non-extensive entropies). An estimation of the multiplicity can be derived with simple models for e.g. ideal gases or Einstein solids and by using an approximation of large numbers, see Schroeder [2000]. To become more comfortable with the very large multiplicity of real systems, its logarithm can be used, and when multiplied with the Boltzmann constant for historical reasons, the statistical or microscopic definition of entropy in a system in equilibrium reads

$$\int_{\mathcal{B}} s dV = k \log[\Omega]. \tag{2.56}$$

The first obvious trait is that (2.56) contradicts the expectation that the entropy spatially varies, but this is only due to (2.56) holding for an equilibrium state, which should be linked to length scales as before (the number of microstates is counted within this length scale). The second characteristic reveals how the entropy splits into parts for composite systems: the number of possible microstates is given by the product of each multiplicity. For two partially specified entropies, (2.56) leads to

$$k \log[\Omega_1 \Omega_2] = k \log[\Omega_1] + k \log[\Omega_2], \tag{2.57}$$

the entropies are added. Another advantage of approaching entropy via (2.56) appears together with particles entering or leaving the system: when a particle is added, the number of possible configurations, i.e., arrangements of particles, increases and with it the entropy (it obviously decreases, when the particles leave

the system). The entropy attributed to this mechanism is called *entropy of mixing* and plays a major role in processes with diffusion of solute particles.

Macroscopic Entropy and Temperature. How can a definition of the entropy be linked to the first one (2.56) to better fit the continuum approach of the body? The formal procedure for the conversion of (2.56) into a macroscopic entropy definition is not included here[12], an established definition is favored for simplicity.

To this end, the *thermodynamic* or *absolute temperature*[13] is first introduced as a positive scalar field on $\mathcal{B}$ as

$$\theta : \begin{cases} \mathcal{B} \times T \to \mathcal{R}_+ \\ (\boldsymbol{X}, t) \mapsto \theta(\boldsymbol{X}, t). \end{cases} \tag{2.58}$$

The temperature is a measurable quantity that indicates the tendency of a body to spontaneous energy releases and is equal throughout two bodies in contact with each other and in a thermal equilibrium state – so far, the most basic interpretation in line with what is regularly called the *zeroth law of thermodynamics*. Its link to effects on microscopic scales is the random motion of particles, which gives a hint that temperature is related to entropy as a measure of arrangements of these particles. An additional tie to heat as a transfer of energy caused by temperature differences is obvious.

Now recall the entropy balance law (2.54). The gain in or loss of entropy over the boundary $\partial\mathcal{P}_\mathcal{S}$ and in $\mathcal{P}_\mathcal{S}$ attributed to *reversible processes* is determined by the heat supply divided by the absolute temperature, see Haupt [2002, p. 128], who calls this relation a ≪proven fact≫ for equilibrium systems. The surface contribution in (2.54) can then be substantiated as

$$\int_{\partial\mathcal{P}_\mathcal{S}} h_s da = \int_{\partial\mathcal{P}_\mathcal{S}} \frac{h_q}{\theta} da = \int_{\partial\mathcal{P}_\mathcal{B}} \frac{\boldsymbol{J}_q}{\theta} \cdot \boldsymbol{N} dA, \tag{2.59}$$

while the production and supply term inside of submanifold $\mathcal{P}_\mathcal{S}$ is

$$\int_{\mathcal{P}_\mathcal{S}} \gamma_s + r_s dv = \int_{\mathcal{P}_\mathcal{S}} \gamma_s + \frac{r_q}{\theta} dv = \int_{\mathcal{P}_\mathcal{B}} \Gamma_s + \frac{R_q}{\theta} dV. \tag{2.60}$$

The latter two equations can be understood as a macroscopic definition of entropy, characterized by a separation of supply and production by categorizing the former as an equilibrium process contribution. The divisor θ has been directly applied inside of the integral, which is common practice, but no triviality: the temperature of heat sources at the boundary $\partial\mathcal{P}_\mathcal{S}$ and in $\mathcal{P}$ are not a priori identical to the one at corresponding material points, see Tadmor et al.

[12]It can be found in most textbooks, even the `Wikipedia page` on the fundamental thermodynamic relation has an outline (as of April 2017).

[13]The absolute temperature is not relative to a referential value, but starts at a fix value and can only in- or decrease in one way, see Fermi [1956, p. 35 et sqq.].

[2012, p. 175]. The underlying assumption is that heat sources are in thermal equilibrium with the points at which they occur.

Note again that in pure equilibrium processes, Γ_s vanishes and the integrands are constant, allowing the relations to be expressed in global variables (the more common form to write down this relation between equilibrium entropy and the quotient of heat and temperature).

Combination with the First Law. Insertion of (2.59) and (2.60) into (2.54) leads to a local balance that shows remarkable changes in comparison with (2.55), i.e.,

$$\dot{s} = -\operatorname{Div}\left[\frac{\boldsymbol{J}_q}{\theta}\right] + \Gamma_s + \frac{R_q}{\theta}. \tag{2.61}$$

While this alternate form does not yet reveal unknown characteristics of the entropy, it can be reformulated by evaluation of the divergence term using the chain rule, rearrangements of terms and multiplication by θ, which yields

$$R_q - \operatorname{Div}[\boldsymbol{J}_q] = \theta\dot{s} - \theta\Gamma_s - \frac{1}{\theta}\nabla\theta \cdot \boldsymbol{J}_q. \tag{2.62}$$

The essential advantage of this form is that it can easily be inserted into the local form of the balance of energy (2.53). This eliminates heat source and parts of the heat flux and, after rearrangement for the entropy production, the equation reads

$$\theta\Gamma_s = -\dot{u} + \boldsymbol{P} : \dot{\boldsymbol{F}} + \theta\dot{s} - \frac{1}{\theta}\nabla\theta \cdot \boldsymbol{J}_q + (\mu - \Gamma_\xi)\dot{c} - \boldsymbol{J}_c \cdot \nabla\mu + \boldsymbol{J}_\xi \cdot \nabla\dot{c}. \tag{2.63}$$

Though a bit lengthy, this equation enables for a sound evaluation of the second law of thermodynamics in the following by sharpening which terms are associated with the entropy production Γ_s.

2.6.3 Finally... the Second Law of Thermodynamics

Increase of Entropy. With the notion of entropy at hand, the second law of thermodynamics can be expressed in simple words: ≪the multiplicity tends to increase≫, and equivalently ≪entropy tends to increase≫ (Schroeder [2000]). This does not apply to supply within or transport of entropy between unit cells, because these terms have been identified as sources external to the cell and capture an exchange of entropy with the environment, not an increase across cells, i.e., in the whole system. In other words, in an isolated unit cell or system (the different length scales have no influence when no quantities are exchanged with the environment), the entropy can be constant or increase. This transforms into the elementary formulation of the second law of thermodynamics,

$$\Gamma_s \geq 0. \tag{2.64}$$

For reversible processes, $\Gamma_s = 0$ holds, for irreversible processes it is $\Gamma_s > 0$. When using the results of Section 2.6.2, equation (2.64) expands to

$$\theta\Gamma_s = -\dot{u} + \boldsymbol{P} : \dot{\boldsymbol{F}} + (\mu - \Gamma_\xi)\,\dot{c} - \boldsymbol{J}_c \cdot \nabla\mu + \boldsymbol{J}_\xi \cdot \nabla\dot{c} + \theta\dot{s} - \frac{1}{\theta}\nabla\theta \cdot \boldsymbol{J}_q \geq 0. \quad (2.65)$$

A possible doubt might be the use of this inequality when the entropy production term Γ_s appears along the entropy rate $\dot{s}$, preventing the computation of one of these variables. But keep in mind that the entropy itself is a state variable that can be prescribed by a constitutive function similar to a material law. This is only possible for simple systems like ideal gases (and in general, the determination of state by the entropy pulls in more difficulties than advantages), but while the entropy s can at least be specified in terms of system variables, the entropy production is only determined by (2.65). Additionally, it is possible to find alternative state functions that eliminate the entropy rate in (2.65), which will be addressed later on.

Dissipative Effects. The product of temperature and entropy production is called *dissipation*, cf. e.g. Haupt [2002, p. 130]. This is a convenience definition, making relation (2.65) easier to evaluate. A *dissipation postulate* is regularly proposed instead of the second law, but these two inequalities are interchangeable, they barely differ except the more idiomatic approach the dissipation postulate implies by circumventing the interpretation of entropy. The dissipation postulate is furthermore formulated as a global statement, it reads

$$\int_{\mathcal{P}_\mathcal{S}} \theta\Gamma_s dV \geq 0, \quad (2.66)$$

where the replacement of the integrand by (2.65) is left out for the sake of brevity. *Dissipative effects* are thus irreversible processes in a non-equilibrium state that contribute to the entropy production. In continuum mechanics, this is often associated with the internal friction of inelastic materials and thus heat that leaves the system. When facing fluid transport in an elastic medium however, these effects are negligible, and dissipation or entropy increase is due to particle movement and arrangements on the microscopic scale, e.g. fluid flows along a negative concentration gradient – a process, which is irreversible and produces entropy by means of a more homogeneous distribution of particles: this raises the multiplicity, as there are more possible microstates in a solid less densely filled with solute particles.

Outlook for Constraints. One aspect has been vaguely announced but not really covered so far: the direction of processes and related constraints onto constitutive material laws. An evaluation of the inequality (2.64) with this regard has not been possible up to this point, due to a missing clarification of independent and dependent state variables and state functions, i.e., thermodynamic potentials. Once state functions are introduced as constitutive scalar ingredients that characterize the system and define energetic variables such as the internal

energy u in dependence on other system parameters, an evaluation of the second law of thermodynamics is doable.

2.6.4 Thermal Equilibrium by Isothermal Assumption

Up to this point, no restrictions on the temperature were formulated: θ could vary in time and space. As thermal effects are however not taken into account in this work, *isothermal conditions* must be assumed to rule out certain quantities and to simplify the second law of thermodynamics. Thus, thermal equilibrium is assumed, which, as stated earlier, is understood as a homogeneous temperature throughout the solid and – the thermodynamic system is not adiabatically closed – the environment, too.

Two consequences arise from this assumption. First, the time derivative of θ vanishes, which will be of use soon. And second, the gradient $\nabla\theta$ is equally zero. The latter can be inserted into (2.65), which reads

$$\theta\Gamma_s = -\dot{u} + \boldsymbol{P} : \dot{\boldsymbol{F}} + (\mu - \Gamma_\xi)\,\dot{c} - \boldsymbol{J}_c \cdot \nabla\mu + \boldsymbol{J}_\xi \cdot \nabla\dot{c} + \theta\dot{s} \geq 0. \tag{2.67}$$

Assuming constant temperature in $\mathcal{B}$ and its environment is equal to assuming an immediate thermal equilibrium state once the temperature rises, e.g. due to internal friction of the material. This involves heat conduction in $\mathcal{B}$ and exchange across $\partial\mathcal{P}$ to be fast relative to kinetic and kinematic processes and is of course a strong idealization of real processes. But when no external thermal stimuli are present, the very low temperature differences that arise from mechanisms internal to the material often justify isothermal assumptions.

Chapter 3:
Constraints and Objectivity of Constitutive Functions

The notion of a constitutive material law already came up in the introductory chapters, but the term has not been substantiated by further elaboration. Though the necessity for material laws can be understood as an intuitive ingredient of a continuum-mechanical theory, its necessity shall shortly be motivated, and rationales that limit the dependencies of admissible material laws are proposed and evaluated. These assumptions and restrictions are important enough to dedicate a chapter to the consequences that they impose on the choice of constitutive functions. In fact, textbooks on continuum mechanics often cover these aspects called *constitutive theories* – which suggests that the following outline does not intend to present new findings, but rather acts as a collection of common principles that sharpen the understanding of a material law, tailored to diffusion-deformation problems, phase separation and the phase-field modeling of fracture.

Thus, much more detailed studies of presumptions and constraints related to material theory are available. The reader is referred to Haupt [2002], Tadmor et al. [2012] or Truesdell, C. & Noll [1965] for in-depth analyzes. Scope and limitations of a material law characterized in the following fit into the class of *gradient-extended standard dissipative materials*, where the notion of *standard dissipative materials* is a common and established term (see e.g. Halphen & Nguyen [1975]), and its generalization towards gradient-extended theories with non-local internal variables[14] or order parameter for the description of a microstructural evolution have been developed throughout the last decades, exemplary theories are Svendsen et al. [2009], Maugin [2009] Nguyen [2010], Miehe [2011] or Mariano [2002].

Unknown Variables vs. Governing Equations. Counting the number of variable components involved in a continuum theory and balancing it with the number of equations extracted from balance laws is an exercise often presented as it motivates the need for additional information brought in by constitutive functions. To this end, a focus is put on the material forms of balance equations as in Section 2.5 and Section 2.6, and it is assumed that the material density ρ and the supply terms R_c, $\boldsymbol{R}_{\boldsymbol{\varphi}}$ and R_ξ are known as environmental circumstances as

[14]It should be mentioned that the term *non-local* is subject of discussion. Some authors even understand it as a meaningless attribute.

Unknowns			Equations		
Deformation map	$\boldsymbol{\varphi}$	3	Kinematics	(2.13)	9
Strain measure	$\boldsymbol{F}$	9	Linear momentum	(2.45)	3
Stress tensor	$\boldsymbol{P}$	9	Angular momentum	(2.47)	3
Fluid concentration	c	1	Fluid mass	(2.43)	1
Fluid flux	$\boldsymbol{J}_c$	3	Microforce balance	(2.49)	1
Chemical microforce	Γ_ξ	1	Energy balance	(2.53)	1
Microtraction	$\boldsymbol{J}_\xi$	3	Entropy balance	(2.61)	1
Chemical potential	μ	1			
Entropy	s	1			
Entropy Production	Γ_s	1			
Energy	u	1			
Number of unknowns		33	Number of equations		19

Table 3.1: Overview of balance equations and unknown variables, where all supply terms in the balance equations are assumed to be known. The gap of fourteen components must be provided by material models.

well as not influencing the material properties, cf. Liu [2013, p. 63]. Now consider the fields of interest introduced so far together with the definition of the principal strain measure, material Newtonian balance laws, the material microforce and fluid mass balances as well as the energy balance as gathered in Table 3.1. Thus, 33 unknown variables are combined with 19 governing equations. The fourteen missing components are the subject of constitutive laws.

Prescribed Quantities. Two of the chosen constitutive laws prescribe the first Piola-Kirchhoff stress (six additional equations with (2.47) taken into account) and the fluid flux (three equations). In fact, these two quantities will be determined by two well-known and widely used material laws, namely Fick's law of diffusion and the Neo-Hookean material law for finite elastic deformation. Furthermore, the chemical microforce will be subject to a material law (one equation), as well as the microtractions (three equations) and the scalar internal energy (one equation). This is vaguely in line with Coleman & Gurtin [1967] or Coleman & Noll [1963], when dropping the additional process quantities (internal variables or heat flux and temperature, respectively) and adding fluid transport and related microstructural arrangements[15].

[15]It is thus not really in line with the two theories referred to. On the other hand, mentioning some authorities in the field is considered advantageous.

3.1 Preliminary Assumptions on Constitutive Material Laws

The most general assumptions are briefly outlined in this section together with those of increasing applicability. All assumptions shall be equally valid for the solid material response as well as the material laws for diffusion, microtraction, chemical potential and internal energy. The terms *material law* or *material response*, *constitutive function* or any combination of these terms refer to the same concept: a ≪relation representing the individual response of any material element to a given input process≫, see Haupt [2002, p. 275].

3.1.1 Determinism, Equipresence and Homogeneity

Principle of Determinism. The principle of *determinism* states that what happened in the past uniquely determines present events. In some sense, this is ridiculously trivial for continuum mechanics, because it is the Newtonian methodology that shaped a deterministic understanding of real world processes. While findings of quantum mechanics have continuously eliminated determinism as the higher order principle it once was, the principle of determinism is another way of announcing that the validity of Newtonian mechanics can still be assumed for the length and time scales involved in the present work.

Principle of Equipresence. The principle of *equipresence* has been impressively worded by Truesdell & Toupin [1960, p. 703][16], who demand that ≪constitutive equations, then, should not artificially divert these theories into disjoint channels≫, where the theories referred to are the collection of more or less all field theories. This pragmatic demand simply means that every state variable one constitutive law depends on should also be a dependency of all other constitutive laws.

Homogeneous Material Response. The material shall behave equally throughout $\mathcal{B}$, i.e., a constitutive law considered in this work shall not depend on the position $\boldsymbol{X}$. This is the first restriction that actually reduces the number of dependencies of constitutive functions. This assumption is often left out, though afterwards, the coordinates do not show up in the list of function arguments to the material laws once they are specified. In this case, position-dependent constitutive functions have been tacitly included, because narrowing them down to homogeneous response functions does not facilitate the constitutive theory. Here, utmost generality is no primary goal, and the restriction is postulated.

[16]Their tendency to be bold has been a constant source of joy to the author during the time of writing.

3.1.2 Local Action and Grade One Material Response

Principle of Local Action. The aforementioned determinism goes hand in hand with the principle of *local action* in order to limit the scope of constitutive laws such that they can be formulated for a body $\mathcal{B}$ without taking the state at every point $\boldsymbol{X}$, past and present, into account. More precisely, the constitutive law prescribing a material response at $\boldsymbol{X}$ is influenced by the state at points within a neighborhood of $\boldsymbol{X}$, while this neighborhood can be chosen arbitrarily small (cf. Tadmor et al. [2012, p. 182], where it is also pointed out that the principle of local action is presumed in many continuum theories, but not to be taken as common sense).

Theory of Grade One. In the local neighborhood of $\boldsymbol{X}$, variables can be approximated by a Taylor series expansion. The restriction to a *theory of grade one* limits the number of summands of such a series expansion to two, which makes the approximation of the quantity relative to its value at $\boldsymbol{X}$ of first order. This will apply to all quantities that form a material response, i.e., that appear as a dependency of the constitutive law. The deformation in the neighborhood of $\boldsymbol{X}$ for example is approximated by the values of $\boldsymbol{\varphi}$ and $\boldsymbol{F}$, the concentration by c and ∇c and so on.

3.1.3 Entropy, Substate and its Specification

The principles above give assistance in composing influences on constitutive laws, but do not dictate an exact set of quantities that must be incorporated as dependencies. Truesdell & Toupin's [1960] starting point is the *caloric equation of state* for the internal energy, however, formulated in terms of a *substate*, which is left unspecified not to lose generality. The thermodynamic state is composed of this substate and the entropy, and is assumed to determine the internal energy. Here, a specific set of quantities is chosen as the substate, which is a strong assumption but unavoidable. Before the first version of the caloric equation of state can be drafted, three further restrictions are formulated in the following.

Independence from the Motion. In line with Truesdell & Toupin [1960], the internal energy does not depend on the deformation $\boldsymbol{\varphi}$ at $\boldsymbol{X}$. This assumption is an optional a priori simplification – if a dependency on the deformation is taken into account, it will be canceled out by the demand of material frame invariance.

Exclusion of Memory. Neither memory nor aging of the material is considered. There is no need for a concept of *internal variables* as outlined in the famous work of Coleman & Gurtin [1967] and others, which is a major simplification of the material laws in this work. Solid material laws for plasticity are an example of constitutive approaches that rely on internal variables, remembering earlier state for the determination of the current material response. Note that this waiving of

material memory not only applies to the solid material response for the stresses, but all other constitutive relations, too.

Exclusion of Deformation Rates. It shall be assumed that the constitutive laws do not depend on the rates of deformation. This is an assumption that rules out any form of viscous solid material behavior. Taken together with the previous restriction, not only viscoelastic, but also viscoplastic material behavior of the solid is excluded a priori.

Formulas, Please. The function that prescribes u is denoted by $\widehat{u}$, and it reads, conforming with all previous restrictions and assumptions,

$$u(\boldsymbol{X},t) = \widehat{u}(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu). \tag{3.1}$$

Thus, deformation gradient, concentration and its gradient as well as chemical potential together with its gradient are identified as the substate, influencing the material response in conjunction with the entropy. One discrepancy between (3.1) and its counterpart in Truesdell & Toupin [1960, p. 619] has to be commented on: the homogeneity of the material response has been incorporated in (3.1), which led to canceling out a dependence on material coordinates $\boldsymbol{X}$. Note also that the time dependence is only implicit by entropy and substate fields, which depend on time explicitly.

The only principle that has been introduced above but has not been utilized yet is the principle of equipresence. Its application is straightforward by transferring the structure of (3.1) to all other constitutive functions. The same notational convention is used, i.e., values of the quantities $\boldsymbol{P}$, $\boldsymbol{J}_c$, $\boldsymbol{J}_\xi$ and Γ_ξ are prescribed by the functions $\widehat{\boldsymbol{P}}$, $\widehat{\boldsymbol{J}}_c$, $\widehat{\boldsymbol{J}}_\xi$ and $\widehat{\Gamma}_\xi$. Thus, the complete list of material laws reads

$$\begin{aligned}
u(\boldsymbol{X},t) &= \widehat{u}(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu)\\
\boldsymbol{P}(\boldsymbol{X},t) &= \widehat{\boldsymbol{P}}(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu)\\
\Gamma_\xi(\boldsymbol{X},t) &= \widehat{\Gamma}_\xi(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu)\\
\boldsymbol{J}_c(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_c(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu)\\
\boldsymbol{J}_\xi(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_\xi(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu).
\end{aligned} \tag{3.2}$$

Having laid out this basic list of dependencies, which mainly goes back to Gurtin [1996, 1989], only one further requirement must be addressed before an evaluation of the second law can be performed.

Demand of Regularity. The internal energy function $\widehat{u}(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu)$ shall be differentiable multiple times with respect to all of its arguments, such that these derivatives are invertible (cf. Truesdell & Toupin [1960, p. 622]). The same demand will propagate to similar scalar functions derived from the internal energy.

3.2 Coleman-Noll Evaluation of the Second Law

The isothermal, local version of the second law combined with the energy balance (2.67) is almost in a suitable form for the analysis called *Coleman-Noll procedure* (after their inventors Coleman & Noll [1963]) that has become an established standard in continuum mechanics. The constraint (2.67) is sharpened by a split into two inequalities. This is a common technique to separate different contributions to the entropy production, and here, it reads

$$-\dot{u} + \boldsymbol{P} : \dot{\boldsymbol{F}} + (\mu - \Gamma_\xi)\,\dot{c} + \boldsymbol{J}_\xi \cdot \nabla\dot{c} + \theta\dot{s} \geq 0 \tag{3.3}$$

with the only left-over summand related to solute transport

$$-\boldsymbol{J}_c \cdot \nabla\mu \geq 0. \tag{3.4}$$

For now, the focus lies on (3.3), and the inequality (3.4) is approached later. The internal energy rate in the first inequality is, relying on its regularity claimed before, expanded to

$$\dot{u} = \partial_s\widehat{u}\dot{s} + \partial_{\boldsymbol{F}}\widehat{u} : \dot{\boldsymbol{F}} + \partial_c\widehat{u}\dot{c} + \partial_{\nabla c}\widehat{u} \cdot \nabla\dot{c} + \partial_\mu\widehat{u}\dot{\mu} + \partial_{\nabla\mu}\widehat{u} : \nabla\dot{\mu} \tag{3.5}$$

where the explicit dependencies of $\widehat{u}$ have been dropped for notational convenience. (3.5) is inserted into (3.3), and the resulting expression reads after reordering of terms

$$\begin{aligned}(\theta - \partial_s\widehat{u})\,\dot{s} + (\boldsymbol{P} - \partial_{\boldsymbol{F}}\widehat{u}) : \dot{\boldsymbol{F}} + (\mu - \Gamma_\xi - \partial_c\widehat{u})\,\dot{c} \\ + (\boldsymbol{J}_\xi - \partial_{\nabla c}\widehat{u}) \cdot \nabla\dot{c} - \partial_\mu\widehat{u}\dot{\mu} - \partial_{\nabla\mu}\widehat{u} \cdot \nabla\dot{\mu} \geq 0.\end{aligned} \tag{3.6}$$

The Coleman-Noll reasoning is that inequality (3.6) must hold for every admissible process. These admissible processes can be constructed as special cases by arbitrary choices of the thermodynamic state (recall that thermodynamic state per assumption is s, $\boldsymbol{F}$, c, ∇c, μ and $\nabla\mu$).

First, consider an *isentropic* process (during which the entropy is constant) of time-independent deformation and concentration and with a homogeneous, time-dependent chemical potential. The left hand side of (3.6) boils down to a simple product that has to hold for any positive or negative value of $\dot{\mu}$, which reads

$$-\partial_\mu\widehat{u}\dot{\mu} \geq 0 \quad \Rightarrow \quad \partial_\mu\widehat{u} = 0. \tag{3.7}$$

The function prescribing the internal energy must therefore not depend on the chemical potential.

Next, consider a similar process with an inhomogeneous rate of the chemical potential, while the assumptions on the rates of entropy, deformation and concentration still hold. This process again must not violate the second law.

Using the vanishing derivative of $\widehat{u}$ with respect to μ, the evaluation of (3.6) under these conditions leads to

$$\partial_{\nabla\mu}\widehat{u}\cdot\nabla\dot{\mu}\geq 0 \quad \Rightarrow \quad \partial_{\nabla\mu}\widehat{u}=\boldsymbol{0}. \tag{3.8}$$

A second dependency could thus be eliminated.

The next admissible process in question is characterized by time-independent deformation and concentration. The entropy rate can be chosen arbitrarily, and the second law for such a process yields

$$(\theta-\partial_s\widehat{u})\,\dot{s}\geq 0 \quad \Rightarrow \quad \theta=\partial_s\widehat{u}. \tag{3.9}$$

This finding is worth to be written down explicitly with the functional dependencies for both temperature and internal energy. After elimination of the dependency on chemical potential and its gradient, it reads

$$\theta=\widehat{\theta}(s,\boldsymbol{F},c,\nabla c)=\partial_s\widehat{u}(s,\boldsymbol{F},c,\nabla c). \tag{3.10}$$

Here, the relation between temperature and internal energy has been used to emphasize the new function $\widehat{\theta}(s,\boldsymbol{F},c,\nabla c)$, which is not meant to yield an additional constitutive law, but represents the functional dependence that the temperature obtains from its relation to the internal energy. This detail is of importance, because the regularity assumption for $\widehat{u}$ allows for the inversion of (3.10), such that an entropy function depending on θ, $\boldsymbol{F}$, c and ∇c can be constructed. This will reveal its significance soon, as it is suitable for an isothermal context. Furthermore, note that (3.10) is a local statement and thus has a different meaning than the corresponding global relation, which usually appears in textbooks on thermodynamics. This is a nice reminder of the local equilibrium postulate that enabled for the conventional thermodynamic methodology on local cells with homogeneous state.

Three steps similar to those above are still missing for the complete evaluation of the entropy inequality. Consider an isentropic process with a zero concentration gradient. The entropy inequality is again reduced to a simple term that allows for a functional relation,

$$(\boldsymbol{P}-\partial_{\boldsymbol{F}}\widehat{u}):\dot{\boldsymbol{F}}\geq 0 \quad \Rightarrow \quad \boldsymbol{P}=\partial_{\boldsymbol{F}}\widehat{u}, \tag{3.11}$$

because the inequality has to hold for any rate $\dot{\boldsymbol{F}}$. This implies an expectable, yet important insight: the power contribution of mechanical stresses inside $\mathcal{B}$ does not contribute to an increase of entropy. Or in other words, the mechanical deformation is purely reversible, i.e., stress tensors are state variables. The assumption is justified in the context of elasticity. Just note that at this point, the stresses are often additively decomposed into an elastic stress tensor and one that accounts for viscous effects. Here, such additional components are unnecessary.

Note again that u is related to the unit volume, which is often the case in continuum mechanics because no solid mass supply exists and there is no need

to multiply the right hand side of (3.11) with the density ρ, which was necessary for an internal energy defined per unit mass.

Next, a further isentropic process is constructed in which the deformation and concentration gradients are time-independent. Another constitutive relation can then be formulated by evaluation of (3.6), which reads

$$(\mu - \Gamma_\xi - \partial_c \widehat{u}) \geq 0. \quad \Rightarrow \quad \Gamma_\xi = \mu - \partial_c \widehat{u}. \tag{3.12}$$

This relation might give rise to some doubts: why does the chemical potential that has been introduced as a primary variable in Section 2.3.2 and is thus sought for, appear on the right hand side without any differential operator that could at least make up a similarity to some balance principle? The answer to this question lies in the interpretation of the supply term (the attribute *supply* is rather misleading here), i.e., the chemical microforce Γ_ξ, which is, in contrast to other supply terms, neither zero nor known by assumptions on the environment. Γ_ξ is an unknown variable just like μ or $\boldsymbol{P}$ (see the left columns of Table 3.1), and can be interpreted similarly to a stress tensor: given the deformation map, stresses are computed by the solid material response, and these enter the Newtonian balance laws – given the chemical potential, microforces can be computed by (3.12), and they enter the microforce balance. Only because Γ_ξ and μ are both work-conjugate to c, one on a macroscopic and one on a macroscopic scale, they appear as is in the above relation.

Finally, an isentropic process with time-independent deformation and a time-dependent concentration gradient can be used together with the last relationship to formulate the second law for the last fictitious reversible process,

$$(\boldsymbol{J}_\xi - \partial_{\nabla c} \widehat{u}) \geq 0. \quad \Rightarrow \quad \boldsymbol{J}_\xi = \partial_{\nabla c} \widehat{u}. \tag{3.13}$$

The chemical microtractions are thus related to the internal energy by partial differentiation with respect to the concentration gradient.

While the evaluation of (3.6) has become somewhat repetitive in the end, the way is now paved for two important results. First, the list of constitutive functions must be updated, because μ and $\nabla\mu$ have been eliminated as dependencies,

$$\begin{aligned}
u(\boldsymbol{X},t) &= \widehat{u}(s, \boldsymbol{F}, c, \nabla c) \\
\boldsymbol{P}(\boldsymbol{X},t) &= \widehat{\boldsymbol{P}}(s, \boldsymbol{F}, c, \nabla c) = \partial_{\boldsymbol{F}} \widehat{u}(s, \boldsymbol{F}, c, \nabla c) \\
\Gamma_\xi(\boldsymbol{X},t) &= \widehat{\Gamma}_\xi(s, \boldsymbol{F}, c, \nabla c) = \mu - \partial_c \widehat{u}(s, \boldsymbol{F}, c, \nabla c) \\
\boldsymbol{J}_\xi(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_\xi(s, \boldsymbol{F}, c, \nabla c) = \partial_{\nabla c} \widehat{u}(s, \boldsymbol{F}, c, \nabla c) \\
\boldsymbol{J}_c(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_c(s, \boldsymbol{F}, c, \nabla c, \mu, \nabla\mu).
\end{aligned} \tag{3.14}$$

Note that $\widehat{\boldsymbol{J}}_c$ did not change with respect to (3.2), because it has not been part of the inequality (3.3). Note further that with the above relations between internal energy and thermodynamic state variables, this inequality becomes an equality, and the associated contribution to the entropy production is zero. Hence, after the Coleman-Noll procedure let so many terms disappear, the local second law of thermodynamics simply reads

$$\theta\Gamma_s = -\boldsymbol{J}_c \cdot \nabla\mu \geq 0. \tag{3.15}$$

No other mechanism except the transport of solute particles is thus responsible for an entropy increase. This reduced inequality is identical with (3.4) and will be used to formulate restrictions on the constitutive law for the fluid flux $\widehat{\boldsymbol{J}}_c$.

3.3 Free Energy and (Dual) Dissipation Potential

Helmholtz Free Energy. In (3.14), all material laws depend on the entropy. This is a consequence of the thermodynamic balance equations for isothermal conditions that induced the initial choice of state variables. Depending on the context, it might be acceptable to not pay attention to this dependency: a constitutive law for an elastic material response then has material parameters as functions of the entropy, usually constant over time, cf. Coleman & Noll [1963, p. 177].

In other circumstances, where the entropy has to be taken into account for a material law, a challenge arises from the handling of its value. The microscopic definition (2.56) shows that dealing with this quantity is not straightforward. The entropy may be calculated, but, at least in macroscopic continuum theories, is huge and hardly controllable by assumptions or boundary conditions.

This can, however, even be regarded as a technical issue compared to the following difficulty: when dissipative effects and thus an entropy increase are part of a process with the energetic state captured by $\widehat{u}(s, \boldsymbol{F}, c, \nabla c)$, the material laws derived from $\widehat{u}$ in (3.14) capture the material response *at fixed entropy* as a consequence of the partial derivative operators. It does not violate any physical law, but will almost certainly not yield the quantity of interest. This can be exemplified with the constitutive expression for a stress tensor: during entropy increase due to fluid transport in $\mathcal{B}$, the relation $\boldsymbol{P} = \partial_{\boldsymbol{F}}\widehat{u}(s, \boldsymbol{F}, c, \nabla c)$ yields the solid material response only for the isentropic portion of the process, thus, the state change due to diffusion stays unnoticed.

To circumvent this difficulty inherent to the constitutive internal energy, a new quantity shall be introduced that is independent of the entropy. As mentioned earlier, the demand for regularity of $\widehat{u}$ and its descendants ensures the existence of a differentiable entropy function

$$s = \widehat{s}(\theta, \boldsymbol{F}, c, \nabla c), \tag{3.16}$$

inverse to the temperature (3.10). The entropy can thus be taken as dependent on θ and the thermodynamic substate. When seeking the alternate potential

$$\psi = \widehat{\psi}(\theta, \boldsymbol{F}, c, \nabla c) \tag{3.17}$$

that yields the negative of (3.16) when derived with respect to the temperature (the sign inversion is for historical reasons), it can be shown that this potential, well-known as the *Helmholtz free energy* ψ (often denoted by F or A), is constructed by

$$\widehat{\psi}(\theta, \boldsymbol{F}, c, \nabla c) = \widehat{u}(\widehat{s}(\theta, \boldsymbol{F}, c, \nabla c), \boldsymbol{F}, c, \nabla c) - \theta \widehat{s}(\theta, \boldsymbol{F}, c, \nabla c), \tag{3.18}$$

while the relation between ψ and the entropy is the analogy to (3.10) and reads

$$s = \widehat{s}(\theta, \boldsymbol{F}, c, \nabla c) = -\partial_\theta \widehat{\psi}(\theta, \boldsymbol{F}, c, \nabla c). \tag{3.19}$$

This exchange of dependent and independent variables in (3.18) is called a *Legendre transformation*, best known for its use in thermodynamics and mechanics. The above closed form can be established for monotonous derivatives of $\widehat{u}$ with respect to its argument s and thus convexity of $\widehat{u}$ with respect to s. Note that this convexity condition does not automatically apply to the thermodynamic substate, i.e., deformation gradient, concentration and its gradient.

Note also that ψ is a quantity related to the unit undeformed volume. As mentioned for the internal energy in Section 3.2, this is not in line with many textbooks on thermodynamics, where thermodynamic potentials are usually related to the unit mass. Because solid mass exchange is excluded a priori, potential densities are easier to use, as there is no need to carry the density ρ across equations when relating ψ to e.g. stresses.

The quantity ψ measures the mechanical and non-mechanical work that can be performed by the material body at constant temperature. The constant temperature acts however still as an input parameter to the free energy, which is no discrepancy, but obvious when thinking about mechanical work being performed at two different temperatures that are constant over time – at least the material parameters that govern the solid material response must then be functions of θ. Also, the restriction to constant temperature responses does by no means exclude a temperature evolution. Instead, the free energy captures the portion of the material response which is acts at constant temperature. The material response for time-dependent temperature must be prescribed by other functions, which, in the present case, is unnecessary due to the assumptions of isothermal conditions.

Coleman-Noll, Again. The relation between the internal energy and free energy quantities, not functions, simply is $\psi = u - \theta s$, hence its time derivative $\dot{\psi} = \dot{u} - \dot{\theta} s - \theta \dot{s}$. As a consequence of the isothermal assumption, the temperature rate vanishes, and rearrangement for the internal energy rate reads $\dot{u} = \dot{\psi} + \theta \dot{s}$. This can be inserted into the equation (3.6) for the Coleman-Noll evaluation of

the second law, leading to a simpler representation because the product of entropy rate and temperature rate cancels out,

$$\theta\Gamma_s = -\dot{\psi} + \boldsymbol{P} : \dot{\boldsymbol{F}} + (\mu - \Gamma_\xi)\,\dot{c} + \boldsymbol{J}_\xi \cdot \nabla\dot{c} \geq 0. \tag{3.20}$$

Upon expansion of the time derivative $\dot{\psi}$ as before and reordering, the partial constraint associated with the second law reads

$$\begin{aligned}\theta\Gamma_s = \left(\boldsymbol{P} - \partial_{\boldsymbol{F}}\widehat{\psi}\right) : \dot{\boldsymbol{F}} + \left(\mu - \Gamma_\xi - \partial_c\widehat{\psi}\right)\dot{c} + \left(\boldsymbol{J}_\xi - \partial_{\nabla c}\widehat{\psi}\right) : \nabla\dot{c} \\ - \partial_\mu\widehat{\psi}\dot{\mu} - \partial_{\nabla\mu}\widehat{\psi} : \nabla\dot{\mu} \geq 0.\end{aligned} \tag{3.21}$$

The steps to be taken to evaluate this inequality are identical to those that dealt with (3.6) and thus not repeated here. Instead, the resulting list of constitutive functions based upon the free energy is given as

$$\begin{aligned}
\psi(\boldsymbol{X},t) &= \widehat{\psi}(\theta,\boldsymbol{F},c,\nabla c)\\
\boldsymbol{P}(\boldsymbol{X},t) &= \widehat{\boldsymbol{P}}(\theta,\boldsymbol{F},c,\nabla c) = \partial_{\boldsymbol{F}}\widehat{\psi}(\theta,\boldsymbol{F},c,\nabla c)\\
\Gamma_\xi(\boldsymbol{X},t) &= \widehat{\Gamma}_\xi(\theta,\boldsymbol{F},c,\nabla c) = \mu - \partial_c\widehat{\psi}(\theta,\boldsymbol{F},c,\nabla c)\\
\boldsymbol{J}_\xi(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_\xi(\theta,\boldsymbol{F},c,\nabla c) = \partial_{\nabla c}\widehat{\psi}(\theta,\boldsymbol{F},c,\nabla c)\\
\boldsymbol{J}_c(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_c(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu).
\end{aligned} \tag{3.22}$$

The temperature dependence thus propagated to all material laws related to the free energy.

Temperature Dependence. It is common to omit the temperature as an explicit functional dependency of the material laws under isothermal assumptions. Material parameters that appear in the specific constitutive laws are clearly functions of the homogeneous temperature under which the process in question happens. These parameters are expected to be chosen with the thermal process conditions in mind, which shall justify the elimination of the temperature from the constitutive dependencies. This is a notational convenience, and no formal or rational element of constitutive theory. It simplifies the above list to

$$\begin{aligned}
\psi(\boldsymbol{X},t) &= \widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\boldsymbol{P}(\boldsymbol{X},t) &= \widehat{\boldsymbol{P}}(\boldsymbol{F},c,\nabla c) = \partial_{\boldsymbol{F}}\widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\Gamma_\xi(\boldsymbol{X},t) &= \widehat{\Gamma}_\xi(\boldsymbol{F},c,\nabla c) = \mu - \partial_c\widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\boldsymbol{J}_\xi(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_\xi(\boldsymbol{F},c,\nabla c) = \partial_{\nabla c}\widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\boldsymbol{J}_c(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_c(s,\boldsymbol{F},c,\nabla c,\mu,\nabla\mu).
\end{aligned} \tag{3.23}$$

Note again that the last expression for the fluid flux has not changed during the Coleman-Noll procedure because it has been part of the separated inequality (3.4). The structure of a constitutive fluid flux function will instead be addressed in the following.

Dissipation Potential. The reduced entropy inequality (3.15) (identical to (3.4)) is used to deduce restrictions on the constitutive fluid flux. The inequality is satisfied for a simple flux that linearly depends on the chemical potential gradient, realized through the application of a positive semidefinite tensorial constitutive function on $\nabla\mu$. The loss of generality when demanding such a linear diffusion law seems acceptable, because circumstances that induce nonlinear counterparts are very rare.

