Analysis of constitutive relations in the phase field approach to brittle fracture and its application to Hertzian indentation fracture

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Ing. Michael Strobl
aus Heidenheim an der Brenz

Tag der mündlichen Prüfung: 29.10.2024

Hauptreferent: Prof. Dr.-Ing. habil. Thomas Seelig
Korreferent: Prof. Dr.-Ing. habil. Ralf Müller

Karlsruhe 2025

„Man kann diese wunderbare Theorie nicht studieren, ohne bisweilen die Empfindung zu haben, als wohne den mathematischen Formeln selbständiges Leben und eigener Verstand inne, als seien sie klüger als wir, klüger sogar als ihre Erfinder, als gäben sie mehr heraus, als seinerzeit in sie hineingelegt wurde."

"We cannot study this wonderful theory without at times feeling as if an independent life and a reason of its own dwelt in these mathematical formulas; as if they were wiser than we were, wiser even than their discoverer; as if they gave out more than had been put into them."

Heinrich Rudolf Hertz[1]

[1] In 1889, Hertz remarked in a lecture referring to Maxwell's equations for electromagnetic light theory (Maxwell, 1865), "that an equation, once discovered, sometimes seems wiser than its discoverer" (Shour, 2021). The article containing the quote has been published as Hertz (1890), the English translation of the quote is taken from Boltzmann (1893).

Abstract

Phase field approaches to fracture have steadily gained popularity to compute complex crack patterns in statics and dynamics, including the formation of cracks and determination of unknown crack paths. The smooth approximation of cracks by introducing an additional scalar field, the phase field, and gradual degradation of material stiffness without the need to track the exact position of the crack are key features of the method, which is said to result in mesh independent solutions. However, current approaches inevitably exhibit some limitations which do not seem to be obvious and are therefore neglected in many investigations. The research presented in this dissertation focuses on the analysis of the approach and discusses specific constitutive choices along with the approach. For example, the assumption of isotropic stiffness degradation by the scalar phase field parameter does not capture the anisotropy introduced by a macroscopic crack and severely restricts the scope of application. Widely used tension-compression splits violate the specific conditions at cracks like traction-free crack surfaces and, in addition, reintroduce mesh orientation dependency. In order to satisfy crack boundary conditions the phase field approach has to be modified by taking the orientation of the crack into account.

As a second subject, the problem of Hertzian indentation fracture in brittle solids is analyzed. First, the concept of finite fracture mechanics is applied to investigate crack formation from an initially defect-free surface. The method assumes the spontaneous formation of a small, yet finite, crack and employs a stress-based as well as an energetic criterion. In evaluating the energetic part of the fracture criterion a semi-analytical and a numerical approach, the latter involving standard finite element simulations, are compared. The functionality of the hybrid (two-part) criterion in application to indentation fracture is analyzed in principle and, moreover, its predictive capabilities are illustrated by comparison with experimental findings. Then, the formation and subsequent growth of cracks in brittle solids caused by indentation loading is investigated in a holistic manner using the phase field approach. While this methodology, at first glance, appears to be a promising tool for handling problems like the present one including crack initiation in absence of a pre-crack, its actual application reveals a number of critical issues. Inherent shortcomings of common phase field models for fracture based on a variational structure are illustrated and appropriate modifications are discussed which render the method — at least to some extent — suitable for the simulation of indentation fracture. Particular emphasis is placed on the significance of the material's tensile strength and fracture toughness as independent physical properties necessary for the quantitative description of crack formation. Numerical predictions are compared with results obtained by finite fracture mechanics as well as experimental findings from the literature.

Keywords: fracture, crack evolution, phase field method, tension-compression decomposition, crack boundary condition, crack-induced anisotropy, crack surface function, critical phase field analysis, degradation function, phase field transition, crack formation, complex crack pattern, finite fracture mechanics, hybrid fracture criterion, Hertzian indentation fracture, Auerbach range

Kurzfassung

Die Phasenfeldmethode zur Beschreibung von komplexen statischen und dynamischen Bruchvorgängen hat stetig an Bedeutung gewonnen, einschließlich der Bildung von Rissen und der Bestimmung unbekannter Risspfade. Die glatte Approximation von Rissen durch die Einführung eines zusätzlichen skalarwertigen Feldes, des Phasenfeldes, und die sukzessive Reduktion der Materialsteifigkeit ohne die Notwendigkeit, die genaue Position des Risses zu verfolgen, sind die Hauptmerkmale der Methode, die zu netzunabhängigen Lösungen führen soll. Die derzeitigen Ansätze weisen jedoch zwangsläufig einige Einschränkungen auf, die nicht offensichtlich zu sein scheinen und daher in vielen Untersuchungen vernachlässigt werden. Die in dieser Dissertationsschrift vorgestellten Untersuchungen konzentrieren sich auf die Analyse der Methode und erörtern spezifische konstitutive Ansätze, die im Zusammenhang mit der Methode stehen. Zum Beispiel erfasst die Annahme einer isotropen Degradation der Steifigkeit durch den skalaren Phasenfeldparameter nicht die durch einen makroskopischen Riss erzeugte Anisotropie und schränkt den Anwendungsbereich stark ein. Die weit verbreiteten Zug-Druck-Zerlegungen verletzen die spezifischen Randbedingungen an Rissen, wie z.B., dass Oberflächen frei von Zugspannungen sein müssen, und führen darüber hinaus wieder eine Abhängigkeit von der Ausrichtung des Netzes ein. Um Rissrandbedingungen zu erfüllen muss die Phasenfeldmethode modifiziert werden, indem die Orientierung des Risses berücksichtigt wird.

Als zweites Thema wird das Problem des Hertz'schen Kegelrisses in spröden Festkörpern mit zwei verschiedenen Ansätzen betrachtet. Zunächst wird das Konzept der finiten Bruchmechanik angewendet, um die Rissbildung von einer ursprünglich defektfreien Oberfläche aus zu untersuchen. Die Methode geht von der spontanen Bildung eines kleinen, aber endlich langen Risses aus und verwendet sowohl ein spannungsbasiertes als auch ein energetisches Bruchkriterium. Bei der Bewertung des energetischen Teils des Bruchkriteriums werden ein semi-analytischer und ein numerischer Ansatz verglichen, wobei letzterer auf Simulationen mit einer gewöhnlichen Finite-Elemente-Methode basiert. Die Funktionsweise des hybriden (zweiteiligen) Kriteriums bei der Anwendung auf den Kegelriss wird analysiert, darüber hinaus wird dessen Vorhersagefähigkeit durch den Vergleich mit experimentellen Ergebnissen aufgezeigt. Anschließend wird die Entstehung und das anschließende Wachstum von Rissen, welche durch Eindruckbelastung auf spröde Festkörper verursacht werden, mit Hilfe der Phasenfeldmethode ganzheitlich untersucht. Während diese Methode auf den ersten Blick ein vielversprechendes Werkzeug für die Behandlung von Problemen wie dem vorliegenden zu sein scheint, einschließlich der Simulation der Rissbildung bei fehlendem Anfangsriss, offenbart die tatsächliche Anwendung

eine Reihe von kritischen Punkten. Die Unzulänglichkeiten gängiger Phasenfeldmodelle, welche auf einer variationellen Struktur beruhen, zur Simulation von Bruchvorgängen werden veranschaulicht und geeignete Modifikationen diskutiert, die der Methode – zumindest bis zu einem gewissen Grad – die Simulation von Kegelrissen erlauben. Besonders hervorgehoben wird die Bedeutung der Zugfestigkeit und der Bruchzähigkeit des Materials als unabhängige physikalische Größen, die für die quantitative Beschreibung der Rissbildung maßgeblich sind. Die numerischen Vorhersagen werden mit Ergebnissen der finiten Bruchmechanik sowie mit experimentellen Erkenntnissen aus der Literatur verglichen.

Schlüsselwörter: Bruch, Risswachstum, Phasenfeldmethode, Zug-Druck-Zerlegung, Rissrandbedingung, rissinduzierte Aniotropie, Rissoberflächenfunktion, kritische Phasenfeldanalyse, Degradationsfunktion, Phasenfeldübergang, Rissentstehung, komplexes Rissmuster, finite Bruchmechanik, hybrides Bruchkriterium, Hertz'scher Kegelriss, Auerbach-Bereich

Acknowledgments

This thesis presents the results of my research carried out during my time as a research associate and doctoral candidate at the Institute of Mechanics (IfM) of the Karlsruhe Institute of Technology (KIT). Along the way, I had the great privilege of gaining invaluable insights and meeting remarkable individuals. At this point, I would like to thank a number of people for their support — not only during the periods of intensive work, but throughout the entire journey.

First and foremost, I would like to express my deepest gratitude to Prof. Thomas Seelig for introducing me to the compelling field of fracture mechanics and for supporting my scientific endeavors. I greatly value his vast knowledge and experience, his commitment to precision, and the countless hours of insightful discussions we shared. His encouragement and long-standing dedication to advancing my research have been a constant source of motivation. I am truly thankful for his invaluable scientific guidance, thoughtful suggestions, and infinite patience, which made the successful completion of this thesis possible. Dear Thomas, thank you for your time, your inspiration, and your generous and unwavering support.

My special thanks also go to Prof. Ralf Müller (Technical University of Darmstadt) for his genuine interest in my work and for serving as the second examiner of this thesis. Furthermore, I would like to thank Prof. Werner Wagner, Prof. Frank Schilling, Prof. Thomas Böhlke, and Prof. Steffen Freitag, who kindly agreed to serve on the doctoral examination committee, as well as Prof. Peter Betsch, who additionally acted as chairman. Their ideas, helpful suggestions, and constructive comments contributed to the improvement of my work.

I am sincerely grateful to Prof. Krishnaswamy Ravi-Chandar for giving me the opportunity to work with him and his research group at the Center for Mechanics of Solids, Structures and Materials at the University of Texas at Austin. The discussions with him deepened my understanding of fracture mechanics, broadened my perspective on research, and made my six-month stay a truly memorable experience. I also thank Prof. Chad Landis and Dr. Khai Hong Pham for the stimulating discussions on various aspects of phase field modeling and Dr. Mazen Diab for his insights. In this regard, I gratefully acknowledge the Texas Advanced Computing Center (TACC) for providing access to extensive computing resources, and the Karlsruhe House of Young Scientists (KHYS) for their financial and administrative support.

Furthermore, I am truly thankful to Prof. Karl Schweizerhof for his long-standing interest in my work, for mentoring me during my Civil Engineering studies, and for guiding me

into the challenging and fascinating field of computational mechanics. I would also like to thank apl. Prof. Alexander Konyukhov and Ridvan Izi, especially for the time and effort they have invested in me since 2007.

I extend my gratitude to all my former colleagues at the institute for creating a friendly and pleasant working environment. In particular, I thank Dr. Marlon Franke, Dr. Jonas Hund, Dr. Christoph Schmied, and Dr. Yinping Yang for their excellent collaboration in teaching. I am very grateful to Lukas Morand and Piotr Dowgiałło for contributing to my research through their master's theses. Additionally, I thank the former student assistants Emilia von Fritsch, Martin Utz, Thank Thank Nguyen, and Alexandra Wahn, as well as the numerous student tutors for their work. Special thanks go to my long-time office colleague Dr. Simon Martinez Choy for the enjoyable time we shared, Dr. Jürgen Lenz for his always welcome visits to our office, and Johann Bitzenbauer for his thoughtful engagement in both technical and informal discussions. Moreover, I thank Rosemarie Krikis, Gabriele Herrmann, Klaus Neidhardt, Willi Wendler, Heiko Bendler, and Florian Schrammer for their continuous administrative and technical support.

Finally — and most importantly — I would like to thank my friends and entire family for their unconditional love, steadfast support, and constant encouragement. I dedicate this thesis to my beloved parents — my father Georg, who passed away too soon, and my mother Christa.

Karlsruhe, June 2025 Michael Strobl

Contents

1 Introduction

1.1 Motivation

Flaws of different size and geometry are present in almost all solid bodies. When subjected to loads, flaws may act as a nucleus for fracture, i.e. the separation of a body by one or several cracks. But even in the absence of flaws or initial cracks, fracture may be initiated from stress concentrations. Possible fracture is often a critical and thus a limiting constraint in engineering applications, e.g., in iron-based structures (Kirkaldy, 1863), fluid and gas pipelines (Ives et al., 1974) and reactor pressure vessels (Cheverton et al., 1981). In the worst case, the formation and growth of cracks eventually lead to catastrophic failures like the spectacular loss of T1 and T2 tankers and Liberty ships in the 1940s (Kobayashi and Onoue, 1943; Williams and Meyerson, 1949) and two airplane crashes of the de Havilland Comet in 1954 (Cohen, 1955). These well-known early incidents have promoted research activities in the field of fracture mechanics significantly.

The fundamental question if, when, and how cracks propagate is always present as well as the question whether inherent flaws will become critical or not. While separation of material is a process at the microscopic scale, fracture of bodies in engineering applications is typically treated by geometrically and physically idealized concepts on the macroscopic scale. Therefore, they are based on the well-developed concepts of continuum mechanics. Since the direct application of stress or strain to characterize the loading condition at crack tips usually does not succeed, i.e. they may be infinitely large, a comprehensive theoretical framework including different fracture hypotheses has been established. For some specific cases, for instance, crack propagation of perfect pre-cracks in elastic bodies caused by quasi-static loads, analytical solutions exist. Apart from that, laboratory experiments are still one important source of information (e.g. Ravi-Chandar, 2004). Because experiments, depending on their exact set-up, may be difficult to achieve and are thus time-consuming and expensive, the use of computational methods attracts much interest. Since the first works on this subject (e.g. Chan et al., 1970; Levy et al., 1971; Tracey, 1971) different numerical approaches to simulate fracture have been developed. The ever increasing computing power has had an enormous effect on their development. But, although numerical methods have become a widely used and accepted tool, the modeling of fracture implying discontinuous and singular mechanical fields is a challenge, arising from missing theory, insufficient numerical approaches or missing computing power. Thus, both the methods and their application for solving specific fracture problems are still a very active and attractive field of research.

1.2 Background and state of the art

The field of fracture mechanics, as it is known today, is basically founded on the pioneering works of Griffith (1921) and Irwin (1957). More than one century ago, Alan Arnold Griffith established the fundamental criterion, that the amount of energy required for a crack to propagate, i.e. to separate material and create new surfaces, must be provided by the release of the potential energy of the mechanical system, quantified by the *energy release rate*. While this criterion, known as the *Griffith criterion*, is generally valid regarding the onset of crack growth, it does not provide any specific information about the crack path. Thus all possible configurations must be considered. A few years later, it was George Rankine Irwin who introduced fracture criteria of practical relevance. He separated the loading of cracks according to their opening modes as mode I ("tensile opening"), mode II ("in-plane shear"), and mode III ("out-of-plane shear" or "anti-plane shear"). This allowed to normalize the singular stress field in the vicinity of a stationary crack tip ($1/\sqrt{r}$-singularity with r being the radial distance from the crack tip) to describe the stress distribution in terms of a so-called *crack-tip field* and to quantify the loading by a *stress intensity factor*. Its comparison to the *fracture toughness*, a material parameter, determines the onset of straight (mode I) crack growth. For mixed mode I+II loading, different failure criteria exist. The most prominent ones are the maximum energy release rate criterion, the maximal hoop stress criterion (Erdogan and Sih, 1963), and the principle of local symmetry (Gol'dstein and Salganik, 1974). In the absence of perfect pre-cracks, e.g. at blunt notches or V-notches, weaker stress concentrations are present and the above criteria are not valid anymore. As a remedy, the concept of *finite fracture mechanics* (FFM) has been suggested, which considers the instantaneous formation of a crack of small, yet finite, length instead (e.g. Hashin, 1996; Leguillon, 2001; Taylor et al., 2005; Li and Zhang, 2006). In this context, Leguillon (2002), for example, proposed a hybrid (two-part) fracture initiation criterion based on two well-defined material parameters: the (tensile) strength and the (fracture) toughness (Weißgraeber et al., 2016).

Although in the context of structural mechanical analysis the *finite element method* (FEM) has become a widely used and extremely versatile numerical tool since its practical introduction in the 1950s (Argyris, 1955; Turner et al., 1956), its predictive capabilities regarding fracture processes are still limited and an active research topic. Over the years, a vast number of numerical methods have been developed to simulate the macroscopic behavior of a single or multiple cracks, including the complete separation of material as well as the computation of loading states at the crack front. In this respect, cracks are either introduced as discrete separation or modeled in a "smeared", continuous manner. A discrete crack representation is based on the tracking of the crack to incorporate the discontinuity, e.g. by making use of remeshing techniques. In the *extended finite element method* (X-FEM), mainly introduced by Moës et al. (1999) and named as such by Dolbow (1999), a discontinuous approximation is obtained by locally enriching the displacement field with discontinuous shape functions, still based on the partition of unity (Melenk and Babuška, 1996). In addition, the vicinity of the crack front is enriched with known analytical solutions, e.g. crack-tip fields. This method was very popular for a

while (Belytschko and Black, 1999; Daux et al., 2000; Dolbow et al., 2001; Belytschko et al., 2003; Huang et al., 2003; Zi and Belytschko, 2003), but appears at least from a technical point of view to be very demanding in three dimensions (Sukumar et al., 2000; Moës et al., 2002; Areias and Belytschko, 2005). In addition, as mentioned above, it is based on analytical solutions, thus it cannot predict situations apart from known solutions, e.g. from classical fracture theories. By contrast, a continuous crack representation is typically based on a continuum damage description (Kachanov, 1958; Rabotnov, 1959; Rabotnov, 1969) by introducing one or several additional order parameters that represent the state of damage, i.e. successive stiffness degradation with increased loading. The concept has been further developed for different applications, e.g., anisotropic media (Sidoroff, 1981), concrete (Bažant and Oh, 1983; Mazars and Lemaitre, 1985; Ortiz, 1985; Ramtani et al., 1992) and ductile materials (Coleman and Hodgdon, 1985). Since direct implementation into the finite element method results in mesh dependency (Lasry and Belytschko, 1988; Needleman, 1988), non-local or gradient enhanced continuum damage approaches have been developed (Hall and Hayhurst, 1991; Sluys, 1992; Bažant, 1994; Peerlings et al., 1995).

The *variational theory of fracture* has been introduced by the pioneering work of Francfort and Marigo (1998). It can be regarded as a generalization of Griffith's theory. However, only the spatial regularization of cracks by Bourdin (1998) made the method, nowadays known as the *phase field approach to fracture*, also accessible for a standard numerical treatment, i.e. with the finite element method. With its diffuse representation of the crack by a scalar field, it does not require tracking of cracks and shows some similarities to the non-local continuum damage models mentioned above. Because of dimensional reasons, spatial regularization introduces an additional *internal length* parameter. At the same time, there have been several attempts in the physics community to adapt phase transition models to fracture mechanics (e.g. Aranson et al., 2000; Karma et al., 2001; Henry and Levine, 2004; Karma and Lobkovsky, 2004). Although their structure is similar, they typically do not rely on the Griffith criterion. The phase field method has become a very popular computational technique for analyzing phenomena involving the formation of cracks as well as their propagation along a priori unknown paths including complex topologies such as branching without additional assumptions, simply by the solution of partial differential equations. In recent years, a number of notable modifications of the original method (Bourdin, 2007a; Bourdin, 2007b; Bourdin et al., 2000) have been developed. The quasi-static phase field approach to fracture has been successfully extended to the dynamic setting by including inertia terms (Borden et al., 2012; Hofacker and Miehe, 2012; Schlüter et al., 2014). Extensive studies on the properties of dynamic fracture like limit speeds have been presented by Schlüter et al. (2016) and Bleyer et al. (2017). To account for the different behavior of most brittle materials under tension and compression and to incorporate the effect of crack closure, several modifications, so-called *tension-compression splits* have been proposed. Basically motivated by similar approaches in continuum damage mechanics, the most prominent of them are based on the volumetric-deviatoric decomposition (Amor et al., 2009; Lancioni and Royer-Carfagni, 2009) and the spectral decomposition of the strain tensor (Freddi and Royer-Carfagni, 2010; Miehe et al., 2010c). Since they have

been assumed to be fully working, these splits have been used in phase field analyses all the time (e.g. Hofacker and Miehe, 2013; May et al., 2015; Kiendl et al., 2016; Bilgen et al., 2019; Mang et al., 2019). Also modifications of the spatial regularization of cracks have been proposed in different ways. Regularization by using higher-order phase field derivatives to obtain a better approximation accuracy and a higher rate of convergence has been proposed by Borden (2012) and was further discussed, e.g., by Borden et al. (2014) and Weinberg and Hesch (2015). Other types of the local crack description have been introduced by Pham et al. (2011) and Wu (2017), among others. The treatment of the related bound-constrained problem has been discussed by Pham et al. (2011), Wu (2017) and Gerasimov and De Lorenzis (2019). Modifications of the polynomial stiffness degradation function, mainly to modify the transition between elastic and crack surface energy, have been proposed and analyzed by Borden et al. (2012) and Kuhn et al. (2015). Other types of degradation functions are, for example, exponential functions (Wilson et al., 2013; Sargado et al., 2018; Steinke and Kaliske, 2019), rational functions (Lorentz and Godard, 2011) or trigonometric functions (Yin and Kaliske, 2019). The transition of the phase field variable has been analyzed for spatially homogeneous (e.g. Pham et al., 2011; Borden, 2012; Kuhn, 2013) and spatially inhomogeneous phase field solutions (e.g. Aranson et al., 2000; Karma et al., 2001; Henry and Levine, 2004; Benallal and Marigo, 2007; Pham and Marigo, 2008; Kuhn and Müller, 2013; Kuhn et al., 2015; Wu, 2017). Thereby, the role of the internal length parameter has been revealed to be connected to the strength of the material (e.g. Pham et al., 2011).

The phase field method for fracture has also been extended to, e.g. ductile fracture (Ambati et al., 2015b; Alessi et al., 2018a), rubbery polymers (Miehe and Schänzel, 2014; Schänzel, 2015), heterogeneous materials (Nguyen et al., 2015; Kuhn and Müller, 2016), cohesive fracture (Verhoosel and Borst, 2013; Vignollet et al., 2014; May et al., 2015; Wu, 2017), pressurized or hydraulic fracture (Wheeler et al., 2014; Mikelić et al., 2015; Wilson and Landis, 2016; Ehlers and Luo, 2017; Singh et al., 2018; Luo, 2019; Kienle and Keip, 2019), cracks in porous media (Cajuhi et al., 2017; Cajuhi et al., 2018; Cajuhi, 2019), thermal fracture (Schlüter et al., 2015), electromechanical fracture (Miehe et al., 2010b), and fatigue (Alessi et al., 2018b; Lo et al., 2019). In most of these applications, propagation of already existing cracks is regarded. Crack formation in the absence of a perfect pre-crack has been studied by using a phase field method from circular notches (Kuhn and Müller, 2013) and V-notches (Kuhn, 2013) and has been compared to experimental results and theoretical predictions based on finite fracture mechanics (Leguillon, 2001; Hebel, 2010). Both finite fracture mechanics and the phase field method were able to capture size effects, despite some deviations (Kuhn, 2013). Tanné et al. (2018) used a phase field approach to investigate crack formation at V-notches with varying opening angle (Leguillon and Sanchez-Palencia, 1987) with a good correlation to experimental data.

A far more complex situation of crack formation is obtained for the mixed mode I+III loading, where the crack front breaks up into multiple fragments (Pham and Ravi-Chandar, 2014; Pham and Ravi-Chandar, 2016). In addition to the original "parent" crack, numerous "daughter" cracks, also called echelon cracks, are formed by mode I along the crack front, most likely in the direction perpendicular to the maximum principal stress (Knauss,

1970). The echelon cracks are connected by a second type of daughter crack (Pham and Ravi-Chandar, 2016). In a comprehensive three-dimensional analysis, Pham (2015) successfully simulated with the phase field method the formation of echelon cracks as well as subsequent crack growth, shielding effects and facet coarsening. Although the specific phase field approach has been selected carefully, the spacing between the fragments has not been captured correctly (Pham and Ravi-Chandar, 2017).

Another challenging fracture problem is that of indentation fracture. Indentation fracture describes the formation and further evolution of cracks caused by the compression of a stiff indenter onto the initially defect-free surface of a brittle solid. Since the early investigations by Heinrich Rudolf Hertz at the end of the nineteenth century and later named after him, see e.g. Fischer-Cripps (2007) or Lawn (1993), it has been (and still is to date) subject of numerous experimental and theoretical studies. Much work on the subject, for instance, has been motivated by the practical application of indentation fracture for determining the fracture toughness of brittle materials such as glass and ceramics; see, e.g., the review articles by Ostojic and McPherson (1987), Lawn (1998) or Kocer (2003b). In the case of cylindrical or spherical indenters and isotropic materials, Hertzian fracture proceeds essentially in two stages: At some critical load a ring crack of short extension "spontaneously" forms at the free surface somewhat outside the contact region, while subsequently it grows in a stable manner into a cone-shaped crack of almost constant angle.

Early theoretical studies on indentation fracture, e.g. Frank and Lawn (1967), were essentially based on the analytically known pre-fracture indentation stress field in an elastic half-space. Yet, later numerical investigations have shown the necessity to account for changes of this stress field in the course of fracture in order to correctly predict the crack path. The latter was analyzed in the framework of finite element methods (Lawn et al., 1974; Yingzhi and Hills, 1991; Kocer and Collins, 1998; Kocer, 2003a), boundary element methods (Bush, 1999; Selvadurai, 2000; De Lacerda and Wrobel, 2002) and extended finite element methods (Tumbajoy-Spinel et al., 2013) as well as by semi-analytical techniques such as weight functions (Fett et al., 2004) or ring dislocation representations (Wang et al., 2008). It is important to note that in all these approaches the pre-existence of the initially forming ring crack was assumed so that classical fracture criteria could be employed to analyze its further growth. Conceptually more difficult than tracking the evolution of an existing crack is the prediction of the sudden formation of the ring crack at some critical load. From a fracture mechanics point of view, this is not surprising since crack formation lies beyond the scope of classical fracture hypotheses. Studies on indentation fracture initiation, therefore, either assumed the pre-existence of a statistical distribution of surface flaws, e.g. Fischer-Cripps (2007), or the axial extension of the ring crack was taken as a free parameter in an analysis involving crack propagation stability arguments, e.g. Mouginot and Maugis (1985). A detailed discussion of these approaches is not aimed in the present work and may be found, for instance, in the review article of Kocer (2003b). Yet, it should be mentioned that various researchers have conjectured a likelihood of the formation of shorter cracks closer to the contact region from the strong variation of the indentation stress field with distance from the indenter and depth from the free surface,

e.g. Lawn et al. (1974), Mouginot and Maugis (1985) and Fischer-Cripps (1997). It was also emphasized the prerequisite of a sufficiently strong tensile loading acting on flaws of a certain size in order to cause fracture; a sound criterion, however, remained lacking. Indentation fracture involves further interesting issues such as the influence of geometry and material properties on the location and extension of the initial ring crack as well as on the shape of the subsequently evolving cone crack. For instance, it is found that the ratio of the ring crack radius to the indenter radius decreases with increasing indenter radius, e.g. Mouginot and Maugis (1985), and that the cone crack angle mainly depends on Poisson's ratio of the fracturing solid, e.g. Kocer and Collins (1998). In addition, it was very early observed by Felix Auerbach, a contemporary of Hertz (e.g. Lawn, 1993; Fischer-Cripps, 2007), and later studied experimentally and theoretically (e.g. Tillett, 1956; Roesler, 1956b; Finnie et al., 1981; Mouginot and Maugis, 1985) that the dependence of the critical load on the indenter radius displays two distinct regimes, with one being commonly referred to as the *Auerbach range*.

1.3 Scope and outline of the work

The aim of this thesis is to provide a thorough and thus critical analysis of the phase field approach to brittle fracture and to address the problem of Hertzian indentation fracture — more specifically — the formation and growth of cracks caused by the compression of a stiff cylindrical indenter onto the surface of a brittle solid.

Related challenges are manifold: Initial crack formation in the course of Hertzian indentation fracture cannot be tackled by classical fracture hypotheses. Thus, the quest for the size of the crack ("critical flaw") initially created is to be solved by the concept of finite fracture mechanics (FFM). According to a coupled two-part fracture criterion, e.g. Leguillon (2002), crack formation is assumed to take place once the material's tensile strength is exceeded by the normal stress along this crack *and* the energy released per area by its formation exceeds the fracture toughness. Applied to indentation fracture initiation, FFM provides a unique determination of the size (axial length) and location (radius) of the spontaneously formed ring crack as well as the critical load — without any additional assumptions. The dependence of these quantities on the indenter radius includes a size effect commonly referred to as *Auerbach's law* in the literature on indentation fracture, e.g. Fischer-Cripps (2007), which is shown to be also captured by FFM. In this specific approach, the Auerbach range corresponds to the range where crack formation is controlled solely by the energetic criterion instead of both (energy and strength) criteria being active.

In the existing literature the two stages of indentation fracture, i.e. initial crack formation and subsequent crack propagation, have been analyzed separately. While FFM provides essentially a fracture initiation criterion, the holistic description of the entire process of Hertzian indentation fracture requires numerical simulation techniques. As mentioned in the previous section, the phase field approach is able to reproduce complex phenomena involving the formation of cracks as well as their propagation along a priori unknown

paths. It therefore appears to be well suited for studying situations such as indentation fracture. However, owing to the "smeared" description of fracture in phase field models, the role of fracture related physical parameters is much less clear than, for instance, in FFM. This is particularly so as these parameters occur in conjunction with several further phase field modeling ingredients such as degradation functions, the (non-local) internal length parameter, the crack surface function or the coupling between mechanical fields and crack driving force. These ingredients bring into play some arbitrariness and at the same time may lead to conflicts with physical requirements which unfortunately are seldom addressed in the vast body of literature on phase field fracture modeling. Thus, a profound understanding of the approach is necessary to apply it appropriately to the situation of interest. Only in a few recent publications have the limited predictive capabilities of this methodology been the subject of thorough investigations, e.g. Klinsmann et al. (2015) or Sargado et al. (2018). During the course of this work, it was found out that the application of so-called tension-compression splits in phase field approaches to fracture tends to fail in most situations by not predicting the macroscopic mechanical response of a crack correctly (Strobl and Seelig, 2015; Strobl and Seelig, 2016). Consequently, own results as well as results from literature including phase field induced cracks subjected to compression or shear have to be handled with caution. The purpose of the present study, therefore, is two-fold: on the one hand side, the entire process of indentation fracture is analyzed numerically, including crack initiation at some a priori unknown location as well as the subsequent evolution of the cone-shaped crack, thereby studying the influence of geometrical and material parameters. On the other hand, this non-standard fracture problem serves as a paradigm for a thorough assessment of the phase field methodology. This involves a critical investigation of various modeling aspects as well as the development of modifications necessary for the reliable simulation of indentation fracture. Parts of this work have been published as

> "Strobl M., Dowgiałło P., Seelig T., 2017; Analysis of Hertzian indentation fracture in the framework of finite fracture mechanics; Int. J. Frac. 206 (1), 67–79" ‡

and

> "Strobl M., Seelig T., 2020; Phase field modeling of Hertzian indentation fracture; J. Mech. Phys. Solids 143, 104026". ¶

The dissertation is organized as follows: Recent works on phase field modeling of brittle fracture and previous works on Hertzian indentation fracture are reviewed in **Chapter 1**. As a basis for the present work, the first section of **Chapter 2** provides a brief introduction to continuum mechanics, mainly focused on the small deformation theory which is relevant in subsequent chapters. Besides a summary of the essential balance equations and constitutive requirements of a solid body, the solution of mechanical problems is discussed and different variational principles are introduced. Selected fundamentals of fracture and damage mechanics are subject of the second section. In particular, the classical

‡ Reproduced with permission from Springer Nature.
¶ Reproduced with permission from Elsevier Ltd.

concepts of linear elastic fracture mechanics (LEFM) are presented with special attention to the energetic consideration introduced by Griffith, providing the basis of the phase field approach to fracture. Additionally, topics like cohesive behavior, continuum damage mechanics and the hybrid fracture initiation criterion are introduced. The latter can be used to determine crack initiation in absence of a well established pre-crack, e.g. from a blunt notch, a hole or from other stress concentrations. In addition, a variational treatment of fracture is presented, which provides the basis for the phase field approach to fracture.

Chapter 3 is devoted to the theoretical and numerical treatment of brittle fracture in the framework of phase field modeling. The spatial regularization of the variational approach to fracture allows to set up a general framework of phase field modeling, which is presented in the first two sections of the chapter. Within this framework, different modeling choices are available regarding the degradation of elastic stiffness, the distribution of surface energy in the vicinity of a crack and the incorporation of the irreversibility of fracture. These aspects are investigated and discussed in the subsequent sections. One key aspect of the work is the tension-compression asymmetric response of cracks. Since a crack introduces anisotropy aligned with its (normal) direction, the common isotropic stiffness degradation in the phase field approach fails in reproducing the behavior of cracks. Boundary conditions at cracks are discussed along with their incorporation into phase field models. The last section of the chapter is concerned with additional numerical investigations.

The fascinating problem of Hertzian indentation fracture is analyzed in **Chapter 4**. The specific indentation fracture problem setting along with the analysis of the stress field prior to fracture is provided in the first section. Based on the framework of finite fracture mechanics, crack formation in absence of a pre-crack is analyzed by applying a hybrid fracture initiation criterion in the second section. The study provides, without any additional assumptions, a new insight into the spontaneous formation of the initial ring crack, both qualitatively and quantitatively. In the third section, a holistic description of the entire process of Hertzian indentation fracture from nucleation to ultimate cone crack propagation by employing a phase field method is presented. For this purpose, essential modifications of ingredients of the phase field approach, also based on the previous investigations of the method, are discussed. Studying the influence of geometric and material parameters on the fracture process and a comparison of the results with experimental findings from the literature complement the chapter.

Finally, the main outcomes of this work and their implication on fracture modeling using the phase field method are summarized in **Chapter 5**. Based on the current findings, future research perspectives are identified.

Further details on different phase field ingredients, their computational treatment and numerical implementation are provided in the **Appendix**. Along with analytical solutions used to analyze the indentation fracture problem, such as the pre-fracture stress distribution, the appendix also provides results regarding the variation of the underlying hybrid fracture criterion.

2 Foundations of solid mechanics

This chapter is devoted to provide some fundamental concepts, principles and basic equations of solid mechanics which are used in subsequent chapters. It includes a brief introduction into the subjects of continuum mechanics, the challenges of fracture and damage and some insight into the numerical treatment of mechanical (initial) boundary value problems. In addition, the chapter provides an introduction to a variational approach to fracture.

2.1 Fundamentals of continuum mechanics

In this work, the mechanical behavior of solid bodies is described based on continuum mechanics. A continuum model assumes that bodies consist of an infinite number of material points (implying an at least piecewise continuous distribution of material). It ignores the fact that material is made of discrete atoms with discontinuities in-between. This assumption simplifies the mathematical description of the body and all related physical quantities drastically. Continuum approaches are widely used in physics and engineering approaches and mostly provide the basis of an accurate description as long as the mechanical effects are of length scales much greater than that of inter-atomic distances[1]. Thus, the continuum based approach does not rely on the microscopic and molecular structure and no detailed knowledge on these scales is needed. Besides fundamental physical balance laws as the conservation of mass, the conservation of linear and angular momentum and the conservation of energy, which are shortly introduced in the following, the study of static and dynamic body deformations and stresses is of main interest. The following sections makes reference to the widely used textbook of Holzapfel (2000). In addition, constitutive relations to complete the set of mechanical equations and to incorporate the specific behavior of materials are discussed. Throughout the work mainly the common bold face notation of vectors (usually indicated by lowercase letters) and tensors of second order (indicated by uppercase letters) is used while specific components

[1] Typical distances between the nuclei of atoms in solid materials are on the order of $1 \ldots 10$ Å (1 Å$= 10^{-10}$ m), see e.g. Daams et al. (2006) for metals and alloys. Brittle materials like glass possess an amorphous structure, which are characterized by greatly varying inter-atomic distances. However, they are still of the same order of size as in other solid materials (Shelby, 2005).

are occasionally indicated. Further details on definitions and the notation are provided in App. A.1.