Nevertheless, a generalization of (3.15) is made to enable a classification of the material as gradient-extended standard dissipative. This formal step is the introduction of a *dissipation potential* that incorporates the diffusion law, i.e., the relation between chemical potential gradient and fluid flux. It leads to a simple representation of this constitutive law that could not be related to a thermodynamic potential as the other material responses in (3.23), and can be embedded into a variational formulation with ease. The dissipation potential is denoted by ϕ, relates to the unit undeformed volume, and its value is prescribed by the material response $\widehat{\phi}$, which is assumed to depend on the negative chemical potential and *known* deformation gradient and concentration,

$$\phi = \widehat{\phi}(-\nabla\mu; \boldsymbol{F}, c), \tag{3.24}$$

where the semicolon separates the dependency on $-\nabla\mu$ from the given input $\boldsymbol{F}$ and c. The initial dependencies on entropy, concentration gradient and chemical potential in (3.2) disappeared – which is an assumption that can be shown to be justified by the proven phenomenological validity of the diffusion law that will be inserted, but not by a principle being part of the constitutive theory. The dependency on s can at least be argued to be implicitly shifted to constitute the material parameters in $\widehat{\phi}$ (as it has been done for the temperature dependence of $\widehat{\psi}$ and its related functions)[17]. Note also that the sign inversion of the chemical potential is a notational dodge that enables for a compact representation of the reduced entropy inequality as

$$\theta\Gamma_s = \partial_{-\nabla\mu}\widehat{\phi}(-\nabla\mu; \boldsymbol{F}, c) \cdot \nabla\mu \geq 0. \tag{3.25}$$

The inequality is fulfilled when $\widehat{\phi}$ is normalized (it vanishes for $\nabla\mu = \boldsymbol{0}$), positive for any chemical potential gradient and convex with respect to $-\nabla\mu$. This form of the second law of thermodynamics hence allows for the construction of *thermodynamically consistent* material laws, which is a convex, positive and normalized dissipation potential in the present work, because all other

[17] This rationale comes to grief however when facing the technical difficulties of the entropy as a function parameter.

constitutive laws were shown to prescribe quantities involved in reversible parts of the process.

Now, all constitutive laws can be related to two scalar functions by partial differentiation, and the complete list of constitutive laws reads

$$\begin{aligned}
\psi(\boldsymbol{X},t) &= \widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\boldsymbol{P}(\boldsymbol{X},t) &= \widehat{\boldsymbol{P}}(\boldsymbol{F},c,\nabla c) = \partial_{\boldsymbol{F}}\widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\Gamma_\xi(\boldsymbol{X},t) &= \widehat{\Gamma}_\xi(\boldsymbol{F},c,\nabla c) = \mu - \partial_c\widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\boldsymbol{J}_\xi(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_\xi(\boldsymbol{F},c,\nabla c) = \partial_{\nabla c}\widehat{\psi}(\boldsymbol{F},c,\nabla c)\\
\boldsymbol{J}_c(\boldsymbol{X},t) &= \widehat{\boldsymbol{J}}_c(-\nabla\mu;\boldsymbol{F},c) = \partial_{-\nabla\mu}\widehat{\phi}(-\nabla\mu;\boldsymbol{F},c).
\end{aligned} \tag{3.26}$$

Dual Dissipation Potential. *Dual* dissipation potentials can be constructed by Legendre transformations and are often preferred within continuum theories. The dual counterparts can reveal explicit evolution equations or ship with differentiability by use of the smooth hull that a Legendre transform wraps around non-convex functions. Here, a further exploitation of dual dissipation potentials is postponed to the problem-specific chapters, as the unnecessary generality at this point hinders an instructive outline.

3.4 Frame Invariance, Material Symmetry and Isotropy

Two principles are missing to conclude the application of standard constitutive theory to elasticity with diffusion and phase separation. The first one is concerned with the equivalence of constitutive response functions under different frames of reference. The second one comes along with symmetry properties of materials, which can be a challenge of three-dimensional spatial sense and imagination, but luckily boils down to the simplest case.

3.4.1 Principle of Material Frame Invariance

The principle of *material frame invariance*, also named principle of *objectivity* states that all constitutive material response functions shall be independent on the reference frame used to describe the motion of the material body, and is enforced by the study of transformation behavior of constitutive functions and their dependencies under a change of this frame. This idea has first been formulated by Oldroyd [1950] and emerged as a postulate beyond reproach.

Here, an alternate statement is used that demands unchanged material responses for rigid body motions superimposed onto the deformed body. To this end, consider the rigid body motion

$$\boldsymbol{\varphi}^+(\boldsymbol{X},t) = \widehat{\boldsymbol{\varphi}}^+(\boldsymbol{x},t) \circ \boldsymbol{\varphi}(\boldsymbol{X},t) = \boldsymbol{c}(t) + \boldsymbol{Q}(t)\boldsymbol{\varphi}(\boldsymbol{X},t), \tag{3.27}$$

where $\boldsymbol{c}(t)$ is a time-dependent translation and $\boldsymbol{Q}(t)$ a time-dependent rotation, hence a proper orthogonal tensor. All quantities defined on the configuration $\boldsymbol{\varphi}^+(\boldsymbol{X},t)$ are accentuated with a superscript plus sign. Before studying how the response functions transform with such a superimposed motion, the transformation of functional dependencies can be deduced as

$$\boldsymbol{F}^+ = \boldsymbol{Q}\boldsymbol{F}, \quad c^+ = c, \quad \nabla c^+ = \nabla c \quad \text{and} \quad \nabla \mu^+ = \nabla \mu. \tag{3.28}$$

The relation for the deformation gradient is obtained by application of the chain rule, while scalar quantities (the concentration) as well as material tensors (concentration and chemical potential gradient, both scalars derived with respect to material coordinates) remain unaffected upon superposition of $\boldsymbol{\varphi}^+$. Next, the quantities prescribed by constitutive laws transform according to

$$\psi^+ = \psi, \quad \boldsymbol{P}^+ = \boldsymbol{Q}\boldsymbol{P}, \quad \Gamma_\xi^+ = \Gamma_\xi, \quad \boldsymbol{J}_\xi^+ = \boldsymbol{J}_\xi \quad \text{and} \quad \boldsymbol{J}_c^+ = \boldsymbol{J}_c. \tag{3.29}$$

Here, scalars are again unaffected by the additional rigid body motion, and from (2.34), it is clear that referential fluxes neither change. The transformations (3.28) and (3.29), together with the actual demand of material frame invariance, can be combined into the form

$$\begin{aligned}
\widehat{\psi}(\boldsymbol{F},c,\nabla c) &= \widehat{\psi}(\boldsymbol{Q}\boldsymbol{F},c,\nabla c)\\
\boldsymbol{Q}\widehat{\boldsymbol{P}}(\boldsymbol{F},c,\nabla c) &= \widehat{\boldsymbol{P}}(\boldsymbol{Q}\boldsymbol{F},c,\nabla c)\\
\widehat{\Gamma}_\xi(\boldsymbol{F},c,\nabla c) &= \widehat{\Gamma}_\xi(\boldsymbol{Q}\boldsymbol{F},c,\nabla c)\\
\widehat{\boldsymbol{J}}_\xi(\boldsymbol{F},c,\nabla c) &= \widehat{\boldsymbol{J}}_\xi(\boldsymbol{Q}\boldsymbol{F},c,\nabla c)\\
\widehat{\boldsymbol{J}}_c(-\nabla\mu;\boldsymbol{F},c) &= \widehat{\boldsymbol{J}}_c(-\nabla\mu;\boldsymbol{Q}\boldsymbol{F},c).
\end{aligned} \tag{3.30}$$

Constitutive functions that fulfill the above requirement of frame invariance are called *objective*. This demand is a strong restriction upon the possible choices of admissible material laws. Using the independence of referential tensors with regard to the superimposed motion, it can be shown that formulating the constitutive laws (3.26) in terms of the right Cauchy-Green tensor (2.20) satisfies the demand of objectivity a priori, see e.g. Tadmor et al. [2012, p. 207 et sqq.] for a detailed derivation. The list of constitutive functions (3.26) can now be refined one more time as

$$\begin{aligned}
\psi(\boldsymbol{X},t) &= \widehat{\psi}(\boldsymbol{C},c,\nabla c)\\
\boldsymbol{P}(\boldsymbol{X},t) &= \widehat{\boldsymbol{P}}(\boldsymbol{F},c,\nabla c) = 2\boldsymbol{F}\partial_{\boldsymbol{C}}\widehat{\psi}(\boldsymbol{C},c,\nabla c)
\end{aligned}$$

$$\Gamma_\xi(\boldsymbol{X},t) = \widehat{\Gamma}_\xi(\boldsymbol{C},c,\nabla c) = \mu - \partial_c\widehat{\psi}(\boldsymbol{C},c,\nabla c) \tag{3.31}$$

$$\boldsymbol{J}_\xi(\boldsymbol{X},t) = \widehat{\boldsymbol{J}}_\xi(\boldsymbol{C},c,\nabla c) = \partial_{\nabla c}\widehat{\psi}(\boldsymbol{C},c,\nabla c)$$

$$\boldsymbol{J}_c(\boldsymbol{X},t) = \widehat{\boldsymbol{J}}_c(-\nabla\mu;\boldsymbol{C},c) = \partial_{-\nabla\mu}\widehat{\phi}(-\nabla\mu;\boldsymbol{C},c).$$

Note that the ability to formulate the constitutive functions depending on the right Cauchy-Green tensor instead of the deformation gradient is sufficient to fulfill the principle of material frame invariance; once this is made sure, the material laws can still be written down as functions of $\boldsymbol{F}$.

3.4.2 Symmetry Properties of an Isotropic Material

The last principle imposing restrictions on the construction of valid material laws is concerned with material symmetries, which can be identified when ≪a stress-strain relation remains invariant under certain changes of reference configuration≫ (Haupt [2002, p. 276]). Here, not only identical stress-strain relations are sought for certain orientations, but identical material responses for all quantities that are specified by a constitutive law.

These orientations of the reference configuration are constructed by rigid body motions. Here, the application of material symmetries is comparable to the principle of frame invariance, however, the motion is imposed on the reference, not the current configuration. Another distinctive feature of material symmetries is that they are motivated by microstructural characteristics of the solid.

Symmetries can be studied by considering rotations only. The kinematic manifestation of such a superimposed rotation can be written as a multiplicative decomposition of the deformation gradient into a pure rotation $\boldsymbol{Q}$ and a mixed rotational and translational strain measure $\boldsymbol{F}^\diamond$, where a constitutive law is supposed to be invariant with regard to the first rotation. The decomposition reads

$$\boldsymbol{F} = \boldsymbol{F}^\diamond\boldsymbol{Q}, \tag{3.32}$$

where the rotation $\boldsymbol{Q}$ belongs to a symmetry group that subsumes referential orientations under which the constitutive response remains unchanged in accordance with the microstructure of the material. Common materials with distinct symmetry groups (i.e., showing *anisotropic* behavior) are crystals or fiber-reinforced composites. However, the simplest case is full isotropy, which applies within the present work. The transformation of functional dependencies is computed using the above multiplicative decomposition

$$\boldsymbol{C}^\diamond = \boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T, \quad c^\diamond = c, \quad \nabla c^\diamond = \boldsymbol{Q}\nabla c \quad \text{and} \quad \nabla\mu^\diamond = \boldsymbol{Q}\nabla\mu, \tag{3.33}$$

where the right Cauchy-Green strain measure appears instead of the deformation gradient as induced by the principle of objectivity (3.31). The transformation behavior of prescribed quantities on the other hand is

$$\psi^\diamond = \psi, \quad \boldsymbol{P}^\diamond = \boldsymbol{P}\boldsymbol{Q}^T, \quad \Gamma_\xi^\diamond = \Gamma_\xi, \quad \boldsymbol{J}_\xi^\diamond = \boldsymbol{Q}\boldsymbol{J}_\xi \quad \text{and} \quad \boldsymbol{J}_c^\diamond = \boldsymbol{Q}\boldsymbol{J}_c, \tag{3.34}$$

obtained via push-forward operations. These do not alter scalars, rotate material vectors by $\boldsymbol{Q}$ and transform the material slots of second order tensors[18]. It allows for the formulation of the actual restriction, which reads

$$\begin{aligned}
\widehat{\psi}(\boldsymbol{C}, c, \nabla c) &= \widehat{\psi}(\boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T, c, \boldsymbol{Q}\nabla c) \\
\widehat{\boldsymbol{P}}(\boldsymbol{F}, c, \nabla c)\boldsymbol{Q}^T &= \widehat{\boldsymbol{P}}(\boldsymbol{F}\boldsymbol{Q}^T, c, \boldsymbol{Q}\nabla c) \\
\widehat{\Gamma}_\xi(\boldsymbol{C}, c, \nabla c) &= \widehat{\Gamma}_\xi(\boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T, c, \boldsymbol{Q}\nabla c) \\
\boldsymbol{Q}\widehat{\boldsymbol{J}}_\xi(\boldsymbol{C}, c, \nabla c) &= \widehat{\boldsymbol{J}}_\xi(\boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T, c, \boldsymbol{Q}\nabla c) \\
\boldsymbol{Q}\widehat{\boldsymbol{J}}_c(-\nabla\mu; \boldsymbol{C}, c) &= \widehat{\boldsymbol{J}}_c(-\boldsymbol{Q}\nabla\mu; \boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T, c)
\end{aligned} \tag{3.35}$$

for arbitrary rotations $\boldsymbol{Q}$, cf. Šilhavý [1997, p. 157 et sqq.]. It can be shown that the requirements in the case of isotropy lead to a reduced form of the constitutive free energy, which depends on the right Cauchy-Green tensor through its invariants instead of $\boldsymbol{C}$ itself. The invariants of $\boldsymbol{C}$ are defined as

$$I_c^1 = \operatorname{tr}[\boldsymbol{C}], \quad I_c^2 = \tfrac{1}{2}\left(\left(I_c^1\right)^2 - \operatorname{tr}\left[\boldsymbol{C}^2\right]\right) \quad \text{and} \quad I_c^3 = \det[\boldsymbol{C}], \tag{3.36}$$

while the ensemble of all three invariants shall be denoted by the notational shortcut $\boldsymbol{I}_c = \left(I_c^1, I_c^2, I_c^3\right)$. Note that this a priori satisfaction of the constraints imposed by material symmetry does only affect the constitutive free energy density and its derivatives with respect to strain measures, i.e., the constitutive stresses. The demands of (3.35) must be separately satisfied for the chemical microforce and microtraction as well as for the fluid flux, which is performed upon specification of the constitutive laws. Then, the updated list of constitutive functions takes the invariants of $\boldsymbol{C}$ into account and reads

$$\begin{aligned}
\psi(\boldsymbol{X}, t) &= \widehat{\psi}(\boldsymbol{I}_c, c, \nabla c) \\
\boldsymbol{P}(\boldsymbol{X}, t) &= \widehat{\boldsymbol{P}}(\boldsymbol{F}, c, \nabla c) = 2\boldsymbol{F}\partial_{\boldsymbol{C}}\widehat{\psi}(\boldsymbol{I}_c, c, \nabla c) \\
\Gamma_\xi(\boldsymbol{X}, t) &= \widehat{\Gamma}_\xi(\boldsymbol{I}_c, c, \nabla c) = \mu - \partial_c\widehat{\psi}(\boldsymbol{I}_c, c, \nabla c) \\
\boldsymbol{J}_\xi(\boldsymbol{X}, t) &= \widehat{\boldsymbol{J}}_\xi(\boldsymbol{I}_c, c, \nabla c) = \partial_{\nabla c}\widehat{\psi}(\boldsymbol{I}_c, c, \nabla c)
\end{aligned} \tag{3.37}$$

[18] When compared to the transformation lists (3.29) and (3.28), more quantities are affected by the motion superimposed onto the reference configuration – this results from favoring a Lagrangian description over an Eulerian. Choosing the latter would lead to the opposite.

$$\boldsymbol{J}_c(\boldsymbol{X},t) = \widehat{\boldsymbol{J}}_c(-\nabla\mu;\boldsymbol{C},c) = \partial_{-\nabla\mu}\widehat{\phi}(-\nabla\mu;\boldsymbol{C},c).$$

Again, $\widehat{\psi}$ does not have to be formulated in terms of $\boldsymbol{I}_c$, but it must be convertible into a representation based on these invariants.

Chapter 4:
Phase Separation of Binary Constituents: the Cahn-Hilliard Equation

As a first application of the theory outlined so far, the Cahn-Hilliard equation attributed to Cahn & Hilliard [1958] and Cahn [1961] is studied on a rigid domain (published in similar form as Miehe et al. [2014a]). This fourth-order partial differential equation describes the formation of microstructures that are often a form of phase separation. An in-depth understanding of the PDE (4.3) that characterizes important qualitative features of the evolution of two-phase systems in material science is thus crucial for the understanding of microstructures. To this end, elastic deformations are excluded to simplify the problem by a vanishing deformation map, related strain measures and time derivatives. Fluid and microforce balances (2.43), (2.49) are the governing local equations.

The Cahn-Hilliard equation determines the evolution of an order parameter, here, the spatial and time-dependent evolution of the concentration field (2.24). When the species supply R_c and the microforce supply R_ξ in (2.43) and (2.49) are assumed to be zero, the mass of solute particles is conserved. The phase separation is achieved by a combination of down- and up-hill diffusion, where the fluid moves in and against the direction of the concentration gradient, the latter controlled by the gradient of the chemical potential. These effects are determined by the competition of two terms in the free energy of the system, a non-convex configurational energy function of the concentration, and a surface energy function of the concentration gradient. During initial stages of phase separation, the configurational energy is minimized by favoring the spatial domination of pure phases. Subsequent stages minimize the surface energy by forming larger single-phase domains, which successively absorb smaller phase regions, i.e., the interface area between the two phases are minimized.

Variational Principle for the Phase Field Evolution. The constitutive structure behind the Cahn-Hilliard equation has been outlined in Gurtin [1996] in the context of extended continuum theories based on microforce balances. This comprises a rational approach to the derivation of the Cahn-Hilliard equation by combining the two balances for species content and microforce with constitutive assumptions for species flux, microforce and microtractions. A clear separation of balance laws and constitutive material functions is intended, a premature insertion of specific functions thus avoided. In this sense, the Cahn-Hilliard

equation can be embedded into gradient-extended continuum formulations, such as gradient damage, gradient plasticity or phase field theories – models that can be motivated by microforce balance equations. A variational formulation of these continuum theories based on rate and incremental variational principles was outlined in Miehe [2011], and applied to a wide range of problems.

A difficulty is associated with the construction of a variational principle for the Cahn-Hilliard theory due to its inherent dissipative transport equation, which requires a new variational approach. With regard to this challenge, the Cahn-Hilliard equation on a rigid domain provides the simplest application for the specification of variational potentials for problems that include dissipative effects due to mass transport phenomena.

In this chapter, the Cahn-Hilliard equation is shown to be related to a mixed variational principle for the evolution problem, formulated in terms of the rate of the concentration and the chemical potential. Its Euler-Lagrange equations are two differential equations which incorporate the balances of fluid content and microforce in terms of the constitutive free energy storage and the dissipation potential. The variational foundation highlights a symmetry of the evolution problem, which is considered an important ingredient of the Cahn-Hilliard theory, ready to be exploited within a finite element analysis.

Numerical Implementation of the Saddle Point Principle.A direct numerically approximated solution of the Cahn-Hilliard equation with its fourth-order gradients in the concentration requires higher-order interpolations of the primary field, i.e., basis functions that are globally C^1-continuous. With this regard, finite difference approaches such as applied by Choo et al. [2004] provide straightforward solutions on simple regular and non-moving domains. For non-regular domains or deforming solids however, finite element schemes are the appropriate discretization technique, previously studied by e.g. Stogner et al. [2008] or Rajagopal et al. [2010]. The treatment of C^1-continuous finite element spaces requires an increase of the polynomial degree, isogeometric methods play a key role in this context, as in Gomez et al. [2008], Cottrell et al. [2009] or Anders & Weinberg [2011]. The discontinuous Galerkin method has additionally been successfully applied to the Cahn-Hilliard theory by Wells et al. [2006]. To circumvent the complexities of higher-order finite element spaces, the Cahn-Hilliard equation can also be reformulated as two coupled partial differential equations of second order, which reduces the continuity requirement to C^0, see the works of Barrett et al. [2000], Feng & Prohl [2004], Elliott & French [1987] and Zhang et al. [2013]. An alternative formulation has been presented by Ubachs et al. [2004] and Kuhl & Schmid [2007], which bases on local and nonlocal concentrations as primary unknowns at each node. All of these formulations do however not exploit a variational principle for the coupled problem, and the formulations are not symmetric in the chosen primary fields.

In contrast, the new time-discrete incremental potential based on the continuous rate potentials mentioned above is suitable for the incremental update problem. A space discretization of the problem by C^0 elements with nodal degrees of freedom for the concentration and the chemical potential results in a space-time-discrete potential that determines the finite element residual and the symmetric tangent by its first and second derivatives.

4.1 Summary of Equations and Boundary Conditions

The boundary value problem for Cahn-Hilliard diffusion is a coupled two-field problem with concentration and chemical potential as primary fields. The original definition (2.24) of the former is modified such that c refers to a constant maximal referential value c_{ref}, which is not specified but shifts the concentration domain to $[0,1]$ and renders it dimensionless. The slightly modified definition of the concentration field then reads

$$c : \begin{cases} \mathcal{B} \times T \to [0,1] \\ (\boldsymbol{X}, t) \mapsto c(\boldsymbol{X}, t). \end{cases} \tag{4.1}$$

The chemical potential field (2.29) is left unchanged.

Recall the local balance equations (2.43), (2.49) and the list of constitutive functions (3.37), which can be tailored to the present model problem. First, the external supply terms R_c and R_ξ are assumed to vanish for simplicity. Second, mechanical quantities do not need to be prescribed on a rigid domain, reducing the number of necessary constitutive laws. Third, the right Cauchy-Green strain measure drops from the list of dependencies. Thus, the ensemble of governing equations and constitutive functions reduce to

1.	Balance of solute content	$\dot{c} = -\operatorname{Div}[\boldsymbol{J}_c]$	
2.	Balance of microforce	$\Gamma_\xi = -\operatorname{Div}[\boldsymbol{J}_\xi]$	
3.	Constitutive microforce	$\Gamma_\xi = \mu - \partial_c \widehat{\psi}(c, \nabla c)$	(4.2)
4.	Constitutive microtractions	$\boldsymbol{J}_\xi = \partial_{\nabla c} \widehat{\psi}(c, \nabla c)$	
5.	Constitutive flux	$\boldsymbol{J}_c = \partial_{-\nabla \mu} \widehat{\phi}(-\nabla \mu; c)$	

Combining these five ingredients gives the fourth-order Cahn-Hilliard equation

$$\dot{c} = -\operatorname{Div}\left[\partial_{-\nabla\mu}\widehat{\phi}\left(-\nabla\left[\partial_c\widehat{\psi} - \operatorname{Div}[\partial_{\nabla c}\widehat{\psi}]\right]\right)\right], \tag{4.3}$$

when suitable constitutive functions as defined in (4.37) and (4.40) are inserted for $\widehat{\psi}$ and $\widehat{\phi}$.

In addition, initial and boundary conditions have to be defined. The domain surface $\partial\mathcal{B}$ is decomposed according to the two primary fields by $\partial\mathcal{B} = \partial\mathcal{B}^c \cup$

$\partial\mathcal{B}^{H_\xi} = \partial\mathcal{B}^{\mu} \cup \partial\mathcal{B}^{H_c}$, with $\partial\mathcal{B}^{c} \cap \partial\mathcal{B}^{H_\xi} = \emptyset$ and $\partial\mathcal{B}^{\mu} \cap \partial\mathcal{B}^{H_c} = \emptyset$. We consider Dirichlet and Neumann boundary conditions for the concentration and the microtractions

$$c = \bar{c}(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{c} \quad \text{and} \quad \boldsymbol{J}_\xi \cdot \boldsymbol{N} = \bar{H}_\xi(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{H_\xi}, \tag{4.4}$$

as well as for the chemical potential and the solute flux

$$\mu = \bar{\mu}(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{\mu} \quad \text{and} \quad \boldsymbol{J}_c \cdot \boldsymbol{N} = \bar{H}_c(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{H_c} \tag{4.5}$$

with prescribed concentration $\bar{c}$ and solute outflux $\bar{H}_c$, chemical potential $\bar{\mu}$ and microtraction $\bar{H}_\xi$, respectively. The pairing of Dirichlet and Neumann conditions is chosen as a preparation of the variational principle developed below. The initial condition for the concentration is

$$c(\boldsymbol{X}, t_0) = c_0(\boldsymbol{X}) \text{ in } \mathcal{B}, \tag{4.6}$$

which is the last ingredient for the problem definition necessary to construct a variational principle.

4.2 Variational Principles for Evolution and Update Problem

The governing coupled field equations combine the balance equations with the constitutive equations in (4.2). Insertion of $(4.2)_{3,4}$ into $(4.2)_2$ gives the constitutive representation of the microforce balance

$$\delta_c \widehat{\psi} - \mu = 0 \text{ in } \mathcal{B} \quad \text{with} \quad \left[\partial_{\nabla c} \widehat{\psi}\right] \cdot \boldsymbol{N} = \bar{H}_\xi \text{ on } \mathcal{B}^{H_\xi}, \tag{4.7}$$

where the notation refers to the functional derivative of a scalar quantity as defined in Section 2.1. Next, insertion of $(4.2)_5$ into $(4.2)_1$ gives the constitutive representation for the balance of solute content

$$\dot{c} + \delta_\mu \widehat{\phi} = 0 \text{ in } \mathcal{B} \quad \text{with} \quad \left[\partial_{-\nabla \mu} \widehat{\phi}\right] \cdot \boldsymbol{N} = \bar{H}_c \text{ on } \mathcal{B}^{H_c}. \tag{4.8}$$

The two equations (4.7) and (4.8), together with the Dirichlet conditions (4.4) and (4.5), are the central characteristics for a variational formulation of the Cahn-Hilliard diffusion problem. Their relation to minimization and saddle point principles is shown in the following.

4.2.1 Stored Energy, Dissipation Potential and Load Functionals

Two functionals related to energy storage and dissipative transport mechanisms are introduced, both depending on the primary fields of the problem. The stored energy functional depends on the concentration field c, reads

$$E(c) = \int_{\mathcal{B}} \widehat{\psi}\,(c, \nabla c)\, dV \tag{4.9}$$

and characterizes the chemical energy stored in the domain $\mathcal{B}$. It is constituted by the free Helmholtz energy density $\widehat{\psi}$.

The suitable definition of a dissipation potential functional is a key ingredient to the variational formulation in terms of rate potentials. It is assumed to be determined by the variational statement

$$D(\dot{c}) = \int_{\mathcal{B}} \widehat{\chi}\,(\dot{c}; c)\, dV = \sup_{\mu} \left\{ \int_{\mathcal{B}} \left[-\mu \dot{c} - \widehat{\phi}(-\nabla\mu; c) \right] dV - P^{\mu}_{\text{ext}}\,(\mu; t) \right\}, \tag{4.10}$$

depending on the rate $\dot{c}$ of the concentration, and where the canonical dissipation potential function $\widehat{\chi}(\dot{c}; c)$ is introduced for later use. It contains an external load functional

$$P^{\mu}_{\text{ext}}(\mu; t) = -\int_{\partial\mathcal{B}^{H_c}} \mu \bar{H}_c dA \tag{4.11}$$

of prescribed species flux $\bar{H}_c$ related to the Neumann condition (4.5). Observe that, when taking the variation based on the admissible functions $\delta\mu \in \mathcal{W}_\mu = \{\delta\mu \mid \delta\mu = 0 \text{ on } \partial\mathcal{B}^{\mu}\}$ which satisfy the homogeneous form of the Dirichlet condition (4.5), the Euler equation of the optimization problem (4.10) yields the constitutive form (4.8) for the balance of solute content including the Neumann condition for the solute flux. The above definition (4.10) of the dissipation functional $D(\dot{c})$ provides the key approach for incorporating dissipative transport phenomena into a rate-type variational formulation. The variational problem (4.10) can be interpreted as a Legendre-Fenchel transformation, where the chemical potential plays the role of a mixed variable, dual to the rate of concentration. The formulation extends the local form to a global definition that accounts for boundary conditions of the dual variable μ.

Finally, a second external load functional related to the Neumann boundary condition (4.4) is introduced as

$$P^{c}_{\text{ext}}\,(\dot{c}; t) = \int_{\partial\mathcal{B}^{c}} \dot{c} \bar{H}_{\xi} dA \tag{4.12}$$

that contains a prescribed microtraction $\bar{H}_{\xi}$. For simplicity, all external actions are assumed to be dead loads, i.e., independent of the primary fields c and μ.

4.2.2 The Single-Field Minimization Principle

Based on the energy and the dissipation potential functionals E and D and the external work functional P^c_{ext} above, the rate potential

$$\Pi(\dot{c}) = \frac{d}{dt}E(c) + D(\dot{c}) - P^c_{\text{ext}}(\dot{c};t) \tag{4.13}$$

is postulated at given state $\{c\}$ and time t. Rewritten with its internal and external contributions, it reads

$$\Pi(\dot{c}) = \int_{\mathcal{B}} \pi\left(\dot{c}, \nabla\dot{c}\right) dV - P^c_{\text{ext}}\left(\dot{c};t\right) \tag{4.14}$$

in terms of the internal rate potential per unit volume

$$\pi\left(\dot{c}, \nabla\dot{c}\right) = \frac{d}{dt}\widehat{\psi}(c, \nabla c) + \widehat{\chi}(\dot{c}; c) \tag{4.15}$$

that contains the rate of the energy storage function $\widehat{\psi}$ and the canonical dissipation potential function $\widehat{\chi}$ introduced in (4.10). The evolution of the concentration at a given state is determined by the minimization principle

$$\{\dot{c}\} = \text{Arg}\left\{\inf_{\dot{c}\in\mathcal{W}_{\dot{c}}} \Pi\left(\dot{c}\right)\right\}. \tag{4.16}$$

Note that this variational principle is governed by the convexity of the dissipation potential function $\widehat{\chi}$. Computing the variation of the rate potential (4.13) for admissible virtual rates $\delta\dot{c} \in \mathcal{W}_{\dot{c}} = \{\delta\dot{c} \mid \delta\dot{c} = 0 \text{ on } \partial\mathcal{B}^c\}$ of the concentration, which satisfies the homogeneous form of the Dirichlet boundary conditions, the microforce balance (4.7) is obtained as the Euler equations of the variational principle (4.16), along with the associated Neumann boundary conditions. They determine the concentration field for a given chemical potential.

4.2.3 The Two-Field Saddle Point Principle

The variational principle (4.16) can be combined with the Legendre-Fenchel transformation (4.10) to an extended variational principle that contains the chemical potential as a mixed variable. To this end, the extended dissipation potential functional

$$D^*(\dot{c}, \mu) = \int_{\mathcal{B}} \left\{-\mu\dot{c} - \widehat{\phi}\left(-\nabla\mu; c\right)\right\} dV \tag{4.17}$$

is defined. Note that this functional now explicitly contains the constitutive dissipation potential function $\widehat{\phi}$. Based on the energy E (4.9), the dual dissipation

potential functional D^* above and the external work functional $P^{c,\mu}_{\text{ext}} = P^{c}_{\text{ext}} + P^{\mu}_{\text{ext}}$, the mixed two-field potential is defined as

$$\Pi^*(\dot{c}, \mu) = \frac{d}{dt} E(c) + D^*(\dot{c}, \mu) - P^{c,\mu}_{\text{ext}}(\dot{c}, \mu; t) \tag{4.18}$$

at a given state $\{c\}$. Note that the only modification of (4.13) affects the dissipation, which is now expressed in terms of the rate of the concentration and the chemical potential. Again, this potential is reformulated with its internal and external contributions as

$$\Pi^*(\dot{c}, \mu) = \int_{\mathcal{B}} \pi^* \left(\dot{c}, \nabla\dot{c}, \mu, \nabla\mu\right) dV - P^{c,\mu}_{\text{ext}}(\dot{c}, \mu; t) \tag{4.19}$$

in terms of the mixed internal rate potential per unit volume

$$\pi^* \left(\dot{c}, \nabla\dot{c}, \mu, \nabla\mu\right) = \frac{d}{dt}\widehat{\psi}(c, \nabla c) - \mu\dot{c} - \widehat{\phi}(-\nabla\mu; c) \tag{4.20}$$

Note that the potential π^* is linear with respect to the rates of the state variables and concave with respect to the chemical potential. The latter property stems from the convexity of the dissipation potential function $\widehat{\phi}$. The evolution of the concentration as well as the chemical potential at a given state are determined by the mixed two-field variational principle

$$\{\dot{c}, \mu\} = \text{Arg}\left\{ \inf_{\dot{c}\in\mathcal{W}_{\dot{c}}} \sup_{\mu\in\mathcal{W}_{\mu}} \Pi^*\left(\dot{c}, \mu\right) \right\}, \tag{4.21}$$

where the admissible space $\mathcal{W}_{\dot{c}}$ has been specified above and $\mathcal{W}_{\mu} = \{\mu \in H^1(\mathcal{B}) \mid \mu = \bar{\mu} \text{ on } \partial\mathcal{B}^{\mu}\}$. The statement (4.21) defines the rate $\dot{c}$ of the concentration along with the chemical potential at the given state $\{c\}$ and time t. The variation of the potential (4.19) yields the governing equations (4.7) and (4.8) as Euler equations of the variational statement (4.21), i.e.,

$$\delta_{\dot{c}}\pi^* = -\mu + \partial_c\widehat{\psi} - \text{Div}\left[\partial_{\nabla c}\widehat{\psi}\right] = 0 \tag{4.22}$$

as well as

$$\delta_{\mu}\pi^* = -\dot{c} - \text{Div}\left[\partial_{-\nabla\mu}\widehat{\phi}\right] = 0 \tag{4.23}$$

along with the Neumann boundary conditions. Note that the above two coupled field equations are obtained variational derivatives of the rate potential π^* defined in (4.20).

4.2.4 Time-Discrete Incremental Variational Principle

The rate potentials outlined above form the basis for a time-discrete, incremental formulation of Cahn-Hilliard diffusion problems. Consider a discrete time interval $[t_n, t]$ with step length $\tau = t - t_n > 0$, and assume all fields at time t_n to be known.

The goal is then to determine the fields at the current discrete time t (variables without subscript) based on a variational statement valid for the time increment under consideration. A mixed incremental variational principle of dissipative, gradient-extended solids is based on the modified functional

$$\Pi^{*\tau}(c,\mu) = E^{\tau}(c) + D^{*\tau}(c,\mu) - P_{\text{ext}}^{c,\mu}(c,\mu;t) \tag{4.24}$$

in terms of the incremental energy functional E^{τ}, the incremental mixed dissipative work functional $D^{*\tau}$ and the load functional $P_{\text{ext}}^{c,\mu}$. It can be written in the form

$$\Pi^{*\tau}(c,\mu) = \int_{\mathcal{B}} \pi^{*\tau}\left(c, \nabla c, \mu, \nabla\mu\right) dV - P_{\text{ext}}^{c,\mu}(c,\mu;t), \tag{4.25}$$

where $\pi^{*\tau}$ is called the mixed incremental potential density per unit volume. It is related to the rate potential π^{*} (4.20) by the algorithm

$$\pi^{*\tau}(c,\nabla c,\mu,\nabla\mu) = \mathsf{Algo}\left\{\int_{t_n}^{t} \pi^{*}\left(\dot{c},\nabla\dot{c},\mu,\nabla\mu\right) dt\right\} \tag{4.26}$$

and governed by $\widehat{\psi}$ and $\widehat{\phi}$. Then, the incremental variational principle

$$\{c,\mu\} = \text{Arg}\left\{\inf_{c\in\mathcal{W}_c}\sup_{\mu\in\mathcal{W}_\mu} \Pi^{*\tau}(c,\mu)\right\} \tag{4.27}$$

determines the concentration field as well as the chemical potential at the current time t as a saddle point of $\Pi^{*\tau}$. The algorithm Algo is constructed such that the variation of the potential (4.25) for admissible variations $\delta c \in \mathcal{W}_c = \{\delta c \mid \delta c = 0 \text{ on } \partial\mathcal{B}^c\}$ and $\delta\mu \in \mathcal{W}_\mu$ leads to consistent algorithmic counterparts of the continuous equations (4.22) and (4.23) as the Euler equations. For the sake of a convenient notation, a compact notation can be introduced by arrays of primary and state variables

$$\mathfrak{U} = [c,\mu]^T \quad \text{and} \quad \mathfrak{Q} = [c,\nabla c,\mu,-\nabla\mu]^T. \tag{4.28}$$

With these definitions at hand, consider the algorithmic form of the mixed incremental potential (4.26)

$$\pi^{*\tau}(\mathfrak{Q}) = \widehat{\psi}(c,\nabla c) - \mu(c-c_n) - \tau\widehat{\phi}(-\nabla\mu;c_n) \tag{4.29}$$

based on a backward Euler approximation, which is a function of the constitutive state array (4.28)[19].

[19]Note that the formulation (4.29) employs a dissipation function at frozen state c_n. This ensures variational consistency with the continuous rate potential (4.20).

The finite element treatment is then characterized by the first and second derivatives of this potential. This defines a generalized stress array

$$\mathfrak{S} = \partial_{\mathfrak{Q}} \pi^{*\tau}(\mathfrak{Q}) = \begin{bmatrix} \partial_c \widehat{\psi} - \mu \\ \partial_{\nabla c} \widehat{\psi} \\ -(c - c_n) \\ -\tau \partial_{-\nabla \mu} \widehat{\phi} \end{bmatrix} \tag{4.30}$$

and a symmetric generalized tangent moduli array

$$\mathfrak{C} = \partial^2_{\mathfrak{Q}\mathfrak{Q}} \pi^{*\tau}(\mathfrak{Q}) = \begin{bmatrix} \partial^2_{cc} \widehat{\psi} & \cdot & -1 & \cdot \\ \cdot & \partial^2_{\nabla c \nabla c} \widehat{\psi} & \cdot & \cdot \\ -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -\tau \partial^2_{-\nabla \mu - \nabla \mu} \widehat{\phi} \end{bmatrix}, \tag{4.31}$$

where a decoupled energy function has been assumed as considered in (4.37) below. These two arrays are the constitutive input for the spatial discretization by means of the finite element method. The symmetry of the tangent moduli (4.31) for the two-field representation is an important ingredient of the proposed incremental variational principle (4.27) based on the algorithmic potential density (4.26). Note the indefinite structure of the moduli stemming from the inf-sup problem (4.27). This intrinsic symmetry is not visible in previous treatments of splitting methods reported in Elliott et al. [1989], Kuhl & Schmid [2007] or Zhang et al. [2013].

4.2.5 Finite Element Potential for Dirichlet Problems

Based on the incremental variational principle (4.27), a spatial discretization of the coupled problem is considered. For simplicity, the Neumann boundary conditions are assumed to be zero, $P^{c,\mu}_{\text{ext}} = 0$. Let $\mathfrak{T}^h$ denote a finite element triangulation of the rigid domain $\mathcal{B}$. The index h indicates a typical mesh size based on finite element domains $\mathcal{B}^h_e \in \mathfrak{T}^h$ and N^h global nodal points. Associated with the triangulation $\mathfrak{T}^h$, the finite element interpolations of primary variables and the state variable vector are written in the form

$$\mathfrak{U}^h = \underline{\boldsymbol{N}}(\boldsymbol{X}) \underline{\boldsymbol{d}} \quad \text{and} \quad \mathfrak{Q}^h = \underline{\boldsymbol{B}}(\boldsymbol{X}) \underline{\boldsymbol{d}} \tag{4.32}$$

in terms of the nodal state vector $\boldsymbol{d}$, which contains the primary variables $\mathfrak{U}$ (4.28) at a nodal point of the finite element mesh. $\underline{\boldsymbol{N}}$ and $\underline{\boldsymbol{B}}$ are symbolic representations of the global shape function matrices. In the model problem below, the low-order C^0 basis functions are used for both the concentration and the chemical potential. The spatial discretization of the incremental functional (4.25) then reads

$$\Pi^{*\tau h}(\underline{\boldsymbol{d}}) = \int_{\mathcal{B}} \pi^{*\tau} \left(\underline{\boldsymbol{B}} \underline{\boldsymbol{d}} \right) dV \tag{4.33}$$

for the pure Dirichlet problem with $P_{\text{ext}}^{c,\mu} = 0$. Then, the discrete minimization principle

$$\underline{\boldsymbol{d}} = \text{Arg}\left\{ \underset{\underline{\boldsymbol{d}}}{\text{stat}}\, \Pi^{*\tau h}\left(\underline{\boldsymbol{d}}\right) \right\} \tag{4.34}$$

determines the nodal solution vector $\underline{\boldsymbol{d}}$ at current time t. The necessary condition of this discrete variational statement reads

$$\underline{\boldsymbol{R}} = \Pi^{*\tau h}_{,\underline{\boldsymbol{d}}} = \int_{\mathcal{B}} \underline{\boldsymbol{B}}^T \boldsymbol{\mathcal{S}} dV = \boldsymbol{0} \tag{4.35}$$

and provides a nonlinear algebraic system for the determination of $\underline{\boldsymbol{d}}$. For smooth problems, a standard Newton iteration updates the nodal state vector by the algorithm

$$\underline{\boldsymbol{d}} \Leftarrow \underline{\boldsymbol{d}} - \underline{\boldsymbol{K}}^{-1}\underline{\boldsymbol{R}} \quad \text{with} \quad \underline{\boldsymbol{K}} = \Pi^{*\tau h}_{,\underline{\boldsymbol{d}}\underline{\boldsymbol{d}}} = \int_{\mathcal{B}} \underline{\boldsymbol{B}}^T \mathfrak{C} \underline{\boldsymbol{B}} dV \tag{4.36}$$

in terms of the monolithic tangent $\underline{\boldsymbol{K}}$ that contains the generalized tangent moduli $\mathfrak{C}$ (4.31). Note that the finite element residual $\underline{\boldsymbol{R}}$ and tangent $\underline{\boldsymbol{K}}$ are first and second derivatives of the potential $\Pi^{*\tau h}$ (4.33). Thereby, the symmetry of the tangent $\underline{\boldsymbol{K}}$ is a consequence of the variational principle in the extended space of the concentration and chemical potential fields.