2.1.1 Kinematics, strain and and stress measures

To introduce the basic kinematics, we regard a continuum body which continuously moves in space. The physical reason for the movement is irrelevant for now. The body moves from its initial position at a certain instant of time $t = t_0$ (which we take as reference) to its current position at a later instant of time $t > t_0$ as depicted in Fig. 2.1.

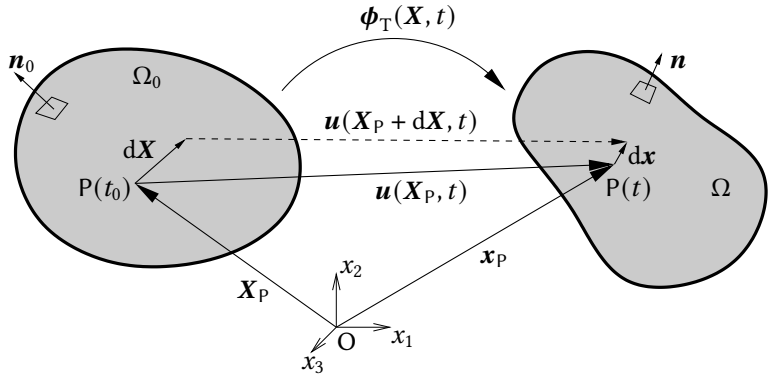


Figure 2.1: Solid body shown in reference placement (at time t_0) and current placement (at $t > t_0$) to illustrate the transformation of a particle $\mathsf{P}(\boldsymbol{X}_\mathrm{P}, t_0)$.

The position of every particle P inside the body is given with respect to a fixed origin O by the position vector $\boldsymbol{X}_\mathrm{P}$ in the reference placement and the position vector $\boldsymbol{x}_\mathrm{P}$ in the current placement. Thus, the current position of all particles is described by the generally non-linear vector field

$$\boldsymbol{x} = \boldsymbol{\phi}_\mathrm{T}(\boldsymbol{X}, t) \tag{2.1}$$

with respect to fixed material coordinates $\boldsymbol{X}$, the so-called material or Lagrangian description, which is favored for solid bodies and used in the following. Alternatively, by presuming a unique mapping between both placements, the inverse relation (spatial/Eulerian description) is given by the inverse motion $\boldsymbol{X} = \boldsymbol{\psi}_\mathrm{T}(\boldsymbol{x}, t)$ which is the description with respect to fixed spatial coordinates $\boldsymbol{x}$. The displacement vector field is defined by the difference between the current and initial positions of particles

$$\boldsymbol{u}(\boldsymbol{X}, t) := \boldsymbol{x} - \boldsymbol{X} = \boldsymbol{\phi}_\mathrm{T}(\boldsymbol{X}, t) - \boldsymbol{X} \tag{2.2}$$

and in general contains rigid body translation (homogeneous displacement) and rotation (homogeneous change of orientation) and, for deformable solids, also the change of shape and size of the body. To study the deformation, the motion of a particle located at an

infinitesimal distance $\mathrm{d}\boldsymbol{X}$ from particle $\boldsymbol{X}$ can be described by means of the Taylor series expansion in material coordinates as

$$\boldsymbol{x}(\boldsymbol{X}+\mathrm{d}\boldsymbol{X},t)=\boldsymbol{x}(\boldsymbol{X},t)+\underbrace{\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}}(\boldsymbol{X},t)}_{=:\boldsymbol{F}}\cdot\mathrm{d}\boldsymbol{X}+\underbrace{\frac{\partial^2 \boldsymbol{x}}{\partial \boldsymbol{X}^2}(\boldsymbol{X},t)\,\mathrm{d}\boldsymbol{X}^2+\ldots}_{\approx\boldsymbol{0}} \quad . \tag{2.3}$$

Since higher order derivatives vanish due to $|\mathrm{d}\boldsymbol{X}|^2 \approx 0$ the deformation related to a particle at position $\boldsymbol{X}$ is given by the first spatial derivative, the deformation gradient

$$\boldsymbol{F}(\boldsymbol{X},t):=\frac{\partial \boldsymbol{\phi}_{\mathrm{T}}(\boldsymbol{X},t)}{\partial \boldsymbol{X}}=\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}} \quad . \tag{2.4}$$

The deformation gradient is a fundamental quantity in nonlinear continuum mechanics and has in general nine distinct components (it is in general not symmetric since the tensor refers to both placements). It maps the infinitesimal line element $\mathrm{d}\boldsymbol{X}$ from the reference placement to its current length and orientation by

$$\mathrm{d}\boldsymbol{x}=\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{X} \quad . \tag{2.5}$$

An infinitesimal surface of area A_0 and orientation $\boldsymbol{n}_0$ (given by the normal of the surface) which is spanned by two infinitesimal line elements according to $\mathrm{d}A_0\,\boldsymbol{n}_0=\mathrm{d}\boldsymbol{X}_{\mathrm{A}}\times\mathrm{d}\boldsymbol{X}_{\mathrm{B}}$ is mapped by "Nanson's formula"

$$\mathrm{d}A\,\boldsymbol{n}=\mathrm{d}\boldsymbol{x}_{\mathrm{A}}\times\mathrm{d}\boldsymbol{x}_{\mathrm{B}}=\cdots=\det(\boldsymbol{F})\,\mathrm{d}A_0\,\boldsymbol{F}^{-\mathrm{T}}\cdot\boldsymbol{n}_0 \tag{2.6}$$

from the reference to the current placement. Accordingly, an infinitesimal volume element spanned by three infinitesimal line elements $\mathrm{d}V_0=(\mathrm{d}\boldsymbol{X}_{\mathrm{A}}\times\mathrm{d}\boldsymbol{X}_{\mathrm{B}})\cdot\mathrm{d}\boldsymbol{X}_{\mathrm{C}}$ is mapped by

$$\mathrm{d}V=(\mathrm{d}\boldsymbol{x}_{\mathrm{A}}\times\mathrm{d}\boldsymbol{x}_{\mathrm{B}})\cdot\mathrm{d}\boldsymbol{x}_{\mathrm{C}}=\cdots=\mathrm{d}V_0\,\underbrace{\det(\boldsymbol{F})}_{:=J} \tag{2.7}$$

from the reference to the current placement. The volume ratio, the Jacobian determinant, is

$$J:=\det(\boldsymbol{F})=\frac{\mathrm{d}V}{\mathrm{d}V_0} \quad . \tag{2.8}$$

To guarantee unique mapping between both placements (requires invertible functions as mentioned above), to preserve the orientation and to prevent unphysical self-penetration of material points the requirement $J>0$ (no negative volumes) must hold for every point. In addition to the deformation gradient (2.4), the closely-related displacement gradient

$$\boldsymbol{H}(\boldsymbol{X},t):=\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}}=\frac{\partial(\boldsymbol{x}-\boldsymbol{X})}{\partial \boldsymbol{X}}=\boldsymbol{F}-\boldsymbol{I} \quad . \tag{2.9}$$

is introduced. By the polar decomposition

$$\boldsymbol{F}=\boldsymbol{R}\cdot\boldsymbol{U}=\boldsymbol{V}\cdot\boldsymbol{R} \quad , \tag{2.10}$$

the deformation gradient can be multiplicatively decomposed into a proper orthogonal rotation tensor $\boldsymbol{R}$ and either the right stretch tensor $\boldsymbol{U}$ (which corresponds to stretch before rotation) or the left stretch tensor $\boldsymbol{V}$ (corresponds to stretch after rotation). Since $\boldsymbol{U}^2 = \boldsymbol{F}^{\mathrm{T}} \cdot \boldsymbol{F}$ and $\boldsymbol{V}^2 = \boldsymbol{F} \cdot \boldsymbol{F}^{\mathrm{T}}$ both stretch tensors are symmetric and positive definite. The positive (real) eigenvalues, the principal stretches $\lambda_{1\ldots3}$, are equal for both stretch tensors. Accordingly, the spectral decomposition of the right stretch tensor is

$$\boldsymbol{U} = \sum_{\alpha=1}^{3} \lambda_\alpha \, \tilde{\boldsymbol{n}}_{0\,\alpha} \otimes \tilde{\boldsymbol{n}}_{0\,\alpha} \qquad \text{or} \qquad \boldsymbol{U} \cdot \tilde{\boldsymbol{n}}_{0\,\alpha} = \lambda_\alpha \, \tilde{\boldsymbol{n}}_{0\,\alpha} \quad . \tag{2.11}$$

The meaning of the principal stretches is found by regarding the deformation of a material line element $\mathrm{d}\tilde{\boldsymbol{X}}_1 := \mathrm{d}S_1 \, \tilde{\boldsymbol{n}}_{0\,1}$ with length $\mathrm{d}S_1$ in the direction of the first principal eigenvalue $\tilde{\boldsymbol{n}}_{0\,1}$. The line element in the current placement $\mathrm{d}\tilde{\boldsymbol{x}}_1 := \mathrm{d}s_1 \, \tilde{\boldsymbol{n}}_1$ with the length $\mathrm{d}s_1$ and direction $\tilde{\boldsymbol{n}}_1$ yields, by using (2.5), (2.11) and the property $\tilde{\boldsymbol{n}}_\alpha = \boldsymbol{R} \cdot \tilde{\boldsymbol{n}}_{0\,\alpha}$, the relation

$$\mathrm{d}s_1 \, \tilde{\boldsymbol{n}}_1 = \mathrm{d}\tilde{\boldsymbol{x}}_1 = \boldsymbol{F} \cdot \mathrm{d}\tilde{\boldsymbol{X}}_1 = \mathrm{d}S_1 \, \boldsymbol{R} \cdot \boldsymbol{U} \cdot \tilde{\boldsymbol{n}}_{0\,1} = \lambda_1 \, \mathrm{d}S_1 \, \boldsymbol{R} \cdot \tilde{\boldsymbol{n}}_{0\,1} = \lambda_1 \, \mathrm{d}S_1 \, \tilde{\boldsymbol{n}} \quad . \tag{2.12}$$

Thus, the principal stretches in (2.11) correspond to the ratios between current and initial lengths $\lambda_{1\ldots3} = \mathrm{d}s_{1\ldots3} / \mathrm{d}S_{1\ldots3}$ with respect to the principal directions $\tilde{\boldsymbol{n}}_{0\,1\ldots3}$.

Strain measure

While the displacement is a quantity defined by the difference between two configurations as introduced above, the definition of strain measures always relies on a strain concept. According to e.g. Bažant and Cedolin (2010), a valid strain tensor must

- be symmetric,
- vanish for all rigid body motions (transitions and rotations),
- be a function of displacement gradients, which is continuous, continuously differentiable and monotonic,
- approach $\nabla_{\mathrm{s}} \boldsymbol{u} = \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{T}} \right) / 2$ as $||\nabla_{\mathrm{s}} \boldsymbol{u}|| \to 0$.

Considering these requirements, an infinite number of different strain tensors could be introduced, but only some became practically relevant. Hill (1968)[2], for example, introduced a generalized strain concept with a family of strain tensors according to

$$\boldsymbol{E}_{(\mathrm{m})} := \frac{1}{2m} (\boldsymbol{U}^{2m} - \boldsymbol{I}) \quad . \tag{2.13}$$

[2] Hill's general strain concept is based on the one-dimensional strain measure proposed by Seth (1964). However, in Doyle and Ericksen (1956) the same relation had been published a few years earlier.

They depend on the choice of the parameter $m \in \mathbb{R}$ (including the Hencky strain $m = 0$). Due to its formulation in terms of $\boldsymbol{U}$, all strain tensors are symmetric. A common strain measure in the material description is the Green-Lagrange strain tensor

$$\boldsymbol{E}_{\mathrm{G}}(\boldsymbol{X}, t) := \frac{1}{2}(\boldsymbol{F}^{\mathrm{T}} \cdot \boldsymbol{F} - \boldsymbol{I}) = \frac{1}{2}(\boldsymbol{U}^2 - \boldsymbol{I}) = \frac{1}{2}(\boldsymbol{H} + \boldsymbol{H}^{\mathrm{T}} + \boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H}) \quad , \tag{2.14}$$

which is obtained from (2.13) by setting $m = 1$. The Green-Lagrange strain measure is based on the change of the squared length of the line element

$$\begin{aligned} \mathrm{d}s^2 - \mathrm{d}S^2 &= \mathrm{d}\boldsymbol{x} \cdot \mathrm{d}\boldsymbol{x} - \mathrm{d}\boldsymbol{X} \cdot \mathrm{d}\boldsymbol{X} = \mathrm{d}\boldsymbol{X} \cdot \boldsymbol{F}^{\mathrm{T}} \cdot \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X} - \mathrm{d}\boldsymbol{X} \cdot \mathrm{d}\boldsymbol{X} \\ &= \mathrm{d}\boldsymbol{X} \cdot (\boldsymbol{F}^{\mathrm{T}} \cdot \boldsymbol{F} - \boldsymbol{I}) \cdot \mathrm{d}\boldsymbol{X} \quad . \end{aligned} \tag{2.15}$$

Small deformation theory

Brittle materials often undergo only small stretches $\lambda_{1\ldots3} \approx 1$ before failure occurs. Thus, the strain of interest is small. If, in addition, rigid body rotations are considered to be infinitesimally small, i.e. $\boldsymbol{R} \approx \boldsymbol{I}$, the norm of the displacement gradient (2.9) becomes

$$\begin{aligned} ||\boldsymbol{H}|| &= \sqrt{\boldsymbol{H}^2} = \sqrt{(\boldsymbol{R} \cdot \boldsymbol{U} - \boldsymbol{I})^2} \\ &\approx \sqrt{(\boldsymbol{U} - \boldsymbol{I})^2} = \sqrt{(\lambda_1 - 1)^2 + (\lambda_2 - 1)^2 + (\lambda_3 - 1)^2} \ll 1 \quad . \end{aligned} \tag{2.16}$$

The Green-Lagrange strain tensor (2.14) can be linearized ($\boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{H} \to \boldsymbol{0}$) to the small (infinitesimal) strain tensor

$$\boldsymbol{\varepsilon} := \frac{1}{2}(\boldsymbol{H} + \boldsymbol{H}^{\mathrm{T}}) = \frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{T}}\right) = \nabla_{\mathrm{s}} \boldsymbol{u} \quad , \tag{2.17}$$

which obviously is symmetric. In fact, this was already a prerequisite for the formulation of a valid strain measure. Consequently, the linearization of all other strain measures at $||\nabla_{\mathrm{s}} \boldsymbol{u}|| \to 0$ gives the same result, thus $\boldsymbol{\varepsilon}$ is also called the linearized strain tensor. Its main diagonal elements describe the change in length. Thus, the trace is equal to the volume change of an infinitesimal volume element $\mathrm{d}V$, i.e. the volumetric strain

$$\varepsilon_{\mathrm{vol}} := \frac{\mathrm{d}V - \mathrm{d}V_0}{\mathrm{d}V_0} = \frac{\mathrm{d}V}{\mathrm{d}V_0} - 1 = \mathrm{tr}(\boldsymbol{\varepsilon}) = \boldsymbol{I} : \boldsymbol{\varepsilon} \quad . \tag{2.18}$$

The off-diagonal elements of the infinitesimal strain tensor are related to the change in angle (shearing). The infinitesimal strain tensor can be additionally decomposed into a traceless symmetric tensor, the deviatoric strain tensor

$$\boldsymbol{\varepsilon}_{\mathrm{dev}} := \boldsymbol{\varepsilon} - \frac{1}{3}\,\mathrm{tr}(\boldsymbol{\varepsilon})\,\boldsymbol{I} = \boldsymbol{\varepsilon} - \frac{1}{3}\varepsilon_{\mathrm{vol}}\,\boldsymbol{I} \tag{2.19}$$

and the volumetric strain tensor

$$\boldsymbol{\varepsilon}_{\mathrm{vol}} := \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_{\mathrm{dev}} = \frac{1}{3}\,\mathrm{tr}(\boldsymbol{\varepsilon})\,\boldsymbol{I} = \frac{1}{3}\varepsilon_{\mathrm{vol}}\,\boldsymbol{I} \quad . \tag{2.20}$$

The volumetric (or spherical) term describes the volume change while the deviatoric strain describes a pure distortion at constant volume. Additionally, we define fourth order projection tensors

$$\mathbb{P}_{\text{vol}} := \frac{1}{3}\boldsymbol{I}\otimes\boldsymbol{I} \quad \text{and} \quad \mathbb{P}_{\text{dev}} := \mathbb{I}_{\text{sym}} - \mathbb{P}_{\text{vol}} = \mathbb{I}_{\text{sym}} - \frac{1}{3}\boldsymbol{I}\otimes\boldsymbol{I} \tag{2.21}$$

which are orthogonal to each other ($\mathbb{P}_{\text{dev}} : \mathbb{P}_{\text{vol}} = \mathbb{O}$) to map the strain tensor to the corresponding volumetric and deviatoric strain tensors

$$\boldsymbol{\varepsilon}_{\text{vol}} = \mathbb{P}_{\text{vol}} : \boldsymbol{\varepsilon} \qquad \text{and} \qquad \boldsymbol{\varepsilon}_{\text{dev}} = \mathbb{P}_{\text{dev}} : \boldsymbol{\varepsilon} \quad , \tag{2.22}$$

respectively. Due to its symmetry, the infinitesimal strain tensor can be represented by real eigenvalues ε_α and corresponding eigenvectors $\boldsymbol{\hat{n}}_\alpha$ according to

$$\boldsymbol{\varepsilon} = \sum_{\alpha=1}^{3} \varepsilon_\alpha \; \boldsymbol{\hat{n}}_\alpha \otimes \boldsymbol{\hat{n}}_\alpha \quad , \tag{2.23}$$

see App. A.5 for further details. Since eigenvectors are not unique in case of equal eigenvalues one might use projection tensors instead. In a Cartesian coordinate system with basis $\boldsymbol{e}_i$ and in the principal coordinate system with base vector $\boldsymbol{n}_i$ the infinitesimal strain tensor is

$$\left[\boldsymbol{\varepsilon}\right]_{\{\boldsymbol{e}_i\}} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix} \qquad \text{and} \qquad \left[\boldsymbol{\varepsilon}\right]_{\{\boldsymbol{n}_i\}} = \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{bmatrix} \quad , \tag{2.24}$$

respectively, usually with the convention $\varepsilon_1 \geq \varepsilon_2 \geq \varepsilon_3$.