4.3 Constitutive Model and Simulations of the Two-Field Problem

The standard form of the Cahn-Hilliard equation can be obtained by specific choices of the constitutive functions $\widehat{\psi}$ and $\widehat{\phi}$. The free energy decomposes into a configurational contribution and an interface term,

$$\widehat{\psi}(c, \nabla c) = \widehat{\psi}_{\text{con}}(c) + \widehat{\psi}_{\text{int}}(\nabla c). \tag{4.37}$$

The configurational energy is assumed to be identical to that considered in Cahn & Hilliard [1958], i.e.,

$$\widehat{\psi}_{\text{con}}(c) = A\left[c \ln c + (1-c)\ln(1-c)\right] + Bc(1-c). \tag{4.38}$$

The first term (governed by the parameter A) ensures that the concentration remains in the range $[0, 1]$, which is achieved by growth conditions for $c \to 0$ and $c \to 1$ as depicted in the first row of Figure 4.1. The second term induces a non-convexity for $B > 2A$ that favors a phase separation. This double-well non-convexity is visualized in the bottom row of Figure 4.1. Finally, and consistent with the classical Cahn-Hilliard theory, the interface contribution to the free energy reads

$$\widehat{\psi}_{\text{int}}(\nabla c) = \frac{C}{2}|\nabla c|^2. \tag{4.39}$$

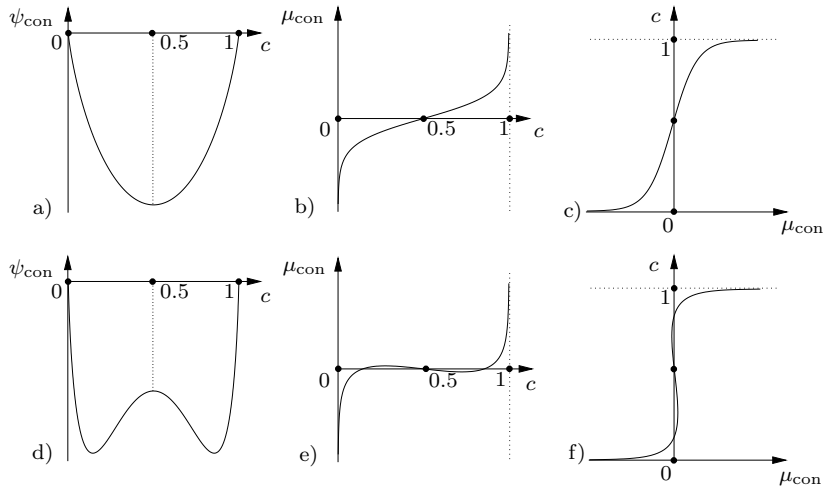


Figure 4.1: Plots related to the configurational energy $\widehat{\psi}_{\text{con}}(c) = A[\, c\ln c + (1-c)\ln(1-c)\,] + Bc(1-c)$, its derivative $\mu_{\text{con}} = \widehat{\psi}'_{\text{con}}(c) = A\ln(c/(1-c)) + B(1-2c)$ with inverse $c = \widehat{\psi}'^{-1}_{\text{con}}(\mu)$. a–c) for $A = 10$ and $B = 0$ and d–f) $A = 10$ and $B = 25$.

The dependence of the dissipation potential on the negative gradient of the chemical potential has already been assumed, its specific convex form reads

$$\widehat{\phi}(-\nabla\mu; c) = c(1-c)\frac{M}{2}|-\nabla\mu|^2. \tag{4.40}$$

at given concentration c (which is cast to c_n when inserted into (4.29)). The parameter $M > 0$ is the species diffusivity. Note that the above constitutive function (4.40) incorporates the linear diffusion law attributed to Fick [1855] (or Darcy [1856] in the case of porous media theories). It is modified such that the normalized concentration field (4.1) is accounted for, i.e., the solute transport decreases not only for low concentration, but also for a saturated state $c \to 1$.

Numerical Benchmark Simulations. In order to explore the modeling capabilities of the proposed symmetric two-field formulation based on the time-space-discrete variational potential (4.33), three simulations are carried out. Residual (4.35) and tangent (4.36) are implemented into the finite element research tool `Feap` (see Taylor [2009]), using its interface for parallel computing based on the `Petsc` library (Balay et al. [2013, 1997]). The Newton-Raphson updates (4.36) are applied together with Jacobian preconditioning and the Generalized Minimal Residual (GMRES) method developed by Saad & Schultz [1986]. The chosen element type is Q_1Q_1.

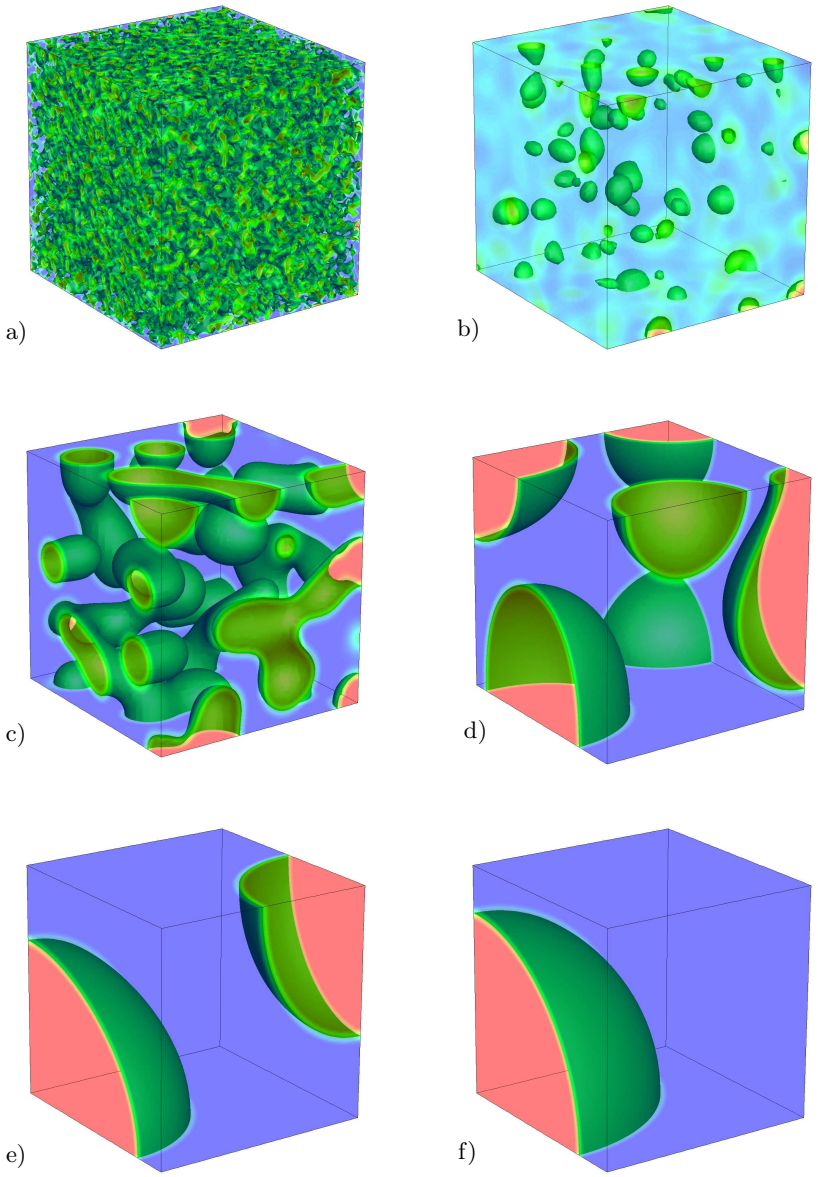


Figure 4.2: Snapshots of the concentration $c(\boldsymbol{x}, t)$ in $\mathcal{B}$ starting from random initial conditions with $p = 2$ and thus $\bar{c} = 1/3$: a) $t = 0.0$, b) $t = 0.4$, c) $t = 7.5$, d) $t = 235.5$, e) $t = 892.0$, f) $t = 4904.8$.

Parameter	Value	Unit
A : Threshold energy parameter	10.00	$\mathrm{J/m^3}$
B : Mixing energy parameter	25.00	$\mathrm{J/m^3}$
C : Gradient parameter	0.001	J/m
M : Diffusivity parameter	0.001	$\mathrm{m^5/Js}$

Table 4.1: Parameter set of the model problem used for the representative boundary value problem shown in Figure 4.2, 4.3 and 4.4.

A cubic domain with side length $L = 1/2$ given by $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^3 | \ \boldsymbol{X} \in [0, L] \times [0, L] \times [0, L]\}$ is considered. On its surface $\partial\mathcal{B}$, zero Neumann boundary conditions are prescribed for both concentration and chemical potential, i.e.,

$$\bar{H}_\xi = 0 \text{ on } \partial\mathcal{B}^{H_\xi} \quad \text{and} \quad \bar{H}_c = 0 \text{ on } \partial\mathcal{B}^{H_c}, \tag{4.41}$$

according to (4.4) and (4.5). The evolution of the concentration starts from a given initial condition (4.6). Note that the boundary condition $(4.41)_2$ enforces a constant average concentration $\bar{c}$ in $\mathcal{B}$, and hence $\bar{c}$ is determined as

$$\bar{c} = \frac{1}{|\mathcal{B}|} \int_{\mathcal{B}} c\,(\boldsymbol{X}, t)\, dV = \frac{1}{|\mathcal{B}|} \int_{\mathcal{B}} c_0\,(\boldsymbol{X})\, dV = \text{const.} \tag{4.42}$$

by the choice of $c_0(\boldsymbol{X})$.

The domain $\mathcal{B}$ is discretized by 50^3 trilinear $\mathrm{Q_1Q_1}$ brick elements, leading to a mesh with $N^h = 51^3$ nodes. The initial conditions $c_0(\boldsymbol{X})$ are determined by the nodal values c_0^I. Starting from random nodal initial conditions leads to coarsening in a mixture with initially fine microstructure. The nodal initial conditions are of the form

$$c_0^I = \left(c_{\mathrm{rand}}^I\right)^p \tag{4.43}$$

where c_{rand}^I are nodal pseudo-random numbers in the range $[0, 1]$ of (4.1) and where $p > 0$ is an exponent that determines the resulting value of $\bar{c}$ for sufficiently large N^h by

$$\bar{c} \approx \frac{1}{p+1}. \tag{4.44}$$

During the monolithic solution of the symmetric linear system of equations, a constant time step of $\tau = 0.1$ s is applied.

Material parameters for the benchmark simulations are given in Table 4.1. The choice of $B = 2.5A$ leads to a double-well separation energy with minima at $c = 0.145$ and $c = 0.855$ (which is visualized in Figure 4.1), such that for

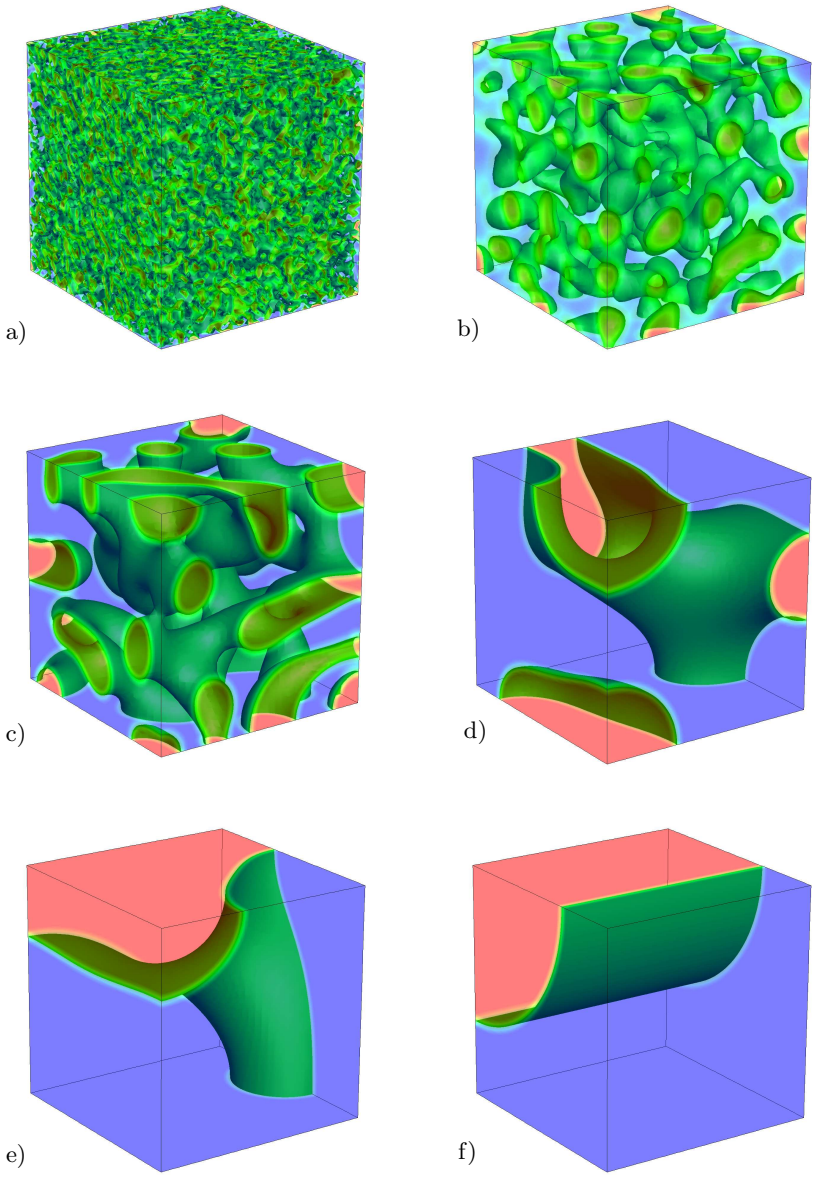


Figure 4.3: Snapshots of the concentration $c(\boldsymbol{x}, t)$ in $\mathcal{B}$ starting from random initial conditions with $p = 3/2$ and thus $\bar{c} = 2/5$: a) $t = 0.0$, b) $t = 0.4$, c) $t = 7.5$, d) $t = 235.5$, e) $t = 892.0$, f) $t = 6190.0$.

$\bar{c} \in (0.145, 0.855)$ the formation and the consequent coarsening of precipitates is expected.

Figures 4.2, 4.3 and 4.4 show the predicted evolution of the concentration from random initial conditions specified by (4.43) with $p \in \{2, 3/2, 1\}$, hence $\bar{c} \in \{1/3, 2/5, 1/2\}$, respectively. Different initial conditions lead to different equilibrium states that are all at least close to local minimizers of the interface surface between the two phases $c = 0.145$ (blue) and $c = 0.855$ (red). This can be related to analytical considerations: for the limit of $C \to 0$ (and hence zero interface thickness), it can be shown that solutions of stationary Cahn-Hilliard type problems indeed minimize the interface surface, see Modica [1987].

In all three initial setups, the initially random distribution quickly evolves to a very fine distribution of non-connected precipitates with $c = 0.855$ in a bulk material with $c = 0.145$. The further evolution then consists of precipitation coarsening. For the given value of C, the interface evolves towards an eighth sphere in the case of $p = 2$. For $p = 3/2$, the average solute content is higher, which is reflected by the higher visual density in Figure 4.3, before an equilibrium state is found in the form of a quarter cylinder. The benchmark starting from initial conditions with $p = 1$ shows an equal mixture of two strongly interwoven bulk phases within the initial stages of the simulation. Later, the surface is minimized for the given volume fraction $\bar{c}$, and a flat interface separating the cube into two equal tetrahedrons is formed.

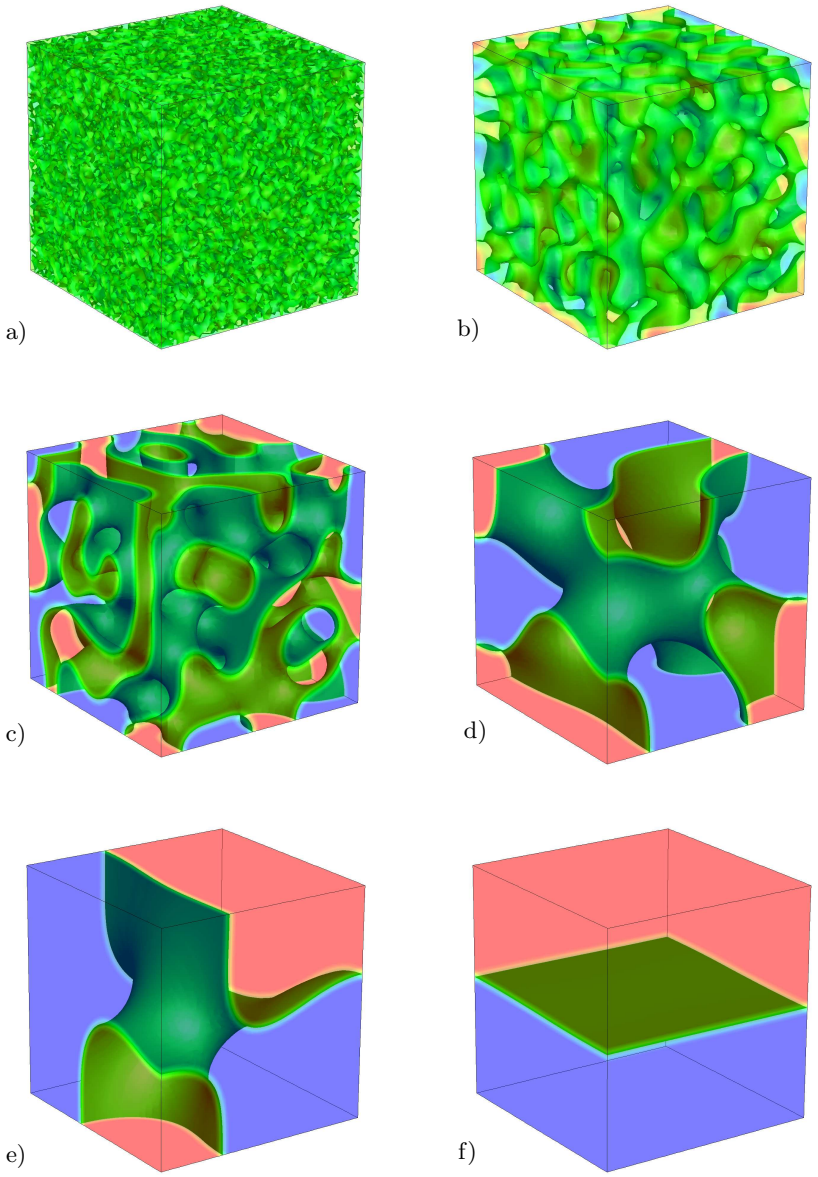


Figure 4.4: Snapshots of the concentration $c(\boldsymbol{x}, t)$ in $\mathcal{B}$ starting from random initial conditions with $p = 1$ and thus $\bar{c} = 1/2$: a) $t = 0.0$, b) $t = 0.4$, c) $t = 7.5$, d) $t = 235.5$, e) $t = 892.0$, f) $t = 6668.0$.

Chapter 5:
Fickian Diffusion in Hydrogels: Swelling, Instabilities and FE Spaces

The second application is devoted to the interplay of possibly large elastic deformation and diffusion along a negative concentration gradient (standard or downhill diffusion). An excellent choice for the study of such effects are polymeric hydrogels. Their strongly hydrophilic and crosslinked network structure of polymer chains allows for a broad range of elastic properties and exhibits large volumetric growth due to fluid influx and shrinkage due to an outflux. Diffusion-deformation phenomena can hence be studied covering a broad variety of related aspects: the incompressibility of polymer structures, the constitutive modeling approach to the energetic contribution of solute content increase related to changes of the entropy of mixing, finite element design and robustness. Here, the application is additionally embedded into a variational formulation.

In contrast to the treatise of the Cahn-Hilliard equation in Chapter 4, the deformation map does not vanish. Its time derivatives however are not accounted for, as quasi-static processes are assumed and hence, inertial terms are neglected. Microstructural arrangements are furthermore not taken into account. This will be realized by a free energy function independent of the concentration gradient: according to (3.37), microtractions vanish for such a constitutive energy storage function, and the microforce balance (2.49) prescribes a zero chemical microforce, when zero external supply term R_ξ is assumed. How gradient-extended diffusion boils down to standard downhill diffusion has also been shown in Miehe et al. [2014b].

In what follows, a short overview of applications and perspectives of polymeric hydrogels is given. Common modeling approaches are referred to, both pure continuum theories and finite element implementations. The boundary value problem is then characterized and its governing equations shortly recapped, before a variational formulation is developed, similar to the one presented in Section 4.2 with regard to the construction of a saddle point principle, yet distinguished by a minimization formulation that can be further elaborated towards an incremental and space-discrete counterpart. Constitutive functions will then be specified, initiating the study of representative finite element simulations. The treatise has been published in similar form as Böger et al. [2017b].

5.1 Applications and Modeling Approaches to Hydrogels

Applicability of Hydrogels. Recent advances in biomedical applications such as drug delivery, tissue engineering and biosensors have raised the interest in hydrogels as a well suited material for such use cases. Their hydraulic permeability is a key mechanism, making commercial employment in agriculture and waste management as well as bio- and nanotechnological tasks feasible. Progressive use of interfacial phenomena, self-assembly of natural polymer structures and controlled diffusion for certain substrates, e.g. antigen solvents, rank among the most seminal future concepts of hydrogels, see Peppas et al. [2006]. Biocompatible polymers are moreover taken as promising materials for artificial organs and implant materials, whereas wound dressings, diapers and contact lenses already show the safe every day use of hydrogels. Stimuli-responsive hydrogels can be employed as smart microactuators, independent of an external power source, as reported by Beebe et al. [2000], see also Ulijn et al. [2007] for an overview of further applications and Bach [2015]; Claßen [2016]; Weber [2014]; Spiecker gen. Döhmann [2014] and Pfeifer [2016] for unrelated topics as well as Wittmund [2015] for an outline of mercantile perspectives, inspired by the seminal work of Golus [2011]. Numerous studies have dealt not only with advanced synthesis of hydrogels, but also with the design of individual material characteristics and features, such as adhesion, refractive index or fracture energy as a function of environmental stimuli, see for example Tokarev & Minko [2009] or Tanaka et al. [2005]. Recently, the mechanical stiffness and toughness of hydrogels have been strengthened up to a remarkable magnitude by means of double-crosslinked polymer chains, opening up new prospects for functional, load-bearing tissue replacements by overcoming diffusion induced damage mechanisms, see Naficy et al. [2011] and Gong [2010].

Constitutive Modeling of Gels. The most prominent and widely used point of departure to capture solute diffusion and volumetric change in polymeric bodies is the Flory-Huggins theory, due to its statistical consideration of polymer structures and the free energy evaluation based on diffusion-induced entropy increase, see Huggins [1942], Flory & Rehner Jr. [1943] or Flory [1950].

Its combination with a linear diffusion law due to Fick [1855] has been used throughout the literature of continuum theories for gels, e.g. by Hong et al. [2008], Duda et al. [2010] or Chester & Anand [2010], where the two latter base their theory upon a multiplicative split of the deformation gradient into chemical and mechanical part. Influences of the elastic properties of gels are discussed by Doi [2009]. Phase separation in gels have been modeled by Hong & Wang [2013] using a phase field method together with different formulations of interface models. For a general review of diffusion in solids undergoing large deformations, see Rajagopal [2003]. Numerous approaches have been presented to capture different effects specific to hydrogels in continuum-mechanical simulations. Zhang et al.

[2009] provide analytical as well as finite element solutions of the coupled problem, while simulations for the recent findings of Holmes et al. [2011] are provided by Lucantonio et al. [2013]. An incorporation of the temperature has been proposed by Chester & Anand [2011] and Chester et al. [2015], outlining both theoretical and finite element foundations. A further focus often lies on the modeling of swelling-induced mechanical instabilities, because in dependency on geometry and geometrical constraints of hydrogel specimen, the exposure to a fluid not only leads to large increase of volume but also to wave-like buckling patterns. Free surfaces expand due to the species influx and are bonded to unswollen inner parts of the gel at the same time. For a high osmotic pressure, this mechanism leads to buckling patterns, that have extensively been analyzed within experimental setups. Tanaka et al. [1987] found evolving instability patterns for gels in Petri-dishes and derived a critical osmotic pressure for the mechanism. Later, Tanaka & Sigehuzi [1994] monitored coarsening cusp formations of geometrically constrained hydrogel specimen. Buckling modes of locally wetted hydrogel beams and circular disks were observed by Holmes et al. [2011], while Mora & Boudaoud [2006] discovered wavelike patterns of rings with different radii, bounded to an inner rigid substrate. For computational modeling approaches, these phenomena may require distinct numerical handling of swelling-induced structural instabilities. In order to capture the mechanical instabilities, Zhang et al. [2009] employ a finite element implementation built upon the modified arc-length method. Within a similar framework, Liu et al. [2010] analyzed wavelike instability patterns and applied the theory to leaf growth.

Common to the modeling approaches mentioned above is the saddle point structure of the underlying theory. Thus, in finite element implementations, the LBB condition might be violated, leading to oscillatory distributions of the chemical potential as the primary variable controlling species diffusion. These challenges can be addressed in several ways, in Bouklas et al. [2015] an in-depth analysis of Taylor-Hood elements for gel modeling is provided, while Krischok & Linder [2016] proposed an enhancement of classical EAS elements for the coupled problem.

Novel Minimization Principle. The problem will be formulated as a three-field saddle point and a two-field minimization principle, which is in line with variational formulations for standard dissipative solids. It furthermore follows Miehe et al. [2015a], where coupled Biot-Darcy fluid transport with an additional fracture phase-field is embedded into a variational formulation, yielding a similar minimization formulation for the coupled problem. The evolution of the coupled problem can be described by construction of appropriate rate potentials, which are then transferred to their time-discrete incremental counterparts. A discretization in space results in potentials suitable for a finite element solution. This variational approach offers the minimization structure with its inherent symmetry as a new formulation for diffusion-deformation coupling in hydrogels.

It can be exploited in the numerical analysis by unconstrained finite element interpolations, as opposed to the saddle point formulation, where the ansatz spaces are constrained by the LBB condition.

5.2 Summary of Equations and Boundary Conditions

Modeling of fluid transport in elastic solids is treated as a three-field problem, characterized by the deformation field (2.3), the chemical potential (2.29) and the concentration field (2.24), modified by multiplication with a constant factor $\bar{v}$ that represents the *volume associated with one solute particle*. The convenience of such a modification is related to the constitute functions to be specified below and the fact that volumetric change is often the field of interest in swelling-dominated problems. The definition of a corresponding field should be accompanied by a new symbol to avoid confusion between the concentration of solute particles and the volume fraction occupied by them. The *swelling volume fraction*[20] is defined as

$$v : \begin{cases} \mathcal{B} \times T \to \mathcal{R}^+ \\ (\boldsymbol{X}, t) \mapsto v(\boldsymbol{X}, t) = \bar{v} c(\boldsymbol{X}, t). \end{cases} \tag{5.1}$$

As a consequence of (5.1), notation and physical meaning of fluid transport kinetics change. Associated spatial and material flux vectors are denoted by $\boldsymbol{j}_v$ and $\boldsymbol{J}_v$, respectively, out- or influxes per unit surface are h_v and H_v, the external supply terms R_v and r_v, and the spatial swelling volume fraction finally is v_s.

Now recall the local balances of fluid content (2.43), linear momentum (2.45) and angular momentum (2.47). For (2.43) multiplied by the constant $\bar{v}$ and $R_v = 0$ per assumption, the relevant local balance equations for quasi-static processes

[20] The term *swelling volume fraction* has previously been used in conjunction with a multiplicative decomposition of the deformation gradient into mechanical and chemical parts, see e.g. the approaches of Chester & Anand [2010] or Duda et al. [2010]. The primary field (5.1) is different from their field, as the present problem statement seeks a solution of (5.1) as a primary field, and there is no direct link to kinematics. Additionally, the range of the variables differ.

can be summarized together with necessary constitutive laws according to (3.37), tailored to downhill diffusion with elastic deformation.

$$
\begin{aligned}
&\text{1. Balance of solute volume} && \dot{v} = -\operatorname{Div}[\boldsymbol{J}_v] \\
&\text{2. Balance of linear momentum} && \operatorname{Div}[\boldsymbol{P}] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \boldsymbol{0} \\
&\text{3. Balance of angular momentum} && \boldsymbol{F}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{F}^T \\
&\text{4. Constituive stresses} && \boldsymbol{P} = \partial_{\boldsymbol{F}}\widehat{\psi}(\boldsymbol{F}, v) \\
&\text{5. Constitutive chemical potential} && \mu = \partial_v \widehat{\psi}(\boldsymbol{F}, v) \\
&\text{6. Constitutive volume flux} && \boldsymbol{J}_v = \partial_{-\nabla\mu}\widehat{\phi}^*(-\nabla\mu; \boldsymbol{C}, v)
\end{aligned}
\tag{5.2}
$$

Note that the balance equations $(5.2)_{1-3}$ are identical in both minimization and saddle point principle, with the only difference that $(5.2)_1$ will be a constraint equation for the minimization and an Euler equation of the saddle point formulation. The constitutive relations $(5.2)_{4,5}$ are valid for both principles, while $(5.2)_6$ will be in an inverse form for the minimization formulation. Note further that the constitutive free energy in $(5.2)_{4,5}$ depends on the deformation and not on the invariants of the right Cauchy-Green tensor as in (3.37) – the material response function $\widehat{\psi}$ is however assumed to be consistent with the constraints formulated in the Sections 3.4.1 and 3.4.2 and *can* be formulated in terms of $\boldsymbol{I}_C$, but are specified in terms of $\boldsymbol{F}$, which simplifies subsequent derivatives.

Boundary conditions have to be specified for enabling a solution of the coupled problem. The referential surface $\partial\mathcal{B}$ is decomposed into mechanical and chemical parts, which reads

$$
\partial\mathcal{B} = \partial\mathcal{B}^{\boldsymbol{\varphi}} \cup \partial\mathcal{B}^{\boldsymbol{T}} \quad \text{and} \quad \partial\mathcal{B} = \partial\mathcal{B}^{\mu} \cup \partial\mathcal{B}^{H_v}, \tag{5.3}
$$

with $\partial\mathcal{B}^{\boldsymbol{\varphi}} \cap \partial\mathcal{B}^{\boldsymbol{T}} = \emptyset$ and $\partial\mathcal{B}^{\mu} \cap \partial\mathcal{B}^{H_v} = \emptyset$. Dirichlet- and Neumann boundary conditions for deformation and stresses are considered,

$$
\boldsymbol{\varphi} = \bar{\boldsymbol{\varphi}}(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{\boldsymbol{\varphi}} \quad \text{and} \quad \boldsymbol{P}\boldsymbol{N} = \bar{\boldsymbol{T}}(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{\boldsymbol{T}}, \tag{5.4}
$$

and for chemical potential and volume transport

$$
\mu = \bar{\mu}(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{\mu} \quad \text{and} \quad \boldsymbol{J}_v \cdot \boldsymbol{N} = \bar{H}_c(\boldsymbol{X}, t) \text{ on } \partial\mathcal{B}^{H_v} \tag{5.5}
$$

with prescribed deformation $\bar{\boldsymbol{\varphi}}$ and traction $\bar{\boldsymbol{T}}$, chemical potential $\bar{\mu}$ and volume transport $\bar{H}_v$. $\bar{\boldsymbol{T}}$ is a spatial vector parameterized by material coordinates and related to the spatial traction vector $\boldsymbol{t}$ via $\boldsymbol{t}da = \boldsymbol{T}dA$. It represents a convenient quantity for the application of mechanical Neumann boundary conditions on the undeformed surface. The initial condition for the swelling volume fraction is

$$
v(\boldsymbol{X}, t_0) = v_0(\boldsymbol{X}) \text{ in } \mathcal{B}. \tag{5.6}
$$

5.3 Continuous and Time-Discrete Incremental Potentials

It is now shown that the set of equations (5.2) can be obtained as the Euler equations of variational principles. On the basis of Miehe et al. [2015a], a minimization and a saddle point principle are developed, first for the evolution problem in terms of rate quantities, then in a time-discrete way in order to obtain actual update equations per time step.

5.3.1 Stored Energy, Dissipation Potential and Load Functionals

In dependency of the primary fields $\boldsymbol{\varphi}$, v and μ, energy storage and dissipative transport mechanisms are characterized by the following functionals. The stored energy functional depends on the deformation field $\boldsymbol{\varphi}$ and the swelling volume fraction v

$$E(\boldsymbol{\varphi}, v) = \int_{\mathcal{B}} \widehat{\psi}(\nabla\boldsymbol{\varphi}, v) dV \tag{5.7}$$

and characterizes the energy stored in the species containing solid. The rate of the stored energy

$$\frac{d}{dt}E(\dot{\boldsymbol{\varphi}}, \dot{v}) = \int_{\mathcal{B}} \left\{ \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} + \partial_v\widehat{\psi}\dot{v} \right\} dV \tag{5.8}$$

is considered a functional of the rates $\dot{\boldsymbol{\varphi}}$ of deformation and $\dot{v}$ of volume fraction at a given state $\{\boldsymbol{\varphi}, v\}$. It can also be rewritten such that the dependency on $\dot{v}$ is replaced by its evolution $(5.2)_1$, injecting a dependency on the volume flux $\boldsymbol{J}_v$. Taking the balance of volume $(5.2)_1$ as a local constraint condition that determines the evolution of the swelling volume fraction, the canonical rate of energy functional is defined at given state $\{\boldsymbol{\varphi}, v\}$ as

$$\frac{d}{dt}E(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) = \int_{\mathcal{B}} \left\{ \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} - \partial_v\widehat{\psi}\,\mathrm{Div}\,[\boldsymbol{J}_v] \right\} dV. \tag{5.9}$$

Now consider a dissipation potential function that depends on the volumetric flux $\boldsymbol{J}_v$. It is related to the dual dissipation potential $\widehat{\phi}^*$ by the Legendre transformation

$$\widehat{\phi}(\boldsymbol{J}_v; \nabla\boldsymbol{\varphi}, v) = \sup_{-\nabla\mu} \left\{ -\nabla\mu \cdot \boldsymbol{J}_v - \widehat{\phi}^*(-\nabla\mu; \nabla\boldsymbol{\varphi}, v) \right\}. \tag{5.10}$$

At given state $\{\boldsymbol{\varphi}, v\}$, the dissipation potential functional is defined by

$$D(\boldsymbol{J}_v) = \int_{\mathcal{B}} \widehat{\phi}\,(\boldsymbol{J}_v; \nabla\boldsymbol{\varphi}, v)\, dV \tag{5.11}$$

and represents a functional of the volume flux vector $\boldsymbol{J}_v$. The external load functional decomposes into a mechanical and a volume swelling contribution

$$P_{\text{ext}}(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) = P^{\boldsymbol{\varphi}}_{\text{ext}}(\dot{\boldsymbol{\varphi}}) + P^{\mu}_{\text{ext}}(\boldsymbol{J}_v), \tag{5.12}$$

where the mechanical part reads

$$P^{\boldsymbol{\varphi}}_{\text{ext}}(\dot{\boldsymbol{\varphi}}) = \int_{\mathcal{B}} \boldsymbol{R}_{\boldsymbol{\varphi}} \cdot \dot{\boldsymbol{\varphi}} dV + \int_{\partial\mathcal{B}^{\boldsymbol{T}}} \bar{\boldsymbol{T}} \cdot \dot{\boldsymbol{\varphi}} dA \tag{5.13}$$

containing known body forces $\boldsymbol{R}_{\boldsymbol{\varphi}}$ and dead surface tractions $\bar{\boldsymbol{T}}$. The external volume swelling power has a single contribution due to volume transport across the surface, which can be pulled back to the reference configuration via (5.5), resulting in a functional of the volume flux

$$P^{\mu}_{\text{ext}}(\boldsymbol{J}_v) = -\int_{\partial\mathcal{B}^{\mu}} \bar{\mu} H_v dA = -\int_{\partial\mathcal{B}^{\mu}} \bar{\mu} \boldsymbol{J}_v \cdot \boldsymbol{N} dA. \tag{5.14}$$

Again, the prescribed chemical potential $\bar{\mu}$ is considered to be independent of the primary fields for simplicity.

5.3.2 Continuous Two-Field Minimization Principle

Based on the rate of energy, dissipation and load functionals defined in (5.9), (5.11) and (5.12), the rate potential

$$\Pi(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) = \frac{d}{dt} E(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) + D(\boldsymbol{J}_v) - P_{\text{ext}}(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) \tag{5.15}$$

is postulated at given state $\{\boldsymbol{\varphi}, v\}$. This potential can be split into internal and external contributions,

$$\Pi(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) = \int_{\mathcal{B}} \pi\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \text{Div}\left[\boldsymbol{J}_v\right]\right) dV - P_{\text{ext}}(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) \tag{5.16}$$

in terms of the internal rate potential density per unit volume

$$\pi\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \text{Div}\left[\boldsymbol{J}_v\right]\right) = \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} - \partial_v\widehat{\psi}\,\text{Div}\left[\boldsymbol{J}_v\right] + \widehat{\phi}(\boldsymbol{J}_v; \nabla\boldsymbol{\varphi}, v) \tag{5.17}$$

that contains the rate of the energy density and the dissipation potential function. Then, the evolution of the deformation and volume flux vector at a given state $\{\boldsymbol{\varphi}, v\}$ is determined by the two-field minimization statement

$$\{\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v\} = \text{Arg}\left\{\inf_{\dot{\boldsymbol{\varphi}}\in\mathcal{W}_{\dot{\boldsymbol{\varphi}}}} \inf_{\boldsymbol{J}_v\in\mathcal{W}_{\boldsymbol{J}_v}} \Pi(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v)\right\}. \tag{5.18}$$

Admissible spaces for the rates of deformation and volume flux are constructed based upon the decomposition of the solid surface (5.3), where $\partial\mathcal{B}^{\boldsymbol{\varphi}}$ and $\partial\mathcal{B}^{H_v}$ are the Dirichlet boundaries in the above variational principle, leading to

$$\mathcal{W}_{\dot{\boldsymbol{\varphi}}} = \left\{\dot{\boldsymbol{\varphi}} \in H^1(\mathcal{B}) \mid \dot{\boldsymbol{\varphi}} = \dot{\bar{\boldsymbol{\varphi}}} \text{ on } \partial\mathcal{B}^{\boldsymbol{\varphi}}\right\} \quad \text{and}$$

$$\mathcal{W}_{\boldsymbol{J}_v} = \left\{ \boldsymbol{J}_v \in H(\mathrm{Div}, \mathcal{B}) \mid \boldsymbol{J}_v \cdot \boldsymbol{N} = \bar{H}_v \text{ on } \partial\mathcal{B}^{H_v} \right\}. \tag{5.19}$$

Calculating the variation of the rate potential (5.16) for admissible virtual rates $\delta\dot{\boldsymbol{\varphi}}$ of deformation and $\delta\boldsymbol{J}_v$ of volume flow, satisfying $\delta\dot{\boldsymbol{\varphi}} = \boldsymbol{0}$ on $\partial\mathcal{B}^{\boldsymbol{\varphi}}$ and $\delta\boldsymbol{J}_v \cdot \boldsymbol{N} = 0$ on $\partial\mathcal{B}^{H_v}$, the following Euler equations are obtained.

1.	Balance of linear momentum	$\mathrm{Div}\left[\partial_{\boldsymbol{F}}\widehat{\psi}\right] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \boldsymbol{0}$	in $\mathcal{B}$
2.	Inverse Fick's law	$\nabla\left[\partial_v\widehat{\psi}\right] + \partial_{\boldsymbol{J}_v}\widehat{\phi} = \boldsymbol{0}$	in $\mathcal{B}$
3.	Prescribed tractions	$\left(\partial_{\boldsymbol{F}}\widehat{\psi}\right)\cdot\boldsymbol{N} - \bar{\boldsymbol{T}} = \boldsymbol{0}$	on $\partial\mathcal{B}^{\boldsymbol{T}}$
4.	Prescribed chemical potential	$\bar{\mu} - \partial_v\widehat{\psi} = 0$	on $\partial\mathcal{B}^{\mu}$

(5.20)

This set of Euler equations include the balance of linear momentum $(5.2)_{2,4}$ and an inverse form of Fick's law $(5.2)_{5,6}$ as well as Neumann boundary conditions. The balance of angular momentum $(5.2)_3$ is automatically satisfied for an objective energy storage function $\widehat{\psi}$. Note that the volumetric solute balance $(5.2)_1$ is incorporated as a constraint condition in the variational principle, i.e., the rate of the swelling volume fraction is determined by $(5.2)_1$ for a given flux $\boldsymbol{J}_v$ obtained as a solution of the two-field minimization principle.