For the small deformation theory, both the material and spatial description are the same and $\boldsymbol{x} \approx \boldsymbol{X}$ applies. Consequently, transformations of line, surface and volume elements (2.5–2.7) become

$$\mathrm{d}\boldsymbol{x} \approx \mathrm{d}\boldsymbol{X} \quad , \qquad \mathrm{d}A\,\boldsymbol{n} \approx \mathrm{d}A_0\,\boldsymbol{n}_0 \quad \text{and} \quad \mathrm{d}V \approx \mathrm{d}V_0 \quad , \tag{2.25}$$

respectively. The velocity and acceleration of a particle are given by the first and second time derivative of the displacement field $\dot{\boldsymbol{u}}(\boldsymbol{x},t) = \mathrm{d}\boldsymbol{u}/\mathrm{d}t$ and $\ddot{\boldsymbol{u}}(\boldsymbol{x},t) = \mathrm{d}^2\boldsymbol{u}/\mathrm{d}t^2$. Due to the assumption of small deformations the velocity and acceleration are

$$\dot{\boldsymbol{u}}(\boldsymbol{x},t) = \frac{\partial \boldsymbol{u}}{\partial t} \qquad \text{and} \qquad \ddot{\boldsymbol{u}}(\boldsymbol{x},t) = \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \quad . \tag{2.26}$$

Stress measure

Another fundamental measure in continuum mechanics is stress, which might be caused by external or internal forces and is a source for/consequence of the deformations discussed

above. To reveal internal stresses, a solid body is (hypothetically) cut by a plane surface (defined by its normal vector $\boldsymbol{n}$) as depicted in Fig. 2.2a. Due to the interaction of both parts, forces are transmitted across the plane. An internal force of size $\Delta \boldsymbol{f}$ acts on a small surface element of the area ΔA on the cutting plane. By relating this force to the area we obtain in the limit state $\Delta A \to 0$ the stress vector

$$\boldsymbol{t} := \lim_{\Delta A \to 0} \frac{\Delta \boldsymbol{f}}{\Delta A} = \frac{\mathrm{d}\boldsymbol{f}}{\mathrm{d}A} \quad . \tag{2.27}$$

It depends on the position as well as on the orientation of the cut surface $\boldsymbol{n}$. The stress vector acting on the opposite surface $-\boldsymbol{n}$ is, due to Newton's law of action and reaction, $-\boldsymbol{t}$. The stress vector can be decomposed into a traction vector perpendicular to the surface (parallel to the normal vector)

$$\boldsymbol{t}_{\parallel} := (\boldsymbol{t} \cdot \boldsymbol{n})\, \boldsymbol{n} = \boldsymbol{N} \cdot \boldsymbol{t} \tag{2.28}$$

and a traction vector in the plane of the surface (perpendicular to the normal vector)

$$\boldsymbol{t}_{\perp} := \boldsymbol{t} - \boldsymbol{t}_{\parallel} = \boldsymbol{t} - (\boldsymbol{t} \cdot \boldsymbol{n})\, \boldsymbol{n} = (\boldsymbol{I} - \boldsymbol{N}) \cdot \boldsymbol{t} \quad , \tag{2.29}$$

see Fig. 2.2b. Corresponding projections are given by the dyads $\boldsymbol{N} := \boldsymbol{n} \otimes \boldsymbol{n}$ and $\boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n}$. The signed normal stress $\sigma = \boldsymbol{t} \cdot \boldsymbol{n} \in \mathbb{R}$ is said to be a tensile stress if $\sigma > 0$ and is compressive if $\sigma < 0$, while the magnitude of the shear stress $\tau = \sqrt{\boldsymbol{t} \cdot \boldsymbol{t} - \sigma^2}$ is always non-negative ($\tau \in \mathbb{R}_{\geq 0}$).

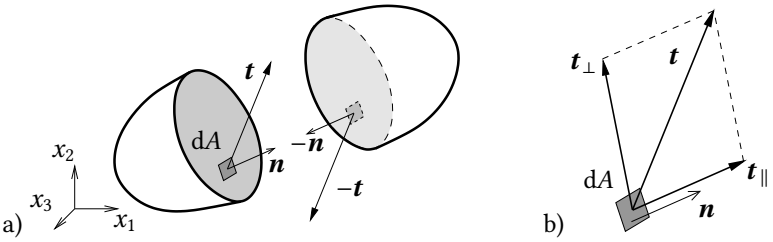


Figure 2.2: a) Solid body hypothetically cut by a plane with normal vector $\boldsymbol{n}$, b) decomposition of stress vector into perpendicular and parallel portions.

The linear dependency between the normal vector $\boldsymbol{n}$ and the corresponding stress vector $\boldsymbol{t}(\boldsymbol{n})$ can be expressed by the Cauchy stress tensor $\boldsymbol{\sigma}$ as

$$\boldsymbol{t} = \boldsymbol{\sigma}^{\mathrm{T}} \cdot \boldsymbol{n} \quad . \tag{2.30}$$

Relation (2.30) is known as the *Cauchy theorem*. In contrast to the stress vector, the stress tensor $\boldsymbol{\sigma}(\boldsymbol{x})$ fully characterizes the stress state at a position. As will be shown in Sect. 2.1.2 the stress tensor is symmetric, thus the Cauchy theorem (2.30) is often simplified to $\boldsymbol{t} = \boldsymbol{\sigma} \cdot \boldsymbol{n}$. Due to its symmetry, the Cauchy stress tensor can be represented by real eigenvalues and corresponding eigenvectors by its spectral decomposition

$$\boldsymbol{\sigma} = \sum_{\alpha=1}^{3} \sigma_\alpha\, \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha \quad , \tag{2.31}$$

see also App. A.5. In a Cartesian coordinate system with basis $\boldsymbol{e}_i$ and in the principal coordinate system with base vector $\boldsymbol{n}_i$ the Cauchy stress tensor is

$$\left[\boldsymbol{\sigma}\right]_{\{\boldsymbol{e}_i\}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \qquad \text{and} \qquad \left[\boldsymbol{\sigma}\right]_{\{\boldsymbol{n}_i\}} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \tag{2.32}$$

respectively, usually with the convention $\sigma_1 \geq \sigma_2 \geq \sigma_3$. The corresponding components are depicted in Fig. 2.3.

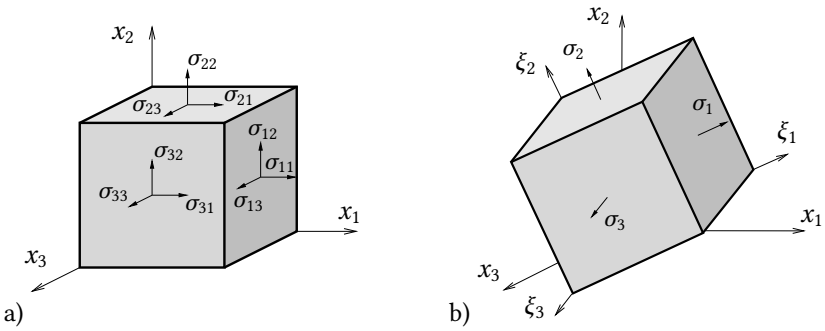


Figure 2.3: Infinitesimal volume element with stress components given in the a) reference Cartesian coordinate system, b) principal axis system.

2.1.2 Balance laws

Balance laws are axioms, which postulate the change of different physical quantities in continuum mechanics. They are required to describe the physical behavior of a continuum body and are valid in the whole body volume Ω, but also in any arbitrary part of the body (described, e.g., by the control volume $P \subseteq \Omega$). A balance law postulates that the change of a physical quantity Ψ (which is described by a smooth tensor field $\psi(\boldsymbol{x}, t)$ of order n) with time is equal to the sum of internal production p_ψ and supply s_ψ of the quantity in the bulk P and flux $\boldsymbol{q}_\psi$ across the boundary ∂P. This fundamental statement is expressed by the global *master balance law* (Holzapfel, 2000)

$$\dot{\Psi} = \frac{\mathrm{d}}{\mathrm{d}t} \int_P \psi(\boldsymbol{x}, t)\, \mathrm{d}P = \int_P p_\psi(\boldsymbol{x}, t)\, \mathrm{d}V + \int_P s_\psi(\boldsymbol{x}, t)\, \mathrm{d}V + \int_{\partial P} \boldsymbol{q}_\psi(\boldsymbol{x}, t) \cdot \boldsymbol{n}(\boldsymbol{x})\, \mathrm{d}A \quad . \tag{2.33}$$

Due to the small deformation theory (2.16), the control volume P is constant in time, thus differentiation with respect to time and integration over the volume can be exchanged according to

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_P \psi(\boldsymbol{x}, t)\, \mathrm{d}V = \int_P \frac{\partial \psi}{\partial t}\, \mathrm{d}V \quad . \tag{2.34}$$

In addition, by using the divergence theorem (A.40), one obtains

$$\int_{\partial P} \boldsymbol{q}_\psi \cdot \boldsymbol{n} \, \mathrm{d}A = \int_P \operatorname{div}(\boldsymbol{q}_\psi) \, \mathrm{d}V \quad , \tag{2.35}$$

with $\boldsymbol{q}_\psi$ being a tensor field of order $n+1$. Finally, inserting (2.34) and (2.35) into (2.33) yields

$$\int_P \left(\frac{\partial \psi}{\partial t} - p_\psi - s_\psi - \operatorname{div}(\boldsymbol{q}_\psi) \right) \mathrm{d}V = 0 \quad . \tag{2.36}$$

Since (2.36) must be valid for any chosen control volume $P \subseteq \Omega$ the local form of the master balance is

$$\dot{\psi} = p_\psi + s_\psi + \operatorname{div}(\boldsymbol{q}_\psi) \quad . \tag{2.37}$$

Balance of mass

Mass is a fundamental physical property of every continuum body. Regarding the macroscopic scale, mass is assumed to be distributed, at least piecewise continuously over all particles. The total mass of a body of volume Ω is

$$m = \int_\Omega \rho(\boldsymbol{x}, t) \, \mathrm{d}V \tag{2.38}$$

with $\rho(\boldsymbol{x}, t)$ being the mass density (mass per unit volume). In classical mechanics[3] mass can neither cross the boundary of a closed system[4], thus the flux term in (2.33) is $\boldsymbol{q}_\rho = \boldsymbol{0}$, nor can be produced ($r_\rho = 0$) or destroyed ($s_\rho = 0$). With the small deformation theory (2.16), thus P is constant in time, the change of the mass in a control volume $P \subseteq \Omega$ is

$$\dot{m} = \frac{\mathrm{d}}{\mathrm{d}t} \int_P \rho(\boldsymbol{x}, t) \, \mathrm{d}V = \int_P \frac{\partial \rho}{\partial t} \, \mathrm{d}V = \int_P \dot{\rho} \, \mathrm{d}V = 0 \tag{2.39}$$

and states the conservation of mass. Due to the arbitrary choice of the control volume P, for which (2.39) must hold, the local form of the mass balance is

$$\dot{\rho} = \frac{\partial \rho}{\partial t} = 0 \quad . \tag{2.40}$$

This specific form of the continuity (mass) equation, where the mass density $\rho(\boldsymbol{x}, t)$ is also a conserved quantity, is the outcome of the small deformation theory.

[3] In special relativity, the mass of an object increases with the velocity with respect to the observer. For small velocities (compared to the speed of light) and small masses, relativistic effects can be neglected and special relativity agrees with Newtonian mechanics.

[4] A closed system is defined as a fixed amount of mass (control mass), which is related to a region P with a boundary ∂P. By contrast, an open system is based on a fixed amount of volume (control volume) with a boundary, across which mass can enter and leave.

Balance of linear momentum

The linear momentum (also translational momentum) of a body of volume Ω is

$$\boldsymbol{p} := \int_{\Omega} \rho\, \dot{\boldsymbol{u}}\, \mathrm{d}V \quad . \tag{2.41}$$

Subjected to surface tractions $\boldsymbol{t}$ acting on the outer surface $\partial\Omega$ of the body and volume forces $\boldsymbol{f}$ (e.g. gravity loading per volume), the resultant force on the body is

$$\boldsymbol{F}(t) = \int_{\partial\Omega} \boldsymbol{t}\, \mathrm{d}A + \int_{\Omega} \boldsymbol{f}\, \mathrm{d}V \quad . \tag{2.42}$$

The linear momentum of a control volume $P \subseteq \Omega$ is

$$\boldsymbol{L}(t) := \int_{P} \boldsymbol{p}\, \mathrm{d}V = \int_{P} \rho\, \dot{\boldsymbol{u}}\, \mathrm{d}V \quad . \tag{2.43}$$

Following Newton's second law, which states that the rate of change of momentum of a body is equal to the applied force and that the directions of the applied force and the change of momentum correspond, the global form of the balance of linear momentum is

$$\dot{\boldsymbol{L}} = \frac{\mathrm{d}}{\mathrm{d}t} \int_{P} \rho\, \dot{\boldsymbol{u}}\, \mathrm{d}V = \int_{\partial P} \boldsymbol{t}\, \mathrm{d}A + \int_{P} \boldsymbol{f}\, \mathrm{d}V \quad . \tag{2.44}$$

With the Cauchy theorem (2.30) and the divergence theorem (A.44) it is

$$\int_{\partial P} \boldsymbol{t}\, \mathrm{d}A = \int_{\partial P} \boldsymbol{\sigma}^{\mathrm{T}} \cdot \boldsymbol{n}\, \mathrm{d}A = \int_{P} \mathrm{div}(\boldsymbol{\sigma}^{\mathrm{T}})\, \mathrm{d}V \quad , \tag{2.45}$$

so (2.44) yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{P} \rho\, \dot{\boldsymbol{u}}\, \mathrm{d}V = \int_{P} \mathrm{div}(\boldsymbol{\sigma}^{\mathrm{T}})\, \mathrm{d}V + \int_{P} \boldsymbol{f}\, \mathrm{d}V \quad . \tag{2.46}$$

By inserting the continuity equation (2.40) and with the requirement that (2.46) must be valid for arbitrary choices of the control volume P, the local form of the linear momentum balance is derived as

$$\rho\, \ddot{\boldsymbol{u}} = \mathrm{div}(\boldsymbol{\sigma}^{\mathrm{T}}) + \boldsymbol{f} \quad . \tag{2.47}$$

The local form of the balance of linear momentum is also known as *equation of motion.*

Balance of angular momentum

The angular momentum (also rotational momentum) of a body of volume Ω is defined as

$$\boldsymbol{J}(t) := \int_{\Omega} \boldsymbol{r}(\boldsymbol{x}) \times (\rho\, \dot{\boldsymbol{u}})\, \mathrm{d}V \quad , \tag{2.48}$$

with the distance vector $\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{x}_{\mathrm{ref}}$ from a reference point $\boldsymbol{x}_{\mathrm{ref}}$. Subjected to surface tractions $\boldsymbol{t}$ acting on the outer surface $\partial\Omega$ of the body and volume forces $\boldsymbol{f}$, the resultant moment (or resultant torque) on the body is

$$\boldsymbol{M}(t) = \int_{\partial\Omega} \boldsymbol{r}(\boldsymbol{x}) \times \boldsymbol{t}(\boldsymbol{x})\, \mathrm{d}A + \int_{\Omega} \boldsymbol{r}(\boldsymbol{x}) \times \boldsymbol{f}(\boldsymbol{x})\, \mathrm{d}V \quad . \tag{2.49}$$

By neglecting distributed resulting couples, the global balance of angular momentum of a control volume $P \subseteq \Omega$ is postulated as

$$\dot{\boldsymbol{J}} = \frac{\mathrm{d}}{\mathrm{d}t} \int_{P} \boldsymbol{r}(\boldsymbol{x}) \times (\rho\, \dot{\boldsymbol{u}})\, \mathrm{d}V = \int_{\partial P} \boldsymbol{r} \times \boldsymbol{t}\, \mathrm{d}A + \int_{P} \boldsymbol{r} \times \boldsymbol{f}\, \mathrm{d}V \quad . \tag{2.50}$$

It states, that the rate of change of angular momentum is equal to the applied momentum, and that the direction of momentum change corresponds to the direction of the applied momentum from an external source. Using the Cauchy theorem (2.30) and the divergence theorem (A.44) it is

$$\begin{aligned} \int_{\partial P} \boldsymbol{r} \times \boldsymbol{t}\, \mathrm{d}A &= \int_{\partial P} \boldsymbol{r} \times (\boldsymbol{\sigma}^{\mathrm{T}} \cdot \boldsymbol{n})\, \mathrm{d}A = \int_{\partial P} \varepsilon_{ijk}\, r_j\, \sigma_{lk}\, n_l\, \boldsymbol{e}_i\, \mathrm{d}A = \int_{P} \frac{\partial}{\partial x_l} \left(\varepsilon_{ijk}\, r_j\, \sigma_{lk}\right) \boldsymbol{e}_i\, \mathrm{d}V \\ &= \int_{P} \varepsilon_{ijk}\, \delta_{jl}\, \sigma_{lk}\, \boldsymbol{e}_i + \varepsilon_{ijk}\, r_j\, \frac{\partial \sigma_{lk}}{\partial x_l}\, \boldsymbol{e}_i\, \mathrm{d}V = \int_{P} \varepsilon : \boldsymbol{\sigma} + \boldsymbol{r} \times \mathrm{div}(\boldsymbol{\sigma}^{\mathrm{T}})\, \mathrm{d}V \quad . \end{aligned} \tag{2.51}$$