5.3.3 Continuous Three-Field Saddle Point Principle

Now consider the dissipation potential function $\widehat{\phi}$ defined by the Legendre transformation (5.10) in terms of its dual function $\widehat{\phi}^*$. Insertion of (5.10) into (5.16) defines the three-field potential

$$\Pi^+\left(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, -\nabla\mu\right) = \int_{\mathcal{B}} \pi^+\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \mathrm{Div}\left[\boldsymbol{J}_v\right], -\nabla\mu\right) dV - P_{\mathrm{ext}}\left(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v\right) \tag{5.21}$$

that introduces a dependence on the vector field $-\nabla\mu$, which is dual to $\boldsymbol{J}_v$ and plays the role of a mixed variable. The mixed rate potential density in terms of $\boldsymbol{J}_v$ and $-\nabla\mu$ then reads

$$\begin{aligned}\pi^+\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \mathrm{Div}\left[\boldsymbol{J}_v\right], -\nabla\mu\right) = \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} - \partial_v\widehat{\psi}\,\mathrm{Div}\left[\boldsymbol{J}_v\right] \\ - \nabla\mu\cdot\boldsymbol{J}_v - \widehat{\phi}^*\left(-\nabla\mu; \nabla\boldsymbol{\varphi}, v\right).\end{aligned} \tag{5.22}$$

This potential can be reduced by replacing $\mathrm{Div}\left[\boldsymbol{J}_v\right]$ by the rate $\dot{v}$ via the elimination equation $(5.2)_1$. An integration by parts and use of the boundary condition $\mu = \bar{\mu}$ on $\partial\mathcal{B}^{\mu}$ results in the reduced mixed potential

$$\Pi^*\left(\dot{\boldsymbol{\varphi}}, \dot{v}, \mu\right) = \int_{\mathcal{B}} \pi^*\left(\nabla\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, -\nabla\mu\right) dV - P^*_{\mathrm{ext}}\left(\dot{\boldsymbol{\varphi}}, \mu\right) \tag{5.23}$$

with the potential density function

$$\pi^*\left(\nabla\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, -\nabla\mu\right) = \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} + \partial_v\widehat{\psi}\dot{v} - \mu\dot{v} - \widehat{\phi}^*(-\nabla\mu; \nabla\boldsymbol{\varphi}, v). \tag{5.24}$$

The modified load functional in (5.23) again decomposes into a mechanical and a chemical, volumetric contribution

$$P^{*}_{\text{ext}}\left(\dot{\boldsymbol{\varphi}}, \mu\right) = P^{\boldsymbol{\varphi}}_{\text{ext}}\left(\dot{\boldsymbol{\varphi}}\right) + P^{H_v}_{\text{ext}}(\mu) \tag{5.25}$$

where the first part is identical to (5.13) and the latter takes the form

$$P^{H_v}_{\text{ext}} = \int_{\partial \mathcal{B}^{H_v}} \mu \bar{H}_v \, dA, \tag{5.26}$$

dual to (5.14) and where $\bar{H}_v$ is a prescribed volume flow on the part $\partial \mathcal{B}^{H_v}$ of the undeformed surface.

Hence, at a given state $\{\boldsymbol{\varphi}, v\}$, the evolution of deformation and swelling volume fraction as well as the chemical potential are determined by the mixed *three-field saddle point principle*

$$\{\dot{\boldsymbol{\varphi}}, \dot{v}, \mu\} = \operatorname{Arg}\left\{ \inf_{\dot{\boldsymbol{\varphi}} \in \mathcal{W}_{\dot{\boldsymbol{\varphi}}}} \inf_{\dot{v} \in L^2} \sup_{\mu \in \mathcal{W}_{\mu}} \Pi^{*}\left(\dot{\boldsymbol{\varphi}}, \dot{v}, \mu\right) \right\}. \tag{5.27}$$

Recall the decomposition of the body surface (5.3), where now $\partial \mathcal{B}^{\boldsymbol{\varphi}}$ and $\partial \mathcal{B}^{\mu}$ are Dirichlet boundaries. The resulting admissible spaces for the rates of deformation and the chemical potential are given by

$$\mathcal{W}_{\dot{\boldsymbol{\varphi}}} = \left\{ \dot{\boldsymbol{\varphi}} \in H^1(\mathcal{B}) \mid \dot{\boldsymbol{\varphi}} = \dot{\bar{\boldsymbol{\varphi}}} \text{ on } \partial \mathcal{B}^{\boldsymbol{\varphi}} \right\} \quad \text{and}$$

$$\mathcal{W}_{\mu} = \left\{ \mu \in H^1(\mathcal{B}) \mid \mu = \bar{\mu} \text{ on } \partial \mathcal{B}^{\mu} \right\}. \tag{5.28}$$

The Euler equations are obtained by the variation of the rate potential (5.23) for admissible virtual rates $\delta\dot{\boldsymbol{\varphi}}$ of deformation and virtual chemical potential $\delta\mu$, while satisfying $\delta\dot{\boldsymbol{\varphi}} = \boldsymbol{0}$ on $\partial \mathcal{B}^{\boldsymbol{\varphi}}$ and $\delta\mu = 0$ on $\partial \mathcal{B}^{\mu}$. They read

1. Balance of linear momentum	$\operatorname{Div}\left[\partial_{\boldsymbol{F}} \widehat{\psi}\right] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \boldsymbol{0}$	in $\mathcal{B}$
2. Constitutive chemical potential	$\partial_v \widehat{\psi} - \mu = 0$	in $\mathcal{B}$
3. Balance of volume	$\dot{v} + \operatorname{Div}\left[\partial_{-\nabla\mu} \widehat{\phi}^{*}\right] = 0$	in $\mathcal{B}$
4. Prescribed tractions	$\left(\partial_{\boldsymbol{F}} \widehat{\psi}\right) \cdot \boldsymbol{N} - \bar{\boldsymbol{T}} = \boldsymbol{0}$	on $\partial \mathcal{B}^{\boldsymbol{T}}$
5. Prescribed volume flow	$\left(\partial_{-\nabla\mu} \widehat{\phi}^{*}\right) \cdot \boldsymbol{N} - \bar{H}_v = 0$	on $\partial \mathcal{B}^{H_v}$

(5.29)

This covers the stress equilibrium $(5.2)_{2,4}$, the constitutive chemical potential $(5.2)_5$ as well as the volumetric solute balance $(5.2)_{1,6}$, along with Neumann boundary conditions.

5.3.4 Incremental Two-Field Minimization Principle

Now consider a discrete time interval $[t_n, t]$ with step length $\tau = t - t_n > 0$, and assume all field variables at time t_n to be known. As in Section 4.2, the construction of time-discrete potentials will determine the fields at current discrete time (again, variables without subscript). For simplicity, the following outline focuses on Dirichlet problems with $P_{\text{ext}} = P^*_{\text{ext}} = 0$ in (5.15) and (5.23).

Time-Discrete Minimization Principle. A time discrete counterpart to (5.16) is the incremental potential in $[t_n, t]$

$$\Pi^\tau(\boldsymbol{\varphi}, \boldsymbol{J}_v) = \int_{\mathcal{B}} \pi^\tau \left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \operatorname{Div}[\boldsymbol{J}_v]\right) dV \tag{5.30}$$

for pure Dirichlet problems, in terms of fields at the discrete time t. π^τ is an incremental potential density. It is related to the rate potential density (5.17) by

$$\pi^\tau(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \operatorname{Div}[\boldsymbol{J}_v]) = \mathsf{Algo}\left\{\int_{t_n}^{t} \pi\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \operatorname{Div}[\boldsymbol{J}_v]\right) dt\right\}. \tag{5.31}$$

Algo represents a time integration that also covers the integration of the elimination equation $(5.2)_1$. Here, the implicit update of $(5.2)_1$ is chosen as

$$v = v_n - \tau \operatorname{Div}[\boldsymbol{J}_v], \tag{5.32}$$

that leads to the closed form of the incremental potential density together with an implicit integration of (5.31), which reads

$$\pi^\tau\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \operatorname{Div}[\boldsymbol{J}_v]\right) = \widehat{\psi}\left(\nabla\boldsymbol{\varphi}, v_n - \tau \operatorname{Div}[\boldsymbol{J}_v]\right) + \tau\widehat{\phi}(\boldsymbol{J}_v; \nabla\boldsymbol{\varphi}_n, v_n). \tag{5.33}$$

To ensure the variational consistency of the incremental formulation, the dissipation function is evaluated in terms of the current volume flux $\boldsymbol{J}_v$ at time t and at a frozen state $\{\boldsymbol{\varphi}_n, v_n\}$ at time t_n. The incremental two-field minimization principle reads

$$\{\boldsymbol{\varphi}, \boldsymbol{J}_v\} = \operatorname{Arg}\left\{\inf_{\boldsymbol{\varphi}\in\mathcal{W}_{\boldsymbol{\varphi}}} \inf_{\boldsymbol{J}_v\in\mathcal{W}_{\boldsymbol{J}_v}} \Pi^\tau(\boldsymbol{\varphi}, \boldsymbol{J}_v)\right\}, \tag{5.34}$$

where the admissible function spaces for the time discrete fields are

$$\mathcal{W}_{\boldsymbol{\varphi}} = \left\{\boldsymbol{\varphi} \in H^1(\mathcal{B}) \mid \boldsymbol{\varphi} = \bar{\boldsymbol{\varphi}} \text{ on } \partial\mathcal{B}^{\boldsymbol{\varphi}}\right\} \quad \text{and} \tag{5.35}$$

$$\mathcal{W}_{\boldsymbol{J}_v} = \left\{\boldsymbol{J}_v \in H(\operatorname{Div}, \mathcal{B}) \mid \boldsymbol{J}_v \cdot \boldsymbol{N} = \bar{H}_v \text{ on } \partial\mathcal{B}^{H_v}\right\}. \tag{5.36}$$

By introduction of the array $\mathfrak{Q} = [\nabla\boldsymbol{\varphi}, \mathrm{Div}\,[\boldsymbol{J}_v]\,, \boldsymbol{J}_v]^T$, the first derivative of π^τ defines a generalized stress array

$$\mathfrak{S} = \partial_{\mathfrak{Q}}\pi^\tau(\mathfrak{Q}) = \begin{bmatrix} \partial_{\boldsymbol{F}}\widehat{\psi} \\ -\tau\partial_v\widehat{\psi} \\ \tau\partial_{\boldsymbol{J}_v}\widehat{\phi} \end{bmatrix} \tag{5.37}$$

and the second derivative a symmetric generalized tangent moduli array

$$\mathfrak{C} = \partial^2_{\mathfrak{Q}\mathfrak{Q}}\pi^\tau(\mathfrak{Q}) = \begin{bmatrix} \partial^2_{\boldsymbol{F}\boldsymbol{F}}\widehat{\psi} & -\tau\partial^2_{\boldsymbol{F}v}\widehat{\psi} & \cdot \\ -\tau\partial^2_{v\boldsymbol{F}}\widehat{\psi} & \tau^2\partial^2_{vv}\widehat{\psi} & \cdot \\ \cdot & \cdot & \tau\partial^2_{\boldsymbol{J}_v\boldsymbol{J}_v}\widehat{\phi} \end{bmatrix}. \tag{5.38}$$

These two arrays serve as an input for the finite element discretization.

Space-Time-Discrete Minimization Principle. Associated with the finite element triangulation $\mathfrak{T}^h$, the interpolations of the local state variables $\mathfrak{Q} = \underline{\boldsymbol{B}}(\boldsymbol{X})\underline{\boldsymbol{d}}$ in terms of the global nodal vector $\underline{\boldsymbol{d}}$ comprising nodal deformation and flux. $\underline{\boldsymbol{B}}$ denotes the finite element interpolation matrix. The spatial discretization of the incremental two-field potential (5.30) then reads

$$\Pi^{\tau h}(\underline{\boldsymbol{d}}) = \int_{\mathcal{B}} \pi^\tau\,(\underline{\boldsymbol{B}}\,\underline{\boldsymbol{d}})\,dV \tag{5.39}$$

for Dirichlet problems with $P_{\mathrm{ext}} = 0$. The algebraic minimization principle

$$\underline{\boldsymbol{d}} = \mathrm{Arg}\left\{ \inf_{\underline{\boldsymbol{d}}\in\mathcal{R}^d} \Pi^{\tau h}(\underline{\boldsymbol{d}}) \right\} \tag{5.40}$$

is the counter of (5.34), discrete in space and time, and determines the nodal state vector $\underline{\boldsymbol{d}}$ of the finite element mesh at time t. The necessary condition of this discrete minimization statement reads

$$\underline{\boldsymbol{R}} = \Pi^{\tau h}_{,\underline{\boldsymbol{d}}} = \int_{\mathcal{B}} \underline{\boldsymbol{B}}^T\mathfrak{S}dV = \boldsymbol{0} \tag{5.41}$$

A Newton-Raphson iteration of this system of equations updates the state vector by

$$\underline{\boldsymbol{d}} \Leftarrow \underline{\boldsymbol{d}} - \underline{\boldsymbol{K}}^{-1}\underline{\boldsymbol{R}} \quad \text{with} \quad \underline{\boldsymbol{K}} = \Pi^{\tau h}_{,\underline{\boldsymbol{d}}\underline{\boldsymbol{d}}} = \int_{\mathcal{B}} \underline{\boldsymbol{B}}^T\mathfrak{C}\underline{\boldsymbol{B}}dV, \tag{5.42}$$

where $\underline{\boldsymbol{K}}$ is the symmetric tangent matrix governed by the tangent moduli (5.38). Updates are performed until an absolute or relative convergence is reached. The symmetry property can be exploited by using solvers for the symmetric, positive definite problems. Concerning the stability of the finite element solution no particular restrictions to the conforming discrete finite element spaces are needed.

This renders a major advantage when compared to the classical saddle point formulation discussed below.

5.3.5 Incremental Three-Field Saddle Point Principle

Incremental Saddle Point Principle. A time-discrete counterpart to (5.23) is the *mixed incremental potential* in the time step $[t_n, t]$

$$\Pi^{*\tau}(\boldsymbol{\varphi}, v, \mu) = \int_{\mathcal{B}} \pi^{*\tau}(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu)\, dV \tag{5.43}$$

for pure Dirichlet problems and in terms of fields evaluated at the discrete time t. Here, π^{τ} is a mixed incremental potential density per unit volume. By analogy with the minimization formulation, it is related to the rate potential density (5.24) by

$$\pi^{*\tau}(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu) = \mathsf{Algo}\left\{ \int_{t_n}^{t} \pi^{*}(\nabla\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \nabla\mu)\, dt \right\}, \tag{5.44}$$

again governed by the energy function and the dual dissipation potential. An implicit time integration of (5.44) is used, which yields the closed form of the mixed incremental potential density

$$\pi^{*\tau}(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu) = \widehat{\psi}(\nabla\boldsymbol{\varphi}, v) - \mu(v - v_n) - \tau\widehat{\phi}^{*}(-\nabla\mu; \nabla\boldsymbol{\varphi}_n, v_n). \tag{5.45}$$

Note again that with regard to a variational consistency of the incremental setting, the dissipation function is partly evaluated for values at time t_n to not violate the variational formalism. The incremental three-field saddle point principle reads

$$\{\boldsymbol{\varphi}, v, \mu\} = \operatorname{Arg}\left\{ \inf_{\boldsymbol{\varphi}\in\mathcal{W}_{\boldsymbol{\varphi}}} \inf_{v\in L^2} \sup_{\mu\in\mathcal{W}_{\mu}} \Pi^{*\tau}(\boldsymbol{\varphi}, v, \mu) \right\}, \tag{5.46}$$

where admissible function spaces for the time discrete fields are

$$\mathcal{W}_{\boldsymbol{\varphi}} = \left\{ \boldsymbol{\varphi} \in H^1(\mathcal{B}) \mid \boldsymbol{\varphi} = \bar{\boldsymbol{\varphi}} \text{ on } \partial\mathcal{B}^{\boldsymbol{\varphi}} \right\} \quad \text{and} \tag{5.47}$$

$$\mathcal{W}_{\mu} = \left\{ \mu \in H^1(\mathcal{B}) \mid \mu = \bar{\mu} \text{ on } \partial\mathcal{B}^{\mu} \right\}. \tag{5.48}$$

Condensation of the Volume Fraction. The saddle point statement (5.46) can be solved in two steps: the swelling volume fraction is locally computed for given deformation and chemical potential by the minimization problem

$$\mathsf{local}: \quad \{v\} = \operatorname{Arg}\left\{ \inf_{v\in L^2} \pi^{*\tau}(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu) \right\}. \tag{5.49}$$

with the nonlinear local equation as a necessary local condition for (5.49),

$$\pi^{*\tau}_{,v} = \partial_v \widehat{\psi}(\boldsymbol{F}, v) - \mu = 0, \tag{5.50}$$

that can be solved for the swelling volume fraction by means of a Newton-Raphson scheme,

$$v \Leftarrow v - k_v^{-1} \pi^{*\tau}_{,v} \quad \text{with} \quad k_v = \pi^{*\tau}_{,vv} = \partial^2_{vv} \widehat{\psi}(\boldsymbol{F}, v) \tag{5.51}$$

until convergence is reached.

Reduced Saddle Point Principle. The above local minimization (5.49) induces the *reduced incremental potential density*

$$\pi^{*\tau}_{\text{red}}(\nabla\boldsymbol{\varphi}, \mu, \nabla\mu) = \inf_{v \in L^2} \pi^{*\tau}(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu) \tag{5.52}$$

that depends only on the deformation and the chemical potential. After the condensation of the swelling volume fraction, the global incremental potential in (5.43) boils down to the reduced two-field potential

$$\Pi^{*\tau}_{\text{red}}(\boldsymbol{\varphi}, \mu) = \int_{\mathcal{B}} \pi^{*\tau}_{\text{red}}\left(\nabla\boldsymbol{\varphi}, \mu, \nabla\mu\right) dV \tag{5.53}$$

for pure Dirichlet problems. With this reduced potential at hand, the second solution step is based on the saddle point problem

$$\text{global}: \quad \{\boldsymbol{\varphi}, \mu\} = \text{Arg}\left\{ \inf_{\boldsymbol{\varphi} \in \mathcal{W}_{\boldsymbol{\varphi}}} \sup_{\mu \in \mathcal{W}_\mu} \Pi^{*\tau}_{\text{red}}(\boldsymbol{\varphi}, \mu) \right\} \tag{5.54}$$

that can be solved for the deformation and the chemical potential. By analogy with the time-discrete minimization formulation, the array $\mathfrak{Q}^* = [\nabla\boldsymbol{\varphi}, \mu, -\nabla\mu]^T$ can be introduced, and the first derivative of $\pi^{*\tau}_{\text{red}}$, and a generalized stress array be defined by

$$\mathfrak{S}^* = \partial_{\mathfrak{Q}^*} \pi^{*\tau}_{\text{red}}(\mathfrak{Q}^*) = \begin{bmatrix} \partial_{\boldsymbol{F}} \widehat{\psi} \\ -(v - v_n) \\ -\tau \partial_{-\nabla\mu} \widehat{\phi}^* \end{bmatrix}. \tag{5.55}$$

The second derivative yields the symmetric generalized tangent moduli array

$$\mathfrak{C}^* = \partial^2_{\mathfrak{Q}^*\mathfrak{Q}^*} \pi^{*\tau}_{\text{red}}(\mathfrak{Q}^*) = \begin{bmatrix} \partial^2_{\boldsymbol{F}\boldsymbol{F}} \widehat{\psi} - \partial^2_{\boldsymbol{F}v} \widehat{\psi} k_v^{-1} \partial^2_{v\boldsymbol{F}} \widehat{\psi} & \partial^2_{\boldsymbol{F}v} \widehat{\psi} k_v^{-1} & \cdot \\ k_v^{-1} \partial^2_{v\boldsymbol{F}} \widehat{\psi} & -k_v^{-1} & \cdot \\ \cdot & \cdot & -\tau \partial^2_{-\nabla\mu - \nabla\mu} \widehat{\phi}^* \end{bmatrix} \tag{5.56}$$

with the tangent $k_v = \partial_{vv}\widehat{\psi}$ of the converged local Newton iteration (5.51). These two arrays are the input for a spatial finite element discretization.

Space-Time-Discrete Saddle Point Principle. Associated with a typical finite element triangulation $\mathfrak{T}^h$, the interpolation of the state variables read $\mathfrak{Q}^* = \underline{\boldsymbol{B}}^*(\boldsymbol{X})\underline{\boldsymbol{d}}^*$ in terms of the global nodal state vector $\underline{\boldsymbol{d}}^*$, which contains nodal values of deformation map and chemical potential. $\underline{\boldsymbol{B}}^*$ is a typical interpolation matrix. The spatial discretization of the incremental two-field functional (5.53) reads

$$\Pi^{*\tau h}_{\text{red}}(\underline{\boldsymbol{d}}^*) = \int_{\mathcal{B}} \pi^{*\tau}_{\text{red}}\,(\underline{\boldsymbol{B}}^*\,\underline{\boldsymbol{d}}^*)\,dV \tag{5.57}$$

for the Dirichlet problem with $P^*_{\text{ext}} = 0$. Then, the saddle point statement

$$\underline{\boldsymbol{d}}^* = \text{Arg}\left\{ \underset{\underline{\boldsymbol{d}}^* \in \mathcal{R}^{d^*}}{\text{stat}}\ \Pi^{*\tau h}_{\text{red}}(\underline{\boldsymbol{d}}^*) \right\} \tag{5.58}$$

provides the spatial discretization of (5.54) and determines the nodal state vector of the finite element mesh at the current time. The necessary condition of this variational principle reads

$$\underline{\boldsymbol{R}}^* = \Pi^{*\tau h}_{red,\underline{\boldsymbol{d}}^*} = \int_{\mathcal{B}} \underline{\boldsymbol{B}}^{*T}\boldsymbol{\mathcal{S}}^* dV = \boldsymbol{0} \tag{5.59}$$

in terms of the generalized stress array (5.55). A standard Newton iteration to solve this system of equations updates the state vector by

$$\underline{\boldsymbol{d}}^* \Leftarrow \underline{\boldsymbol{d}}^* - \underline{\boldsymbol{K}}^{*-1}\underline{\boldsymbol{R}}^* \quad \text{with} \quad \underline{\boldsymbol{K}}^* = \Pi^{*\tau h}_{red,\underline{\boldsymbol{d}}^*\,\underline{\boldsymbol{d}}^*} = \int_{\mathcal{B}} \underline{\boldsymbol{B}}^{*T}\mathfrak{C}^*\underline{\boldsymbol{B}}^* dV \tag{5.60}$$

in terms of the tangent matrix $\underline{\boldsymbol{K}}^*$ governed by the generalized tangent moduli (5.56). Updates are performed until convergence reached, controlled by an absolute or a relative criterion. The symmetry of the tangent $\underline{\boldsymbol{K}}^*$ is a consequence of the reduced two-field variational principle. This symmetry property can be exploited by using solvers for the symmetric, indefinite problem $(5.60)_1$.

A stable finite element design for the saddle point problem must satisfy the LBB condition, see Babuška [1971], Babuška [1973] and Brezzi & Fortin [1991]. A simple condition for the stability, necessary though not sufficient, demands the relationship $n_{\boldsymbol{\varphi}} \geq n_\mu$ between the number of nodal unknowns $n_{\boldsymbol{\varphi}}$ for the deformation and n_μ for the chemical potential. Stable elements in this regard are the Taylor-Hood element applied by Bouklas et al. [2015] or Sandhu & Wilson [1969], the Mini element of Arnold et al. [1984], or the enhanced element designs considered in Papastavrou et al. [1997], Zhou et al. [2007] or Krischok & Linder [2016]. Technically, the implementation of Mini and Taylor-Hood elements is realized by a partition of the finite element interpolation with different polynomial degrees for the deformation and the chemical potential.

5.4 Isotropic Constitutive Model and FE Implementation

Consistent with the constraints in Sections 3.4.2 and 3.4.1, a simple isotropic model for a homogeneous response obtained by the following choices of the two constitutive functions $\widehat{\psi}$ and $\widehat{\phi}$. The dissipation potential function may assume two shapes, one for the saddle point and one for the minimization formulation.

5.4.1 Constitutive Free Energy and Dissipation Potential

The energy storage function depends on the deformation gradient and the swelling volume fraction and is assumed to have contributions from the bulk elasticity stemming from the cross-linked network structure, mixing of solvent molecules with the polymer network and a penalty term that enforces a volumetric constraint

$$\widehat{\psi}(\boldsymbol{F}, v) = \widehat{\psi}_{\mathrm{el}}(\boldsymbol{F}) + \widehat{\psi}_{\mathrm{che}}(v) + \widehat{\psi}_{\mathrm{con}}(\det \boldsymbol{F}, v). \tag{5.61}$$

For the elastic term, a simple Neo-Hookean type free energy is chosen,

$$\widehat{\psi}_{\mathrm{el}}(\boldsymbol{F}) = \frac{\gamma}{2}\left[\boldsymbol{F} : \boldsymbol{F} - 3 - 2\ln(\det \boldsymbol{F})\right], \tag{5.62}$$

in terms of the shear modulus $\gamma = NkT$, where N is the number of monomers in a prototype polymer chain, k represents the Boltzmann constant and T is the absolute temperature. The chemical contribution is of the Flory-Rehner type, see Flory & Rehner Jr. [1943]

$$\widehat{\psi}_{\mathrm{che}}(v) = \alpha\left[v\ln(\frac{v}{1+v}) + \frac{\chi v}{1+v}\right], \tag{5.63}$$

where $\alpha = kT/v$ is introduced as a mixing modulus and χ is a dimensionless parameter called the Flory-Huggins interaction parameter. Note that v is always grater than zero, but without an upper limit, see Figure 5.1. Finally, the penalty term of the free energy takes the quadratic form

$$\widehat{\psi}_{\mathrm{con}}(\det \boldsymbol{F}, v) = \frac{\epsilon}{2}\left[\det \boldsymbol{F} - 1 - v\right]^2 \tag{5.64}$$

and enforces the incompressibility of both the polymer network and the solvent. In this definition, ϵ plays the role of the penalty parameter. Within the minimization formulation, the dissipation potential is coupled to the flux vector $\boldsymbol{J}_v$ and constituted as

$$\widehat{\phi}(\boldsymbol{J}_v; \boldsymbol{C}_n, v_n) = \frac{1}{2Mv_n}\boldsymbol{C}_n : (\boldsymbol{J}_v \otimes \boldsymbol{J}_v), \tag{5.65}$$

at a given state characterized by the metric $\boldsymbol{C}_n$ and the swelling volume fraction v_n at time t_n, and where $M = vD/kT$ is the volumetric mobility. The dissipation

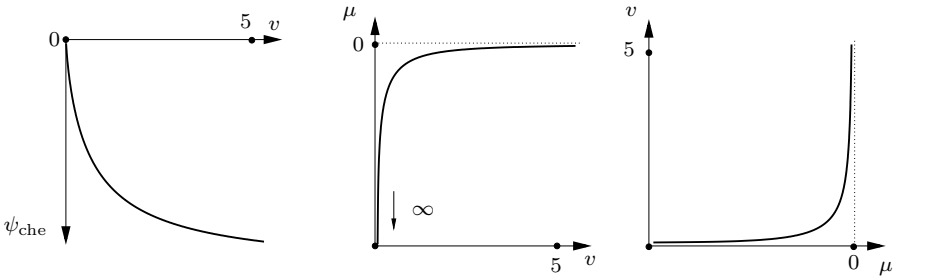


Figure 5.1: The stored chemical energy $\widehat{\psi}_{\text{che}}(v) = \alpha[v\ln(v/(1+v))+(\chi v)/(1+v)]$ and the derivative $\mu = \partial_v \widehat{\psi}_{\text{che}}(v)$ together with its inverse $v = \partial_v \widehat{\psi}^{-1}_{\text{che}}(\mu)$ plotted for $\alpha = 24.2\,\text{N/mm}^2$ and $\chi = 0.2$. The functions are depicted in the range $v \in [0, 5]$, which is a realistic range for the swelling volume fraction.

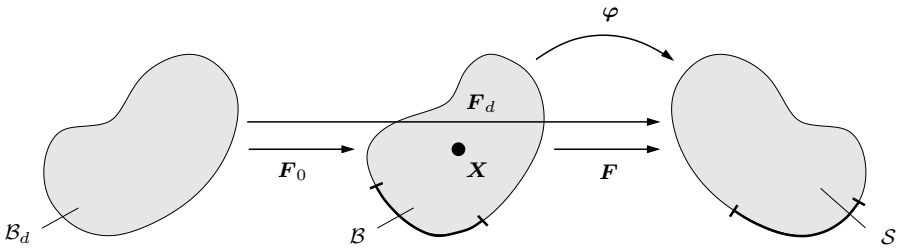


Figure 5.2: Dry reference state $\mathcal{B}_d$ with $v = 0$, the stress-free pre-swollen state $\mathcal{B}$ with $v = v_0$ serving as the reference configuration and the current configuration $\mathcal{S}$. The pre-swelling is characterized by $\boldsymbol{F}_0 = \lambda_0 \boldsymbol{1} = J_0^{1/3} \boldsymbol{1}$.

potential for the saddle point formulation depends on the negative gradient of the chemical potential and is assumed to have the form

$$\widehat{\phi}^*\left(-\nabla\mu; \boldsymbol{C}_n, v_n\right) = \frac{M v_n}{2} \boldsymbol{C}_n^{-1} : (-\nabla\mu \otimes -\nabla\mu)\,, \tag{5.66}$$

again at a given state at time t_n.

5.4.2 The Pre-Swollen, Stabilizing Reference Configuration

A dry polymer network is characterized by a swelling volume fraction $v = 0$. While the state itself represents a reasonable point of departure for boundary value problems, the chemical part of the free energy function ψ_{che} is undefined for zero swelling volume fraction, see (5.63). While its limit for v approaching zero is in fact zero, the chemical potential tends to negative infinity, giving rise to numerical difficulties when a dry reference configuration is in use ($\mathcal{B}_d$ in Figure 5.2 and marked with subscript d in the subsequent notation). To circumvent those, the reference configuration $\mathcal{B}$ is assumed to be pre-swollen and stress-free, as it

has been introduced by Hong et al. [2009] and since then established as common practice in the numerical modeling of hydrogels. A uniform increase of volume $\boldsymbol{F}_0 = \lambda_0 \boldsymbol{1} = J_0^{1/3} \boldsymbol{1}$ is supposed along with $\partial_{\boldsymbol{F}} \psi_d(\boldsymbol{F}_0) = \boldsymbol{0}$ for this pre-swollen state. The latter can be used to obtain an expression for the initial swelling volume fraction v_0,

$$v_0 = \frac{\gamma}{\epsilon} \left(J_0^{-1/3} - \frac{1}{J_0} \right) + J_0 - 1. \tag{5.67}$$

Formulating the deformation gradient of the dry configuration in terms of a multiplicative decomposition $\boldsymbol{F}_d = \boldsymbol{F}\boldsymbol{F}_0$, comparing pre-swollen and dry state and demanding equality of energetic states by $E_d = E$ (as traction-free pre-swelling is assumed), one obtains

$$\int_{\mathcal{B}_d} \widehat{\psi}_d \left(\boldsymbol{F}_d, v \right) dV_d = \int_{\mathcal{B}} \widehat{\psi} \left(\boldsymbol{F}, v \right) dV, \tag{5.68}$$

where the free energy density $\widehat{\psi}(\boldsymbol{F}, v)$ refers to the pre-swollen state and $\widehat{\psi}_d(\boldsymbol{F}_d, v)$ to the dry state. As a consequence of (5.68), they are related to each other via

$$\begin{aligned} \widehat{\psi}(\boldsymbol{F}, v) &= J_0^{-1} \widehat{\psi}_d(\boldsymbol{F}, v)|_{\boldsymbol{F}=\boldsymbol{F}_d=\boldsymbol{F}\boldsymbol{F}_0} \\ &= \frac{\gamma}{2J_0} [J_0^{2/3} \boldsymbol{F} : \boldsymbol{F} - 3 - 2\ln(JJ_0)] \\ &+ \frac{\epsilon}{2J_0} [JJ_0 - 1 - v]^2 + \frac{\alpha}{J_0} \left[v \ln(\frac{v}{1+v}) + \frac{\chi v}{1+v} \right], \end{aligned} \tag{5.69}$$

The free energy function (5.69) is used to obtain the value of the chemical potential at the pre-swollen state

$$\mu_0 = -\frac{\epsilon}{J_0} \left[J_0 - 1 - v_0 \right] + \frac{\alpha}{J_0} \left[\ln(\frac{v_0}{1+v_0}) + \frac{1}{1+v_0} + \frac{\chi}{(1+v_0)^2} \right]. \tag{5.70}$$

Note carefully that fields like deformation map, chemical potential, swelling volume fraction or flux are now functions of material coordinates $\boldsymbol{X}$ of the pre-swollen reference configuration. Note further that for potential terms related to dissipative volume flux, no similar distinction has been made at first sight due to the homogeneous and thus flux-free pre-swelling. The volume flux is intrinsically related to the pre-swollen state, as it is defined by a pull-back of its spatial counterpart, see (2.32). The incremental potential densities for minimization and saddle point principle thus do not change except for the energetic contribution of (5.69), making pre-swelling identical for both saddle point and minimization formulations. The chosen mechanism enables stable convergence rates for a low volumetric pre-swelling J_0 acting as stabilizing parameter for the numerical solution as well as for high values of J_0, necessary for boundary value problems starting with a fully saturated hydrogel body. In both cases, local values of v and v_n have to be initialized to v_0, and within the saddle point problem, nodal quantities of the chemical potential to μ_0, see Table 5.1.

With these modifications, the generalized stress (5.37) array for minimization formulation reads

$$\mathfrak{S} = \begin{bmatrix} \frac{\gamma}{J_0}\left[J_0^{2/3}\boldsymbol{F} - \boldsymbol{F}^{-T}\right] + \epsilon J\left[JJ_0 - 1 - v\right]\boldsymbol{F}^{-T} \\ \tau\left(\epsilon\left[J - \frac{1+v}{J_0}\right] - \frac{\alpha}{J_0}\left[\ln\left(\frac{v}{1+v}\right) + \frac{1}{v+1} + \frac{\chi}{(1+v)^2}\right]\right) \\ \frac{\tau}{Mv_n}\boldsymbol{C}_n\boldsymbol{J}_v \end{bmatrix} \tag{5.71}$$

and does not depend on further reduction by condensation of local variables. The generalized tangent arrays are determined by second derivatives of the incremental potential density. In contrast, the reduced generalized stress array for the saddle point problem is decomposed into a local equation for given deformation $\boldsymbol{F}$ and chemical potential μ, as outlined in Section 5.3.5,

$$\frac{\alpha}{J_0}\left[\ln(\frac{v}{1+v}) + \frac{1}{1+v} + \frac{\chi}{(1+v)^2}\right] - \frac{\epsilon}{J_0}(JJ_0 - 1 - v) - \mu = 0, \tag{5.72}$$

and the reduced generalized stress array (5.55) reads

$$\mathfrak{S}^*_{\text{red}} = \begin{bmatrix} \frac{\gamma}{J_0}\left[J_0^{2/3}\boldsymbol{F} - \boldsymbol{F}^{-T}\right] + \epsilon J\left[JJ_0 - 1 - v\right]\boldsymbol{F}^{-T} \\ -(v - v_n) \\ \tau v_n M\boldsymbol{C}_n^{-1} - \nabla\mu \end{bmatrix}. \tag{5.73}$$

Again, the generalized reduced tangent arrays are constructed via second derivatives as pointed out before.

5.4.3 Conforming and Non-conforming FE Design for the Minimizer

Conforming Raviart-Thomas Elements. A consequence of the non-standard variational structure governing the elastic material response coupled to volumetric species diffusion are restrictions for the choice of finite element spaces. In order to satisfy these conditions, $H(\text{Div}, \mathcal{B})$-conforming finite elements have to be chosen, first constructed by Raviart & Thomas [1977] (RT_k element) or Brezzi et al. [1985] (BDM_k element). In this outline, lowest order quadrilateral RT element ($k = 0$) are considered based upon constant normal fluxes

$$H_v^e = \int_{\partial\mathcal{B}^e} \boldsymbol{J}_v^h \cdot \boldsymbol{N} dA \tag{5.74}$$

across element edges as degrees of freedom. The conformity of this element is ensured by the continuity of H_v^e over the element boundaries. The approximation of the flux $\boldsymbol{J}_v$ on a quadrilateral reference element $\mathcal{A} = [-1, 1] \times [-1, 1]$ reads

$$\boldsymbol{J}_v^h = \begin{bmatrix} a \\ c \end{bmatrix} + \operatorname{diag}[\boldsymbol{\xi}] \cdot \begin{bmatrix} b \\ d \end{bmatrix} \quad \text{with} \quad a, b, c, d \in \mathcal{R} \quad \text{and} \quad \boldsymbol{\xi} \in \mathcal{A}, \tag{5.75}$$

leading to the normal trace $\boldsymbol{J}_v^h \cdot \boldsymbol{N}$ that is constant on each edge of the element. The evaluation of the normal flux H_v^e on the boundary $\partial\mathcal{A}$ of the reference element via (5.74) allows reformulating (5.75) in terms of vectorial shape functions $\boldsymbol{N}^K$ and fluxes H_{vK}^e across the edges K of the boundary of the reference element,

$$\boldsymbol{J}_v^h(\boldsymbol{\xi}) = \sum_{K=1}^{4} \boldsymbol{N}^K(\boldsymbol{\xi})\, H_{vK}^e, \tag{5.76}$$

where the vectorial shape functions $\boldsymbol{N}^K$ take the specific form

$$\boldsymbol{N}^1 = \begin{bmatrix} 0 \\ \frac{1}{4}(\xi_2 - 1) \end{bmatrix}, \quad \boldsymbol{N}^2 = \begin{bmatrix} \frac{1}{4}(\xi_1 + 1) \\ 0 \end{bmatrix}, \tag{5.77}$$

$$\boldsymbol{N}^3 = \begin{bmatrix} 0 \\ \frac{1}{4}(\xi_2 - 1) \end{bmatrix} \quad \text{and} \quad \boldsymbol{N}^4 = \begin{bmatrix} \frac{1}{4}(\xi_1 - 1) \\ 0 \end{bmatrix}. \tag{5.78}$$

Care must be taken when transforming the vectorial quantities from parameter to physical space and back, as vector components normal to an element edge shall keep its properties in both spaces. The contravariant Piola transformation $\mathcal{P}\colon \mathcal{A} \to \mathcal{B}^e$ is well suited for this purpose, since it does not change the component of a vector normal to the element edge up to a constant built by lengths of the same element edges in parameter and physical space. Thus, vectorial shape functions $\boldsymbol{N}^K$ defined in (5.78) are mapped to physical space by

$$\mathcal{P}[\cdot] = \frac{1}{\widehat{J}(\boldsymbol{\xi})} \widehat{\boldsymbol{J}}(\boldsymbol{\xi})\, [\cdot](\boldsymbol{\xi}), \tag{5.79}$$

where $\widehat{\boldsymbol{J}}$ denotes the classical Jacobian matrix of the invertible mapping from $\mathcal{A}$ to $\mathcal{B}^e$ (isoparametric in this case) and $\widehat{J} = \det[\widehat{\boldsymbol{J}}]$.

In addition to the correct mapping between spaces, the positive flux orientation over element edges needs to be defined by an appropriate decision criterion to obtain globally consistent shape functions. Following Anjam & Valdman [2015], a positive flux is evaluated as follows: fix a direction to check the numbering of global nodes per element (clockwise or counter clockwise), and select the positive flux direction on the edge by checking if a higher or lower numbered node follows the first edge node. This ensures the consistency of the flux direction for two neighboring elements. Therefore, a shape function $\boldsymbol{N}^K$ associated with that

edge changes to negative sign for one of the two elements sharing the same edge K.