The application of the conservation of mass (2.40) on the global balance of angular momentum (2.50) yields

$$\int_{P} \boldsymbol{r} \times \left(\dot{\boldsymbol{p}} - \boldsymbol{f} - \mathrm{div}(\boldsymbol{\sigma}^{\mathrm{T}})\right) \mathrm{d}V = \int_{P} \varepsilon : \boldsymbol{\sigma}\, \mathrm{d}V \quad . \tag{2.52}$$

Finally, since the balance must hold for an arbitrary control volume P, the local balance of angular momentum is

$$\boldsymbol{r} \times \left(\dot{\boldsymbol{p}} - \boldsymbol{f} - \mathrm{div}(\boldsymbol{\sigma}^{\mathrm{T}})\right) = \varepsilon : \boldsymbol{\sigma} \quad . \tag{2.53}$$

By inserting the equation of motion (2.47) one obtains

$$\varepsilon : \boldsymbol{\sigma} = \begin{bmatrix} \sigma_{23} - \sigma_{32} \\ \sigma_{31} - \sigma_{13} \\ \sigma_{12} - \sigma_{21} \end{bmatrix} = \boldsymbol{0} \quad , \tag{2.54}$$

which states the symmetry of the Cauchy stress tensor[5]

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathrm{T}} \quad . \tag{2.55}$$

Balance of energy

Energy can exist in several forms in a continuum body. The kinetic energy of a body of volume Ω is

$$K := \int_{\Omega} \frac{1}{2} \rho \, \dot{\boldsymbol{u}} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}V \quad . \tag{2.56}$$

The sum of all forms of energy stored inside the body is given by the internal energy[6]

$$E := \int_{\Omega} e \, \mathrm{d}V \quad , \tag{2.57}$$

with $e(\boldsymbol{x}, t)$ being the internal energy per unit volume. The rate of work done by stresses inside the continuum body is called internal stress power and given by

$$P_{\mathrm{int}} := \int_{\partial\Omega} \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} \, \mathrm{d}V \quad . \tag{2.58}$$

In the case that only mechanical energy exists, it is $\dot{E} = P_{\mathrm{int}}$. The sum of kinetic and internal energy is called total energy of the body. The external mechanical power, i.e. the mechanical energy transport into the body, is

$$P_{\mathrm{ext}} := \int_{\partial\Omega} \boldsymbol{t} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}A + \int_{\Omega} \boldsymbol{f} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}V \tag{2.59}$$

and contains terms related to external tractions $\boldsymbol{t}$ acting on the surface of the body $\partial\Omega$ and volume forces $\boldsymbol{f}$. In general, there are also non-mechanical (e.g. thermal, chemical, electric) energy fluxes into the body, which are denoted by Q here. With the change of total energy being equal to the energy transport into the body, the global form of the energy balance of a control volume $P \subseteq \Omega$ is

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{P} \frac{1}{2} \rho \, \dot{\boldsymbol{u}} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}V + \frac{\mathrm{d}}{\mathrm{d}t} \int_{P} e \, \mathrm{d}V = \int_{\partial P} \boldsymbol{t} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}A + \int_{P} \boldsymbol{f} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}V + Q \tag{2.60}$$

[5] The symmetry property of the stress tensor is a consequence of the model assumption, that no distributed resulting couples in the body exist. In micropolar continua, e.g. for the Cosserat continuum, the symmetry (2.55) is not fulfilled and might be introduced additionally as constitutive assumption (Altenbach and Eremeyev, 2013).

[6] The concept of internal energy has been introduced by Rudolf J.E. Clausius and William J.M. Rankine (Truesdell, 1980).

or, in short notation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(K+E) = P_{\mathrm{ext}} + Q \quad , \tag{2.61}$$

also known as the *first law of thermodynamics*. It should be noted, that by considering only mechanical energy, i.e. $Q = 0$, the energy balance is a consequence of the equation of motion and no additional postulate. This can be shown by inserting the relation

$$\begin{aligned}\int_{\partial P} \boldsymbol{t}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A &= \int_{\partial P} (\boldsymbol{\sigma}^{\mathrm{T}}\cdot\boldsymbol{n})\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A = \int_{P} \operatorname{div}(\boldsymbol{\sigma}^{\mathrm{T}}\cdot\dot{\boldsymbol{u}})\,\mathrm{d}V \\ &= \int_{P} \operatorname{div}(\boldsymbol{\sigma}^{\mathrm{T}})\cdot\dot{\boldsymbol{u}} + \boldsymbol{\sigma}^{\mathrm{T}} : \underbrace{\nabla_{\mathrm{s}}\dot{\boldsymbol{u}}}_{=\dot{\boldsymbol{\varepsilon}}} + \underbrace{\boldsymbol{\sigma}^{\mathrm{T}} : \nabla_{\mathrm{a}}\dot{\boldsymbol{u}}}_{=0}\,\mathrm{d}V\end{aligned} \tag{2.62}$$

into

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,(K+E) &= \int_{P} \rho\,\ddot{\boldsymbol{u}}\cdot\dot{\boldsymbol{u}} + \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V \overset{(2.47)}{=} \int_{P} \boldsymbol{f}\cdot\dot{\boldsymbol{u}} + \operatorname{div}(\boldsymbol{\sigma}^{\mathrm{T}})\cdot\dot{\boldsymbol{u}} + \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V \\ &= \int_{P} \boldsymbol{f}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}V + \int_{\partial P} \boldsymbol{t}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A \overset{(2.59)}{=} P_{\mathrm{ext}} \quad .\end{aligned} \tag{2.63}$$

In case of $\dot{K} = 0$, the acceleration is $\ddot{\boldsymbol{u}} = \boldsymbol{0}$ and the problem is called quasi-static.

By taking also thermal energy into account (other energies are beyond the scope of this work), the total heat supply is

$$Q := \int_{\Omega} s_\theta\,\mathrm{d}V - \int_{\partial\Omega} \boldsymbol{q}_\theta\cdot\boldsymbol{n}\,\mathrm{d}A \quad , \tag{2.64}$$

with the heat supply in the bulk per unit volume $s_\theta(\boldsymbol{x}, t)$ and the heat flux $\boldsymbol{q}_\theta(\boldsymbol{x}, t)$ across the boundary. The last term in (2.64) has a negative sign since the normal vector $\boldsymbol{n}$ points outward, but heat flux into the body is assumed to be positive. By using the conservation of mass (2.40), the global form of the energy balance (2.60) can then be written as

$$\int_{P} (\rho\,\ddot{\boldsymbol{u}}\cdot\dot{\boldsymbol{u}} + \dot{e})\,\mathrm{d}V = \int_{\partial P} \boldsymbol{t}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A + \int_{P} \boldsymbol{f}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}V + \int_{P} s_\theta\,\mathrm{d}V - \int_{\partial P} \boldsymbol{q}_\theta\cdot\boldsymbol{n}\,\mathrm{d}A \quad . \tag{2.65}$$

Application of the Cauchy theorem (2.30), the divergence theorem (A.40) and the chain rule gives

$$\int_{\partial P} \boldsymbol{t}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A = \int_{\partial P} (\boldsymbol{\sigma}^{\mathrm{T}}\cdot\boldsymbol{n})\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A = \int_{P} \operatorname{div}(\boldsymbol{\sigma}\cdot\dot{\boldsymbol{u}})\,\mathrm{d}V = \int_{P} \operatorname{div}(\boldsymbol{\sigma})\cdot\dot{\boldsymbol{u}} + \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V \tag{2.66}$$

and

$$\int_{\partial P} \boldsymbol{q}_\theta\cdot\boldsymbol{n}\,\mathrm{d}A = \int_{P} \operatorname{div}(\boldsymbol{q}_\theta)\,\mathrm{d}V \quad . \tag{2.67}$$

The application of the equation of motion (2.47) yields

$$\int_P \underbrace{\left(\rho\,\ddot{\boldsymbol{u}} - \boldsymbol{f} - \operatorname{div}(\boldsymbol{\sigma}^{\mathrm{T}})\right)}_{=0} \cdot \dot{\boldsymbol{u}}\,\mathrm{d}V + \int_P \left(\dot{e} - \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - s_\theta + \operatorname{div}(\boldsymbol{q}_\theta)\right)\,\mathrm{d}V = 0 \tag{2.68}$$

Finally, since (2.68) is independent of the choice of P, the local form of the energy balance

$$\dot{e} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} + s_\theta - \operatorname{div}(\boldsymbol{q}_\theta) \tag{2.69}$$

is obtained.

Balance of entropy

While the energy balances (2.65) and (2.69) govern the energy transfer in a thermomechanical body, they do not account for the direction of energy transfer. However, as it can be observed, in many physical processes a certain direction of energy transfer exists. Focusing on thermomechanical problems, heat always flows from regions of higher temperature to regions of lower temperature. Thus, the direction of the thermal energy flow is governed by the entropy $\eta_{\mathrm{e}}(\boldsymbol{x}, t)$. The entropy S is a thermodynamic state variable and is often associated with the amount of disorder at the atomic or molecular scale. Alternatively, it can be seen as a measure of the unavailability of energy to do work. The rate of change of entropy of a body of volume Ω is

$$\dot{S} := \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \eta_{\mathrm{e}}\,\mathrm{d}V = \int_\Omega r_\eta\,\mathrm{d}V + \int_\Omega \frac{s_\theta}{\theta}\,\mathrm{d}V - \int_{\partial\Omega} \frac{\boldsymbol{q}_\theta}{\theta}\cdot\boldsymbol{n}\,\mathrm{d}A \tag{2.70}$$

with the entropy production r_η. The heat supply s_θ and heat flux are normalized by the absolute temperature $\theta > 0$, a time dependent scalar field. The *second law of thermodynamics* states, that the total production of entropy

$$\int_P r_\eta\,\mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t}\int_P \eta_{\mathrm{e}}\,\mathrm{d}V - \int_P \frac{s_\theta}{\theta}\,\mathrm{d}V + \int_{\partial P} \frac{\boldsymbol{q}_\theta}{\theta}\cdot\boldsymbol{n}\,\mathrm{d}A \geq 0 \tag{2.71}$$

must be non-negative. It is also known as the *Clausius-Duhem inequality*. By applying the Gauss's divergence theorem (A.40) in combination with the product rule (A.45), i.e.

$$\int_{\partial P} \frac{\boldsymbol{q}_\theta}{\theta}\cdot\boldsymbol{n}\,\mathrm{d}A = \int_P \operatorname{div}\!\left(\frac{\boldsymbol{q}_\theta}{\theta}\right)\,\mathrm{d}V = \int_P \frac{1}{\theta}\,\operatorname{div}(\boldsymbol{q}_\theta)\,\mathrm{d}V - \int_P \frac{1}{\theta^2}\,\nabla\theta\cdot\boldsymbol{q}_\theta\,\mathrm{d}V \quad , \tag{2.72}$$

and due to the arbitrary choice of the control volume P, the local form of (2.71) is

$$r_\eta = \dot{\eta}_{\mathrm{e}} - \frac{s_\theta}{\theta} + \frac{1}{\theta}\,\operatorname{div}(\boldsymbol{q}_\theta) - \frac{1}{\theta^2}\,\nabla\theta\cdot\boldsymbol{q}_\theta \geq 0 \quad . \tag{2.73}$$

The heat supply term in (2.71) is eliminated either by using the energy balance (2.61) in conjunction with the definition of the total heat supply (2.64) to reveal the dissipation inequality

$$\mathcal{D}_{\text{diss}} := P_{\text{ext}} + \int_P \theta\, \dot{\eta}_{\text{e}}\, \mathrm{d}V - \int_P \frac{1}{\theta}\, \nabla\theta \cdot \boldsymbol{q}_\theta\, \mathrm{d}V - \frac{\mathrm{d}}{\mathrm{d}t}(K + E) \geq 0 \tag{2.74}$$

or by using (2.68), which results in the equivalent expression

$$\mathcal{D}_{\text{diss}} := \int_P \theta\, \dot{\eta}_{\text{e}} - (\dot{e} - \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}})\, \mathrm{d}V - \int_P \frac{1}{\theta}\, \nabla\theta \cdot \boldsymbol{q}_\theta\, \mathrm{d}V \geq 0 \quad , \tag{2.75}$$

consisting of a local and a non-local term. By introducing the *Helmholtz free energy*

$$\psi := e - \theta\, \eta_{\text{e}} \tag{2.76}$$

(per unit volume) with $\dot{\psi} = \dot{e} - \dot{\theta}\, \eta_{\text{e}} - \theta\, \dot{\eta}_{\text{e}}$, it is

$$\mathcal{D}_{\text{diss}} := \int_P \dot{\theta}\, \eta_{\text{e}} - \psi + \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}\, \mathrm{d}V - \int_P \frac{1}{\theta}\, \nabla\theta \cdot \boldsymbol{q}_\theta\, \mathrm{d}V \geq 0 \quad . \tag{2.77}$$

The dissipation inequality must hold for the local (Clausius-Planck inequality) and non-local term (heat conduction inequality) independently, i.e.

$$\int_P \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - (\psi + \dot{\theta}\, \eta_{\text{e}})\, \mathrm{d}V \geq 0 \qquad \text{and} \qquad -\int_P \frac{1}{\theta}\, \nabla\theta \cdot \boldsymbol{q}_\theta\, \mathrm{d}V \geq 0 \quad . \tag{2.78}$$

While the heat conduction inequality can be satisfied by an adequate choice of the relation between the heat flux $\boldsymbol{q}_\theta$ and the temperature gradient $\nabla\theta$, e.g. by the Fourier law $\boldsymbol{q}_\theta = -\kappa\nabla\theta$ with the heat conductivity $\kappa > 0$, the local term is often used to derive restrictions on the constitutive relation. Its local form gives the internal dissipation or local production of entropy

$$\mathscr{D}_{\text{int}} := \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \psi + \dot{\theta}\, \eta_{\text{e}} \geq 0 \quad . \tag{2.79}$$

2.1.3 Constitutive relations

The preceding section yields a number of universal equations to describe the thermomechanical behavior of continuum bodies. To illustrate the necessity of additional equations, we focus on isothermal problems, which contain 15 unknowns: the displacement field $\boldsymbol{u}(\boldsymbol{x}, t)$, the strain field $\boldsymbol{\varepsilon}(\boldsymbol{x}, t)$ and the stress field $\boldsymbol{\sigma}(\boldsymbol{x}, t)$. On the other hand, the kinematic relation (2.17) yields six equations. With the equation of motion (2.47) there are three more equations available to determine the unknowns. Note, that the density

is constant because of the assumption of small deformations (2.16). Thus the mass balance (2.40) does not provide an additional equation. The momentum balance provides the symmetry of the stress tensor (2.55) as already used above. According to (2.63) the energy balance (2.69) does not provide any additional information. Thus six additional equations are required, which describe the specific material behavior by giving the relations between stresses and strains.

In general, there are different classes of material behavior such as elastic, plastic, rate-dependent (viscous) or damaging behavior and different combinations thereof. Elasticity, for instance, is governed by reversibility of strain energy. Consequently, the stress is only a function of strain without taking into account, e.g., the history of strain application. The specific relations usually depend on experimental observations and idealize the real material behavior so that it can be represented mathematically by so-called *constitutive relations*. However, they have to fulfill some restrictions, e.g. those provided by the second law of thermodynamics (2.79).

With the assumption of thermoelastic material, described by the Helmholtz free energy $\psi(\boldsymbol{\varepsilon}, \theta)$ depending on the strain field $\boldsymbol{\varepsilon}$ and the absolute temperature θ, the Clausius-Planck inequality is

$$\left(\boldsymbol{\sigma} - \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}}\right) : \boldsymbol{\varepsilon} + \left(\eta_\mathrm{e} - \frac{\partial \psi}{\partial \theta}\right) \theta \geq 0 \quad . \tag{2.80}$$

It implies the constitutive relations

$$\boldsymbol{\sigma} := \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} \qquad \text{and} \qquad \eta_\mathrm{e} := \frac{\partial \psi}{\partial \theta} \quad , \tag{2.81}$$

where the first equation yields the relation between stresses and strains depending on the choice of the potential ψ.