Non-Conforming $\mathbf{Q_1Q_1}$ Elements with Reduced Integration. Interpolation of the volume flux $\boldsymbol{J}_v$ by bilinear shape functions yields a non-conforming approximation of the variational space $H(\mathrm{Div}, \mathcal{B})$. Yet it provides an ansatz easy to implement due to the uniform treatment of primary variables (not separately handled as edge and nodal degrees of freedom). Hence, the interpolation of the volume flux vector $\boldsymbol{J}_v$ is given by

$$\boldsymbol{J}_v^h(\boldsymbol{\xi}) = \sum_{i=1}^{4} N^i(\boldsymbol{\xi}) \boldsymbol{J}_v^i, \tag{5.80}$$

where N^i are the bilinear shape functions employed for volume flux as well as for the approximation of the deformation $\boldsymbol{\varphi}^h$. Note that these shape functions are scalar and the nodal degrees $\boldsymbol{J}_v^i$ vectorial, leading to the double amount of degrees of freedom for the flux approximation compared to the RT_0 ansatz. However, combined with a reduced numerical integration of the coupling part contribution ψ_{che} and ψ_{con}, this approach yields promising results. To this end, the time-discrete potential density can be rewritten as

$$\pi^\tau\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \mathrm{Div}\left[\boldsymbol{J}_v\right]\right) = \pi^\tau_{\mathrm{full}}\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v\right) + \pi^\tau_{\mathrm{red}}\left(\nabla\boldsymbol{\varphi}, \mathrm{Div}\left[\boldsymbol{J}_v\right]\right), \tag{5.81}$$

where a standard numerical integration is employed for the evaluation of the first term π^τ_{full} (2x2 Gauss points for $\mathrm{Q_1Q_1}$ elements), and for the evaluation of the second term π^τ_{red}, a reduced numerical integration is used (1 Gauss point for $\mathrm{Q_1Q_1}$ elements). Explicitly, the potential density π^τ is split into the two contributions

$$\pi^\tau_{\mathrm{full}}\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v\right) = \widehat{\psi}_{\mathrm{el}}(\boldsymbol{F}) + \tau\widehat{\phi}(\boldsymbol{J}_v; \boldsymbol{F}_n, v_n)$$

$$\pi^\tau_{\mathrm{red}}\left(\nabla\boldsymbol{\varphi}, \mathrm{Div}[\boldsymbol{J}_v]\right) = \widehat{\psi}_{\mathrm{che}}\left(\boldsymbol{F}, v_n - \tau\,\mathrm{Div}[\boldsymbol{J}_v]\right) \tag{5.82}$$

$$+ \widehat{\psi}_{\mathrm{con}}\left(\det\boldsymbol{F}, v_n - \tau\,\mathrm{Div}[\boldsymbol{J}_v]\right). \tag{5.83}$$

This reduced integration technique for coupling terms together with bilinear interpolation functions for $\boldsymbol{\varphi}$ and $\boldsymbol{J}_v$ is denoted as $\mathrm{Q_1}$-$\mathrm{Q_1}$-red and yields satisfactory results in numerical simulations. When implementing this scheme, special care has to be taken for a correct handling of history variables v_n – while the evolutionary constraint equation for the swelling volume fraction v at current time t is used within the reduced integration scheme, the swelling volume fraction v_n at time t_n shows up in the derivatives of the dissipation potential, being fully integrated. It is crucial to use quantities v_n here that have formerly been evaluated at reduced Gauss points[21].

[21] Special thanks to Steffen Mauthe for figuring this out in a couple of minutes (after the author had stared at the `f77` code and weird contour plot mandala for hours).

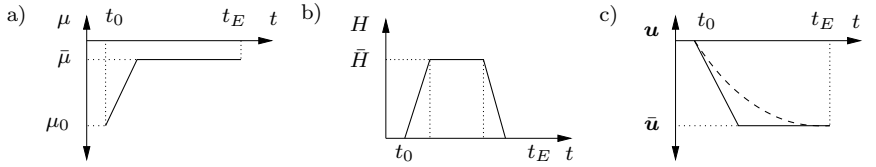


Figure 5.3: Common loading tables: a) Prescribed chemical potential: μ is increased starting from μ_0 and then held constant, while species migration continues until a homogeneous state is reached. b) Prescribed flux: there is no natural saturation state for prescribed fluxes, thus after applying $\bar{H}$, this boundary condition should decrease to zero again. c) Displacement driven boundary value problems are often consolidation processes of saturated bodies.

5.4.4 Summary of Initialization, Boundary Conditions and FE Spaces

With the free energy function (5.61) shifted according to (5.69) and one of the dissipation potentials (5.65) or (5.66) plugged into the minimization or the saddle point formulation, a thorough analysis of representative boundary value problems is possible by using one of the proposed finite element ansatz spaces.

Those boundary value problems may be driven mechanically (e.g. Figure 5.3c), in which case it is often desirable to start from a partially or fully saturated hydrogel body. The latter is associated with a parameter J_0 not only stabilizing the solution (by a choice of e.g. $J_0 \approx 1.01$), but indicating a higher ratio between the pre-swollen volume and the dry volume of the hydrogel block. Note the necessity to initialize the history field v_n in both formulations and global nodal quantities of the chemical potential when employing the saddle point principle.

When dealing with diffusion controlled boundary value problems, a choice can be made between flux-driven processes and those governed by a change of the fluid environment. Prescribed fluxes (see Figure 5.3b for a common loading curve) are applied as Dirichlet boundary conditions in the minimization principle. Note that RT_k elements use flux degrees of freedom integrated over the element surface, such that an application of $\int_{\mathcal{B}^e} \boldsymbol{J}_v \cdot \boldsymbol{N} dA = \int_{\mathcal{B}^e} \bar{H} dA$ on $\partial\mathcal{B}^{H_v}$ suits the associated assembly when using conforming FE ansatz spaces. Some care must be taken when choosing a sign for these absolute flux values, as they should match the orientation of vectorial shape functions associated with the edge on which the boundary condition is applied. Non-conforming $\mathrm{Q}_1\mathrm{Q}_1$ elements employ distinct nodal degrees of freedom for each spatial flux component, the associated boundary condition is best represented by $\boldsymbol{J}_v = \bar{H}\boldsymbol{N}$ on $\partial\mathcal{B}^{H_v}$. Within the saddle point principle, fluxes are applied as Neumann boundary conditions $\bar{H}$. Due to the structure of (5.73) however, applying a flux multiplied with the time step size is convenient within a standard finite element solution procedure, such that $\tau H = \tau\bar{H}$ on $\partial\mathcal{B}^{H_v}$. Thus, when the time step size is increased during

	Minimization	Saddle point
Dirichlet BC	$\bar{\boldsymbol{\varphi}}$, $\pm\int_{\partial\mathcal{B}^e}\bar{H}dA$ (RT_k), $\bar{H}\boldsymbol{N}$ ($\mathrm{Q_1Q_1}$)	$\bar{\boldsymbol{\varphi}}$, $\bar{\mu}$
Neumann BC	$\bar{\boldsymbol{T}}$, $\pm\tau\bar{\mu}$ (RT_k), $\tau\bar{\mu}\boldsymbol{N}$ ($\mathrm{Q_1Q_1}$)	$\bar{\boldsymbol{T}}$, $\tau\bar{H}$
Initialization	$v_n = v_0(J_0)$	$v_n = v_0$, $\mu = \mu_0$
Conf. FE	$\mathrm{Q_1RT_0}$, $\mathrm{Q_2RT_0}$, $\mathrm{P_1RT_0}$, $\mathrm{P_2RT_0}$	$\mathrm{Q_2Q_1}$
Non-conf. FE	$\mathrm{Q_1Q_1}$-red	$\mathrm{Q_1Q_1}$

Table 5.1: Summary of boundary conditions, initialization and selected finite element implementations for minimization and saddle point principle.

a simulation to capture a long term response of intrinsically lengthy diffusion processes, boundary conditions have to be updated accordingly.

Changes of the fluid environment result in boundary conditions $\bar{\mu} \neq \mu_0$ (see Figure 5.3a). Prescribing the chemical potential as a Dirichlet boundary conditions for the minimization problem is most convenient as $\tau\mu = \tau\bar{\mu}$ on $\partial\mathcal{B}^\mu$ due to (5.73), though a dependency on a changing time step size must again be tracked. As before, the prerequisite of a sign adjustment according to the edge-related vectorial shape function comes into play. Using $\mathrm{Q_1Q_1}$ elements, the prescribed chemical potential additionally splits into spatial components according to the discrete flux approximation, as mentioned before.

Prescribing e.g. a constant equilibrium chemical potential is a significant need, and the challenging application of boundary conditions for this quantity within the minimization principle renders a drawback of the proposed formulation. While such a boundary condition can also be realized through fluxes related to the chemical potential by a reaction rate, as pointed out by Di Leo et al. [2014], the application of Dirichlet-type flux boundary conditions still suffers from the additional complexities described above. With regard to boundary conditions, no such complications are inherent to the saddle point formulation: chemical potentials and fluxes, e.g. due to reactions at the boundary, can be prescribed without difficulties.

A summary of the details mentioned above is given in Table 5.1. They have to be taken into account during post-processing of finite element simulations, too, as analysis of nodal values may reflect the desired physical quantities only indirectly. Besides, contour plots for chemical potential or flux components can be produced by evaluating nodal variables at Gauss points and using appropriate projection methods. Inaccuracies of e.g. L2 projections have to be born in mind for an analysis.

Parameter	Value	Unit
γ : Shear modulus	0.1	$\mathrm{N/mm^2}$
α : Mixing modulus	24.2	$\mathrm{N/mm^2}$
χ : Mixing control parameter	0.2	–
M : Volumetric diffusivity parameter	10^{-4}	$\mathrm{mm^4/N\,s}$
J_0 : Pre-swollen Jacobian	1.01	–
ϵ : Volumetric penalty parameter	10	$\mathrm{N/mm^2}$

Table 5.2: Material parameters for the boundary value problems. The parameters are specific to the formulation in terms of the swelling volume fraction as a primary variable, but deduced from experimentally validated data sets as e.g. in Chester & Anand [2010].

5.5 Analysis of Representative Boundary Value Problems

Both minimization and saddle point formulations are evaluated by means of several numerical examples. Swelling is studied as an effect that is characteristic for hydrogels due to the tight coupling of fluid content and volumetric expansion. When mechanical Dirichlet boundary conditions constrain this expansion, structural instability patterns can be observed, challenging the robustness of the underlying implementation.

5.5.1 Solvent Point Source in a 2-Dimensional Block

In order to establish an elementary example for the new minimization formulation together with a validation of the proposed finite element discretization, the following boundary value problem is considered: a two-dimensional square block of hydrogel is given as $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^2 \mid \boldsymbol{X} \in [0, a] \times [0, a]\}$ with $a = 20\,\mathrm{mm}$. It is subjected to no external load and in equilibrium with its external environment. All boundaries are mechanically fixed in both directions, while a stimulus is provided as a volumetric, constant source term $R_v > 0$ in the middle of the body, injecting solvent molecules and thus enforcing swelling. Material parameters in use are shown in Table 5.2, the source term is additionally set to $r = 0.001\ \mathrm{s^{-1}}$ and applied in the central element of the square. 31×31 quadrilateral elements have been used to analyze the boundary value problem (the odd number is due to the geometrically centered point source).

Diffusion of solvent molecules from the center towards the edges is expected to take place, the associated volume change is of insignificant quantity because of

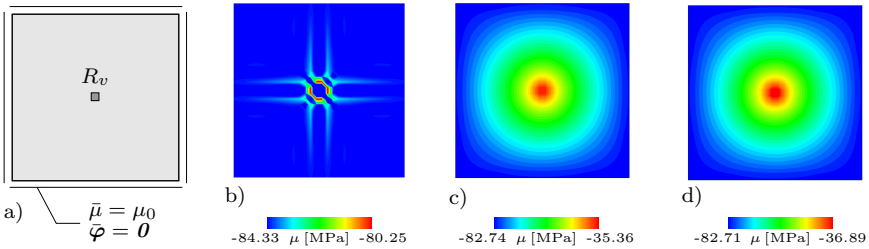


Figure 5.4: Boundary conditions in a) and chemical potential plots for a square hydrogel block subjected to a centered source term. The plots show snapshots of the simulation at time $t = 10^4$ s for 31×31 quadrilateral b) Q_1Q_1, c) Q_1RT_0 and d) Q_1Q_1-red elements. Clearly, a Q_1Q_1 discretization results in a non-physical behavior.

the low value of the prescribed source R_c. A time step size of $\tau = 1$ s has been chosen, distributions of the chemical potential for the three FE discretizations are shown in Figure 5.4 at time $t = 10^4$ s. Clear superior performances of Q_1RT_0 and Q_1Q_1-red compared to Q_1Q_1 elements are observable, the latter producing an nonphysical prediction of the chemical potential due to the non-conforming nature of its underlying finite element ansatz space. Note the very satisfactory agreement between non-conforming Q_1Q_1-red and Q_1RT_0 elements that justifies an employment of the reduced integration scheme.

5.5.2 Finite Element Solutions of the Free Swelling Problem

Again, a square block of hydrogel $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^2 \mid \boldsymbol{X} \in [0, a] \times [0, a]\}$ with $a = 20$ mm is considered. A diffusion process is forced by application of an external chemical potential higher than the initial value μ_0 at every edge of a square, which reflects a change of fluid environment: at time $t = 0$ s, the chemical potential is equal to μ_0, it then increases up to the value $\bar{\mu} = 10^{-5}\mu_0$ at time $t_l = 1000$ s and is finally kept constant, see Figure 5.3a for a sketch of the piece-wise linear load function. There are no mechanical constraints, only rigid body motions are prevented. These boundary conditions will create a chemical potential gradient which drives the diffusion of solvent molecules into the gel sample causing it to swell.

Both saddle point and minimization formulation are used for the analysis. Spatial discretization of the problem is achieved by regular meshes with 40×40 triangular P_1RT_0 elements for the minimization formulation and 20×20 Q_1Q_1 elements with full integration as the non-conforming counterpart. For the saddle point formulation, 20×20 Taylor-Hood elements are used. The set of material parameters used in the simulations are given in Table 5.2, though a penalty

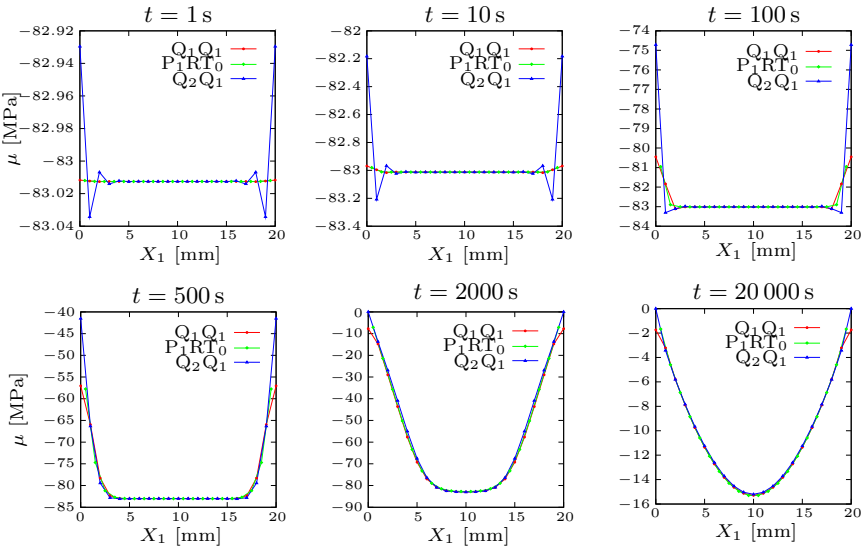


Figure 5.5: Chemical potential profiles on the mid-line of the square at different times. The results of three different finite element formulations are compared – only the minimization formulation discretized by P_1RT_0 elements show completely oscillation-free behavior.

parameter $\epsilon = 10\gamma$ has been selected for the present example. A time step size of $\tau = 1$ s has been chosen.

A profile of the chemical potential is plotted at the horizontal mid-line of the square at different instances of time $t = 1$, 50, 500, 1000, 10000 and 20000 s in Figure 5.5. Figure 5.6 shows the corresponding contour plots of the chemical potential. It can be observed that the non-conforming discretization of the minimization formulation as well as the saddle point formulation lead to oscillatory chemical potential near the edges during initial stages. These effects become less pronounced later and finally vanish at the later steps. A minimization formulation discretized by triangular Raviart-Thomas elements however results in smooth curves during all phases of the simulation.

Note that since the chemical potential is a degree of freedom for the saddle point formulation, Taylor-Hood elements give exact values of the prescribed chemical potential on boundary nodes, which is not the case for Q_1Q_1 and P_1RT_0 elements. The chemical potential is evaluated as an interpolated value at Gauss points – a curve of nodal quantities of μ is thus assembled from L2-projected Gauss point values, which leads to an inaccuracy in regions of steep gradients of the chemical potential. Though nodal values of the chemical potential can be analyzed at

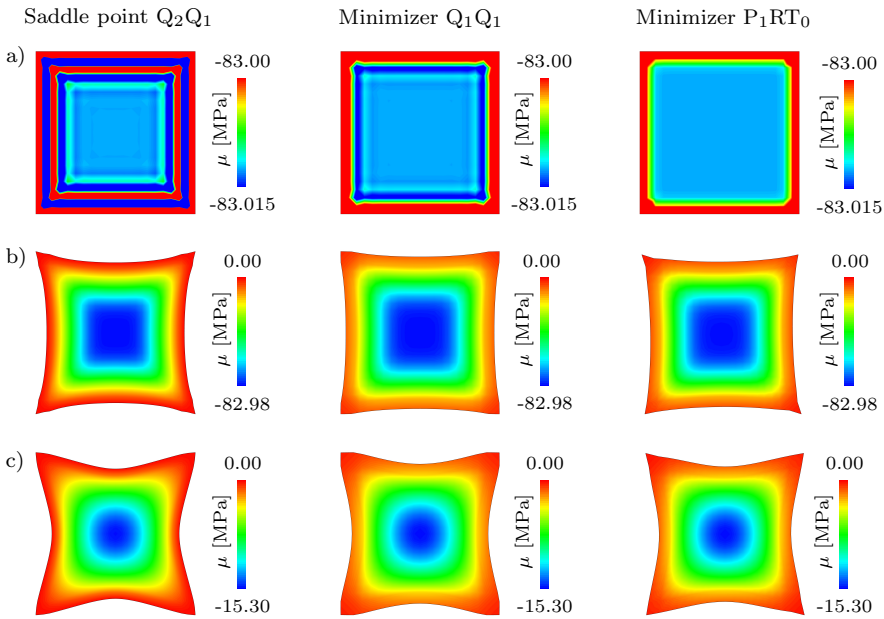


Figure 5.6: Chemical potential contours for free swelling of a square, using saddle point and minimization formulations with different FE discretizations. The plots reflect states at a) $t = 10\,\mathrm{s}$, b) $t = 2000\,\mathrm{s}$ and c) $t = 20\,000\,\mathrm{s}$. The minimization formulation together with P_1RT_0 elements leads to non-oscillatory behavior.

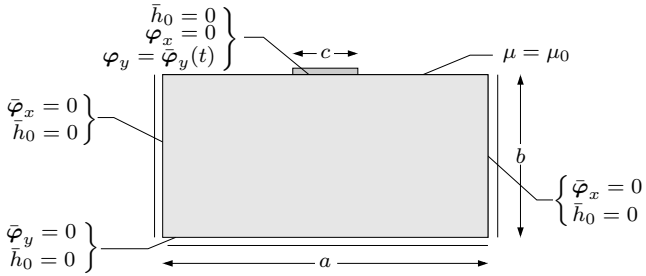


Figure 5.7: The rectangular block is subject to a flat rigid indenter on top. In- or outflux is only allowed at the top face next to the indenter, while the vertical displacement is fixed at the bottom and the horizontal displacement at the left and right faces.

the boundary leading to the prescribed value $\bar{\mu}$, this would obscure the plot by mixing quantities obtained via different methods.

It can be concluded that a simple boundary value problem such as free swelling of a square of hydrogel can lead to oscillations of the chemical potential when employing the saddle point formulation discretized with classical Taylor-Hood elements. While a non-conforming FE space for the minimization formulation does not lead to satisfactory results either, usage of Raviart-Thomas spaces results in non-oscillatory profiles during early and later stages of swelling.

5.5.3 Comparison of Element Designs for an Indentation Test

In what follows, a rectangular block of hydrogel $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^2 \mid \boldsymbol{X} \in [0, a] \times [0, b]\}$ with $a = 200$ mm and $b = 100$ mm is considered under a centered rigid indenter allowing no horizontal slip at the contact surface. This example has been extensively studied by Bouklas et al. [2015], who compared immediate and equilibrium material response of a saddle point formulation discretized by Taylor-Hood finite elements with analytical representations of linear elastic limit responses obtained for an elastic halfspace (after Bower [2009] and Johnson [1985]) and by a theory proposing different material parameters for the two limit cases, see Hui et al. [2006].

In contrast to previously analyzed boundary value problems, this benchmark problem is mechanically driven and exposes the finite element model to deformation states challenging the robustness of the underlying implementation due to the nearly incompressible nature of the polymeric network of hydrogels. A singularity manifests under the indenter edge and demands fine-grained meshes as well discretizations capable of non-trivial deformation modes in the critical region. The body is assumed to be fully saturated, which demands for a high pre-swelling parameter J_0. The indentation length c is small compared to the

dimensions of the body, here $c/b = 0.1$ has been chosen. Mechanical boundary conditions are imposed at the bottom of the specimen (vertical displacement fixed), at the left and right faces (horizontal displacement fixed) as well as under the rigid indenter, where no slip is allowed and where the indentation is controlled (zero horizontal displacement, vertical displacement as a function of time $\bar{\varphi}_2(t)$). Fluxes are prescribed to be zero everywhere except at the top face besides the indenter, thus allowing for a solvent flow away from the indented region towards and out of the free upper surface, see Figure 5.7.

Material parameters coincide with those in Table 5.2, except for the pre-swelling parameter, which is set to $J_0 = 8.0$. The choice of parameter for the piecewise linear loading function (Figure 5.3c) plays a critical role, as the highest challenge of the finite element discretization lies in correctly capturing the material response during a fast loading process. During the loading phase and shortly thereafter, the overall material response can be characterized as nearly elastic and incompressible, as the solvent particles migrate slowly (cf. Bouklas et al. [2015]). Thus, locking effects interplay with solvent diffusion, possibly accompanied by oscillations of chemical potential (saddle point principle) or fluxes (minimization principle) due to improper finite element interpolations. After application of the indentation load or displacement, solvent molecules migrate from the region below the indenter to the outer zone and even cause an outflux at the top edge, where no zero Dirichlet boundary conditions are prescribed for the flux. The process ends, when the chemical potential is homogeneous: this process is intrinsically long-lasting, and after the loading phase with a small time step size, these effects shall be captured within reasonable computing time. To overcome this difficulty, the time step size is augmented after the loading phase. Hence, the load is linearly applied in 2.5 s, with an initial time step size $\tau = 0.1$ s. Afterwards, the time step size is incremented up to 60 times, yielding values between 10^4 s and 10^5 s at the end of the simulation, which is terminated at time $t_{\mathrm{End}} = 10^9$ s. As described in Section 5.4.4, boundary conditions have to be adjusted to the current time step size within the minimization formulations.

The boundary value problem was analyzed by implementations of the saddle point principle using $\mathrm{Q_1Q_1}$ and mixed $\mathrm{Q_2Q_1}$ elements. Meanwhile, simulations based upon the minimization principle were conducted with $\mathrm{Q_1RT_0}$, $\mathrm{Q_2RT_0}$, $\mathrm{P_1RT_0}$, $\mathrm{P_2RT_0}$ as conforming formulations as well as with their non-conforming counterparts $\mathrm{Q_1Q_1}$ and $\mathrm{Q_1Q_1red}$.

To begin with, the initial material response is examined for a force-driven loading process (which differs from the description above) with conforming finite element formulations: the top mid-node is loaded with a linearly increasing force, while all other mechanical degrees of freedom under the indenter are rigidly coupled to the mid-node. Horizontal displacement profiles and load-deflection curves at the end of the loading phase ($t = 2.5$ s) are depicted in Figure 5.8. It can be seen that linear interpolation functions for mechanical degrees of freedom result in stiffer material responses and therefore less deflections at the same force level.

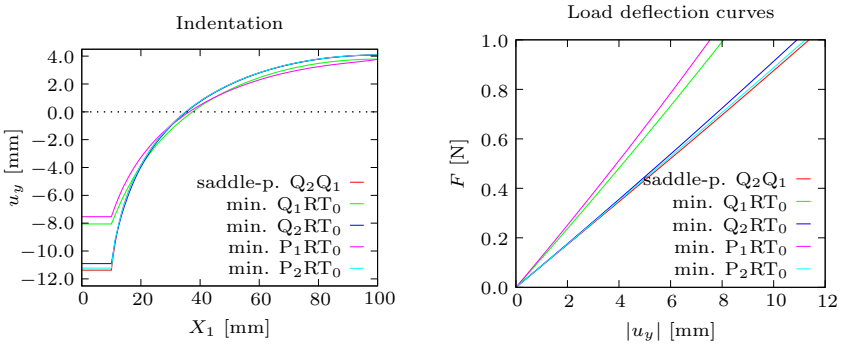


Figure 5.8: Instantaneous material response to a force-driven indentation with rigidly coupled mechanical degrees of freedom under the indenter. Nodal vertical displacement at the top and over half of the body (left) for different element formulations as well as the load deflection curve (right) for the same selection of nodes.

This well-known, locking-related overestimation of elastic stiffness is expected and indicates the deficient approximation of the deformation mode by linear shape functions. All implementations with higher order polynomials predict very similar displacement profiles, independently of the different interpolations of chemical potential or volume flux.

Next, the displacement driven boundary value problem as outlined above is further analyzed. In the left and mid column of Figure 5.9, contour plots of the chemical potential are shown for different formulations and FE implementations, again at the end of the loading phase. Obviously, the aforementioned effects lead to oscillatory (saddle point) or mesh dependent (minimization) distributions of the chemical potential together with an insufficient resolution of the singularity under the indenter vertex. For the Q_1Q_1 saddle point formulation, the poor solution accuracy cannot be traced back unambiguously to a violation of the LBB condition or locking phenomena, they might even interfere. The minimization formulation allows for a clearer separation of concerns – when using conforming elements, nonphysical solution states go back to an insufficient approximation of the deformation modes. Consequently, using the P_2RT_0 element together with the minimizer leads to good results, as well as the mixed Q_2Q_1 element in the saddle point principle. In addition, the non-conforming Q_1Q_1red implementation is an alternative for capturing the immediate material response, as shown in Figure 5.9f.

The issues named above become less important at later instances of time, as the chemical potential slowly approaches homogeneity. In the right column of Figure 5.9, the chemical potential is plotted as a nodal quantity at the top edge, horizontally starting from the mid-point and ending slightly across the indenter.

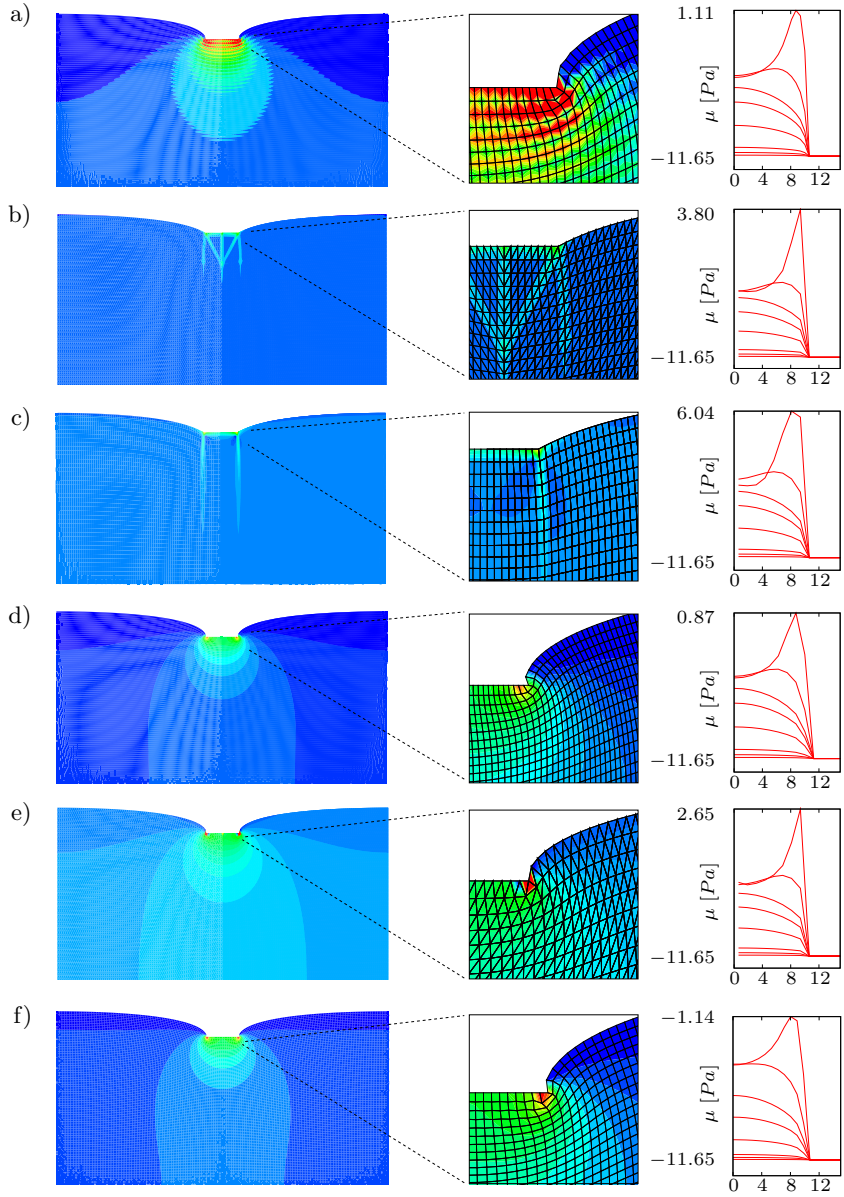


Figure 5.9: Contour plots of the chemical potential at the end of the loading process (left/middle) and top edge nodal values of μ evolving over time (right). a) Saddle point principle with Q_1Q_1, b) minimization principle with P_1RT_0, c) minimization principle with Q_1RT_0, d) saddle point principle with Q_2Q_1, e) minimization principle with P_2RT_0 and f) minimization principle with Q_1Q_1red.

In each individual plot, curves from top to bottom show the development over time: while the upper curve represents the state right after the loading phase, graphs below depict the chemical potential at 0.1%, 0.5%, 1%, 2.5%, 10%, 25% and 100% of the total duration t_{End}. The homogeneous value is equal to $\mu_0 = -11.65$ Pa, which is prescribed as a boundary condition on the right side of the indenter. Note that the maximum values of the chemical potential can be positive (which is in line with results obtained by Bouklas et al. [2015]) and substantially differ in magnitude between all implementations. This indicates that no consistent peak can be determined among all alternatives. Both effects are caused by the singularity at the indenter edge.

A comparison of the computational efficiency of the different implementations has been carried out for the initial loading phase of the problem from $t = 0$ s to $t = 2.5$ s. The time per iteration necessary to solve the system of linear equations has been recorded, which led to approximately 100 temporal data points per implementation (with 4 iteration steps per time increment $\tau = 0.1$ s), based upon which a mean value is determined and listed in Table 5.3. The association of scalar and vectorial degrees of freedom is depicted in Figure 5.10. Element side lengths are identical in all runs. Note that five out of eight element implementations have unused degrees of freedom in order to ease memory management and matrix indexing for an equal numbers of degrees of freedom per node[22], which reduces the definiteness of the comparison. It can be concluded, that the minimization formulation performs slightly better than its saddle point counterpart. Interestingly, this even holds for both $\mathrm{Q_1Q_1}$ implementations, which suggests that the conjunction of scalar nodal chemical potential and a local Newton iteration (5.51) at each integration point is computationally more expensive than an additional degree of freedom per node to capture the vector-valued flux. Recall that (5.32) enters the minimization principle as a constraint equation, which eliminates the need for a local solution step. Comparing the conforming FE interpolations $\mathrm{Q_2RT_0}$ and $\mathrm{P_2RT_0}$ for the minimization principle with the Taylor-Hood element, the same tendency can be observed.

In summary, quadratic shape functions for the mechanical degrees of freedom are indispensable for the indentation problem presented in this section. Together with Raviart-Thomas interpolations of lowest degree, a conforming finite element discretization can be used for the minimization principle, leading to sound solutions even in the complex interplay between fast loading and slow species diffusion. As a further alternative, the reduced integration technique yields reasonable results, too. An implementation of the saddle point principle with Taylor-Hood elements additionally led to good results in the vicinity of the singularity for both short term and long term responses.

[22]The corresponding tangent component is set to one, which leads to zero values of the unused nodal quantity. This has to be taken into account during post-processing, but has no additional effect except for an increase of the computational effort which could be optimized out in future studies.

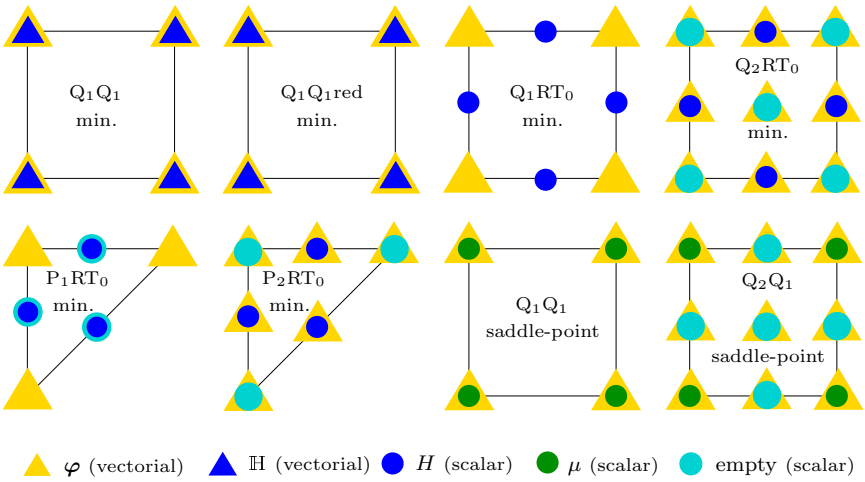


Figure 5.10: Degrees of freedom of the two saddle-point and six minimization implementations. Scalar degrees of freedom are visualized by circles, vectorial ones by triangles.

Minimization	Q_1Q_1	Q_1Q_1red	Q_1RT_0	P_1RT_0	Q_2RT_0	P_2RT_0
$\bar{t}_{\text{iter.}}$ [s]	0.130	0.120	0.188	0.377	0.915	0.779
Dof	16/0	16/0	16/4	12/3	27/5	18/3
Saddle point	Q_1Q_1				Q_2Q_1	
$\bar{t}_{\text{iter.}}$ [s]	0.144				1.036	
Dof	12/0				27/5	

Table 5.3: Computational cost for different implementations during the loading phase of the problem in Section 5.5.3 with a constant number of elements. The average computing time of one iteration is shown together with the number of degrees of freedom per element and unused degrees of freedom separated by a slash.

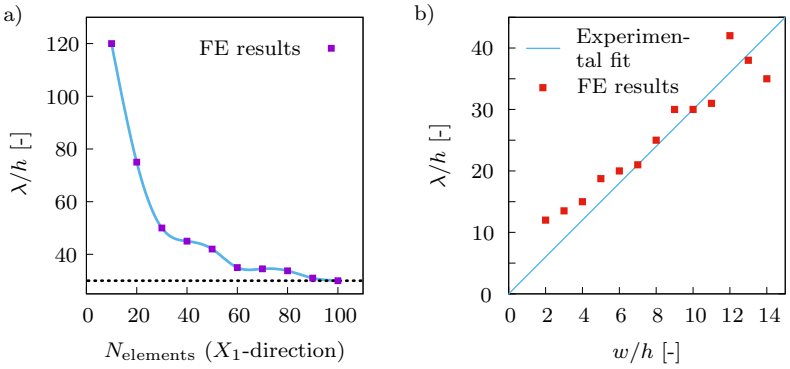


Figure 5.11: Convergence study for regular mesh refinement with constant element edge length in X_2- and X_3-direction in a), instability wavelength depending on the ratio of width to height in b). The straight line is a fit of experimental data taken from Mora & Boudaoud [2006].

5.5.4 Structural Instabilities of a Constrained Strip

The next boundary value problem is a step towards three-dimensional modeling of large, swelling-induced deformation and has been reported by Arnaudov [2014]. A simple thin strip $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^3 \mid \boldsymbol{X} \in [0, l] \times [0, w] \times [0, h]\}$ of length $l = 300$ mm, height $h = 1$ mm and variable width w is considered, matching the specimen of the experimental setup of Mora & Boudaoud [2006], who attached two strips with different elastic properties to each other. They observed that the softer part swells significantly after the specimen is immersed into water, while the stiffer one does not show remarkable volumetric growth. In line with Liu et al. [2010] who tackled the modeling of the above mentioned experiments, the stiffer strip is approximately accounted for by rigid mechanical boundary conditions. Further geometric constraints are applied at $X_1 = 0$ and $X_1 = l$, where the horizontal displacement in the X_1-direction is assumed to be zero in order to capture the macroscopic behavior of an infinitely long strip (the rigidity of these mechanical constraint must however be considered a strong simplification).

The boundary value problem is realized by an implementation of the saddle point principle with trilinear $\mathrm{Q_1Q_1}$ brick elements. While Taylor-Hood elements would clearly be a better choice to circumvent possible violations of the LBB condition, numerical efficiency and a large time step size (τ is chosen to be 1000 s, which relaxes the conditioning of global tangent matrices with respect to this stability criterion) are the rationale behind this interpolation scheme.

A prescribed chemical potential mimics the environmental fluid bath, the corresponding linear loading scheme is visualized in Figure 5.3a. This gives rise to a chemical potential gradient in X_2-direction, which is expected to cause the

volumetric growth. Due to the thin height of the specimen and the horizontal boundary condition, the gel can only expand in one direction, and compressive P_{11} components of the Piola-Kirchhoff stress will cause an out of plane buckling, as observed by Mora & Boudaoud [2006].

A regular mesh is used. The number of elements in X_1 direction is subject to a convergence study with fixed ratio of width to height of 9. Each simulation is evaluated by the wavelength characteristic of the instability pattern, which can easily be computed by an identification of nodal coordinates that are subject to the maximal displacement in X_3-direction. Results are depicted in Figure 5.11a, where a clear asymptotic behavior towards a wavelength of 30 mm is observable and suggests a choice of the mesh density for subsequent simulations.

The wavelength of 30 mm coincides with the experimental observations for the ratio $w/h = 9$, which is a starting point to vary this ratio in both directions. Each time, the wave length is computed and gathered in Figure 5.11b, see also Figure 5.12. An analytical solution that Mora & Boudaoud [2006] found to match their experimental results is plotted as a linear function for the wavelength depending on the ratio w/h. It predicts an increasing wavelength for increasing strip width w. Clearly, the finite element simulation is in very satisfactory agreement with this referential curve.

Note that the robustness of these simulations is no matter of course. The pre-swelling parameter J_0, time step size and the boundary conditions have to be chosen *with care* (which is the most common paraphrase for many hours of dull trial-and-error). Here, dynamic terms are additionally taken into account. To this end, the Newmark β time integration is used for velocity and acceleration together with standard parameters, see Newmark [1959]. While the resulting additional mass/acceleration and damping/velocity terms on local and global stiffness matrix and residual are quite small due to the slow nature of the diffusion process and the high time step size and no macroscopic effect related to the non-quasi-static conditions could be observed, they stabilize the simulations such that neither the critical moment of buckling nor the subsequent swelling harm the iterative solution procedure. Additionally, ideal incompressibility has been assumed by $v = J - 1$ according to Hong et al. [2008], which results in a simpler finite element analysis as the local Newton iteration to solve (5.49) is obsolete. This variant furthermore leads to stronger swelling effects, because the solute influx is directly (and not approximately) bond to the volumetric deformation.

5.5.5 Wavelength Analysis of a Buckling Hydrogel Corona

The next boundary value problem follows experiments of Mora & Boudaoud [2006], who prepared soft hydrogel coronas fixed to a disk of stiff and non-swelling hydrogel in the internal face. The entire structure was immersed in water, and a swelling of the soft gel was observed up to the point where the corona buckles

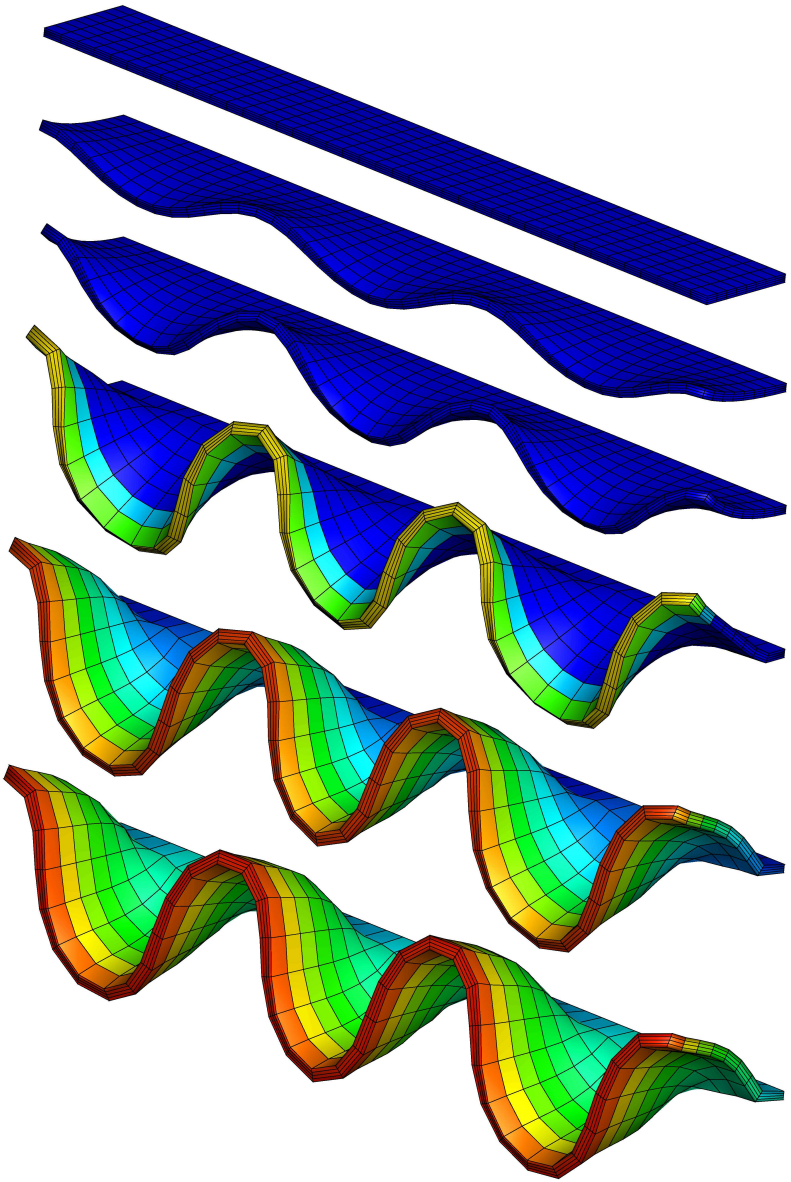


Figure 5.12: Structural instabilities due to strong diffusion-induced volumetric growth at 0 s, $4.621 \cdot 10^6$ s, $5.116 \cdot 10^6$ s, $1.139 \cdot 10^6$ s, $3.102 \cdot 10^6$ s and $4.538 \cdot 10^6$ s (top to bottom). The color depicts the chemical potential (blue color for $\mu = \mu_0$ and red color for $\mu \approx 0$ Mpa).