In the following, only elastic material behavior (at isothermal conditions $\dot{\theta} = 0$) is discussed, which serves as the basis for brittle damage description. The specific choice of a quadratic free energy function

$$\psi = \frac{1}{2} \boldsymbol{\varepsilon} : \mathbb{C} : \boldsymbol{\varepsilon} \tag{2.82}$$

provides, by using relation (2.81), stresses

$$\boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\varepsilon} \tag{2.83}$$

linearly depending on the infinitesimal strain tensor. This choice is known as the linear elastic constitutive law and represents *Hooke's law* of elasticity. The stiffness tensor (or elasticity tensor) $\mathbb{C} = C_{ijkl}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l$ is of fourth order and contains only constant (material specific) parameters. Due to the symmetry of stresses and strains, the elasticity tensor possesses left and right subsymmetries (regarding the first two and last two indices), also called left and right minor symmetry. By assuming the existence of a strain energy

density (as already used above) the elasticity tensor also possesses a main/major symmetry, i.e.

$$C_{ijkl} = C_{klij} = C_{jikl} = C_{ijlk} \quad . \tag{2.84}$$

To obtain non-negative strain energies for all deformations the elasticity tensor must be positive definite, which means, that all eigenvalues of $\mathbb{C}$ have to be positive. As a consequence of the symmetries, the number of different possible elasticity tensor components in the three-dimensional case is reduced from $3^4 = 81$ to 21 for the most general anisotropic case. For the isotropic case, i.e. the material behavior in all directions is identical, the number of material parameters reduces to two. In terms of the Lamé constants $\lambda \in \mathbb{R}$ and $\mu > 0$ the elasticity tensor reads

$$C_{ijkl} = \lambda\, \delta_{ij}\, \delta_{kl} + \mu\, (\delta_{ik}\, \delta_{jl} + \delta_{il}\, \delta_{jk}) \tag{2.85}$$

or, by using the second order identity tensor (A.6) and symmetric fourth order identity tensor (A.21),

$$\mathbb{C} = \lambda\, \boldsymbol{I} \otimes \boldsymbol{I} + 2\mu\, \mathbb{I}_{\text{sym}} \quad . \tag{2.86}$$

By using its spectral decomposition, see (A.26), the eigenvalues of the isotropic stiffness tensor are

$$\lambda_1 = 3\lambda + 2\mu \qquad \text{and} \qquad \lambda_{2\ldots6} = 2\mu \quad , \tag{2.87}$$

with the first principal direction represented by $\boldsymbol{N}_1 = 1/\sqrt{3}\, \boldsymbol{I}$ (all others are perpendicular to this one). The inverse relation of the stiffness tensor is the compliance tensor

$$S_{ijkl} := (C_{ijkl})^{-1} = \frac{1}{\mu}\, (\delta_{ik}\, \delta_{jl} + \delta_{il}\, \delta_{jk}) - \frac{\lambda}{2\mu\, (3\lambda + 2\mu)}\, \delta_{ij}\, \delta_{kl} \tag{2.88}$$

or, by using (A.6) and (A.21),

$$\mathbb{S} := \mathbb{C}^{-1} = \frac{2}{\mu}\, \mathbb{I}_{\text{sym}} - \frac{\lambda}{2\mu\, (3\lambda + 2\mu)}\, \boldsymbol{I} \otimes \boldsymbol{I} \quad , \tag{2.89}$$

which possesses the same symmetry restrictions. Alternative representations of the elasticity tensor are, by using Young's modulus $E > 0$ and Poisson's ratio $-1 < \nu < 1/2$ (materials with $\nu < 0$ are called auxetics and are out of the scope of this work),

$$\mathbb{C} = \frac{E\, \nu}{(1+\nu)(1-2\nu)}\, \boldsymbol{I} \otimes \boldsymbol{I} + \frac{E}{1+\nu}\, \mathbb{I}_{\text{sym}} \tag{2.90}$$

or, by using the bulk modulus

$$K := \lambda + \frac{2}{3}\mu = \frac{E}{3(1-2\nu)} > 0 \tag{2.91}$$

and the shear modulus $G := \mu > 0$,

$$\mathbb{C} = K\, \boldsymbol{I} \otimes \boldsymbol{I} + 2G \left(\mathbb{I}_{\text{sym}} - \frac{1}{3} \boldsymbol{I} \otimes \boldsymbol{I} \right) \quad . \tag{2.92}$$

2.1.4 Variational principles

The solution of mechanical problems as illustrated in Fig. 2.4 is governed by three relations, the equation of motion (2.47) with a given volume force $\boldsymbol{f}(\boldsymbol{x}, t)$, the kinematic relations (2.17) and a constitutive law, e.g., the linear elastic law (2.83). As already mentioned earlier, these field equations are valid for any arbitrary control volume $P \subseteq \Omega$ in the continuum body Ω. Since strains and stresses are given by kinematic relations and a constitutive law, the displacement vector field $\boldsymbol{u}(\boldsymbol{x}, t)$ is the only unknown. On the surface of the body $\partial\Omega$, either displacements

$$\boldsymbol{u}(\boldsymbol{x}_\mathrm{u}, t) = \bar{\boldsymbol{u}}(\boldsymbol{x}_\mathrm{u}, t) \qquad \text{for } \boldsymbol{x}_\mathrm{u} \in \partial\Omega_\mathrm{u} \tag{2.93}$$

or tractions

$$\boldsymbol{t}(\boldsymbol{x}_\mathrm{t}, t) = \boldsymbol{\sigma}^\mathrm{T} \cdot \boldsymbol{n} = \bar{\boldsymbol{t}}(\boldsymbol{x}_\mathrm{t}, t) \qquad \text{for } \boldsymbol{x}_\mathrm{t} \in \partial\Omega_\mathrm{t} \quad , \tag{2.94}$$

with the outer normal vector $\boldsymbol{n}$, have to be given. Note, that the boundary conditions are defined on disjoint parts of the surface, i.e. $\partial\Omega_\mathrm{u} \cup \partial\Omega_\mathrm{t} = \partial\Omega$ and $\partial\Omega_\mathrm{u} \cap \partial\Omega_\mathrm{t} = \emptyset$. Commonly, for the type of mechanical problems as given above, displacement boundary conditions are also called *Dirichlet boundary* conditions or essential/kinematic boundary conditions, while traction boundary conditions are also called *von Neumann boundary* conditions. External constraints like (2.93) are called holonomic, in the case they are given by an inequality, they are called nonholonomic (Holzapfel, 2000).

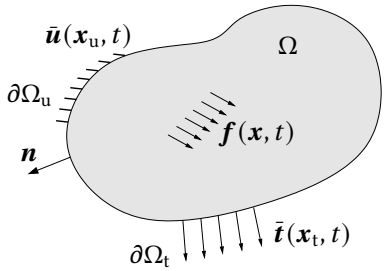


Figure 2.4: Elastic body with displacement boundary conditions on $\partial\Omega_\mathrm{u}$ and traction boundary conditions $\partial\Omega_\mathrm{t}$.

Due to the presence of time derivatives in (2.47), also initial conditions for the displacement field $\boldsymbol{u}(\boldsymbol{x}, t=0) = \boldsymbol{u}_0(\boldsymbol{x})$ and the velocity field $\dot{\boldsymbol{u}}(\boldsymbol{x}, t=0) = \dot{\boldsymbol{u}}_0(\boldsymbol{x})$ must be given. Note, that for compatibility reasons, the initial conditions on the displacement boundary must be $\bar{\boldsymbol{u}}(\boldsymbol{x}_\mathrm{u}, t=0) = \boldsymbol{u}_0(\boldsymbol{x}_\mathrm{u})$ and $\dot{\bar{\boldsymbol{u}}}(\boldsymbol{x}_\mathrm{u}, t=0) = \dot{\boldsymbol{u}}_0(\boldsymbol{x}_\mathrm{u})$. The field equations in conjunction with the boundary conditions and initial conditions are the strong form of the initial boundary-value problem (IBVP). If the problem is independent of time, it is called to be static and no initial conditions are required. In this case, the field equations in conjunction with the boundary conditions are the strong form of the boundary-value problem (BVP). Analytical solutions of the (in general) non-linear, partial differential equations are only found for some simple initial boundary value problems. In all other cases, numerical

techniques, such as the *finite element method* (FEM), can be used to approximate the solution. Numerical solution techniques are often based on the weak form of the governing equations, which is introduced in the following. Note, that the existence of an unique solution cannot be guaranteed for non-linear problems.

The first variation of the displacement field (at a fixed instant of time) is

$$\delta \boldsymbol{u} := (\boldsymbol{u} + \epsilon\, \boldsymbol{w}) - \boldsymbol{u} = \epsilon\, \boldsymbol{w} \quad , \tag{2.95}$$

with a scalar parameter of infinitesimal size $\epsilon \longrightarrow 0$ and an arbitrary vector field $\boldsymbol{w}$. Also referred as to virtual displacement field, $\delta \boldsymbol{u}$ is only virtual (no real displacement) and independent of the displacement vector $\boldsymbol{u}$. The first variation of a smooth function $F(\boldsymbol{u})$ depending only on the displacement is obtained by using the Gâteaux derivative (A.39), i.e.

$$\delta F(\boldsymbol{u}, \delta \boldsymbol{u}) = D_{\delta \boldsymbol{u}} F(\boldsymbol{u}) = \left. \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(\boldsymbol{u} + \epsilon\, \delta \boldsymbol{u}) \right|_{\epsilon = 0} \quad . \tag{2.96}$$

Principle of virtual work

In the following, the initial boundary value problem introduced at the beginning of this section, see Fig. 2.4, is regarded. By multiplying the strong form, i.e. the equation of motion (2.47), with an vector-valued smooth field $\boldsymbol{\eta}_\mathrm{u}(\boldsymbol{x}, t)$ and integrating over the whole domain of the body Ω, the scalar-valued function

$$f(\boldsymbol{u}, \boldsymbol{\eta}_\mathrm{u}) = \int_\Omega \left(\rho\, \ddot{\boldsymbol{u}} - \mathrm{div}(\boldsymbol{\sigma}^\mathrm{T}) - \boldsymbol{f} \right) \cdot \boldsymbol{\eta}_\mathrm{u} \, \mathrm{d}V = 0 \tag{2.97}$$

is obtained, which must vanish. The test function $\boldsymbol{\eta}_\mathrm{u}$ is arbitrary, but must vanish at the Dirichlet boundary $\partial\Omega_\mathrm{u}$, i.e. $\boldsymbol{\eta}_\mathrm{u}(\boldsymbol{x}_\mathrm{u}, t) = \boldsymbol{0}$. Application of the product rule (A.47) and the divergence theorem (A.44) yields

$$f(\boldsymbol{u}, \boldsymbol{\eta}_\mathrm{u}) = \int_\Omega \left[\rho\, \ddot{\boldsymbol{u}} \cdot \boldsymbol{\eta}_\mathrm{u} + \boldsymbol{\sigma} : \nabla \boldsymbol{\eta}_\mathrm{u} - \boldsymbol{f} \cdot \boldsymbol{\eta}_\mathrm{u} \right] \mathrm{d}V - \int_{\partial\Omega_\mathrm{t}} \boldsymbol{t} \cdot \boldsymbol{\eta}_\mathrm{u} \, \mathrm{d}A = 0 \quad . \tag{2.98}$$

Equation (2.98), together with the boundary conditions (2.93) and (2.94) and the slightly modified initial displacement condition

$$\int_\Omega \boldsymbol{u}(\boldsymbol{x}, t=0) \cdot \boldsymbol{\eta}_\mathrm{u} \, \mathrm{d}V = \int_\Omega \boldsymbol{u}_0(\boldsymbol{x}) \cdot \boldsymbol{\eta}_\mathrm{u} \, \mathrm{d}V \tag{2.99}$$

and initial velocity condition

$$\int_\Omega \dot{\boldsymbol{u}}(\boldsymbol{x}, t=0) \cdot \boldsymbol{\eta}_\mathrm{u} \, \mathrm{d}V = \int_\Omega \dot{\boldsymbol{u}}_0(\boldsymbol{x}) \cdot \boldsymbol{\eta}_\mathrm{u} \, \mathrm{d}V \quad , \tag{2.100}$$

are the variational form or weak form of the initial boundary value problem. Since the test function is arbitrary, the equations in the weak form (2.98) and strong form (2.47) are identical. Accordingly, also their solutions are equal as discussed, e.g., by Hughes (1987). By replacing the arbitrary test function $\boldsymbol{\eta}_\mathrm{u}$ in (2.98) with the virtual displacement field $\delta\boldsymbol{u}$, i.e.

$$f(\boldsymbol{u}, \delta\boldsymbol{u}) = \int_\Omega \rho\, \ddot{\boldsymbol{u}} \cdot \delta\boldsymbol{u}\, \mathrm{d}V + \underbrace{\int_\Omega \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon}\, \mathrm{d}V}_{=:\, \delta W_\mathrm{int}} - \underbrace{\left[\int_\Omega \boldsymbol{f} \cdot \delta\boldsymbol{u}\, \mathrm{d}V + \int_{\partial\Omega_\mathrm{t}} \boldsymbol{t} \cdot \delta\boldsymbol{u}\, \mathrm{d}A \right]}_{=:\, \delta W_\mathrm{ext}} = 0 \quad , \tag{2.101}$$

the principle of virtual work[7] (or principle of virtual displacement) is obtained. The virtual displacement field $\delta\boldsymbol{u}$ must be kinematically admissible, i.e. $\delta\boldsymbol{u}(\boldsymbol{x}_\mathrm{u}, t) = 0$. For vanishing accelerations ($\ddot{\boldsymbol{u}} = 0$) it states, that the external virtual work δW_ext equals the internal virtual work δW_int. To be more precise, the external virtual work δW_ext contains the total amount work done by the body force $\boldsymbol{f}$ and the surface traction $\boldsymbol{t}$ if they are subjected to a virtual displacement $\delta\boldsymbol{u}$. The internal virtual work δW_int represents the total amount of virtual stress work $\boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon}$ done by a fixed stress $\boldsymbol{\sigma}$ if it is subjected to a virtual strain $\delta\boldsymbol{\varepsilon}$. It should be noted, that the principle of virtual work is not based on further assumptions, e.g. conservative loads or the existence of a potential. Thus it is applicable to many problems, including problems of friction and inelastic materials.

Principle of stationary potential energy

By assuming a conservative mechanical system, i.e. the existence of energy functionals for both the internal potential

$$\Pi_\mathrm{int}(\boldsymbol{u}) = \int_\Omega \psi(\boldsymbol{\varepsilon})\, \mathrm{d}V \tag{2.102}$$

and the external potential

$$\Pi_\mathrm{ext}(\boldsymbol{u}) = -\int_\Omega \boldsymbol{f} \cdot \boldsymbol{u}\, \mathrm{d}V - \int_{\partial\Omega} \boldsymbol{t} \cdot \boldsymbol{u}\, \mathrm{d}A \quad , \tag{2.103}$$

the total potential energy of the system in static equilibrium is $\Pi = \Pi_\mathrm{int}(\boldsymbol{u}) + \Pi_\mathrm{ext}(\boldsymbol{u})$. It contains the strain energy function $\psi(\boldsymbol{\varepsilon})$ as introduced in (2.82), which describes the elastic energy stored in the body. Both the body force $\boldsymbol{f}$ and the traction $\boldsymbol{t}$ are assumed to be "dead" loads, i.e. independent of $\boldsymbol{u}$. The equilibrium state is obtained for the displacement

[7] One should not be confused with the principle of virtual forces or the principle of complementary virtual work, which is based on the variation of forces (and stresses) instead of displacements (and strains), as used here.

field, for which the total potential energy Π is stationary. For this purpose, the directional derivative as introduced in (2.96) is applied to the total potential energy, i.e.

$$\delta\Pi(\boldsymbol{u}, \delta\boldsymbol{u}) = D_{\delta\boldsymbol{u}}\Pi(\boldsymbol{u}) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Pi(\boldsymbol{u} + \epsilon\,\delta\boldsymbol{u})\right|_{\epsilon=0} = 0 \quad , \tag{2.104}$$

and must vanish in all directions $\delta\boldsymbol{u}$. This is known as the *principle of stationary potential energy*. For elastic problems, the solution obtained by (2.104) typically yields the state of minimum potential energy.

Three-field variational principle

Instead of using the kinematic relations (2.17) and a constitutive law to obtain the strain and stress field as function of an unknown displacement field $\boldsymbol{u}$, these relations can be incorporated by separate variations, similar to the solution of the equilibrium by the principle of virtual work. Thus, the displacements $\boldsymbol{u}$, strains $\boldsymbol{\varepsilon}$ and stresses $\boldsymbol{\sigma}$ are independent unknowns. This so-called three-field variational principle dates back to the works of Hai-Chang (1954) and Washizu (1955) and is a generalization of the principle of virtual work. The Hu-Washizu functional is

$$\delta\Pi^{\mathrm{HW}}(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \int_\Omega \psi(\boldsymbol{\varepsilon}) + \boldsymbol{\sigma} : (\nabla_{\mathrm{s}}\boldsymbol{u} - \boldsymbol{\varepsilon})\,\mathrm{d}V - \int_\Omega \boldsymbol{b}\cdot\boldsymbol{u}\,\mathrm{d}V - \int_{\partial\Omega} \boldsymbol{t}\cdot\boldsymbol{u}\,\mathrm{d}A \quad . \tag{2.105}$$

By applying the directional derivative (A.39), three independent equations are obtained, which represent the constitutive relation

$$D_{\delta\boldsymbol{\varepsilon}}\Pi^{\mathrm{HW}}(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \int_\Omega \left(\boldsymbol{\sigma} - \frac{\partial\psi(\boldsymbol{\varepsilon})}{\partial\boldsymbol{\varepsilon}}\right) : \delta\boldsymbol{\varepsilon}\,\mathrm{d}V = 0 \quad , \tag{2.106}$$

the kinematic relation

$$D_{\delta\boldsymbol{\sigma}}\Pi^{\mathrm{HW}}(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \int_\Omega (\boldsymbol{\varepsilon} - \nabla_{\mathrm{s}}\boldsymbol{u}) : \delta\boldsymbol{\sigma}\,\mathrm{d}V = 0 \tag{2.107}$$

and the (weak form of the) static equilibrium

$$D_{\delta\boldsymbol{u}}\Pi^{\mathrm{HW}}(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}) = \int_\Omega \boldsymbol{\sigma} : \delta(\nabla_{\mathrm{s}}\boldsymbol{u}) - \boldsymbol{f}\cdot\delta\boldsymbol{u}\,\mathrm{d}V - \int_{\partial\Omega} \boldsymbol{t}\cdot\delta\boldsymbol{u}\,\mathrm{d}A = 0 \quad . \tag{2.108}$$

Commonly, by applying appropriate assumptions and using static condensation, the expressions can be reduced to a problem which, again, depends only on the unknown displacements. For more details, see e.g. Washizu (1975).