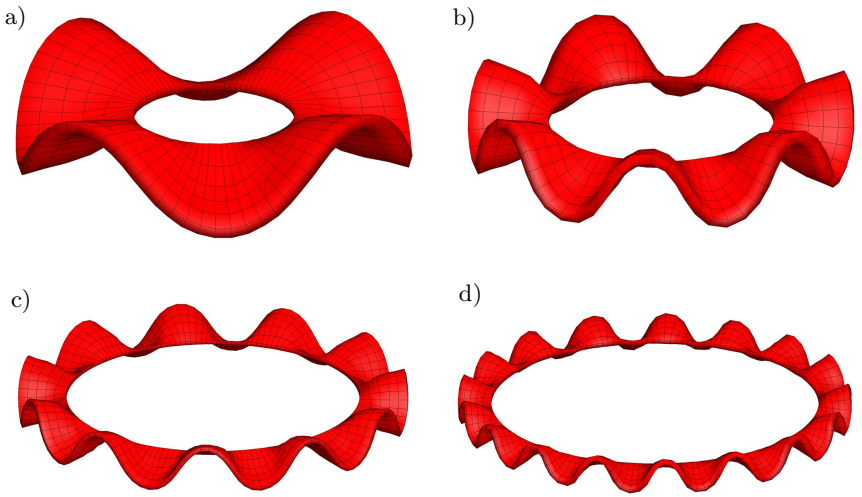


Figure 5.13: Buckling corona specimen according to the setup in Mora & Boudaoud [2006], simulated by using a saddle-point implementation. The ratios of inner to outer radius and wave numbers are a) 0.5 and 4, b) 0.7 and 7, c) 0.81 and 11 d) 0.88 and 17. This is in good agreement with the experimental findings.

out of plane with a wave-like pattern at the outer edge. Increasing wave numbers m (complete waves that appear in the corona) were measured for increasing ratio of internal and external radii and compared to an analytical solution based on a linear stability analysis. This experiment has been modeled with a finite element analysis by Liu et al. [2010] using a classical saddle point principle. In what follows, the minimization formulation is used to simulate this boundary value problem and to validate its robustness with respect to structural instabilities as first reported by Nateghi [2014]. The results are compared to those obtained by an implementation of the saddle point formulation and to experimental results of Mora & Boudaoud [2006].

The corona is mechanically constrained to zero displacements on its internal face. Influx of solvent molecules is forced at its top and bottom faces by appropriate Dirichlet boundary conditions. The outer radius is kept constant at 50 mm, the inner radius varies from 25 mm, 35 mm, 40 mm to 44 mm, while the height remains at 1 mm. The material parameters in use are given in Table 5.2, the volumetric constraint parameter has a value of $\epsilon = 10\gamma$ and the mixing modulus is set to $\alpha = 40.0$ N/mm^2. In order to ensure non-singular solutions at later instances of time, the stabilizing pre-swelling parameter must be set to a higher value (here, $J_0 = 2.0$ has been chosen).

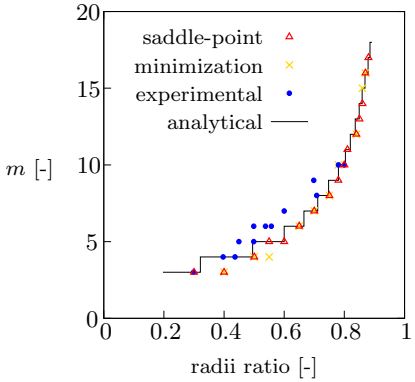


Figure 5.14: Comparison of experimental and analytical results Mora & Boudaoud [2006] with results for both minimization and saddle point formulation, showing very good agreement.

Initially, the expected in-plane volume increase of the corona can be observed. Due to fast swelling at the outer edge and slow species migration towards the inner face, an unbalanced volume change establishes up to the point where an in-plane equilibrium state becomes impossible. Within only few time steps the structure buckles out of plane at the outer edge, showing oscillatory orientation of vertical displacements. Afterwards the buckles grow with time, allowing an unambiguous count of the wave number m. These results are depicted in Figure 5.13 for different aspect ratios of inner to outer radius and compared to experimental and analytical results from Mora & Boudaoud [2006]. These wave numbers are in proper agreement with those obtained from our finite element simulations, for both saddle point and minimization formulation.

Chapter 6:
Fracture Phase-Field Equations for a Regularized Crack Surface

Several methods coexist in computational continuum mechanics that share an identical goal: to model the growth of one or multiple cracks and possibly its initiation. Cracks are discontinuities inside of the material body, and when taken into account as such, fields defined on the reference or the current configuration suffer from jumps across this discontinuity. This brings up two issues: first, the necessary incorporation of interface jump conditions into the description of motion and balance equations, and second, its numerical handling within the finite element method. Particularly the latter issue requires substantial effort and sometimes modification of existing FE solvers beyond the convenient environment of local elements and integration points. The modeling of crack propagation and the static study of crack tips, e.g. stress distribution or accurate element interpolation strategies, can be considered two separate concerns linked by a crack onset criterion. Here, a restriction is made to *brittle fracture*, where an appropriate criterion can be established by comparison of local energetic contributions by an elastic material response under isothermal conditions and the energy release due to new surfaces by a crack tip advancement, see Griffith [1921].

One approach to capture the mechanics at crack tips goes back to Eshelby [1951] which inspired countless studies on defect mechanics and led to a numerical method based on configurational forces, the spatially discrete counterparts of the Eshelby [1970] tensor, see e.g. Müller et al. [2002]; Müller & Maugin [2002] or Steinmann et al. [2005] for a broader context. Crack propagation can be implemented via nodal enrichment strategies, see Miehe & Gürses [2007], but has certain drawbacks. The additional material balance equation, a consequence of imposing varying material coordinates, cannot be solved monolithically with the physical balances, because the problem looses its well-posedness. Hence, the thrill of updating material structures together with the deformed configuration is marginalized into post processing. While the method is computationally efficient (even though the data structure behind the doubling of nodes and element edges yields a fair level of technical sophistication, see Pandolfi & Ortiz [2002]), a certain mesh dependency and additional issues caused the evanescence of the method[23].

[23]This paragraph should have been shorter. But the author has worked on configurational force driven crack propagation for one and a half years (cf. Böger & Miehe

Further approaches to model sharp discontinuities are related to the XFEM method (Belytschko & Black [1999]; Belytschko et al. [2003]), interface elements (Xu & Needleman [1993]), cohesive zone models (de Borst [2003]; Nguyen & Govindjee [2006]), the discontinuous Galerkin method (Hansbo & Hansbo [2002]) or the strong discontinuity approach (Simo et al. [1993]; Linder & Armero [2007]).

Here, the *phase-field approach to fracture mechanics* is outlined to prepare the modeling of failure mechanisms in hydrogels. This method has recently gained much attention due to its simplicity and the ability to capture crack initiation and branching. It relies upon an order parameter that describes the fracture state of the solid (akin to a damage variable) and *regularizes* the sharp crack interface. This regularization can approximately be solved by means of the finite element method, and any discontinuity handling on the global or element level becomes obsolete.

Origins of this regularization of the sharp discontinuity can be traced back to energy minimization concepts of brittle fracture mechanics as suggested by Francfort & Marigo [1998]. A regularized reformulation of their theory in Bourdin et al. [2000] leads to conceptually similar approaches to brittle fracture, see Hakim & Karma [2009], Kuhn & Müller [2010] or Kuhn et al. [2015], where the phase-field evolution appears as a classical Ginzburg-Landau equation. This is very close to the model adapted here, which follows Miehe et al. [2010b] and Miehe et al. [2010a], where the evolution equation is the Euler equation of a minimization statement, and dissipative and energetic contributions of the order parameter are separated, satisfying the entropy inequality a priori by a non-smooth irreversibility condition. This constitutes a gradient damage theory, but can be linked to the Ginzburg-Landau evolution equation, too (cf. Miehe et al. [2010b, 2015b]), though it cannot be derived by using Gurtin's [1996] formalism. For higher order approximations of the crack discontinuity, see the work of Borden et al. [2012], and for a review on the different approaches to phase-field fracture, the reader is referred to Ambati et al. [2015].

The Pure Regularization Problem. The order parameter is denoted by d and interpolates between an intact state ($d = 0$) and a fully broken state ($d = 1$). Its mapping characteristics read

$$d : \begin{cases} \mathcal{B} \times T \to [0, 1] \\ (\boldsymbol{X}, t) \mapsto d = d(\boldsymbol{X}, t). \end{cases} \tag{6.1}$$

[2013]), which dramatically reduces the self-control for keeping these remarks as brief as possible.

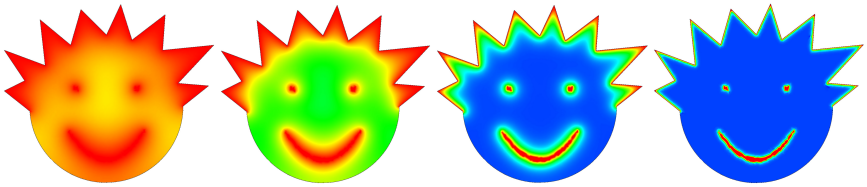


Figure 6.1: Influence of the length scale parameter l_f on the regularization of cracks prescribed by $d = 1$ on $\Gamma(t)$. l_f decreases from left to right, and finally reaches $l_f \approx 2h_e$, where h_e is the charactistic element edge length of the unstructured mesh. This showcase example depicts the stationary solution of (6.4), no deformation has been taken into account.

The parameter d can in general be obtained by minimization of a *crack surface functional*. This functional regularizes a sharp crack surface topology $\Gamma \to \Gamma_l$,

$$\Gamma_l(d) = \int_{\mathcal{B}} \gamma\left(d, \nabla d\right) dV, \tag{6.2}$$

where the integrand is called the crack surface density function γ. A linear evolution of the fracture phase-field can be prescribed by

$$\gamma(d, \nabla d) = \frac{1}{2l} d^2 + \frac{l}{2} \nabla d \cdot \nabla d, \tag{6.3}$$

which also coincides with terms obtained in Γ-convergent regularizations of free discontinuity problems outlined in Ambrosio & Tortorelli [1990]. The parameter l is a length scale controlling the regularization and a central ingredient to the method. A combination of the above equations into the minimization statement reads

$$d = \operatorname{Arg}\left\{ \inf_{d \in \mathcal{W}_d} \Gamma_l(d, \nabla d) \right\}. \tag{6.4}$$

with the admissible space

$$\mathcal{W}_d = \left\{ d \in H^1(\mathcal{B}) \mid d = 1 \text{ on } \Gamma(t) \right\}. \tag{6.5}$$

Note that minimization statement and crack surface density function can be derived and motivated from the study of a one-dimensional bar with a crack. An analytical regularization function can be found that fulfills the Dirichlet condition $d = 1$ on $\Gamma(t)$, and when this function is identified as the solution of a corresponding one-dimensional differential equation, an extension to three dimensions is obvious.

Irreversibility Constraint. The above problem tackles the regularization of a discontinuity by an interpolation between two scalar values, but a connection to cracks in solids and their material response is pending. A crack shall be able to grow as a consequence of time-dependent stress states. To this end, the first

ingredient of the model Miehe et al. [2010b] is an irreversibility condition. This demand postulates that the dissipative creation of crack surfaces cannot violate the entropy inequality, which reads

$$\frac{d}{dt}\Gamma_l(d) = \int_{\mathcal{B}} \delta_d\gamma(d, \nabla d)\dot{d}dV \geq 0. \tag{6.6}$$

The inequality can be satisfied locally and separately with

$$\delta_d\gamma(d, \nabla d) \geq 0 \quad \text{and} \quad \dot{d} \geq 0. \tag{6.7}$$

Both restrictions will be fulfilled, which is shown below. They are main characteristics of this phase-field approach and differ from other models, where the irreversibility is often ensured by application of time-dependent boundary conditions. The latter method does satisfy the irreversibility condition, but rather as a post-processing technique after the finite element solution has been computed, and not as part of the model. The downside of requiring (6.6) is that the Gurtin [1996] formalism, which fits nicely into variational principles (see Miehe [2011]), cannot be applied to obtain the phase-field evolution as a Ginzburg-Landau equation.

Energetic Contributions for Elasticity. Consider an elastic material body undergoing finite deformation as described in Chapter 2. Its stress state is determined by the constitutive free energy function $\widehat{\psi}_{\text{bulk}}$, which must account for the strain as well as the damage state of the solid. By leaving room for further additive energetic contributions, the constitutive free energy is assumed to have the form

$$\widehat{\psi}(\boldsymbol{F}, d, \nabla d) = \widehat{\psi}_{\text{bulk}}(\boldsymbol{F}, d) + \widehat{\psi}_{\text{frac}}(d, \nabla d), \tag{6.8}$$

where both summands fulfill the objectivity and symmetry demands in Sections 3.4.1 and 3.4.2. The bulk contribution is specified to

$$\widehat{\psi}_{\text{bulk}}(\boldsymbol{F}, d) = g(d)\widehat{\psi}_{\text{el}}(\boldsymbol{F}), \tag{6.9}$$

where $g(d)$ is a scalar degradation function with $g(0) = 1$, $g(1) = 0$ and $\partial_d g(1) = 0$, and $\widehat{\psi}_{\text{el}}(\boldsymbol{F})$ is a constitutive free energy function for an elastic material response, e.g. (5.62). Note that no split of the elastic energy into compression and tensile part is made. While this allows for a correct isolation of the tensile elastic energy storage that is responsible for crack onsets, the boundary value problems in Section 7.4 do not show compressive stress states. When not induced by mechanical boundary conditions, diffusion and swelling mechanisms will lead to tensile, not to compressive stresses, which justifies the simplification for the present work. The energetic fracture summand reads

$$\widehat{\psi}_{\text{frac}}(d, \nabla d) = g_c\gamma(d, \nabla d) = \frac{g_c}{2l}\left(d^2 + l^2\nabla d \cdot \nabla d\right), \tag{6.10}$$

where the material parameter g_c is Griffith's critical energy release rate of the solid. At this point, the constitutive character of the crack density function γ becomes visible. Combined with the Griffith parameter, even two constitutive approaches can be seen in (6.10).

The Dissipation Potential. With the constitutive free Helmholtz energy at hand, two ingredients of the model presented by Miehe et al. [2010b] are still missing. The irreversibility $(6.7)_2$ must be incorporated into a material law, and a viscous term shall optionally be included. Both terms are formulated as functions of rate quantities, which is evident for the viscous term and visible in (6.7) for the irreversibility. Hereto, consider the dissipation potential

$$\widehat{\phi}(\dot{d}) = I(\dot{d}) + \frac{\eta}{2}\dot{d}^2 \tag{6.11}$$

based on the non-smooth indicator function and its derivative

$$I(\dot{d}) = \begin{cases} 0 \text{ for } \dot{d} \geq 0 \\ \infty \text{ otherwise} \end{cases} \qquad \text{with} \quad \partial_{\dot{d}} I(\dot{d}) = \begin{cases} 0 \text{ for } \dot{d} \geq 0 \\ \mathcal{R}_- \text{ otherwise,} \end{cases} \tag{6.12}$$

cf. Miehe et al. [2015a]. This indicator function guarantees $(6.7)_2$ and hence constitutes a cornerstone of the phase-field approach considered in this work. Its non-smoothness can however not be treated as is, but requires additional handling techniques as shown below.

Minimization Statement with Fixed Deformation. The governing equation for the order parameter shall now be embedded into a minimization statement in terms of a rate potential at given deformation state, which does not anticipate the variational formulation in Section 7.2, but concludes the outline of this phase-field approach. To this end, the stored energy functional is defined by

$$E(d) = \int_{\mathcal{B}} \widehat{\psi}\,(d, \nabla d; \nabla \boldsymbol{\varphi})\, dV \tag{6.13}$$

with the free energy function (6.8). The time derivative of this functional is considered a function of the phase-field rate (keep in mind that the deformation is assumed to be fixed) and reads

$$\frac{d}{dt} E(\dot{d}) = \int_{\mathcal{B}} \left\{ \partial_d \widehat{\psi} \dot{d} + \partial_{\nabla d} \widehat{\psi} \cdot \nabla \dot{d} \right\} dV. \tag{6.14}$$

A suitable dissipation potential function is introduced as the integral of (6.11), leading to

$$D(\dot{d}) = \int_{\mathcal{B}} \widehat{\phi}(\dot{d}) dV. \tag{6.15}$$

It is straightforward to construct a rate potential based on the above two functionals, because external loads are obsolete for the exemplary case of frozen deformation,

$$\Pi(\dot{d}) = \frac{d}{dt}E(\dot{d}) + D(\dot{d}). \tag{6.16}$$

The fracture phase-field rate is then subject to the minimization statement

$$\{\dot{d}\} = \operatorname{Arg}\left\{ \min_{\dot{d}\in\mathcal{W}_{\dot{d}}} \Pi(\dot{d}) \right\} \tag{6.17}$$

with the admissible space for virtual rates of the order parameter

$$\mathcal{W}_{\dot{d}} = \left\{ \dot{d} \in H^1(\mathcal{B}) \mid \dot{d} = \dot{\bar{d}} \text{ on } \Gamma(t) \right\}. \tag{6.18}$$

When evaluating the necessary conditions of the statement (6.17), the phase-field evolution is obtained as the Euler equation, together with Neumann boundary conditions,

$$\begin{aligned} &1.\ \text{Phase-field evolution} && \partial_d\widehat{\psi} - \operatorname{Div}\left[\partial_{\nabla d}\widehat{\psi}\right] + \partial_{\dot{d}}\widehat{\phi} \ni 0 && \text{in } \mathcal{B} \\ &2.\ \text{Phase-field outflux} && \left(\partial_{\nabla d}\widehat{\psi}\right)\cdot\boldsymbol{N} = 0 && \text{on } \partial\mathcal{B} \end{aligned} \tag{6.19}$$

The Euler equation can be further specified by insertions of the constitutive functions (6.10) and (6.11). This reads

$$\frac{g_c}{l}\left(d - l^2\Delta d\right) - \partial_d g(d)\widehat{\psi}_{\text{el}} + \partial_{\dot{d}} I(\dot{d}) + \eta\dot{d} \ni 0 \tag{6.20}$$

and reveals the relation of elastic energy storage $\widehat{\psi}_{\text{el}}$ to the phase-field evolution.

Reformulation and Stress Criterion. It is possible to bring this evolution equation into a dimensionless form that incorporates the non-smoothness of (6.12) differently and furthermore allows for the formulation of alternative criteria for the crack onset. First, (6.20) is reformulated as

$$d - l^2\Delta d - \frac{l}{g_c}\partial_d g(d)\widehat{\psi}_{\text{el}}(\boldsymbol{F}) + \frac{l}{g_c}\eta\dot{d} + \frac{l}{g_c}\partial_{\dot{d}} I(\dot{d}) \ni 0. \tag{6.21}$$

Then, an energetic history field $\mathcal{H}$ is introduced as in Miehe et al. [2010a],

$$\mathcal{H}(\boldsymbol{X}, t) = \max_{w\in[0,t]} D\left(\boldsymbol{X}, w\right), \tag{6.22}$$

where the supplemental field D is considered a constitutive choice for fracture criteria called *crack driving function*, possibly but not necessarily identical to the Griffith criterion. Based upon (6.22), equation (6.21) can be changed to

$$d - l^2\Delta d - \partial_d g(d)\mathcal{H} + \tilde{\eta}\dot{d} = 0, \tag{6.23}$$

where a modification of the viscosity parameter $\tilde{\eta} = \eta l / g_c$ is accompanied by the more severe elimination of the subdifferential by replacing $\widehat{\psi}_{\text{el}}$ by the accumulated history, see Miehe et al. [2010a] or Miehe et al. [2015b] for a motivation of this step. A form of the crack driving function consistent with the development so far is

$$D(\boldsymbol{X}, t) = \widehat{D}(\boldsymbol{F}) = \frac{\widehat{\psi}_{\text{el}}(\boldsymbol{F}) l}{g_c}. \tag{6.24}$$

The irreversibility of $\dot{d}$ is hence enforced by collecting the maximum of the crack driving state function. A series of alternative constitutive specifications of $\widehat{D}$ is proposed in Miehe et al. [2015b] together with an in-depth discussion of their motivation and characteristics. Here, an isotropic form is chosen that depends on principle tensile stresses and a critical tensile fracture stress $\sigma_c > 0$. The corresponding crack driving field reads

$$D(\boldsymbol{X}, t) = \widehat{D}\left(\frac{\boldsymbol{P}\boldsymbol{F}}{\det[\boldsymbol{F}]}\right) = \left\langle \sum_{i=1}^{3} \left(\frac{\langle \sigma^i \rangle}{\sigma_c}\right)^2 - 1 \right\rangle \tag{6.25}$$

where the principal tensile stresses $\langle \sigma^i \rangle$ are obtained by an additive split of the Cauchy stress based on its spectral decomposition, see Miehe et al. [2015b]. The fracture criterion (6.25) represents an isotropic failure surface in the space of principal tensile stresses and prevents an unrealistic crack evolution in undamaged regions by the threshold contained in the Macaulay bracket in (6.25). Note that the dependence of $\widehat{D}$ on the stresses makes the choice of such a criterion unsuitable for variational formulations. Its simplicity and the exclusion of compressive stress states as crack triggers nevertheless mark this criterion a sound choice for the benchmark simulations in Section 7.4.

Staggered Solution. The energetic history field (6.22) can be combined with a *staggered solution procedure* that solves separately for d while keeping all other degrees of freedom fixed and vice versa. This scheme has been proposed by Miehe et al. [2010a]. It does not resolve issues like the necessity of fine mesh densities or a certain sensitivity to parameter choices (time step size, length scale parameter l), but has proven to be an effective and simple method for solving coupled problems with a fracture order parameter.

Chapter 7:
Mechanically and Diffusion-Driven Crack Growth in Hydrogels

The third and last application of the theory outlined so far is the coupling of elastic deformation, downhill diffusion and the phase-field modeling of fracture – a methodological combination of the model described in Chapter 5 and the regularized crack approximation outlined in Chapter 6. With this contribution, which has been published in similar form as Böger et al. [2017a], a step towards reliable failure prediction of hydrogels shall be made. It may serve as a feasibility study and collection of aspects important for future analysis of crack evolution in elastically deforming solids with dissipative fluid transport, as quite in contrary to the deployment prospects of hydrogels, no finite element based method for the study of the failure mechanism in hydrogels has been established as best practice. To this end, the phase-field approach to fracture is proposed as the simplest way of dealing with cracks in solids. It has recently been used for akin applications, see Wu & Lorenzis [2016] for an analysis of cement paste with evolving cracks and fluid transport. A further related topic that has attracted much attention is hydraulic fracturing. Appropriate models use the phase-field method for crack growth in conjunction with Biot's theory, first proposed by Mikelic et al. [2013], or based upon the theory of porous media, see Heider & Markert [2017]. Stokes flow in cracks has additionally been taken into account by Wilson & Landis [2016]. The coupled problem could also be related to a new minimization principle complementary to classical saddle point formulations by Miehe et al. [2015a].

In the following, a first implementation of the phase-field modeling of fracture in hydrogels is presented, formulated in terms of both minimization and saddle point principles (outlined without the additional damage order parameter field in Section 5.3). First, the boundary value problem is characterized by a summary of field quantities, governing equations and initial as well as boundary conditions. These aspects are then incorporated into a variational formulation and analyzed by means of three two-dimensional boundary value problems. Attempts to validate the model with experimental data are beyond the scope of this treatise and left for future studies, which renders a weak point of this last part of the thesis. Another flaw is the simple influence of damage state on the fluid transport mechanism: Fick's law is assumed to be valid with an identical diffusivity parameter in intact *and* cracked regions, neither Stokes flow nor the simplified Poiseuille law are accounted for as in Wilson & Landis [2016] and Miehe et al. [2015a], respectively. This simplification might be justified by the very slow

solute transport compared to the crack evolution, together with the influence of the degradation of the elastic energy storage (see (7.33) below) on the solute concentration. The latter mechanism can be understood as follows: the degraded elastic stiffness of cracked regions allows for large deformation, and the volumetric strain part is bonded to the solute content by (5.64), causing an influx into fractured zones. Thus, cracks do adhere larger amounts of fluid, and when the related flux rate is still assumed to be very slow, additional mechanisms for this phenomenon might indeed be neglected here and left for future work. Irrespectively of this detail, the minimization principle behind the governing equations can be constructed analogously to Section 5.3, and this innovative formulation is believed to justify this treatise by balancing the shortcomings mentioned above.

7.1 Summary of Equations and Boundary Conditions

The boundary value problem of coupled fracture, solute diffusion elastic deformation is treated as a four-field problem, governed by the deformation field (2.3), the swelling volume fraction (5.1) and the fracture phase-field (6.1). Again, the governing equations are balance of volumetric solute content, which is the balance of fluid mass (2.43) multiplied by the constant $\bar{v}$, the balances of linear and angular momentum (2.45) and (2.47) and the fracture phase-field evolution equation (6.20). The constitutive material laws are composed of (3.37) by cancellation of concentration gradient terms and by (6.9), (6.10) and (6.11). In a Lagrangian description and assuming the external supply R_v to be zero, the list of balance equations and constitutive relations then reads

$$
\begin{aligned}
&1.\ \text{Balance of volumetric solute} && \dot{v} = -\operatorname{Div}\left[\boldsymbol{J}_v\right] \\
&2.\ \text{Balance of linear momentum} && \operatorname{Div}\left[\boldsymbol{P}\right] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \boldsymbol{0} \\
&3.\ \text{Balance of angular momentum} && \boldsymbol{F}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{F}^T \\
&4.\ \text{Phase-field evolution} && \tfrac{g_c}{l}\left(d - l^2\Delta d\right) - 2\left(1-d\right)\mathcal{E} + \mathcal{K} \ni 0 \\
&5.\ \text{Constitutive stresses} && \boldsymbol{P} = \partial_{\boldsymbol{F}}\widehat{\psi}\left(\boldsymbol{F}, v, d\right) \\
&6.\ \text{Constitutive chemical potential} && \mu = \partial_v\widehat{\psi}\left(\boldsymbol{F}, v, d\right) \\
&7.\ \text{Constitutive volumetric flux} && \boldsymbol{J}_v = \partial_{-\nabla\mu}\widehat{\phi}^*(-\nabla\mu, \dot{d}; \boldsymbol{F}, v) \\
&8.\ \text{Constitutive crack driving quantity} && \mathcal{E} = \widehat{\psi}_{\mathrm{el}}\left(\boldsymbol{F}\right) \\
&9.\ \text{Constitutive crack resistance} && \mathcal{K} = \partial_{\dot{d}}\widehat{\phi}^*(-\nabla\mu, \dot{d}; \boldsymbol{F}, v).
\end{aligned}
\tag{7.1}
$$

Here, the degradation function $g(d) = (1-d)^2 + \zeta$ (where ζ is a stabilizing parameter, see below), has been specified a priori and is inserted into the phase-field evolution equation $(7.1)_4$. The quantities $\mathcal{E}$ and $\mathcal{K}$ have furthermore been

introduced to separate constitutive choices and balance laws. As can be seen from $(7.1)_2$, the quasi-static assumptions made for Chapter 5 persist. An additional assumption that influence the above list is related to the additive composition of the free energy storage; $\widehat{\psi}$ (specified in Section 7.4) is assumed to have one summand exclusively describing the elastic energy storage of the undamaged solid, denoted as $\psi_{\text{el}}(\boldsymbol{F})$ (see also Chapter 6). Again, the balance equations above are identical for both minimization and saddle point formulation.

The surface of the solid is decomposed according to (5.3), with mechanical boundary conditions (5.4) and those related to diffusion (5.5). The initial condition for the swelling volume fraction is (5.6). In addition, and in line with the admissible spaces for the phase-field regularization problem (6.5), the fracture order parameter can be subject to Dirichlet boundary conditions on a sharp crack surface $\boldsymbol{\Gamma}(t)$ on the reference configuration $\mathcal{B}$, while the phase-field outflux is assumed to be zero, which reads

$$d = 1 \text{ on } \Gamma(t) \quad \text{and} \quad \partial_{\nabla d}\widehat{\psi} \cdot \boldsymbol{N} = 0 \text{ on } \partial\mathcal{B}. \tag{7.2}$$

Hereby, the boundary value problem is entirely specified. Besides balance equations and boundary conditions, nominal stresses, chemical potential and the crack driving quantity can be constituted according to $(7.1)_{5,6,8}$ from the free energy. The dissipation potential will prescribe a kinetic transport law (again, this will be Fick's diffusion law) and the irreversibility constraint preventing self-healing of cracks by $(7.1)_{7,9}$. Free energy and dissipation potential are assumed to be consistent with the principles outlined in Section 3, which can be shown upon specialization of $\widehat{\psi}$ and $\widehat{\phi}$.

7.2 Continuous and Time-Discrete Incremental Potentials

In the following, continuous and time-discrete minimization and saddle point principles are outlined, on the basis of Section 5.3 and Miehe et al. [2015a]. The resulting potentials will imply the set of equations $(7.1)_{1,2,4}$ as Euler equations.

7.2.1 Stored Energy, Dissipation Potential and Load Functionals

The stored energy functional is constructed in dependency of the primary variables $\boldsymbol{\varphi}$, v and d, and is governed by the free energy,

$$E(\boldsymbol{\varphi}, v, d) = \int_{\mathcal{B}} \widehat{\psi}(\nabla\boldsymbol{\varphi}, v, d, \nabla d) dV. \tag{7.3}$$

The rate of E is called rate-of-energy functional, depending on the rates of the primary fields at a given state $\{\boldsymbol{\varphi}, v, d\}$, which reads

$$\frac{d}{dt}E(\dot{\boldsymbol{\varphi}}, \dot{v}, \dot{d}) = \int_{\mathcal{B}} \left\{ \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} - \partial_v\widehat{\psi}\,\mathrm{Div}\,[\boldsymbol{J}_v] + \partial_d\widehat{\psi}\dot{d} + \partial_{\nabla d}\widehat{\psi}\cdot\nabla\dot{d} \right\} dV. \tag{7.4}$$

Here, the evolution $(7.1)_1$ has again injected a dependency on the volume flux $\boldsymbol{J}_v$. This balance equation is thus taken as a local constraint condition yielding the evolution $\dot{v}$. Next, a dissipation potential function is formulated in terms of $\boldsymbol{J}_v$ and $\dot{d}$ that extends the previous (5.10). It can in analogy be related to its dual $\widehat{\phi}^*$ by the Legendre transformation

$$\widehat{\phi}(\boldsymbol{J}_v, \dot{d}; \nabla\boldsymbol{\varphi}, v) = \sup_{-\nabla\mu} \left\{ -\nabla\mu\cdot\boldsymbol{J}_v - \widehat{\phi}^*(-\nabla\mu, \dot{d}; \nabla\boldsymbol{\varphi}, v) \right\}. \tag{7.5}$$

Hence, at given $\{\boldsymbol{\varphi}, v, d\}$, the definition of the canonical dissipation potential functional reads

$$D(\boldsymbol{J}_v, \dot{d}) = \int_{\mathcal{B}} \widehat{\phi}(\boldsymbol{J}_v, \dot{d}; \nabla\boldsymbol{\varphi}, v) dV. \tag{7.6}$$

The functional for external loads splits into a mechanical contribution and one associated with volume fluxes over the boundary,

$$P_{\mathrm{ext}}\,(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) = \int_{\mathcal{B}} \boldsymbol{R}_{\boldsymbol{\varphi}}\cdot\dot{\boldsymbol{\varphi}} dV + \int_{\partial\mathcal{B}^{\boldsymbol{T}}} \bar{\boldsymbol{T}}\cdot\dot{\boldsymbol{\varphi}} dA - \int_{\partial\mathcal{B}^{\mu}} \bar{\mu}\boldsymbol{J}_v\cdot\boldsymbol{N} dA \tag{7.7}$$

where the external body forces $\boldsymbol{R}_{\boldsymbol{\varphi}}$ are assumed to be known being independent of the deformation, which holds for the surface tractions $\bar{\boldsymbol{T}}$, too. The prescribed chemical potential $\bar{\mu}$ is similarly assumed to not depend on the primary fields.

7.2.2 Continuous Three-Field Minimization Principle

Equipped with the rate-of-energy functional and the dissipation as well as load functionals (7.4), (7.6) and (7.7), the rate potential

$$\Pi(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \dot{d}) = \frac{d}{dt}E(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \dot{d}) + D(\boldsymbol{J}_v, \dot{d}) - P_{\mathrm{ext}}\,(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) \tag{7.8}$$

can be introduced at a given state $\{\boldsymbol{\varphi}, v, d\}$. It may be split up into an internal and an external term,

$$\Pi(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \dot{d}) = \int_{\mathcal{B}} \pi\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \mathrm{Div}\,[\boldsymbol{J}_v], \dot{d}, \nabla\dot{d}\right) dV - P_{\mathrm{ext}}\,(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v), \tag{7.9}$$

where the internal rate potential density appears as a quantity per unit volume, evaluated as

$$\begin{aligned}\pi(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \mathrm{Div}\,[\boldsymbol{J}_v], \dot{d}, \nabla\dot{d}) = {}& \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} - \partial_v\widehat{\psi}\,\mathrm{Div}\,[\boldsymbol{J}_v] \\ & + \partial_d\widehat{\psi}\dot{d} + \partial_{\nabla d}\widehat{\psi}\cdot\nabla\dot{d} + \widehat{\phi}(\boldsymbol{J}_v, \dot{d}; \nabla\boldsymbol{\varphi}, v).\end{aligned} \tag{7.10}$$

Then, deformation rate, volume flux and fracture phase-field rate at given variables $\{\boldsymbol{\varphi}, d, v\}$ are the solution of the three-field minimization principle

$$\left\{\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \dot{d}\right\} = \operatorname{Arg}\left\{\inf_{\dot{\boldsymbol{\varphi}} \in \mathcal{W}_{\dot{\boldsymbol{\varphi}}}} \inf_{\boldsymbol{J}_v \in \mathcal{W}_{\boldsymbol{J}_v}} \inf_{\dot{d} \in \mathcal{W}_{\dot{d}}} \Pi\left(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \dot{d}\right)\right\} \tag{7.11}$$

with admissible spaces $\mathcal{W}_{\dot{\boldsymbol{\varphi}}}$, $\mathcal{W}_{\boldsymbol{J}_v}$ as in (5.19) for the rates of deformation and volume flux, and the additional admissible space $\mathcal{W}_{\dot{d}}$ defined in (6.18) for the phase-field rate $\dot{d}$. Evaluating the variation of (7.9) for virtual rates of deformation and fracture phase-field and virtual volume flow results in the Euler equations

$$\begin{aligned}
&1.\ \text{Balance of linear momentum} && \operatorname{Div}\left[\partial_{\boldsymbol{F}}\widehat{\psi}\right] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \boldsymbol{0} && \text{in } \mathcal{B}\\
&2.\ \text{Inverse Fick's law} && \nabla\left[\partial_v\widehat{\psi}\right] + \partial_{\boldsymbol{J}_v}\widehat{\phi} = \boldsymbol{0} && \text{in } \mathcal{B}\\
&3.\ \text{Phase-field evolution} && \partial_d\widehat{\psi} - \operatorname{Div}\left[\partial_{\nabla d}\widehat{\psi}\right] + \partial_{\dot{d}}\widehat{\phi} \ni 0 && \text{in } \mathcal{B}\\
&4.\ \text{Prescribed tractions} && \left(\partial_{\boldsymbol{F}}\widehat{\psi}\right)\cdot\boldsymbol{N} - \bar{\boldsymbol{T}}_0 = \boldsymbol{0} && \text{on } \partial\mathcal{B}^{\boldsymbol{T}}\\
&5.\ \text{Prescribed chemical potential} && \bar{\mu} - \partial_v\widehat{\psi} = 0 && \text{on } \partial\mathcal{B}^{\mu}\\
&6.\ \text{Phase-field outflux} && \left(\partial_{\nabla d}\widehat{\psi}\right)\cdot\boldsymbol{N} = 0 && \text{on } \partial\mathcal{B}
\end{aligned} \tag{7.12}$$

These equations include the local equation of motion $(7.1)_{2,5}$ and an inverse form of the linear diffusion law attributed to Fick $(7.1)_{6,7}$ as well as the phase-field evolution $(7.1)_4$ and Neumann boundary conditions. The balance of volume $(7.1)_1$ is part of the variational statement as a constraint condition, i.e., $\dot{v}$ is given by $(7.1)_1$ for volume fluxes $\boldsymbol{J}_v$ that stem from the solution of the minimization principle.

7.2.3 The Extended Four-Field Saddle Point Principle

Consider the dissipation potential function $\widehat{\phi}$ resulting from the generalized Legendre transformation (7.5) in terms of the dual function $\widehat{\phi}^*$. Deploying (7.5) in (7.9) leads to an extended four-field potential

$$\Pi^+(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \dot{d}, -\nabla\mu) = \int_{\mathcal{B}} \pi^+\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \operatorname{Div}\left[\boldsymbol{J}_v\right], \dot{d}, \nabla\dot{d}, -\nabla\mu\right) dV - P_{\text{ext}}(\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v) \tag{7.13}$$

that introduces a dependency on $-\nabla\mu$ as a mixed variable dual to $\boldsymbol{J}_v$. The mixed rate potential density then reads

$$\pi^+\left(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \operatorname{Div}[\boldsymbol{J}_v], \dot{d}, \nabla\dot{d}, -\nabla\mu\right) = \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} - \partial_v\widehat{\psi}\operatorname{Div}[\boldsymbol{J}_v]$$
$$+\partial_d\widehat{\psi}\dot{d} + \partial_{\nabla d}\widehat{\psi}\cdot\nabla\dot{d} + -\nabla\mu\cdot\boldsymbol{J}_v - \widehat{\phi}^*\left(-\nabla\mu, \dot{d}; \nabla\boldsymbol{\varphi}, v\right). \tag{7.14}$$

In this potential, $\boldsymbol{J}_v$ is replaced by its relationship to the rate $\dot{v}$ via $(7.1)_1$. After insertion of both terms into (7.13), integration by parts and use of the boundary condition $\mu = \bar{\mu}$ on $\partial\mathcal{B}^{\mu}$, the reduced mixed potential is obtained,

$$\Pi^*(\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \dot{d}) = \int_{\mathcal{B}} \pi^*\left(\nabla\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \nabla\mu, \dot{d}, \nabla\dot{d}\right) dV - P^*_{\text{ext}}(\dot{\boldsymbol{\varphi}}, \mu), \tag{7.15}$$

where the integrand of the first term is the potential density

$$\pi^*(\nabla\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \nabla\mu, \dot{d}, \nabla\dot{d}) = \partial_{\boldsymbol{F}}\widehat{\psi} : \nabla\dot{\boldsymbol{\varphi}} + \partial_v\widehat{\psi}\dot{v} + \partial_d\widehat{\psi}\dot{d}$$
$$+\partial_{\nabla d}\widehat{\psi}\cdot\nabla\dot{d} - \mu\dot{v} - \widehat{\phi}^*(-\nabla\mu, \dot{d}; \nabla\boldsymbol{\varphi}, v). \tag{7.16}$$

The load functional in (7.15) is again composed of a mechanical and a chemical term,

$$P^*_{\text{ext}}(\dot{\boldsymbol{\varphi}}, \mu) = \int_{\mathcal{B}} \boldsymbol{R}_{\boldsymbol{\varphi}}\cdot\dot{\boldsymbol{\varphi}}dV + \int_{\partial\mathcal{B}^{\boldsymbol{T}}} \bar{\boldsymbol{T}}\cdot\dot{\boldsymbol{\varphi}}dA + \int_{\partial\mathcal{B}^{H_v}} \mu\bar{H}_v dA \tag{7.17}$$

where the first two summands can identically be found in (7.7) and $\bar{H}_v$ denotes a prescribed volume flow on the referential surface part $\partial\mathcal{B}^{H_v}$. Hence, $\{\boldsymbol{\varphi}, d, v\}$ given, the four-field saddle point principle determines the deformation rate, the evolution of the swelling volume fraction, the chemical potential as well as the fracture phase-field by

$$\left\{\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \dot{d}\right\} = \operatorname{Arg}\left\{\inf_{\dot{\boldsymbol{\varphi}}\in\mathcal{W}_{\boldsymbol{\varphi}}}\inf_{\dot{v}\in L^2}\sup_{\mu\in\mathcal{W}_{\mu}}\inf_{\dot{d}\in\mathcal{W}_{\dot{d}}}\Pi^*(\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \dot{d})\right\}. \tag{7.18}$$

The admissible spaces are defined in (5.19), (6.18) and $(5.28)_2$. The Euler equations are an outcome of the first variation of (7.15) for admissible virtual

rates $\delta\dot{\boldsymbol{\varphi}}$, virtual chemical potential $\delta\mu$ and swelling volume fraction rate $\delta\dot{v}$ as well as the virtual fracture phase-field rate $\delta\dot{d}$. They read

$$
\begin{aligned}
&\text{Balance of linear momentum} && \operatorname{Div}\left[\partial_{\boldsymbol{F}}\widehat{\psi}\right] + \boldsymbol{R}_{\boldsymbol{\varphi}} = \boldsymbol{0} && \text{in } \mathcal{B}\\
&\text{Constitutive chemical potential} && \partial_v\widehat{\psi} - \mu = 0 && \text{in } \mathcal{B}\\
&\text{Balance of volume} && \dot{v} + \operatorname{Div}\left[\partial_{-\nabla\mu}\widehat{\phi}^*\right] = 0 && \text{in } \mathcal{B}\\
&\text{Phase-field evolution} && \partial_d\widehat{\psi} - \operatorname{Div}\left[\partial_{\nabla d}\widehat{\psi}\right] + \partial_{\dot{d}}\widehat{\phi} \ni 0 && \text{in } \mathcal{B}\\
&\text{Prescribed tractions} && \left(\partial_{\boldsymbol{F}}\widehat{\psi}\right)\cdot\boldsymbol{N} - \bar{\boldsymbol{T}} = \boldsymbol{0} && \text{on } \partial\mathcal{B}^{\boldsymbol{T}}\\
&\text{Prescribed volume flow} && \left(\partial_{-\nabla\mu}\widehat{\phi}^*\right)\cdot\boldsymbol{N} - \bar{H}_v = 0 && \text{on } \partial\mathcal{B}^{H_v}\\
&\text{Phase-field outflux} && \left(\partial_{\nabla d}\widehat{\psi}\right)\cdot\boldsymbol{N} = 0 && \text{on } \partial\mathcal{B}
\end{aligned}
\tag{7.19}
$$

This again covers the local equation of motion $(7.1)_{2,5}$, a constitutive relation for the chemical potential $(7.1)_6$, the volume balance $(7.1)_{1,7}$ and the phase-field evolution $(7.1)_4$ together with Neumann boundary conditions.