Hamilton's principle

Hamilton's principle[8] of stationary action states, that the evolution of a discrete mechanical system, which is described in terms of a finite number of generalized coordinates $\boldsymbol{q} = \{q_1, q_2 \ \dots \ q_n\}$, between two states at specified instants of time t_1 and t_2 is a stationary point of the action functional

$$H = \int_{t_1}^{t_2} L\big(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t)\big)\,\mathrm{d}t = \text{stat.} \quad , \tag{2.109}$$

with $L = K(\dot{\boldsymbol{q}}) - V(\boldsymbol{q})$ being the Lagrangian of the system including the kinetic energy $K(\dot{\boldsymbol{q}})$ and the potential energy $V(\boldsymbol{q})$. For its application to continuum bodies, we follow the concept described by Bedford (1985). The definition of an admissible motion $q(\boldsymbol{x}, t)$ is introduced as well as the corresponding comparison motion

$$q^*(\boldsymbol{x}, t, \epsilon) := q(\boldsymbol{x}, t) + \epsilon\, \eta_\mathrm{q}(\boldsymbol{x}, t) \quad , \tag{2.110}$$

with ϵ being a scalar parameter. The scalar field η_q vanishes on the Dirichlet boundary. In addition $\eta_\mathrm{q}(\boldsymbol{x}, t = t_1) = 0$ and $\eta_\mathrm{q}(\boldsymbol{x}, t = t_2) = 0$ must hold. The kinetic and potential energy in a continuum body can then be considered to be functions of multiple comparison motions q_i^* and their spatial and time derivatives, i.e.

$$K^*(\epsilon) = K(q_1^*, q_2^*, \dots, \nabla q_1^*, \nabla q_2^*, \dots, \dot{q}_1^*, \dot{q}_2^*, \dots) \tag{2.111}$$

and

$$\begin{aligned} \Pi^*(\epsilon) &= \Pi_\mathrm{int}^*(\epsilon) + \Pi_\mathrm{ext}^*(\epsilon) \\ &= \Pi(q_1^*, q_2^*, \dots, \nabla q_1^*, \nabla q_2^*, \dots, \dot{q}_1^*, \dot{q}_2^*, \dots) \quad . \end{aligned} \tag{2.112}$$

Since the expressions based on comparison motions approach the solution for $\epsilon \to 0$, the variational operator (A.38) is introduced to determine

$$\delta K = \left.\frac{\partial K^*}{\partial \epsilon}\right|_{\epsilon=0} \qquad \text{and} \qquad \delta\Pi = \left.\frac{\partial \Pi^*}{\partial \epsilon}\right|_{\epsilon=0} \quad . \tag{2.113}$$

Although the principle is only applicable to conservative systems, it is common practice (e.g. Bedford, 1985) to include also non-conservative generalized forces Q, which are assumed to be work-conjugate to q. Accordingly, their virtual work is

$$\delta W = Q\ \delta q \quad . \tag{2.114}$$

[8] In analogy to Fermat's principle (Born, 1980), that light traveling between two points will follow the path that requires the least time, Hamilton (1834; 1835) formulated a theory for dynamic systems consisting of particles.

The extended principle for an elastic material states (Bedford, 1985): "among admissible comparison motions, the actual motion of the material is such that"

$$\int_{t_1}^{t_2} \Big(\delta \underbrace{(K - \Pi)}_{=: L} + \delta W \Big) \mathrm{d}t = 0 \quad . \tag{2.115}$$

The functional

$$L = \int_{\Omega} \mathcal{L}(q_1, q_2, \dots) \, \mathrm{d}V = K - \Pi \tag{2.116}$$

is called Lagrangian and includes the Lagrangian density $\mathcal{L}$, which depends on admissible fields q_i. In addition, Bedford (1985) extends the fundamental lemma of variational calculus as follows:

"**Lemma 1**: Let $\mathcal{W}$ be an inner product space, and consider a C^0 field $\boldsymbol{f} : \Omega \times [t_1, t_2] \to \mathcal{W}$. If the equation

$$\int_{t_1}^{t_2} \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}V \, \mathrm{d}t = 0 \tag{2.117}$$

holds for every C^∞ field $\boldsymbol{w} : \Omega \times [t_1, t_2] \to \mathcal{W}$ that vanishes at time t_1, at time t_2, and on $\partial\Omega$, then $\boldsymbol{f} = \boldsymbol{0}$ on $\Omega \times [t_1, t_2]$."

"**Lemma 2**: Suppose that $\partial\Omega$ consists of complementary regular subsurfaces $\partial\Omega_1$ and $\partial\Omega_2$. Let $\mathcal{W}$ be an inner product space, and consider a function $\boldsymbol{f} : \Omega_2 \times [t_1, t_2] \to \mathcal{W}$ that is piecewise regular and continuous in time. If the equation

$$\int_{t_1}^{t_2} \int_{\partial\Omega_2} \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}V \, \mathrm{d}t = 0 \tag{2.118}$$

holds for every C^∞ field $\boldsymbol{w} : \Omega \times [t_1, t_2] \to \mathcal{W}$ that vanishes at time t_1, at time t_2, and on $\partial\Omega_1$, then $\boldsymbol{f} = \boldsymbol{0}$ on $\partial\Omega_2 \times [t_1, t_2]$."

Hamilton's principle is also known as principle of least action and can be seen as a generalization of the principle of stationary potential energy (2.104) to continuum dynamics. As demonstrated, e.g. by Willner (2003), previous results like the Hu-Washizu functional (2.105) can be applied to dynamic problems in continuum mechanics by using Hamilton's principle, i.e.

$$H_{\mathrm{HW}} = \int_{t_1}^{t_2} (K - \Pi_{HW}) \, \mathrm{d}t = \text{stat.} \tag{2.119}$$

2.1.5 Linear elastodynamics

The governing equations of linear elastodynamics are analyzed by following the comprehensive textbook of Ravi-Chandar (2004). Substitution of the kinematic relations (2.17) and the elasticity law (2.83) in the equation of motion (2.47), see (A.51), yields a hyperbolic system of three partial differential equations

$$\rho\,\ddot{\boldsymbol{u}} = (\lambda+\mu)\,\nabla(\operatorname{div}\boldsymbol{u}) + \mu\,\Delta\boldsymbol{u} + \boldsymbol{f} \quad . \tag{2.120}$$

This set of equations, also known as *Navier equations*, provides the basic equations of linear elastodynamics. It depends only on the displacement field, but contains derivatives of high order. By neglecting body forces ($\boldsymbol{f} = \boldsymbol{0}$) and making use of the identity $\nabla(\operatorname{div}\boldsymbol{u}) - \Delta\boldsymbol{u} = \operatorname{curl}(\operatorname{curl}\boldsymbol{u})$, see (A.52), it is

$$\begin{aligned}\ddot{\boldsymbol{u}} &= \frac{\lambda+2\mu}{\rho}\,\nabla(\operatorname{div}\boldsymbol{u}) + \frac{\mu}{\rho}\,\big(\Delta\boldsymbol{u} - \nabla(\operatorname{div}\boldsymbol{u})\big)\\ &\overset{(A.53)}{=} \frac{\lambda+2\mu}{\rho}\,\nabla(\operatorname{div}\boldsymbol{u}) - \frac{\mu}{\rho}\,\operatorname{curl}(\operatorname{curl}\boldsymbol{u}) \quad .\end{aligned} \tag{2.121}$$

Application of the divergence operator on (2.121) with (A.36) extracts the equation for irrotational deformation ($\operatorname{curl}\boldsymbol{u} = \boldsymbol{0}$)

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\operatorname{div}\boldsymbol{u}) = c_{\mathrm{d}}^2\,\Delta(\operatorname{div}\boldsymbol{u}) \quad , \tag{2.122}$$

while the application of the curl operator on (2.121) with (A.37) extracts the equation for equivoluminal deformations ($\operatorname{div}\boldsymbol{u} = 0$)

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\operatorname{curl}\boldsymbol{u}) = c_{\mathrm{s}}^2\,\Delta(\operatorname{curl}\boldsymbol{u}) \quad . \tag{2.123}$$

Both relations are known as the *Helmholtz wave equations*. Alternatively, by introducing the auxiliary functions $\phi_{\mathrm{L}}(\boldsymbol{x},t)$ and $\boldsymbol{\psi}_{\mathrm{L}}(\boldsymbol{x},t)$ according to

$$\boldsymbol{u} = \nabla\phi_{\mathrm{L}} + \operatorname{curl}\boldsymbol{\psi}_{\mathrm{L}} \tag{2.124}$$

and $\operatorname{div}\boldsymbol{\psi}_{\mathrm{L}} = 0$, one obtains the wave equations

$$\ddot{\phi}_{\mathrm{L}} = c_{\mathrm{d}}^2\,\Delta\phi_{\mathrm{L}} \qquad \text{and} \qquad \ddot{\boldsymbol{\psi}}_{\mathrm{L}} = c_{\mathrm{s}}^2\,\Delta\boldsymbol{\psi}_{\mathrm{L}} \quad , \tag{2.125}$$

also known as *Lamé's solution*. Note, that $\operatorname{div}\boldsymbol{u} = \Delta\phi_{\mathrm{L}} - 0$ and $\operatorname{curl}\boldsymbol{u} = \boldsymbol{0} - \operatorname{curl}(\operatorname{curl}\boldsymbol{\psi}_{\mathrm{L}})$. The *Lamé potentials* ϕ_{L} and $\boldsymbol{\psi}_{\mathrm{L}}$ correspond to the dilatational wave and shear waves, respectively. For plane strain problems, e.g. with $u_3 = 0$, $\psi_{\mathrm{L}\,1} = \psi_{\mathrm{L}\,2} = 0$, the second relation in (2.125) is simplified to $c_{\mathrm{s}}^2\,\psi_{\mathrm{L}\,3,ii} = \ddot{\psi}_{\mathrm{L}\,3}$.

The coefficients c_{d} and c_{s} consist of material parameters only and can be identified as the speed of dilatational (longitudinal) waves

$$c_{\mathrm{d}} = \sqrt{\frac{\lambda+2\mu}{\rho}} = \sqrt{\frac{E\,(1-\nu)}{\rho\,(1+\nu)(1-2\nu)}} \tag{2.126}$$

and the speed of shear (distortional, transversal, equivoluminal) waves

$$c_\mathrm{s} = \sqrt{\frac{\mu}{\rho}} = \sqrt{\frac{E}{2\rho\,(1+\nu)}} \quad . \tag{2.127}$$

In silica glass, for instance, the wave speeds are about $c_\mathrm{d} = 5500\,\mathrm{m/s}$ and $c_\mathrm{s} = 3300\,\mathrm{m/s}$ (Ravi-Chandar, 2004). It is interesting to note that the ratio

$$\frac{{c_\mathrm{s}}^2}{{c_\mathrm{d}}^2} = \frac{\mu}{\lambda + 2\mu} = \frac{1-2\nu}{2(1-\nu)} < 1 \tag{2.128}$$

depends only on Poisson's ratio ν. Typically, dilatational waves are about two-times faster than shear waves, especially in brittle materials. In seismology, these waves are called primary and secondary waves due to their order of occurrence to an observer.

Dilatational and distortional waves exist inside a solid body. At surfaces or at discontinuities inside the body, these waves are reflected or partially reflected, so that the boundary conditions are fulfilled at any time. A dilatational wave, for instance, is reflected at a boundary with fixed displacements with a change of sign. On surfaces, however, a third form of waves exists, discovered by Rayleigh. They travel only along the free surface and decay rapidly in the body with increasing distance to the surface. Regarding fracture mechanics, they play a crucial role, since fracture creates new surfaces. First, a semi-infinite body occupying the upper half space with $x_2 = 0$ being a traction-free surface under plane strain conditions is regarded. For a wave traveling along the x_1-direction the only solution is, see e.g. Ravi-Chandar (2004),

$$\phi_\mathrm{L} = A\,\exp(-\alpha\,x_2)\,\cos[k\,(x_1 - c_\mathrm{R}\,t)] \tag{2.129}$$

and

$$\psi_{\mathrm{L}\,3} = B\,\exp(-\beta\,x_2)\,\cos[k\,(x_1 - c_\mathrm{R}\,t)] \quad , \tag{2.130}$$

with the wave number k and the unknown Rayleigh wave speed c_R. For this particular case, the Rayleigh function governs the Rayleigh wave speed with

$$R(c_\mathrm{R}) = 4\sqrt{1-\left(c_\mathrm{R}/c_\mathrm{d}\right)^2}\,\sqrt{1-\left(c_\mathrm{R}/c_\mathrm{s}\right)^2} - \left(2-\left(c_\mathrm{R}/c_\mathrm{s}\right)^2\right)^2 = 0 \quad , \tag{2.131}$$

which has only one real root in the speed range $0 < c_\mathrm{R} < c_\mathrm{s}$ (Achenbach, 1973). As the Poisson's ratio ν varies from 0 to 0.5, the Rayleigh wave speed c_R increases monotonically from about $0.8740\,c_\mathrm{s}$ to $0.9553\,c_\mathrm{s}$ and is clearly below c_s. The early approximation of the solution

$$c_\mathrm{R} \approx \frac{0.87 + 1.12\,\nu}{1+\nu}\,c_\mathrm{s} \tag{2.132}$$

proposed by Bergmann (1948) is only accurate for $\nu \geq 0$. A polynomial curve fit of data in the range $\nu \in (-1,\ 0.5)$ in a least-squares sense provides the more accurate solution

$$c_\mathrm{R} \approx \frac{1}{1000}\left(873 + 192\nu - 35\nu^2 - 43\nu^3\right)c_\mathrm{s} \quad . \tag{2.133}$$

2.2 Fundamentals of fracture and damage mechanics

Fracture, in general, describes the separation of a solid body by introducing or extending crack surfaces within the body domain. From a mechanical point of view, a crack typically leads to discontinuous mechanical fields across the crack. Brittle fracture is a sudden process with (almost) no inelastic deformation before separation starts whereas ductile fracture occurs after notable inelastic deformations. Some failure hypotheses, based on simple stress or strain states (before failure), are discussed in Sect. 2.2.1. Classical fracture mechanics, by contrast, presumes macroscopic cracks causing discontinuous mechanical fields. Additionally, the crack front/tip leads to singular fields where the application of classical failure criteria fails. Moreover, the creation of new cracks cannot be predicted. In case of linear elastic material behavior and further assumptions discussed in Sect. 2.2.2 a number of fracture criteria are presented (Sect. 2.2.3-Sect. 2.2.5). A continuous separation is modeled by cohesive fracture models, as introduced in Sect. 2.2.7. By contrast, *damage mechanics* as discussed in Sect. 2.2.8 is a rather general expression which describes the ongoing change of material behavior due to the initiation and evolution of voids and microcracks usually leading to loss of load carrying ability on the macroscopic scale. The formulation of damage approaches can be strongly motivated by microscopic mechanisms, but can also be of phenomenological nature. If a damage mechanics approach leads to the situation that the material looses all strength localized on a plane (in 3D) or line (2D) the overall mechanical behavior is similar to a discrete crack. In this case, the continuous material degradation up to total failure can be used to model crack initiation. Following the comprehensive books on fracture (Gross and Seelig, 2011) and damage mechanics (Kachanov, 1986; Lemaitre and Desmorat, 2005; Murakami, 2012), a short introduction of some definitions and the basic concepts is provided in the following.

2.2.1 Failure hypotheses

Failure hypotheses are a concept to quantify the onset of non-elastic effects like plastic deformations or material failure. They characterize a limit state, at which the material behavior typically changes significantly. Strength hypotheses are motivated by experimental observations and can be used to estimate failure loads for simple loading scenarios. Corresponding material parameters are the tensile, compressive or yield strength, but also the failure strain. It should, however, be noted, that the onset of failure does not only depend on the material parameters but also on the loading conditions. While, e.g, for plastic deformations it is well known that hydrostatic stresses often do not lead to inelastic states, brittle materials are essentially influenced by these stresses. By assuming that the current material state can be characterized by a simple function of the current stress state (and does not depend, e.g. on the loading history) a classical strength hypothesis is formulated according to

$$F(\boldsymbol{\sigma}) = 0 \quad . \tag{2.134}$$

Here, $F(\boldsymbol{\sigma})$ is interpreted as a failure surface in the six-dimensional stress space or three-dimensional principal stress space. Alternatively, failure criteria can also be formulated in the strain space.

Failure criterion

Among the infinite number of possible failure criteria, simple failure criteria are the principal stress hypothesis, the principal strain hypothesis, and the strain energy hypothesis. Yield criteria, such as the von Mises or Tresca condition, and more advanced failure criteria, like the Coulomb-Mohr or Drucker-Prager hypotheses, which are mainly used to describe failure of granular or geological materials subjected to compression, and the Johnson-Cook criterion (Johnson and Cook, 1985), which describes ductile failure after pronounced plastic deformation, are out of the scope of this work.

According to the principal stress hypothesis

$$F(\boldsymbol{\sigma}) = \max\left\{\sigma_1 - \sigma_\mathrm{t},\ -\sigma_3 - \sigma_\mathrm{c}\right\} = 0 \tag{2.135}$$

(by assuming $\sigma_1 \geq \sigma_2 \geq \sigma_3$), failure is expected if the maximum principal stress reaches the material tensile strength $\sigma_\mathrm{t} > 0$ or the minimum principal stress reaches the (negative) compressive strength $\sigma_\mathrm{c} > 0$. This criterion, however, neglects the influence of the other stress components.

The principal strain hypothesis postulates that failure occurs if the maximum principal strain reaches a critical value $\varepsilon_\mathrm{crit} > 0$. Assuming linear elastic behavior prior to failure, the failure criterion might be related to the tensile strength $\sigma_\mathrm{t} = E\,\varepsilon_\mathrm{crit}$. The corresponding failure criterion reads

$$F(\boldsymbol{\sigma}) = \max\left\{\sigma_1 - \nu\,(\sigma_2 + \sigma_3),\ \sigma_2 - \nu\,(\sigma_1 + \sigma_3),\ \sigma_3 - \nu\,(\sigma_1 + \sigma_2)\right\} - \sigma_\mathrm{t} = 0 \quad . \tag{2.136}$$

For hydrostatic loading ($\sigma = \sigma_1 = \sigma_2 = \sigma_3$), in case of tension the criterion predicts failure for $\sigma = \sigma_\mathrm{t}/(1-2\nu)$, in case of compression no failure is predicted. In case of uniaxial compression, however, the criterion predicts failure for $\sigma_3 = \sigma_\mathrm{t}/\nu$ which is not in line with experimental findings (Gross and Seelig, 2011).

According to the strain energy hypothesis, failure is expected if the strain energy reaches a critical value ε_t. In case of linear elastic behavior prior to failure, the critical strain energy might be related to the tensile strength σ_t with $\psi_\mathrm{t} = \sigma_\mathrm{t}^2/(2E)$. By using the stress (2.83) the failure criterion reads

$$\begin{aligned} F(\boldsymbol{\sigma}) = {} & (1+\nu)\left((\sigma_1-\sigma_2)^2 + (\sigma_2-\sigma_3)^2 + (\sigma_1-\sigma_3)^2\right) \\ & + (1-2\nu)(\sigma_1+\sigma_2+\sigma_3)^2 - 3\sigma_c^2 = 0 \quad . \end{aligned} \tag{2.137}$$

Due to the presence of $\mathrm{tr}(\boldsymbol{\sigma})^2$ failure is predicted for hydrostatic compression which contradicts experimental findings.