7.2.4 Incremental Three-Field Minimization Principle

Consider a time interval with step size τ, with the fields known at time t_n. The objective is to obtain these quantities at time t (associated variables have no subscript) by means of variational principles formulated for the discrete time increment. In what follows, a restriction to pure Dirichlet problems is presumed for simplicity, i.e., $P_{\text{ext}} = 0$ in (7.8). The incremental potential in $[t_n, t]$ is discrete in time and reads

$$
\Pi^{\tau}\left(\boldsymbol{\varphi}, \boldsymbol{J}_v, d\right) = \int_{\mathcal{B}} \pi^{\tau}\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \operatorname{Div}\left[\boldsymbol{J}_v\right], d, \nabla d\right) dV, \tag{7.20}
$$

without Neumann boundary conditions. π^{τ} is an incremental potential density per unit volume linked to the rate potential density (7.10) by

$$
\pi^{\tau}\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \operatorname{Div}\left[\boldsymbol{J}_v\right], d, \nabla d\right) = \mathsf{Algo}\left\{\int_{t_n}^{t} \pi(\nabla\dot{\boldsymbol{\varphi}}, \boldsymbol{J}_v, \operatorname{Div}\left[\boldsymbol{J}_v\right], \dot{d}, \nabla\dot{d}) dt\right\}. \tag{7.21}
$$

Algo represents a time integration algorithm of the argument. It includes an integration of $(7.1)_1$ as $v = v_n - \tau \operatorname{Div}\left[\boldsymbol{J}_v\right]$ and the time discrete fracture phase-

field rate $\dot{d} = (d - d_n)/\tau$. Combined with an implicit integration of the continuous potential density π, the incremental potential density reads

$$\pi^\tau\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, \operatorname{Div}\left[\boldsymbol{J}_v\right], d, \nabla d\right) = \widehat{\psi}\left(\nabla\boldsymbol{\varphi}, \boldsymbol{J}_v, v_n - \tau \operatorname{Div}\left[\boldsymbol{J}_v\right], d, \nabla d\right) + \tau\widehat{\phi}\left(\boldsymbol{J}_v, d; \nabla\boldsymbol{\varphi}_n, v_n, d_n\right). \tag{7.22}$$

Then, the incremental three-field minimization statement reads

$$\{\boldsymbol{\varphi}, \boldsymbol{J}_v, d\} = \operatorname{Arg}\left\{\inf_{\boldsymbol{\varphi}\in\mathcal{W}_{\boldsymbol{\varphi}}}\inf_{\boldsymbol{J}_v\in\mathcal{W}_{\boldsymbol{J}_v}}\inf_{d\in\mathcal{W}_d}\Pi^\tau\left(\boldsymbol{\varphi}, \boldsymbol{J}_v, d\right)\right\}, \tag{7.23}$$

with $\mathcal{W}_{\boldsymbol{\varphi}}$ and $\mathcal{W}_{\boldsymbol{J}_v}$ according to (5.36) and $\mathcal{W}_d$ being the admissible spaces of the fracture phase-field (6.5). With the incremental potential Π^τ, a spatial discretization can be setup by approximatively interpolating primary variables and a set of state variables. The generalized stresses and generalized tangent arrays can then be defined as first and second derivatives of (7.22), and they are in turn used to determine a finite element solution procedure for the non-linear global residual. This procedure is identical to the outline in Section 5.3.4, only the state variables are enriched by the order parameter d and its gradient. For the sake of brevity, the formal construction of finite element matrices are omitted.

7.2.5 Incremental Four-Field Saddle Point Principle

The time-discrete incremental potential in $[t_n, t]$ is an analog counterpart to (7.15) and, for pure Dirichlet problems ($P^*_{\text{ext}} = 0$), reads

$$\Pi^{*\tau}\left(\boldsymbol{\varphi}, v, \mu, d\right) = \int_{\mathcal{B}} \pi^{*\tau}\left(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu, d, \nabla d\right) dV. \tag{7.24}$$

Here, incremental potential density $\pi^{*\tau}$ per unit volume is linked to the rate potential density π^* (7.16) by

$$\pi^{*\tau}\left(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu, d, \nabla d\right) = \mathsf{Algo}\left\{\int_{t_n}^{t} \pi^*\left(\nabla\dot{\boldsymbol{\varphi}}, \dot{v}, \mu, \nabla\mu, d, \nabla d\right) dt\right\}. \tag{7.25}$$

Using an implicit integration of (7.25) yields

$$\pi^{*\tau}\left(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu, d, \nabla d\right) = \widehat{\psi}(\nabla\boldsymbol{\varphi}, v, d, \nabla d) - \mu\left(v - v_n\right) - \tau\widehat{\phi}^*(-\nabla\mu, d; \nabla\boldsymbol{\varphi}_n, v_n, d_n). \tag{7.26}$$

Then, the incremental four-field saddle point principle can finally be formulated as

$$\{\boldsymbol{\varphi}, v, \mu, d\} = \operatorname{Arg}\left\{\inf_{\boldsymbol{\varphi}\in\mathcal{W}_{\boldsymbol{\varphi}}}\inf_{v\in L^2}\sup_{\mu\in\mathcal{W}_\mu}\inf_{d\in\mathcal{W}_d}\Pi^{*\tau}\left(\boldsymbol{\varphi}, v, \mu, d\right)\right\}, \tag{7.27}$$

with admissible spaces (5.48), and (6.5). In line with Section 5.3.5, the stationary condition can be matched in two steps. First, v is determined locally for given state $\{\boldsymbol{\varphi}, \mu, d\}$ by the local minimization problem

$$\text{local}: \quad \{v\} = \operatorname{Arg}\left\{ \inf_{v \in L^2} \pi^{*\tau}\left(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu, d, \nabla d\right) \right\}, \tag{7.28}$$

with a standard Newton-Raphson method. This minimization induces a reduced incremental potential density

$$\pi^{*\tau}_{\text{red}}(\nabla\boldsymbol{\varphi}, \mu, \nabla\mu, d, \nabla d) = \inf_{v \in L^2} \pi^{*\tau}\left(\nabla\boldsymbol{\varphi}, v, \mu, \nabla\mu, d, \nabla d\right), \tag{7.29}$$

and the global incremental potential $\Pi^{*\tau}$ (7.24) boils down to the reduced three-field potential

$$\Pi^{*\tau}_{\text{red}}\left(\boldsymbol{\varphi}, \mu, d\right) = \int_{\mathcal{B}} \pi^{*\tau}_{\text{red}}\left(\nabla\boldsymbol{\varphi}, \mu, \nabla\mu, d, \nabla d\right) dV. \tag{7.30}$$

The second, global solution step is based on the saddle point problem for admissible spaces defined in (5.48) and (6.5),

$$\text{global}: \quad \{\boldsymbol{\varphi}, \mu, d\} = \operatorname{Arg}\left\{ \inf_{\boldsymbol{\varphi} \in \mathcal{W}_{\boldsymbol{\varphi}}} \sup_{\mu \in \mathcal{W}_{\mu}} \inf_{d \in \mathcal{W}_d} \Pi^{*\tau}_{\text{red}}\left(\boldsymbol{\varphi}, \mu, d\right) \right\}. \tag{7.31}$$

Again, the spatial discretization is an analogy to the previous development in Section 5.3.5, the extension of the set of state variables by d and ∇d is as straightforward as the introduction of generalized stresses and tangent moduli, which can then be shown to determine global finite element residual and tangent, respectively.

7.3 Isotropic Constitutive Model and Algorithmic Treatment

The isotropic model for the material response extends the constitutive specifications made in Section 5.4.1 by terms associated with the phase-field approach to fracture. These additional ingredients ship as summands of the free energy function and the dissipation potential, specified in Chapter 6. A pre-swollen reference state is necessary by analogy with Section 5.4.2 and must be extended to account for the additional terms in the constitutive free energy storage and the dissipation potential, which is straightforward due to the methodological resemblance to the model free of the order parameter. Then, three numerical benchmark simulations are presented that test the coupling mechanism of fluid transport and crack growth in hydrogels as proposed by the present model.

7.3.1 Constitutive Energy Storage and Dissipation Potential

The energy storage function is now composed by an elastic bulk response, solvent molecule mixing, a penalty summand incorporating a volumetric constraint and a fracture contribution that contains an accumulated dissipative work density,

$$\widehat{\psi}\left(\boldsymbol{F}, v, d, \nabla d\right) = \widehat{\psi}_{\text{bulk}}\left(\boldsymbol{F}, d\right) + \widehat{\psi}_{\text{che}}\left(v\right) + \widehat{\psi}_{\text{con}}\left(\det \boldsymbol{F}, v\right) + \widehat{\psi}_{\text{frac}}\left(d, \nabla d\right). \quad (7.32)$$

The bulk response portion $\widehat{\psi}_{\text{bulk}}$ is related to the Neo-Hookean material response, degraded however by a prefactor that represents the fracture state of the solid and still conserves a stabilizing elastic rest energy density,

$$\widehat{\psi}_{\text{bulk}}\left(\boldsymbol{F}, d\right) = \left[(1-d)^2 + \zeta\right] \widehat{\psi}_{\text{el}}(\boldsymbol{F}) \quad (7.33)$$

where $\widehat{\psi}_{\text{el}}$ is defined in (5.62). The stabilizing parameter $\zeta \approx 0$ is chosen as small as possible. (7.33) reveals that no additional decomposition of the elastic energy into tensile and compressive parts has been made for simplicity. The free energy terms $\widehat{\psi}_{\text{che}}$, $\widehat{\psi}_{\text{con}}$ and $\widehat{\psi}_{\text{frac}}$ are identical to (5.63), (5.64) and (6.10), respectively. Again, the volumetric penalty term provides an approximate constraint for the incompressibility of polymer chain network and the interstitial solvent. The dissipation potential is additively composed of a bulk part incorporating the constitutive relation between volume flux and chemical potential and a fracture part that ensures the crack phase-field irreversibility and has an additional viscous term,

$$\widehat{\phi}\left(\boldsymbol{J}_v, d; \nabla\boldsymbol{\varphi}_n, v_n, d_n\right) = \widehat{\phi}_{\text{bulk}}(\boldsymbol{J}_v; \nabla\boldsymbol{\varphi}_n, v_n) + \widehat{\phi}_{\text{frac}}\left(\frac{d-d_n}{\tau}\right) \quad (7.34)$$

and

$$\widehat{\phi}^*(-\nabla\mu, d; \nabla\boldsymbol{\varphi}_n, v_n, d_n) = \widehat{\phi}^*_{\text{bulk}}(-\nabla\mu; \nabla\boldsymbol{\varphi}_n, v_n) + \widehat{\phi}_{\text{frac}}\left(\frac{d-d_n}{\tau}\right) \quad (7.35)$$

Here, $\widehat{\phi}_{\text{bulk}}$ and $\widehat{\phi}^*_{\text{bulk}}$ refer to the dissipation potentials specified in (5.65) and (5.66) for minimization and saddle-point principle, respectively. For the minimization principle at a given state $\{\nabla\boldsymbol{\varphi}_n, v_n, d_n\}$, the dissipation potential reads

$$\widehat{\phi}(\boldsymbol{J}_v, d; \boldsymbol{C}_n, v_n, d_n) = \frac{1}{2Mv_n}\boldsymbol{C}_n : (\boldsymbol{J}_v \otimes \boldsymbol{J}_v) + I\left(\frac{d-d_n}{\tau}\right) + \frac{\eta}{2}\left(\frac{d-d_n}{\tau}\right)^2, \quad (7.36)$$

with the dissipation potential contribution (6.11) inserted with time-discrete phase-field rates. Within the saddle point principle, the dissipation potential depends on $-\nabla\mu$ and the approximated rate $\dot{d}$. It is defined as

$$\widehat{\phi}^*\left(-\nabla\mu, \dot{d}; \boldsymbol{C}_n, v_n\right) = \frac{Mv_n}{2}\boldsymbol{C}_n^{-1} : (-\nabla\mu \otimes -\nabla\mu)$$

$$+ I\left(\frac{d-d_n}{\tau}\right) + \frac{\eta}{2}\left(\frac{d-d_n}{\tau}\right)^2, \tag{7.37}$$

again at given chemo-mechanical state. The fracture phase-field enters the dissipation potentials only via the irreversibility condition and a quadratic viscous term – the bulk diffusivity is thus independent of the crack evolution. This assumption is valid for the boundary value problems of this contribution, as the slow fluid transport would not cause remarkable changes of the macroscopic material response. A more fine-grained mechanism that couples the fracture state to the diffusivity is nevertheless intended for future studies.

7.3.2 Revisiting the Pre-Swollen Reference State

In Section 5.4.2, the challenges due to a zero swelling volume fraction have been addressed. The identical technique is necessary to enable stable convergence rates for problem with an addition damage order parameter. Notation (subscript d for the dry state), zero stress assumption and multiplicative decomposition of the deformation gradient are adopted, and a comparison of energetic states reads

$$\int_{\mathcal{B}_d} \widehat{\psi}_d\left(\boldsymbol{F}_d, v, d, \nabla d\right) dV_d = \int_{\mathcal{B}} \widehat{\psi}\left(\boldsymbol{F}, v, d, \nabla d\right) dV. \tag{7.38}$$

As a consequence, the two free energy densities can be related to each other via

$$\widehat{\psi}\left(\boldsymbol{F}, v, d, \nabla d\right) = J_0^{-1} \widehat{\psi}_d\left(\boldsymbol{F}, v, d, \nabla d\right)|_{\boldsymbol{F}=\boldsymbol{F}_d=\boldsymbol{F}\boldsymbol{F}_0}. \tag{7.39}$$

From these rationales, initial conditions for the swelling volume fraction and the chemical potential can be constructed, see (5.67) and (5.70). For the dissipation potential contribution related to volumetric solute fluxes, pre-swelling is automatically accounted for, see Section 5.4.2. This differs from the summand accounting for the crack evolution, $\widehat{\phi}_{\mathrm{frac}}$, the dissipation potential portion must be related to the pre-swollen state similarly to (7.38),

$$\int_{\mathcal{B}_d} \widehat{\phi}_{\mathrm{frac},d}(\dot{d}) dV_d = \int_{\mathcal{B}} \widehat{\phi}_{\mathrm{frac}}(\dot{d}) dV, \tag{7.40}$$

which leads to

$$\widehat{\phi}_{\mathrm{frac}}(\dot{d}) = J_0^{-1} \widehat{\phi}_{\mathrm{frac},d}(\dot{d}) = \frac{1}{J_0} I\left(\frac{d-d_n}{\tau}\right) + \frac{\eta}{2J_0}\left(\frac{d-d_n}{\tau}\right)^2. \tag{7.41}$$

This relation turns out to have no effect on the fracture phase-field evolution equation, because the pre-swelling factor can be easily eliminated in (7.12) or (7.19) by multiplication with J_0.

With this pre-swollen state at hand, all fields (chemical potential, volume flux, deformation, swelling volume fraction and fracture phase-field) depend on pre-swollen positions $\boldsymbol{X}$. Local variables v and v_n must be initially set in the finite

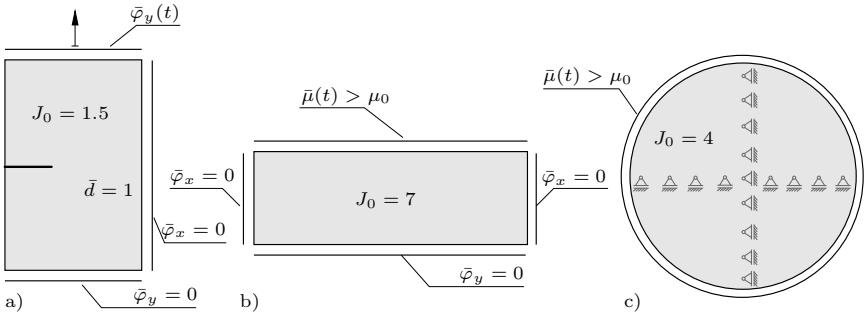


Figure 7.1: The three boundary value problems, where a) is mechanically driven with a prescribed notch and b), c) are drying processes without any defects.

element context, and for the saddle point formulation, the chemical potential at the nodes has to be initialized to μ_0, see Section 5.4.4 for a detailed comment on technical aspects of boundary conditions.

7.4 Analysis of Representative Boundary Value Problems

In the following, the capabilities of both minimization and saddle point principle shall be analyzed in terms of three representative boundary value problems (see Figure 7.1). The solution scheme is based on an operator split and the energetic history field (6.22) as described in Chapter 6. It has been implemented in conjunction with a sequential direct solver, which turned out to be the most reliable option for the coefficient matrices suffering from a mediocre conditioning due to the highly nonlinear Flory-Rehner term (5.63) and different time scales of diffusion and fracture. The latter stems from the slow nature of non-equilibrium diffusion processes that competes with small time step sizes necessary for the solution of phase-field fracture simulations. Fortunately, the following boundary value problems can be designed for a relatively slow crack evolution, which permits a solution even for large time step sizes (it must however be accepted that the large time increment has a bad influence on convergence rates, and artificially makes the crack evolution even slower). The criterion based on effective stresses (6.25) is chosen as the constitutive input for (6.22) in all three numerical benchmark simulations.

7.4.1 Delayed Crack Propagation in a Notched Block

An implementation of the minimization formulation is used for the following boundary value problem, in which a failure mechanism characteristic to hydrogels as described in Bonn et al. [1998] is captured. A two-dimensional square block

Parameter	Value	Unit
γ : Shear modulus	1.0	$\mathrm{N/mm^2}$
α : Mixing modulus	24.2	$\mathrm{N/mm^2}$
χ : Mixing control parameter	0.1	–
M : Volume diffusivity parameter	10^{-4}	$\mathrm{mm^4/N\,s}$
J_0 : Pre-swelling parameter	1.5	–
ϵ : Penalty parameter	10	$\mathrm{N/mm^2}$
σ_c : Critical stress	3.0	$\mathrm{N/mm^2}$
η : Crack viscosity	0	$\mathrm{N/mm^2}s$
l : Length scale parameter	$2h_e$	mm

Table 7.1: Material parameters for boundary value problems, where h_e is the side length of a typical finite element triangulation.

of hydrogel is given as $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^2 \mid \boldsymbol{X} \in [0, a] \times [0, b]\}$ with $a = 75$ mm and $b = 150$ mm. It is subject to no external load and in equilibrium with its external environment, i.e., zero tractions for the mechanical part, and μ_0 being prescribed as a Neumann boundary condition on the whole boundary. The specimen is horizontally fixed at the right edge by zero Dirichlet boundary conditions. At the top and the bottom faces, a non-zero displacement is applied, linearly increasing from zero to a finite value $\pm\Delta\bar{\varphi}_2$ at time t^* and then being held constant. The loading curve is depicted in Figure 7.3. The horizontal notch is realized by Dirichlet boundary conditions for the phase-field, $\bar{d}(\boldsymbol{X}, t) = 1$ for $X_2 = b/2$ and $0 \leq X_1 \leq a/5$. A non-uniform mesh is chosen to ensure a sufficient resolution of the regularized crack surface where the crack is expected to grow while keeping the computational cost of the simulation affordable by non-parallel processing. In addition, the symmetry of the specimen is exploited and only one half of the problem is discretized, together with proper Dirichlet boundary conditions that render the symmetry condition. In terms of finite element spaces, $\mathrm{Q_1Q_1Q_1}$ elements with linear shape functions for displacement, flux and fracture phase-field have been chosen to provide a simple, efficient approach. Note that these elements do not fulfill the $H(\mathrm{Div}, \mathcal{B})$ conformity required by the non-standard variational structure of the minimization formulation, see Section 5.4.3. The implementation of conforming (e.g. Raviart-Thomas) elements is left for future investigations, as the linear interpolations turned out to provide satisfactory results as a start. Material parameters for the simulations are listed in Table 7.1.

The maximum displacement $\Delta\bar{\varphi}_2$ and the loading speed controlled by t^* (the time from which on the applied displacement is held constant) influence the

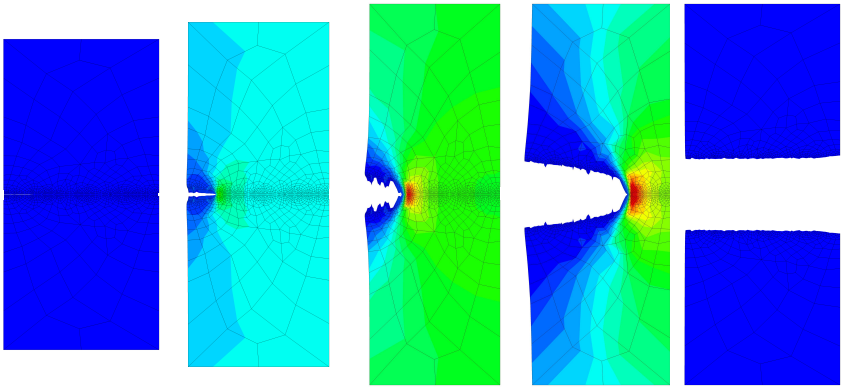


Figure 7.2: Contour of the nominal stress P_{22} for 5 different time steps. Parts of the specimen with $d > 0.98$ are blanked out for better visibility of the crack pattern. As symmetry has been exploited, the lower half of the contour plots are mirrored.

onset of crack growth. Here, a time step size of $\tau = 10^{-1}$ s has been chosen together with $t^* = 10$ s and $\Delta\bar{\varphi}_2 = 0.11b$. The crack does not propagate during loading. Instead, in the phase with constant displacement, a diffusion-driven, viscous effect slowly increases the tension at the crack tip, until the specimen cracks relatively quick. Snapshots taken during the simulation are depicted in Figure 7.2, while loading curve and the normalized sum of mechanical reaction forces in x_2-direction at the top of the specimen are shown in Figure 7.3. Clearly, the relaxation of nodal forces keep record of the diffusion impact and strongly change during crack evolution. As can be seen in Figure 7.3, the crack suddenly starts propagating at $t \approx 93$ s, which is a large delay compared to the loading of the hydrogel specimen.

Next, performance and robustness of the solution scheme based on minimization formulation, operator split and energetic history variable are analyzed in the following. To this end, the loading is modified such that the crack does not initiate after a delay, but when the maximum displacement is applied or slightly thereafter. The mechanical reaction forces in vertical direction are analyzed over time at the top of the specimen together with the crack length (measured via X_1-coordinates of nodes with $d > 0.95$), the number of iterations per time step and the average duration of one iteration step for varying parameters known to be critical to the solution of phase-field fracture growth: mesh density, regularization length scale l and time step size τ. While variations of the latter two parameters are straightforward, the density of the mesh is only approximately altered due to the non-uniformity of the mesh; the characteristic element edge length h_e at $X_2 = 0$ (in the vicinity of the crack) is controlled during triangulation.

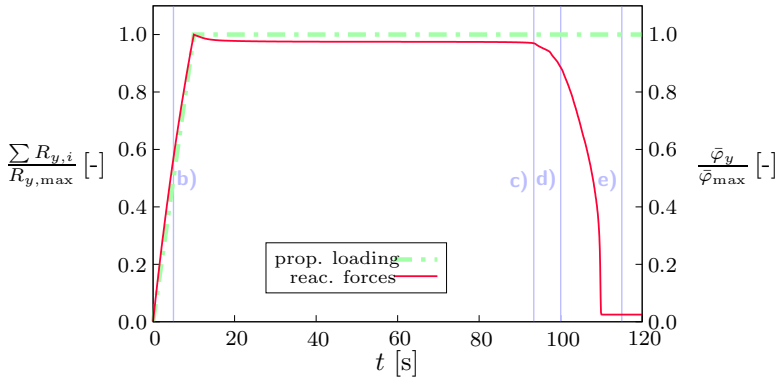


Figure 7.3: Plots of loading scheme and normalized nodal reaction forces of the upper boundary in $y-$direction over time. The rapid decrease of reaction forces clearly indicates the delayed and viscous-like failure of the specimen. Points in time at which snapshots in Fig. 7.2 are taken are marked by the vertical lines, i.e., b) 5 s, c) 93.4 s, d) 100 s and e) 115 s.

l/h_e	$\bar{t}_{\text{iter./node}}$ [ms]	h_e [mm]	$\bar{t}_{\text{iter./node}}$ [ms]	τ [s]	$\bar{t}_{\text{iter./node}}$ [ms]
1.0	0.150	0.75	0.195	0.005	0.184
1.5	0.158	1.00	0.165	0.010	0.171
2.0	0.151	1.25	0.158	0.050	0.162
3.0	0.147	1.50	0.141	0.100	0.148
2.5	0.152	1.75	0.134		

Table 7.2: Average duration of one iteration step per node for different crack regularization, mesh density and time step size. A pronounced influence of these parameters on the computational cost cannot be observed.

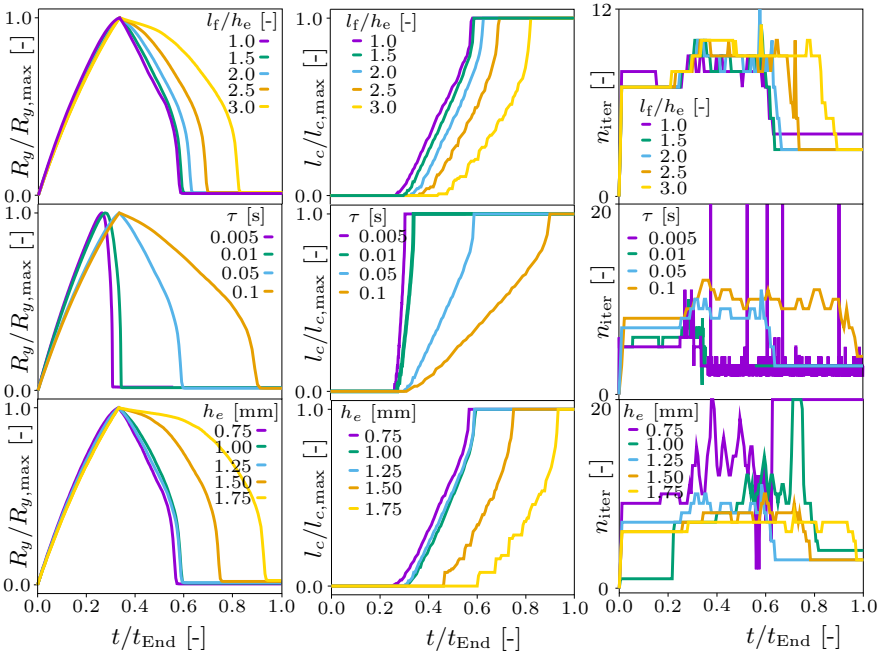


Figure 7.4: Parameter study of the material response by normalized vertical reaction forces at the top (left), normalized crack length (middle) and number of iterations per time step (right) until an absolute convergence criterion is reached. When not varied, the base values are $l_f/h_e = 1.5$, $h_e = 1.25\,\text{mm}$ and $\tau = 0.05\,\text{s}$. The solution scheme exhibits clear sensitivities and an asymptotic behavior for crack length and reaction forces.

The results are visualized in Figure 7.4 and gathered in Table 7.2. A clear dependence of reaction forces and crack length on time step size and spatial resolution of both mesh and crack regularization can be observed, with asymptotic behavior for smaller values of τ, h_e and l – which is a typical finding in studies of the phase-field approach to fracture. The number of iterations per time step has been recorded together with an absolute convergence criterion. While the iteration count turns out to be only slightly dependent on the crack regularization parameter, stronger peaks and fluctuations can be witnessed when the time step size and the characteristic element edge vary. This is a difficulty associated with the non-smoothness of the irreversibility constraint (6.12) that is still inherent to the approach based upon the energetic history (6.22). Finally, the duration of one iteration step per node in Table 7.2 shows that no distinct in- or decrease of the computational cost is related to the parameter variations.

7.4.2 Drying Rectangle with Geometrical Constraints

While the first example has been mechanically driven and realized through an implementation of the minimization formulation, now saddle point principle and boundary value problems with constant mechanical constraints come to the fore, where the crack evolution is exclusively caused by stresses that arise from a drying process.

To this end, consider a rectangular block $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^2 \mid \boldsymbol{X} \in [0, a] \times [0, b]\}$ with $a = 300$ cm and $b = 60$ cm, in equilibrium with its external environment at time $t = 0$. Fluid will be forced to diffuse out of the gel at the top edge by application of a linearly increasing draining condition on the top surface $\bar{\mu}(t) > \mu_0$, which is a loading scenario inverse to Figure 5.3a. Additionally, a fluid-saturated initial configuration is required, where the preswelling comes in handy: choosing a high parameter J_0 renders a homogeneous swollen state of the hydrogel. Vertical displacement is blocked at the bottom, horizontal displacement at left and right faces. Material parameters are as before (see Table 7.1), however with $J_0 = 4$ and a critical stress $\sigma_c = 1.0$ N/mm^2. A uniform mesh is used to discretize the specimen, and for simplicity, linear $\mathrm{Q_1Q_1Q_1}$ elements have been chosen that bilinearly interpolating all nodal primary values, i.e., displacement, chemical potential and fracture phase-field within the saddle point principle. While usage of Taylor-Hood elements for displacement and chemical potential clearly is a better choice, again the usage of more advanced finite element techniques is beyond the scope of this treatise.

It can be observed that multiple crack paths develop from top to bottom due to draining-induced tractions, which is a similar observation that can be made for water-saturated soils during drying. As no defects are taken into consideration, the crack initiation has instability-like characteristics – how many cracks develop and where they are placed over the width of the specimen cannot be controlled, see Figure 7.6 for multiple crack patterns as a consequence of deviations of

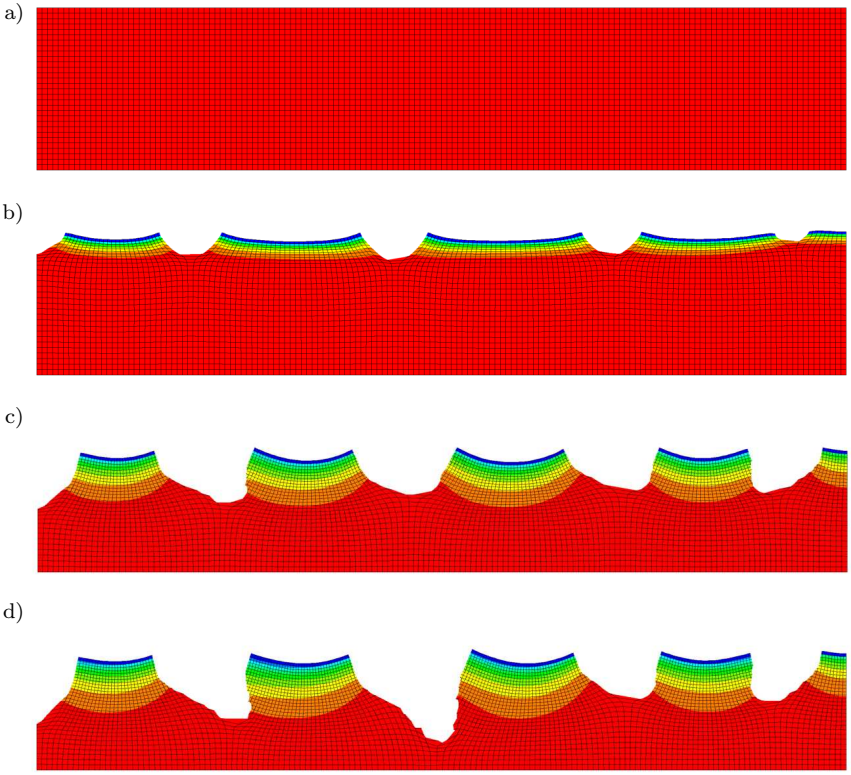


Figure 7.5: Drying-induced crack evolution in a hydrogel rectangle horizontally fixed at both vertical edges. The chemical potential contour lies between $\mu_0(J_0)$ (red) and $\bar{\mu}$ (blue). Snapshots are taken at a) 0 s, b) $1.85 \cdot 10^5$ s, c) $1.30 \cdot 10^6$ s s and d) $1.97 \cdot 10^6$ s, regions with $d > 0.91$ are blanked out for the sake of better crack pattern visibility.

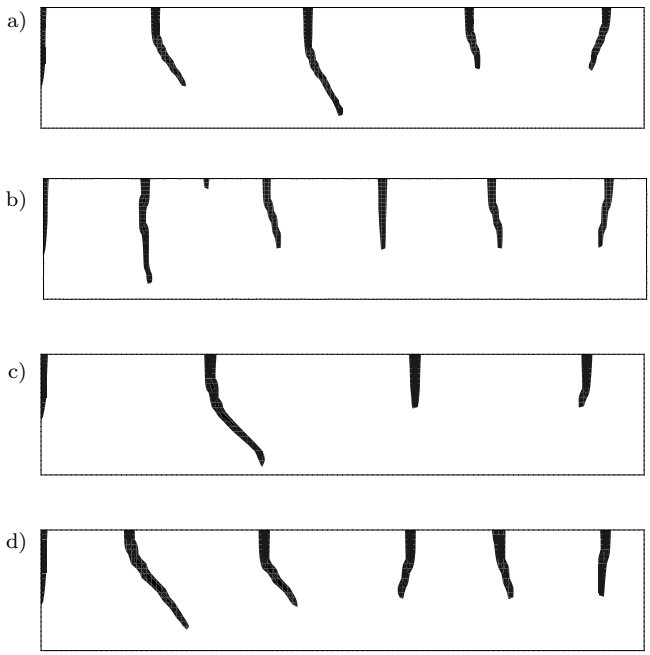


Figure 7.6: Different crack patterns of the same block of hydrogel and visualized in the undeformed configuration for one varying parameter or a perturbed discretization: a) parameter as described in Section 7.4.2 b) parameter as in Section 7.4.2, but with slightly perturbed nodal coordinates c) penalty parameter ϵ reduced by 25% d) pre-swelling parameter $J_0 = 5$ (previously, $J_0 = 4$).

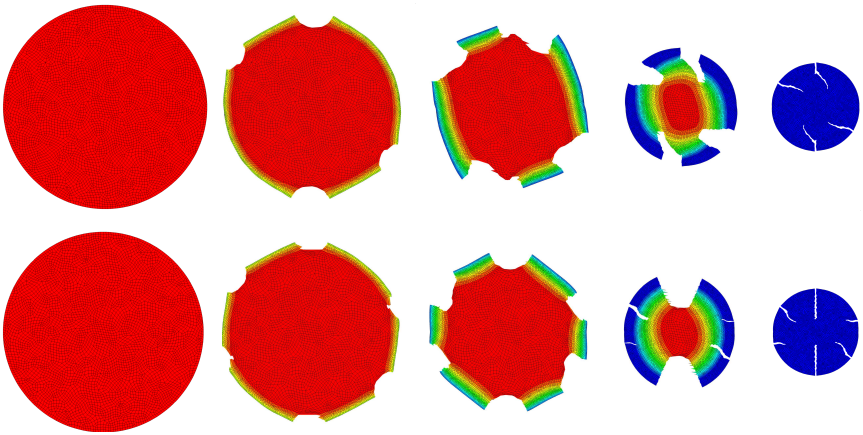


Figure 7.7: Geometrically unconstrained diffusion-driven crack pattern for two different discretizations at different instances of time. The instability-like crack initiation is clearly mesh-dependent. Contour plots for the chemical potential are shown with a range between $\mu_0(J_0)$ (red) and $\bar{\mu}$ (blue), illustrating the homogeneous initial and end states of the disk. Again, elements with $d > 0.95$ are blanked out for the sake of crack pattern visibility.

material parameter or discretization. Between the first visible crack evolution and the arrival of cracks at the bottom of the specimen $1.97 \cdot 10^6$ s pass by. This again underlines the slow nature of fluid transport and diffusion-driven fracture in the solid (that would crack in one or at most in a few time steps without the consideration of species diffusion). For this reason, the simulation could be run with a large time step size of $7 \cdot 10^2$ s.

7.4.3 2-Dimensional Circle with Prescribed Draining Boundary

The third example is again implemented as the saddle point formulation and similar to the previous one, however without the harsh geometrical constraints that force crack-driving tension. Instead, a disk $\mathcal{B} = \{\boldsymbol{X} \in \mathcal{R}^2 \mid ||\boldsymbol{X}|| \leq d/2\}$ with $d = 4$ mm is considered, subject to mechanical Dirichlet boundary conditions allowing for unconstrained homogeneous shrinking: nodes on the vertical and horizontal axes through the circle center are fixed horizontally and vertically, respectively. Quadrilateral elements with bilinear shape functions for all degrees of freedom have be chosen for simplicity. The chemical potential $\bar{\mu}(t) > \mu_0$ is applied everywhere on $\partial\mathcal{B}$ and again linearly increasing starting from μ_0. Here, a heavily swollen initial configuration is assumed by setting $J_0 = 7$. While σ_c is 0.1 N/mm^2 and $M = 10^{-2}$ mm^4/N s, all other material parameters can be found in Table 7.1.

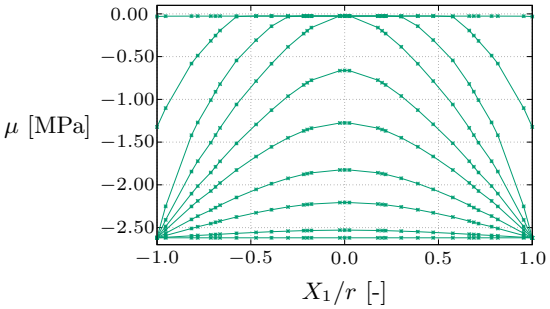


Figure 7.8: Profile of the chemical potential for nodes with $X_2 \approx 0$ and at 0 s, 25 s, 50 s, 75 s, 90 s, 95 s, 100 s, 105 s, 115 s and 150 s (top to bottom). It starts at μ_0 and increases at the outer edges as prescribed by boundary conditions until a homogeneous state is reached. The gradient of μ causes the circle to shrink.

The volume decrease induces pressure in the inner and tension in the outer regions of the disk. As soon as the tension reaches a critical value, radial crack initiation occurs at various locations at the surface. Again, the onset of crack growth is an instability-like mechanism, and thus the simulation turned out to be mesh-dependent, which is illustrated in Figure 7.7 for two different discretization. In this plot, the chemical potential is depicted as a contour plot, showing the gradient from outer to inner regions as the non-equilibrium state variable responsible for shrinking. At the end ($t \approx$1500 s), the hydrogel disk reaches a homogeneous state. Crack evolution however stopped earlier, because the critical stress was not reached anymore. This boundary value problem illustrates that the possibly large volume change characteristic to hydrogels can not only lead to the instabilities studied in Sections 5.5.4 or 5.5.5 in the presence of geometrical constraints, but also to crack initiation and growth in unconstrained specimen.

Chapter 8: Conclusion

In the present work, a variational formulation for phase separation and diffusion-deformation processes with and without crack initiation and growth in solids is presented along with representative model problems implemented by means of the finite element method.

Large deformation kinematics are outlined to build a terminological basis for the subsequent treatment; notions of point maps and strain measures are introduced with some details on how these quantities are embedded into differential geometry in modern continuum mechanics. Classical Newtonian balance laws are formulated for the mechanical part, along with micro-force and fluid mass balances that lead to an evolution of fluid concentration and possibly regularized phase boundaries, driven by the chemical potential. Thermodynamic rationales are outlined with an emphasis on concepts for non-equilibrium states. These are approached by adjusting the length scale of equilibrium postulates to infinitesimal unit cells. These cells allow for piecewise homogeneous state variables with global inhomogeneities, the latter being responsible for kinetic rate processes like solute particle transport.

Constraints of the entropy inequality under isothermal conditions are then applied to isotropic material laws together with other classical principles of constitutive theory. Further sensible restrictions are made to prepare for an elastic Neo-Hooke solid material response at large strains and the linear diffusion model named Fick's law.

Aside, the phase-field approach to brittle fracture mechanics is introduced as a separate variational statement that seeks for the minimum of a crack surface functional. Regularization of the sharp crack is the key feature that circumvents the handling of discontinuities and enables a simple implementation in standard finite element solvers. In terms of constitutive functions, a quadratic crack surface density function is chosen together with an irreversibility constraint that a priori prevents self-healing. This approach represents a gradient damage theory applied to brittle fracture.

Three specifications of the above theory are presented. First, the Cahn-Hilliard equations is studied in a rigid, undamaged domain, i.e., in the absence of elastic deformation and cracks. The fourth order differential equation is shown to be the Euler equation of a minimization formulation. A related saddle point principle is furthermore shown to yield the common split into two second order

differential equations with chemical potential and normalized concentration as primary fields. The underlying continuous potential is transformed into its time-discrete incremental counterpart and discretized in space. Using standard linear shape functions, a finite element solution scheme is presented for binary mixture coarsening as a representative numerical example.

Second, the problem statement is altered towards standard downhill diffusion (by exclusion of concentration gradient terms in the constitutive free energy) and coupled to large elastic deformation. Again, the governing equations appear as Euler equations of variational statements, both continuous and discrete in time and space. Complementary to the classical saddle point principle, a new minimization formulation is presented that implies the fluid mass balance as a constraint equation and is suitable for an inverse form of Fick's law. As a consequence, the chemical potential is replaced by the flux as a primary variable, which imposes challenges when it comes to the application of boundary conditions. An additional complexity stems from non-standard finite element spaces that match the required $H(\mathrm{Div}, \mathcal{B})$ conformity. Despite these technical challenges, the proposed formulation is independent of the LBB condition that constrains conventional approaches. In order to study the capabilities of the theory, hydrogels are chosen due to the tight coupling of possibly huge volumetric deformation to the solute particle content. Swelling is hence studied extensively, along with different conforming and non-conforming element designs.

Third, crack growth in hydrogels is analyzed by a combination of the above theory with the phase-field method for brittle fracture. Not surprisingly, the problem is embedded into a variational structure, and both implementations of minimization and saddle point formulation are presented by means of mechanically and diffusion-driven boundary value problems. Large delay of crack evolution and initiation and growth of fracture surfaces due to volumetric shrinkage can be captured. The solution scheme relies upon the efficient combination of an energetic history field and an operator split, with well-known sensitivities to parameter choices being a shortcoming of the method.

In summary, the distinct novelties of these contributions are: existence and specification of incremental potentials for the Cahn-Hilliard equation and a minimization formulation for diffusion-deformation processes in hydrogels, by itself and combined with the phase-field method for fracture. The most complete aspect is the second one, as the additional complexities of the new minimization method can largely be solved, and experimental data is referred to for an elementary model validation.

The other two aspects allow for further improvements. Most urgently, symmetric tangent matrices should be exploited as one of the main advantages of variational formulations. While this benefit is often announced (see e.g. Miehe et al. [2014a]), its exploitation is implemented on the element level by computing only half of the *local* tangent, which is a computational shortcut and not comparable to

the prospective speed-up of iterative solvers for sparse, symmetric and possibly positive definite *global* tangent matrices[24] The Cahn-Hilliard equation is an excellent candidate for such an uncompromising reduction of the computational cost enabled by the symmetry of variational formulations.

Besides, the staggered solution strategy for phase-field fracture problems deserves amendments, since the parameter sensitivity is unacceptable in the long run (and will be kept in bad memory by the author). Luckily, this known issue is subject to current research and might substantially be improved soon, see e.g. Gerasimov & Lorenzis [2016] for a line search approach. A future detail for the diffusive crack modeling should also be the implementation of conforming finite element spaces for the minimization formulation and interpolations more suitable for the saddle point structure and the associated LBB condition. While this is realized for the crack-free diffusion-deformation problem, incorporation of the damage phase-field into these element formulations is an open task. One further important aspect for future work is the consideration of modified fluid transport in cracks. Related models were proposed quite recently for the modeling of hydraulic fracturing in porous media (see e.g. Wilson & Landis [2016] or Miehe et al. [2015a]), well suited for the failure modeling of gels as well.

Last, experimental data on hydrogel specimen that fail during mechanical loading, swelling or shrinkage can hardly be found in the literature. The boundary value problems in Section 7.4 can hence be understood as a feasibility study capturing certain mechanisms characteristic to gels, but do not represent a thorough validation of a reliable computational model. Future work should ideally incorporate both experiments and modeling in order to justify the advanced model for multiple effects and relate its predictive capabilities to realistic scenarios.

Despite these shortcomings, the models proposed in this work are a step towards the reliable computational modeling of diffusion phenomena in elastic, highly deformable and possibly failing solids. Numerically challenging buckling modes as well as crack evolution can be captured, and the minimization formulation shows a direct, advantageous influence of the abstraction that ships with variational principles on actual finite element simulations.

[24] The finite element tool `Feap` might need to be sacrificed for the sake of including innovations from related disciplines. Whether such a step appears feasible or too disruptive is an open question.

Bibliography

Alberty, R. A.: Legendre transforms in chemical thermodynamics. *The Journal of Chemical Thermodynamics* **29** (1997), 501 – 516.

Ambati, M.; Gerasimov, T. & De Lorenzis, L.: A review on phase-field models of brittle fracture and a new fast hybrid formulation. *Computational Mechanics* **55** (2015), 383–405.

Ambati, M.; Kruse, R. & De Lorenzis, L.: A phase-field model for ductile fracture at finite strains and its experimental verification. *Computational Mechanics* **57** (2016), 149–167, ISSN 1432-0924.

Ambrosio, L. & Tortorelli, V.: Approximation of functionals depending on jumps by elliptic functionals via gamma-convergence. *Communications on Pure and Applied Mathematics* **43** (1990), 999–1036.

Anand, L.: A Cahn-Hilliard-type theory for species diffusion coupled with large elastic-plastic deformations. *Journal of the Mechanics and Physics of Solids* **60** (2012), 1983–2002.

Anders, D.; Hesch, C. & Weinberg, K.: Computational modeling of phase separation and coarsening in solder alloys. *International Journal of Solids and Structures* **49** (2012), 1557 – 1572.

Anders, D. & Weinberg, K.: Numerical simulation of diffusion induced phase separation and coarsening in binary alloys. *Computational Materials Science* **50** (2011), 1359 – 1364.

Anjam, I. & Valdman, J.: Fast MATLAB assembly of FEM matrices in 2d and 3d: Edge elements. *Applied Mathematics and Computation* **267** (2015), 252–263.

Arnaudov, N.: *Modeling of Swelling-Induced Mechanical Instabilities in Polymeric Hydrogels.* Bachelor thesis, Institut of Applied Mechanics (CE) Chair I, University of Stuttgart 2014.

Arnold, D. N.; Brezzi, F. & Fortin, M.: A stable finite element for the Stokes equations. *Calcolo* **21** (1984), 337–344.

Babuška, I.: Error-bounds for finite element method. *Numer. Math.* **16** (1971), 322–333.

Babuška, I.: The finite element method with Lagrangian multipliers. *Numer. Math.* **20** (1973), 179–192.

Bach, A.: *Stahlbetonbauteile unter kombinierten statischen und detonativen Belastungen in Experiment, Simulation und Bemessung*. Dissertation, Technische Universität Dresden 2015.

Başar, Y. & Weichert, D.: *Nonlinear Continuum Mechanics of Solids*. Springer 2000.

Bai, G. & Suzuki, A.: Phase separation of weakly ionized polymer gels during shrinking phase transition. *The Journal of Chemical Physics* **111** (1999), 10338–10346.

Bai, G. & Suzuki, A.: Shrinking patterns and phase coexistence of polymer gels. *Materials & Design* **21** (2000), 547–550.

Baierlein, R.: The elusive chemical potential. *American Journal of Physics* **69** (2001), 423–434.

Balay, S.; Brown, J.; Buschelman, K.; Eijkhout, V.; Gropp, W.; Kaushik, D.; Knepley, M.; McInnes, L.; Smith, B. & Zhang, H.: *PETSc Users Manual*. Tech. Rep. ANL-95/11 - Revision 3.4, Argonne National Laboratory 2013.

Balay, S.; Gropp, W. D.; McInnes, L. C. & Smith, B. F.: Efficient management of parallelism in object oriented numerical software libraries. In Arge, E.; Bruaset, A. M. & Langtangen, H. P. (eds.): *Modern Software Tools in Scientific Computing*, Birkhä user, Boston 1997, pp. 163–202.

Ballhause, D. & Wallmersperger, T.: Coupled chemo-electro-mechanical finite element simulation of hydrogels: I. chemical stimulation. *Smart Materials and Structures* **17** (2008).

Balluffi, R.; Allen, S. & Carter, W.: *Kinetics of materials*. Wiley 2005.

Barrer, R. M.: *Diffusion in and through solids*. Cambridge University Press 1941.

Barrett, J.; Blowey, J. F. & Garcke, H.: Finite element approximation of the cahn-hilliard equation with degenerate mobility. *SIAM Journal on Numerical Analysis* **37** (2000), 286–318.

Baumberger, T.; Caroli, C. & Martina, D.: Solvent control of crack dynamics in a reversible hydrogel. *Nat Mater* **5** (2006), 552–555.

Baumberger, T.; Caroli, C.; Martina, D. & Ronsin, O.: Magic angles and cross-hatching instability in hydrogel fracture. *Phys. Rev. Lett.* **100** (2008), 178303.

Beebe, D. J.; Moore, J. S.; Bauer, J. M.; Yu, Q.; Liu, R. H.; Devadoss, C. & Jo, B.-H.: Functional hydrogel structures for autonomous flow control inside microfluidic. *Nature* **404** (2000), 588–590.

Belytschko, T. & Black, T.: Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering* **45** (1999), 601–620.

Belytschko, T.; Chen, H.; Xu, J. & Zi, G.: Dynamic crack propagation based on loss of hyperbolicity and a new discontinuous enrichment. *International Journal for Numerical Methods in Engineering* **58** (2003), 1873–1905.

Biot, M.: General theory of three-dimensional consolidation. *Journal of applied physics* **12** (1941), 155–164.

Böger, L.; Keip, M.-A. & Miehe, C.: Minimization and saddle-point principles for the phase-field modeling of fracture in hydrogels (2017a), dOI: 10.1016/j.commatsci.2017.06.010.

Böger, L. & Miehe, C.: On configurational-force-driven crack propagation and adaptive mesh refinement in finite inelasticity. *PAMM* **13** (2013), 77–78, ISSN 1617-7061.

Böger, L.; Nateghi, A. & Miehe, C.: Minimization- and saddle-point-based modeling of diffusion-deformation-processes in hydrogels. *PAMM* **16** (2016), 307–308, ISSN 1617-7061.

Böger, L.; Nateghi, A. & Miehe, C.: A minimization principle for deformation-diffusion processes in polymeric hydrogels: Constitutive modeling and fe implementation. *International Journal of Solids and Structures* **121** (2017b), 257–274, ISSN 0020-7683.

Bokstein, B. S.; Mendelev, M. I. & Srolovitz, D. J.: *Thermodynamics and kinetics in materials science: a short course*. Oxford university press, USA 2005.

Bonet, J. & Wood, R. D.: *Nonlinear continuum mechanics for finite element analysis*. Cambridge university press 1997.

Bonn, D.; Kellay, H.; Prochnow, M.; Ben-Djemiaa, K. & Meunier, J.: Delayed fracture of an inhomogeneous soft solid. *Science* **280** (1998), 265–267.

Borden, M. J.; Verhoosel, C. V.; Scott, M. A.; Hughes, T. J. R. & Landis, C. M.: A phase-field description of dynamic brittle fracture. *Computer Methods in Applied Mechanics and Engineering* **217-220** (2012), 77–95.

Bouklas, N.; Landis, C. M. & Huang, R.: A nonlinear, transient finite element method for coupled solvent diffusion and large deformation of hydrogels. *Journal of the Mechanics and Physics of Solids* **79** (2015), 21–43.

Bourdin, B.; Francfort, G. & Marigo, J.: Numerical experiments in revisited brittle fracture. *Journal of the Mechanics and Physics of Solids* **48** (2000), 797–826.

Bower, A. F.: *Applied mechanics of solids*. CRC press 2009.

Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique* **8** (1974), 129–151.

Brezzi, F.; Douglas, J. & Marini, L.: Two families of mixed finite elements for second order elliptic problems. *Numer. Math.* **47** (1985), 217–235.

Brezzi, F. & Fortin, M.: *Mixed and hybrid finite element methods*. Springer-Verlag 1991.

Bröcker, T. & Jänich, K.: *Einführung in die Differentialtopologie*. Springer Verlag 1973.

Bruck, S. D.: Aspects of three types of hydrogels for biomedical applications. *Journal of Biomedical Materials Research* **7** (1973), 387–404.

Cahn, J.: On spinoidal decomposition. *Acta Metallurgica* **9** (1961), 795–801.

Cahn, J. W. & Hilliard, J. E.: Free energy of a nonuniform system I. Interfacial free energy. *The Journal of Chemical Physics* **28** (1958), 258–267.

Calvert, P.: Gel sensors and actuators. *MRS Bulletin* **33** (2008), 207–212.

Carroll, S. M.: Lecture notes on general relativity. *arXiv preprint gr-qc/9712019* (1997).

Chen, J. & Park, K.: Synthesis and characterization of superporous hydrogel composites. *Journal of Controlled Release* **65** (2000), 73–82.

Chester, S. A. & Anand, L.: A coupled theory of fluid permeation and large deformations for elastomeric materials. *Journal of the Mechanics and Physics of Solids* **58** (2010), 1879–1906.

Chester, S. A. & Anand, L.: A thermo-mechanically coupled theory for fluid permeation in elastomeric materials: Application to thermally responsive gels. *Journal of the Mechanics and Physics of Solids* **59** (2011), 1978–2006.

Chester, S. A.; Di Leo, C. V. & Anand, L.: A finite element implementation of a coupled diffusion-deformation theory for elastomeric gels. *International Journal of Solids and Structures* **52** (2015), 1–18.

Choo, S.; Chung, S. & Lee, Y.: A conservative difference scheme for the viscous cahn-hilliard equation with a nonconstant gradient energy coefficient. *Applied Numerical Mathematics* **51** (2004), 207 – 219.

Claßen, M.: *Zum Trag- und Verformungsverhalten von Verbundträgern mit Verbunddübelleisten und großen Stegöffnungen*. Dissertation, Institut für Massivbau, RWTH Aachen 2016.

Coleman, B. D. & Gurtin, M. E.: Thermodynamics with internal state variables. *The Journal of Chemical Physics* **47** (1967), 597–613.

Coleman, B. D. & Noll, W.: The thermodynamics of elastic materials with heat conduction and viscosity. *Archive for Rational Mechanics and Analysis* **13** (1963), 167–178.

Cook, G. & Dickerson, R. H.: Understanding the chemical potential. *American Journal of Physics* **63** (1995), 737–742.

Cottrell, J.; Hughes, T. & Bazilevs, Y.: *Isogeometric Analysis: Toward Integration of CAD and FEA*. John Wiley & Sons 2009.

Crank, J.: *The mathematics of diffusion*. Oxford university press 1975.

Darcy, H.: *Les fontaines publiques de la ville de Dijon*. Dalmont, Paris 1856.

de Borst, R.: Numerical aspects of cohesive-zone models. *Engineering Fracture Machanics* **70** (2003), 1743–1757.

Di Leo, C. V.; Rejovitzky, E. & Anand, L.: A cahn–hilliard-type phase-field theory for species diffusion coupled with large elastic deformations: Application to phase-separating li-ion electrode materials. *Journal of the Mechanics and Physics of Solids* **70** (2014), 1 – 29, ISSN 0022-5096, URL `http://www.sciencedirect.com/science/article/pii/S0022509614000842`.

Doi, M.: Gel dynamics. *Journal of the Physical Society of Japan* **78** (2009).

Dolbow, J.; Fried, E. & Ji, H.: Chemically induced swelling of hydrogels. *Journal of the Mechanics and Physics of Solids* **52** (2004), 51–84.

Duda, F. P.; Souza, A. C. & Fried, E.: A theory for species migration in a finitely strained solid with application to polymer network swelling. *Journal of the Mechanics and Physics of Solids* **58** (2010), 515–529.

Spiecker gen. Döhmann, F.: *Clinical and radiological risk stratification of patients with acute pulmonary embolism*. Dissertation, Georg-August-Universität Göttingen 2014.

Ehlers, W. & Acartürk, A.: The role of weakly imposed dirichlet boundary conditions for numerically stable computations of swelling phenomena. *Computational Mechanics* **43** (2008), 545–557.

Elliott, C. & French, D.: Numerical studies of the cahn-hillard equation for phase-separation. *IMA Journal of Applied Mathematics* **38** (1987), 97–128.

Elliott, C.; French, D. & Milner, F.: A second order splitting method for the cahn-hilliard equation. *Numerische Mathematik* **54** (1989), 575–590.

Elliott, C. & Zheng, S.: On the cahn-hilliard equation. *Archive for Rational Mechanics and Analysis* **96** (1986), 339–357.

Eshelby, J. D.: The force on an elastic singularity. *Philosophical Transactions of the Royal Society London A* **224** (1951), 87–112.

Eshelby, J. D.: Energy relations and the energy-momentum tensor in continuum mechanics. *Inelastic Behavior of Solids* (1970), 77–115.

Feng, X. & Prohl, A.: Analysis of a fully discrete finite element method for the phase field model and approximation of its sharp interface limits. *Mathematics of Computation* **73** (2004), pp. 541–567.

Fermi, E.: *Thermodynamics*. Dover Publications 1956.

Fick, A.: Ueber diffusion. *Annalen der Physik* **170** (1855), 59–86.

Flory, P. J.: Statistical mechanics of swelling of network structures. *The Journal of Chemical Physics* **18** (1950), 108–111.

Flory, P. J. & Rehner Jr., J.: Statistical mechanics of cross-linked polymer networks ii. swelling. *The Journal of Chemical Physics* **11** (1943), 521–526.

Francfort, G. A. & Marigo, J. J.: Revisiting brittle fracture as an energy minimization problem. *Journal of the Mechanics and Physics of Solids* **46** (1998), 1319–1342.

Galaev, I. Y. & Mattiasson, B.: Smart polymers and what they could do in biotechnology and medicine. *Trends in Biotechnology* **17** (1999), 335–340.

Gerasimov, T. & Lorenzis, L. D.: A line search assisted monolithic approach for phase-field computing of brittle fracture. *Computer Methods in Applied Mechanics and Engineering* **312** (2016), 276 – 303, ISSN 0045-7825, URL `http://www.sciencedirect.com/science/article/pii/S0045782515004235`.

Ghez, R.: *A primer of diffusion problems*. Wiley 1988.

Göküzüm, F.: *Configurational-Force-Driven Crack Propagation in Viscoelastic Solids*. Bachelor thesis, Institut of Applied Mechanics (CE) Chair I, University of Stuttgart 2013.

Golus, P.: *Numerische Untersuchungen zur mitwirkenden Plattenbreite für Querkraft von Fahrbahnplatten ohne Querkraftbewehrung*. Diploma thesis, Institut für Massivbau, RWTH Aachen 2011.

Gomez, H.; Calo, V. M.; Bazilevs, Y. & Hughes, T.: Isogeometric analysis of the cahn-hilliard phase-field model. *Computer Methods in Applied Mechanics and Engineering* **197** (2008), 4333 – 4352.

Gong, J. P.: Why are double network hydrogels so tough? *Soft Matter* **6** (2010), 2583–2590.

Graham, T.: Liquid diffusion applied to analysis. *Philosophical Transactions of the Royal Society of London* **151** (1861), 183–224.

Griffith, A. A.: The phenomena of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London. Seria A, Containing Papers of a Mathematical or Physical Character* **221** (1921), 582–593.

de Groot, S. R. & Mazur, P.: *Non-equilibrium thermodynamics*. Dover 1984.

Gurtin, M. E.: On a nonequilibrium thermodynamics of capillarity and phase. *Quarterly of applied mathematics* **47** (1989), 129–145.

Gurtin, M. E.: Generalized ginzburg-landau and cahn-hilliard equations based on a microforce balance. *Physica D: Nonlinear Phenomena* **92** (1996), 178–192.

Gyarmati, I.: *Non-equilibrium thermodynamics*. Springer 1967.

Haase, R.: *Thermodynamik der irreversiblen Prozesse*. Dr. Dietrich Steinkopff 1964.

Hakim, V. & Karma, A.: Laws of crack motion and phase-field models of fracture. *Journal of the Mechanics and Physics of Solids* **57** (2009), 342–368.

Halphen, B. & Nguyen, Q. S.: Sur les matéraux standards généralisés. *Journal de Mécanique* **40** (1975), 39–63.

Hansbo, A. & Hansbo, P.: An unfitted finite element method, based on nitsche's method, for elliptic interface problems. *Computer Methods in Applied Mechanics and Engineering* **191** (2002), 5537–5552.

Haupt, P.: *Continuum Mechanics and Theory of Materials*. Springer 2002.

Heider, Y. & Markert, B.: A phase-field modeling approach of hydraulic fracture in saturated porous media. *Mechanics Research Communications* **80** (2017), 38 – 46, ISSN 0093-6413.

Henry, W.: Experiments on the quantity of gases absorbed by water, at different temperatures, and under different pressures. *Philosophical Transactions of the Royal Society of London* **93** (1803), 29–274.

Hesch, C. & Weinberg, K.: Thermodynamically consistent algorithms for a finite-deformation phase-field approach to fracture. *International Journal for Numerical Methods in Engineering* (2014).

Hoffman, A. S.: Hydrogels for biomedical applications. *Advanced Drug Delivery Reviews* **64** (2012), 18–23.

Holmes, D. P.; Roché, M.; Sinha, T. & Stone, H. A.: Bending and twisting of soft materials by non-homogenous swelling. *Soft Matter* **7** (2011), 5188–5193.

Hong, W.; Liu, Z. & Suo, Z.: Inhomogeneous swelling of a gel in equilibrium with a solvent and mechanical load. *International Journal of Solids and Structures* **46** (2009), 3282–3289.

Hong, W. & Wang, X.: A phase-field model for systems with coupled large deformation and mass transport. *Journal of the Mechanics and Physics of Solids* **61** (2013), 1281–1294.

Hong, W.; Zhao, X.; Zhou, J. & Suo, Z.: A theory of coupled diffusion and large deformation in polymeric gels. *Journal of the Mechanics and Physics of Solids* **56** (2008), 1779–1793.

Huggins, M. L.: Thermodynamic properties of solutions of long-chain compounds. *Annals of the New York Academy of Sciences* **43** (1942), 1–32.

Hui, C.-Y.; Lin, Y. Y.; Chuang, F.-C.; Shull, K. R. & Lin, W.-C.: A contact mechanics method for characterizing the elastic properties and permeability of gels. *Journal of Polymer Science Part B: Polymer Physics* **44** (2006), 359–370.

Irgens, F.: *Continuum mechanics*. Springer Science & Business Media 2008.

Itskov, M.: *Tensor algebra and tensor analysis for engineers*. Springer 2007.

Jänich, K.: *Vektoranalysis*. Springer 1993.

Job, G. & Herrmann, F.: Chemical potential—a quantity in search of recognition. *European Journal of Physics* **27** (2006), 353.

Jog, C. & Hutter, K.: Foundations and applications of mechanics, volume i: Continuum mechanics. *Applied Mechanics Reviews* **56** (2003), 51.

Johnson, K.: *Contact mechanics*. Cambridge University Press 1985.

Kestin, J.: Local-equilibrium formalism applied to mechanics of solids. *International Journal of Solids and Structures* **29** (1992), 1827 – 1836, ISSN 0020-7683.

Kong, H. J.; Wong, E. & Mooney, D. J.: Independent control of rigidity and toughness of polymeric hydrogels. *Macromolecules* **36** (2003), 4582–4588.

Kopecek, J.: Hydrogels: From soft contact lenses and implants to self-assembled nanomaterials. *Journal of Polymer Science Part A: Polymer Chemistry* **47** (2009), 5929–5946.

Kowalski, S. J.: *Thermomechanics of drying processes*. Springer, Berlin 2003, ISBN 3-540-00412-2.

Krischok, A. & Linder, C.: On the enhancement of low-order mixed finite element methods for the large deformation analysis of diffusion in solids. *International Journal for Numerical Methods in Engineering* **106** (2016), 278–297, ISSN 1097-0207.

Kuhl, E.: *Theory and Numerics of Open System Thermodynamics*. Habilitation, Technische Universität Kaiserslautern 2004.

Kuhl, E. & Schmid, D.: Computational modeling of mineral unmixing and growth. *Computational Mechanics* **39** (2007), 439–451.

Kuhn, C. & Müller, R.: A continuum phase field model for fracture. *Engineering Fracture Machanics* **77** (2010), 3625–3634.

Kuhn, C.; Schlüter, A. & Müller, R.: On degradation functions in phase field fracture models. *Computational Materials Science* **108** (2015), 374 – 384, ISSN 0927-0256.

Lebon, G.; Jou, D. & Casas-Vázquez, J.: *Understanding non-equilibrium thermodynamics*. Springer 2008.

Linder, C. & Armero, F.: Finite elements with embedded strong discontinuities for the modeling of failure in solids. *International Journal for Numerical Methods in Engineering* **72** (2007), 1391–1433.

Liu, I.-S.: *Continuum mechanics*. Springer Science & Business Media 2013.

Liu, Z.; Hong, W.; Suo, Z.; Swaddiwudhipong, S. & Zhang, Y.: Modeling and simulation of buckling of polymeric membrane thin film gel. *Computational Materials Science* **49** (2010), S60–S64.

Lucantonio, A.; Nardinocchi, P. & Teresi, L.: Transient analysis of swelling-induced large deformations in polymer gels. *Journal of the Mechanics and Physics of Solids* **61** (2013), 205–218.

Mariano, P. M.: Multifield theories in mechanics of solids. *Advances in Applied Mechanics* **38** (2002), 1 – 93, ISSN 0065-2156.

Marsden, J. & Hughes, T.: *Mathematical Foundations of Elasticity*. Dover Publications 1994.

Maugin, G.: Internal variables and dissipative structures. *Journal of Non-Equilibrium Thermodynamics* **15** (2009), 173–192.

Maugin, G. A.: *The Thermomechanics of Nonlinear Irreversible Behaviors*. World Scientific Publishing 1999.

Mehrer, H.: *Diffusion in Solids*, vol. 155 of *Springer Series in Solid State Science*. Springer Series 2007.

Miehe, C.: A multi-field incremental variational framework for gradient-extended standard dissipative solids. *Journal of the Mechanics and Physics of Solids* **59** (2011), 898–923.

Miehe, C. & Gürses, E.: A robust algorithm for configurational-force-driven brittle crack propagation with r-adaptive mesh alignment. *International Journal for Numerical Methods in Engineering* **72** (2007), 127–155.

Miehe, C.; Hildebrand, F. & Böger, L.: Mixed variational potentials and inherent symmetries of the cahn–hilliard theory of diffusive phase separation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* **470** (2014a), 20130641.

Miehe, C.; Hofacker, M. & Welschinger, F.: A phase field model for rate-independent crack propagation: Robust algorithmic implementation based on operator splits. *Computer Methods in Applied Mechanics and Engineering* **199** (2010a), 2765–2778.

Miehe, C.; Mauthe, S. & Teichtmeister, S.: Minimization principles for the coupled problem of darcy–biot-type fluid transport in porous media linked to phase field modeling of fracture. *Journal of the Mechanics and Physics of Solids* **82** (2015a), 186–217.

Miehe, C.; Mauthe, S. & Ulmer, H.: Formulation and numerical exploitation of mixed variational principles for coupled problems of Cahn-Hilliard-type and standard diffusion in elastic solids. *International Journal for Numerical Methods in Engineering* **99** (2014b), 737–762.

Miehe, C.; Schänzel, L. & Ulmer, H.: Phase field modeling of fracture in multi-physics problems. Part I. balance of crack surface and failure criteria for brittle crack propagation in thermo-elastic solids. *Computer Methods in Applied Mechanics and Engineering* **294** (2015b), 449–485.

Miehe, C.; Teichtmeister, S. & Aldakheel, F.: Phase-field modelling of ductile fracture: a variational gradient-extended plasticity-damage theory and its micromorphic regularization. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **374** (2016), ISSN 1364-503X, URL `http://rsta.royalsocietypublishing.org/content/374/2066/20150170`.

Miehe, C.; Welschinger, F. & Hofacker, M.: Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field fe implementations. *International Journal for Numerical Methods in Engineering* **83** (2010b), 1273–1311.

Mikelic, A.; Wheeler, M. F. & Wick, T.: A phase field approach to the fluid filled fracture surrounded by a poroelastic medium. *ICES Report* **13-15** (2013).

Modica, L.: The gradient theory of phase transitions and the minimal interface criterion. *Annales de l'Institut Henri Poincare - Analyses Non Lineaire* **4** (1987), 487–512.

Mora, T. & Boudaoud, A.: Buckling of swelling gels. *The European Physical Journal E* **20** (2006), 119–124.

Müller, I. & Weiss, W.: *Entropy and Energy – A Universal Competition.* Springer 2005.

Müller, R.; Kolling, S. & Gross, D.: On configurational forces in the context of the finite element method. *International Journal for Numerical Methods in Engineering* **53** (2002), 1557–1574.

Müller, R. & Maugin, G. A.: On material forces and finite element discretizations. *Computational Mechanics* **29** (2002), 52–60.

Murdan, S.: Electro-responsive drug delivery from hydrogels. *Journal of Controlled Release* **92** (2003), 1–17.

Naficy, S.; Brown, H. R.; Razal, J. M.; Spinks, G. M. & Whitten, P. G.: Progress toward robust polymer hydrogels. *Australian Journal of Chemistry* **64** (2011), 1007–1025.

Nateghi, A.: *Variational Diffusion Coupled with Elasto-Dynamics and itsApplication to Simulate Hydrogels.* Master thesis, Institut of Applied Mechanics (CE) Chair I, University of Stuttgart 2014.

Newmark, N.: A method of computation for structural dynamics. *ASCE Journal of Engineering Mechanics Division* **85** (1959), 67–94.

Nguyen, Q.-S.: On standard dissipative gradient models. *Annals of Solid and Structural Mechanics* **1** (2010), 79–86, ISSN 1867-6944.

Nguyen, T. & Govindjee, S.: Numerical study of geometric constraint and cohesive parameters in steady-state viscoelastic crack growth. *International Journal of Fracture* **141** (2006), 255–268.

Oldroyd, J.: On the formulation of rheological equations of state. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, The Royal Society 1950, vol. 200, pp. 523–541.

Pandolfi, A. & Ortiz, M.: An efficient adaptive procedure for three-dimensional fragmentation simulations. *Engineering with Computers* **18** (2002), 148–159.

Papastavrou, A.; Steinmann, P. & Stein, E.: Enhanced finite element formulation for geometrically linear fluid saturated porous media. *Mechanics of Cohesive-Frictonal Materials* **2** (1997), 185–203.

Peppas, N. A.; Hilt, J. Z.; Khademhosseini, A. & Langer, R.: Hydrogels in biology and medicine: from molecular principles to bionanotechnology. *Advanced Materials* **18** (2006), 1345–1360.

Pfeifer, V.: *Identification of reactive oxygen species in iridium-based OER catalysts by in situ photoemission and absorption spectroscopy*. Dissertation, Technische Universität Berlin 2016.

Pizzocolo, F.; Huyghe, J. & Ito, K.: Mode i crack propagation in hydrogels is step wise. *Engineering Fracture Mechanics* **97** (2013), 72–79.

Rajagopal, A.; Fischer, P.; Kuhl, E. & Steinmann, P.: Natural element analysis of the cahn-hilliard phase-field model. *Computational Mechanics* **46** (2010), 471–493.

Rajagopal, K. R.: Diffusion through polymeric solids undergoing large deformations. *Materials Science and Technology* **19** (2003), 1175–1180.

Raoult, F.-M.: Loi générale des tensions de vapeur des dissolvants. *CR Hebd, Seances Acad. Sci* **104** (1887), 1430–1433.

Raviart, P. A. & Thomas, J. M.: Primal hybrid finite element methods for 2nd order elliptic equations. *Mathematics of computation* **31** (1977), 391–413.

Saad, Y. & Schultz, M.: Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* **7** (1986), 856–869.

Sandhu, R. & Wilson, E.: Finite-element analysis of seepage in elastic media. *Journal of the Engineering Mechanics Division* **95** (1969), 641–652.

Sato, N.: *Chemical energy and exergy: an introduction to chemical thermodynamics for engineers*. Elsevier 2004.

Schade, H. & Neemann, K.: *Tensoranalysis*. Walter de Gruyter 2009.

Schroeder, D. V.: *An introduction to thermal physics*, vol. 60. Addison Wesley New York 2000.

Simo, J. C.; Oliver, J. & Armero, F.: An analysis of strong discontinuities induced by strain-softening in rate-independent inelastic solids. *Computational Mechanics* **12** (1993), 277–296.

Sokolnikoff, I. S.: *Tensor analysis: theory and applications*. Wiley 1951.

Steinmann, P.: *Geometrical foundations of continuum mechanics*. Springer 2015.

Steinmann, P.; Maugin, G. A. & (Editors): *Mechanics of Material Forces*. Springer-Verlag 2005.

Stogner, R.; Carey, G. & Murray, B.: Approximation of cahn-hilliard diffuse interface models using parallel adaptive mesh refinement and coarsening with c1 elements. *International Journal for Numerical Methods in Engineering* **76** (2008), 636–661.

Sultan, E. & Boudaoud, A.: The buckling of a swollen thin gel layer bound to a compliant substrate. *Journal of Applied Mechanics* **75** (2008).

Svendsen, B.; Neff, P. & Menzel, A.: On constitutive and configurational aspects of models for gradient continua with microstructure. *ZAMM - Journal of Applied Mathematics and Mechanics* **89** (2009), 687–697, ISSN 1521-4001.

Tadmor, E. B.; Miller, R. E. & Elliott, R. S.: *Continuum mechanics and thermodynamics: from fundamental concepts to governing equations*. Cambridge University Press 2012.

Tanaka, H. & Sigehuzi, T.: Surface-pattern evolution in a swelling gel under a geometrical constraint: Direct observation of fold structure and its coarsening dynamics. *Physical Reviews E* **49** (1994), R39–R42.

Tanaka, T.; Sun, S.-T.; Hirokawa, Y.; Katayama, S.; Kucera, J.; Hirose, Y. & Amiya, T.: Mechanical instability of gels at the phase transition. *Nature* **325** (1987), 796–798.

Tanaka, Y.; Kuwabara, R.; Na, Y.-H.; Kurokawa, T.; Gong, J. P. & Osada, Y.: Determination of fracture energy of high strength double network hydrogels. *The Journal of Physical Chemistry B* **109** (2005), 11559–11562.

Taylor, R.: *FEAP – A Finite Element Analysis Program, Version 8.3 User Manual*. Tech. rep., University of California at Berkeley 2009.

Toh, K.-C. & Phoon, K.-K.: Comparison between iterative solution of symmetric and non-symmetric forms of biot's fem equations using the generalized jacobi preconditioner. *International Journal for Numerical and Analytical Methods in Geomechanics* **32** (2008), 1131–1146.

Tokarev, I. & Minko, S.: Stimuli-responsive hydrogel thin films. *Soft Matter* **5** (2009), 511–524.

Truesdell, C. & Toupin, R.: The classical field theories. In Flügge, S. (ed.): *Principles of Classical Mechanics and Field Theory*. Springer-Verlag, Berlin 1960, vol. III/1 of *Encyclopedia of Physics*, pp. 226–795.

Truesdell, W., C. & Noll: The nonlinear field theories of mechanics. In Fluegge, S. (ed.): *Handbuch der Physik Bd. III/3*. Springer-Verlag 1965.

Tu, L. W.: *An introduction to manifolds*. Springer Science & Business Media 2010.

Ubachs, R. L. J. M.; Schreurs, P. J. G. & Geers, M. G. D.: A nonlocal diffuseuse interface model for microstructure evolution of tin-lead solder. *Journal of the Mechanics and Physics of Solids* **52** (2004), 1763–1792.

Ulijn, R. V.; Bibi, N.; Jayawarna, V.; Thornton, P. D.; Todd, S. J.; Mart, R. J.; Smith, A. M. & Gough, J. E.: Bioresponsive hydrogels. *Materials Today* **10** (2007), 40–48.

Verhoosel, C. V. & de Borst, R.: A phase-field model for cohesive fracture. *International Journal for Numerical Methods in Engineering* **96** (2013), 43–62, ISSN 1097-0207.

Šilhavý, M.: *The Mechanics and Thermodynamics of Continuous Media*. Springer 1997.

Wallmersperger, T. & Ballhause, D.: Coupled chemo-electro-mechanical finite element simulation of hydrogels: II. electrical stimulation. *Smart Materials and Structures* **17** (2008).

Weber, L.: *Automatic Activation Spread during Heuristic Inference from Memory*. Diploma thesis, Philipps Universität Marburg 2014.

Wells, G.; Kuhl, E. & Garikipati, K.: A discontinuous galerkin method for the cahn-hilliard equation. *Journal of Computational Physics* **218** (2006), 860 – 877.

Wiberg, E.: *Die chemische Affinität: eine erste Einführung in die Lehre von der Triebkraft chemischer Reaktionen*. Walter de Gruyter 1972.

Wilmanski, K.: *Thermomechanics of continua*. Springer Science & Business Media 1998.

Wilson, Z. A. & Landis, C. M.: Phase-field modeling of hydraulic fracture. *Journal of the Mechanics and Physics of Solids* **96** (2016), 264–290.

Wittmund, W.: *Analysis and definition of optimization potentials for the implementation of derivatives, follow-up projects and facelifts of the S-Class derived from the lead-project.* Master thesis, Steinbeis Center of Management and Technology, Berlin 2015.

Wriggers, P.: *Nonlinear Finite Element Methods*. Springer-Verlag 2008.

Wu, T. & Lorenzis, L. D.: A phase-field approach to fracture coupled with diffusion. *Computer Methods in Applied Mechanics and Engineering* **312** (2016), 196 – 223, ISSN 0045-7825.

Xu, X. P. & Needleman, A.: Void nucleation by inclusion debonding in a crystal matrix. *Modelling and Simulation in Materials Science and Engineering* **1** (1993), 111–132.

Zhang, J.; Zhao, X.; Suo, Z. & Jiang, H.: A finite element method for transient analysis of concurrent large deformation and mass transport in gels. *Journal of Applied Physics* **105** (2009), 1–9.

Zhang, L.; Tonks, M.; Gaston, D.; Peterson, J.; Andrs, D.; Millett, P. & Biner, B.: A quantitative comparison between and elements for solving the cahn-hilliard equation. *Journal of Computational Physics* **236** (2013), 74 – 80.

Zhao, Y.; Stein, P. & Xu, B.-X.: Isogeometric analysis of mechanically coupled cahn–hilliard phase segregation in hyperelastic electrodes of li-ion batteries. *Computer Methods in Applied Mechanics and Engineering* **297** (2015), 325 – 347, ISSN 0045-7825, URL `http://www.sciencedirect.com/science/article/pii/S0045782515003023`.

Zhou, X. X.; Chow, Y. K. & Leung, C. F.: Hybrid and enhanced finite element methods for problems of soil consolidation. *International Journal for numerical methods in engineering* **69** (2007), 221–249.