Simple hypotheses, as given above, typically fail regarding some specific loading conditions. In case of extensive experimental data, more advanced criteria can be formulated. However, failure and strength hypotheses can only be applied to continuum problems which are not influenced by macroscopic defects like cracks.

Failure kinematics

The failure criterion determines the load state when failure occurs. In addition, failure is usually associated with specific deformations, e.g. plastic yielding or creation of fracture surfaces, and only takes place, if the deformations are kinematically admissible. Failure kinematics might be introduced by physically meaningful assumption based on experimental results. The creation of fracture surfaces, i.e. the separation of material within specific planes, and the corresponding kinematics can only be determined for some simple scenarios, where usually both the material behavior and the loading play a role. If fracture is dominated by normal stresses, the normal of the fracture surface coincides with the maximum principle (tensile) stress, as depicted by Fig. 2.5a. In other cases, the fracture surface is aligned with the direction of critical shear stresses, see Fig. 2.5b, but also mixed forms can occur.

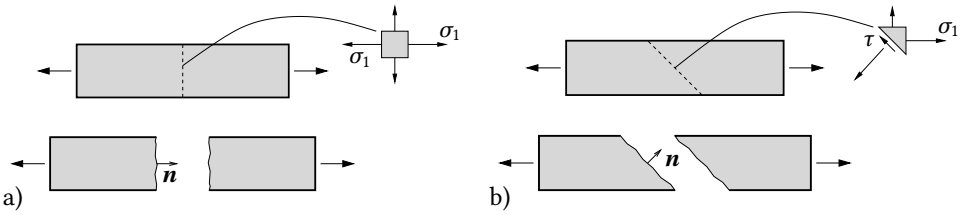


Figure 2.5: Creation of a fracture surface a) perpendicular to the maximum principle tensile stress, b) aligned with the direction of a critical shear stress.

2.2.2 Linear elastic fracture mechancis

In fracture mechanics, the influence of a crack and its growth are analyzed based on macroscopic (continuum) quantities only. Neither microscopic aspects, like the breaking of atomic bonds and the associated energies, nor defects on the mesoscale are considered. Instead a crack is regarded as a (perfect) cut in a continuous body. The crack's dimension is one dimension lower than the dimension of the body. Regarding a three-dimensional structure the crack introduces a surface with opposite crack flanks or crack faces which end at the crack front as depicted in Fig. 2.6a. In a two-dimensional problem the crack forms a line with opposite sides, also denoted as crack faces, which end at the crack tip (Fig. 2.6b). All fracture processes which cannot be described in terms of continuum mechanics, such as bond breaking, are comprised in a process zone, which encloses the crack front or crack tip. To be able to apply a continuum description (using the relations

provided in Sect. 2.1), so that fracture depends only on macroscopic quantities like strain, stress and strain energy, the size of the process zone must be negligibly small compared to all characteristic dimensions of the body (including the size of the crack). For pre-existing macroscopic cracks, this is fulfilled in many brittle materials, such as glass. Since the crack is treated as a perfect separation and no external forces are applied at the crack, the crack flanks are assumed to be traction-free. Obviously, this is only valid for an open crack, i.e. there is no contact between closed crack faces.

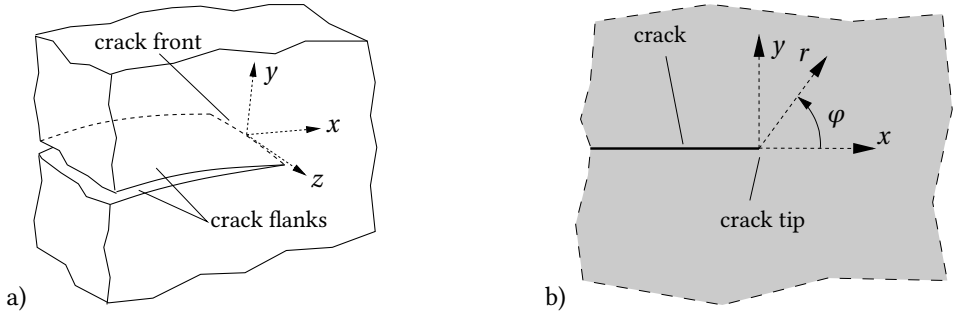


Figure 2.6: a) Crack in a three-dimensional body, b) local polar coordinate system (in 2D) at the crack tip.

The loading of cracks is typically classified according to their opening mode. It was Irwin (1957), who found that any load state on a crack can be decomposed into three independent modes as depicted in Fig. 2.7. The first type, mode I, describes symmetric crack opening with (positive) separation of the crack faces perpendicular to the crack plane ("tensile opening"). Mode II denotes an antisymmetric separation of the crack faces in the crack plane normal to the crack front ("in-plane shear"). Mode III is characterized by an antisymmetric separation in the crack plane parallel to the crack front ("out-of-plane shear" or "anti-plane shear").

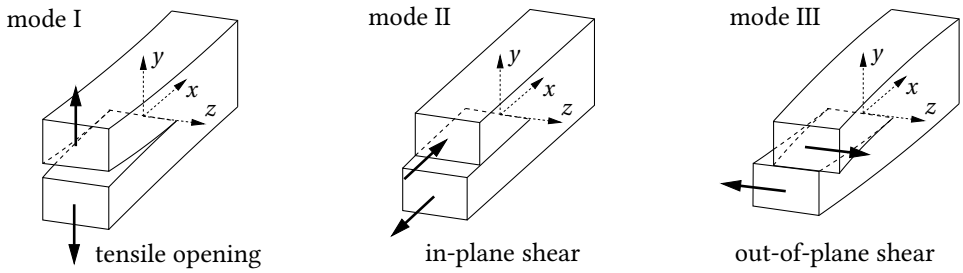


Figure 2.7: Crack opening modes as introduced by Irwin (1957).

As already mentioned, brittle fracture is characterized by almost no inelastic deformation before fracture occurs. For linear elastic fracture, i.e. the material in the entire region of the body is assumed to behave linear elastic prior to fracture, inelastic processes are restricted to a negligibly small region at the crack tip. Also plastic deformations, which occur for

high loads in real materials and blunt the crack tip, are assumed to be limited to a small region. This effect is known as *small scale yielding*. Consequently, stresses and strains outside the process zone are analyzed by using linear-elastic relations in order to gain some insight into the mechanical behavior at the crack tip. In addition, due to the linearity of the problem solutions can be obtained by superposition of multiple sub-solutions.

2.2.3 Stress intensity factors and K-concept

Investigation of the linear elastic strain and stress distribution in the vicinity of a stationary crack tip leads to the so-called crack-tip field. Without going into details, some relevant findings are summarized. The local concentration of stresses at the vertices of an elliptic hole (with the major and minor axes $2a$ and $2b$) in an infinite plate was investigated by Inglis (1913)[9]. Subjected to an uniaxial (tensile) stress σ_∞ normal to the major axis at infinity the maximum stress at the vertex is $\sigma_{\max} = \sigma_\infty(1 + 2a/b)$. While for a circular hole (with $a = b$) the well-known solution is $\sigma_{\max} = 3\sigma_\infty$ the extreme limit of increasing eccentricity[10] $b/a \to 0$ results in a fine straight crack where the maximum stress becomes singular, i.e. $\sigma_{\max} \to \infty$. The application of a strength hypothesis, e.g. the principal stress hypothesis (2.135), predicts that even a very small tensile load would start crack growth and tear the material apart. More powerful and elegant than the analysis of an ellipse (in elliptical coordinates) is the direct application of the complex variable technique[11] to a straight crack (in Cartesian coordinates) as done by Westergaard (1939).

Stress intensity factors

By introducing a local polar coordinate system (with the radial coordinate r and the polar angle φ measured relative to a straight crack extension) at the crack tip according to Fig. 2.6b one observes that the complex stress field and displacement field are dominated by a singular expression as r approaches zero. Additionally, inspired by the results obtained by Sneddon (1946b) and Williams (1952)[12], it was Irwin (1957) who showed that the magnitude of the singular stress field is only controlled by three independent stress

[9] Wieghardt (1907) discussed the problem of finding the linear elastic stress field in a wedge-shaped domain loaded by concentrated forces. He firstly observed the essential factorization of stress fields into radial terms and angular functions and demonstrated the square-root singularity of stress components at the crack tip, see the historical overview written by Rossmanith (1995). However, his findings did not expand into a fracture concept.

[10] To be precise, the eccentricity of an ellipse is the ratio of the distance between center and focus and the distance between center and vertex, i.e. $e = \sqrt{1 - (b/a)^2}$ with $e \in [0, 1]$.

[11] The analytic solution of any planar elastic problem can be obtained by two complex functions $\Phi(z)$ and $\Psi(z)$ where stress and displacement field are considered to be functions a of complex variable $z = r\ \exp(i\ \varphi)$ or its conjugate $\bar{z} = r\ \exp(-i\ \varphi)$. By means of Kolosov's formulas (Kolosov, 1909), these functions can be simply applied to problems in Cartesian (or polar) coordinates.

[12] Williams (1952) used the Airy stress function to derive the linear elastic stress distribution in the vicinity of a notch. He showed that the degree of the stress singularity depends on the notch opening angle.

intensity factors K_I, K_II and K_III[13]. They normalize the singular stress field, depend on the sample geometry (including the crack length) and the applied load and can be obtained by complex analyses or found in the literature (e.g. Tada et al., 2000). Thus, the stress of the crack-tip (near) field is given by

$$\sigma_{ij}(r, \varphi) = \frac{1}{\sqrt{2\pi\, r}} \left(K_\mathrm{I}\, \hat{\sigma}_{ij}(\varphi) + K_\mathrm{II}\, \hat{\sigma}_{ij}(\varphi) + K_\mathrm{III}\, \hat{\sigma}_{ij}(\varphi) \right) \tag{2.138}$$

where $\hat{\sigma}_{ij}$ are (dimensionless) angular distributions describing the spatial variation of the stress field only depending on the polar angle φ, but independent of external loads and the geometry. Explicit representations of the stress functions (2.138) as well as the displacement and strain field are provided in App. A.8. They are assumed to be valid in a region around the crack tip (also called K-determined field) which is larger than the process zone, but much smaller than the characteristic dimensions of the problem, i.e. for a radius $r \ll \min\{a, L\}$ with the crack length a and the characteristic body size L.

K-concept

Since stress intensity factors fully represent the loading at the crack tip, crack propagation is assumed if a critical material value is reached, the so-called K-concept. For a pure mode I loading scenario, i.e. applied tension perpendicular to the crack plane, the criterion is

$$K_\mathrm{I} = K_\mathrm{Ic} \tag{2.139}$$

where K_Ic is the fracture toughness of the material and has to be determined experimentally[14]. In case of pure mode II or mode III loading fracture criteria can be formulated accordingly, i.e.

$$K_\mathrm{II} = K_\mathrm{IIc} \qquad \text{or} \qquad K_\mathrm{III} = K_\mathrm{IIIc} \quad . \tag{2.140}$$

While this concept is easily applicable to a wide class of problems, two issues should be mentioned. First, in mixed mode problems, the formulation of a reasonable criterion is more complex and might be of the format

$$f(K_\mathrm{I}, K_\mathrm{II}, K_\mathrm{III}) = 0 \quad . \tag{2.141}$$

Besides the critical load to determine crack initiation, also the direction of crack extension must be determined. Only in the case of pure mode I loading a straight crack can be assumed to propagate in its longitudinal direction. Second, the stress field given in (2.138) exhibits the typical $1/\sqrt{r}$-singularity of a crack tip. In case of a V-notch with opening

[13] Irwin allegedly chose the letter K to honor his co-worker Joseph A. Kies (Paris, 2014).

[14] A material's fracture toughness K_Ic is determined experimentally by correlating stress intensity factor, loading and crack length of a standardized test specimen (ASTM, 1997), which is initially notched and prepared with a starter crack caused by oscillating load. The fracture toughness of glass is about $20 \ldots 40\,\mathrm{MPa\,mm^{1/2}}$.

angle $2\pi - 2\alpha$ as depicted in Fig. 2.8a the stress field at the notch can be represented in a similar manner, i.e.

$$\sigma_{ij}(r, \phi) = r^{\lambda-1}\, \hat{\sigma}_{ij}(\phi) \quad , \tag{2.142}$$

however, with a different type of singularity, represented by $\lambda \in (0,\, 1]$. Complex analysis yields the expression

$$\sin(2\lambda\, \alpha) = \lambda\, \sin(2\alpha) \tag{2.143}$$

from which $\lambda \in (0,\, 1]$ can be calculated. Results depending on the opening angle are presented in Fig. 2.8b. The limit states $\alpha = \pi/2$ and $\alpha = \pi$ represent a straight outer edge (without stress concentration) and the (perfect) crack, respectively. It should be noted that, due to the different type of singularity, the fracture toughness (e.g. K_{Ic}) determined as material parameter at a perfect pre-crack is no longer applicable. A concept to treat this class of problems, as well as to treat, e.g. blunt notches with high but non-singular stress concentrations, is discussed in Sect. 2.2.6.

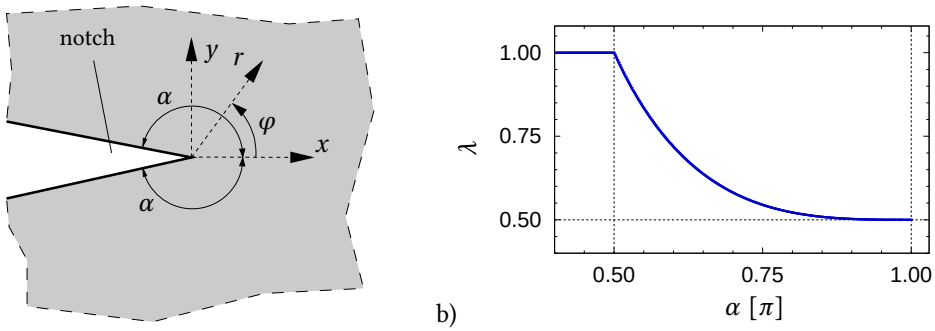


Figure 2.8: a) Specimen with V-notch of opening angle $2\pi - 2\alpha$, b) dependency of stress singularity type $\sigma \sim r^{\lambda-1}$ on the opening angle of the notch (λ corresponds to smallest eigenvalue).

2.2.4 Energetic fracture criteria

As mentioned in Section 2.2.1, inelastic processes (such as bond breaking) to create separated crack surfaces demand a specific amount of energy. On the other hand, crack advance might reduce stresses along the newly formed crack to zero which implicates the release of elastic energy.

Griffith's energy release rate

By regarding the energetic balance including both portions Griffith (1921) developed a general criterion to determine crack initiation. In the following, the fracture surface energy Γ is assumed to be proportional to the crack area A, i.e.

$$\Gamma = \mathcal{G}_{\mathrm{c}}\, A \quad . \tag{2.144}$$

The proportionality factor $\mathcal{G}_c$ is a material constant and called crack resistance. In contrast to common surface energies, Γ comprises all inelastic processes related to fracture. Often the specific fracture surface energy γ_c is given in the literature[15]. The simple proportionality assumption (2.144) contains the assumptions that the surface energy does neither depend on the position (homogeneity) nor on the direction (isotropy) of crack surface creation. Additionally, it implies that $\mathcal{G}_c$ is independent of, e.g., the speed of crack propagation. The energy balance (2.61) supplemented by the surface energy term is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(K + E + \Gamma\right) = P_{\mathrm{ext}} + Q \quad . \tag{2.145}$$

By assuming quasi-static crack advance and neglecting thermal processes the kinetic energy (including its rate) as well the thermal power are zero and the energy balance is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(E + \Gamma\right) = P_{\mathrm{ext}} \quad . \tag{2.146}$$

Conservative external forces possess the potential $P_{\mathrm{ext}} = -\mathrm{d}\Pi_{\mathrm{ext}}/\mathrm{d}t$ while for purely elastic material the internal energy is $E = \Pi_{\mathrm{int}}$. With the total potential $\Pi = \Pi_{\mathrm{int}} + \Pi_{\mathrm{ext}}$ the energetic balance is

$$\frac{\mathrm{d}\Pi}{\mathrm{d}t} + \frac{\mathrm{d}\Gamma}{\mathrm{d}t} = 0 \tag{2.147}$$

or equivalently, by regarding the infinitesimal advance of crack area $\mathrm{d}A$ during $\mathrm{d}t$,

$$\left(\frac{\mathrm{d}\Pi}{\mathrm{d}A} + \frac{\mathrm{d}\Gamma}{\mathrm{d}A}\right)\frac{\mathrm{d}A}{\mathrm{d}t} = 0 \quad . \tag{2.148}$$

The energy release rate is defined as the release of potential energy at an infinitesimal crack extension

$$\mathcal{G} := -\frac{\mathrm{d}\Pi}{\mathrm{d}A} \quad . \tag{2.149}$$

The negative sign is required since the energy release $\mathrm{d}\Pi < 0$ for infinitesimal crack advance $\mathrm{d}A > 0$ should be positive. In conjunction with (2.144) the balance of energy (2.148) using $\dot{A} := \mathrm{d}A/\mathrm{d}t$ is

$$\left(\mathcal{G}_c - \mathcal{G}\right)\dot{A} = 0 \quad . \tag{2.150}$$

Griffith originally derived only the energetic equivalence

$$\mathcal{G} = \mathcal{G}_c \tag{2.151}$$

to determine crack initiation. Regarding (2.150), as long as $\mathcal{G} < \mathcal{G}_c$ the extension of crack surface $\dot{A}$ must be zero. Thus, according to the Griffith criterion (2.151) fracture is initiated if the energy release during infinitesimal crack growth is equal to the energy

[15] Particular care must be taken since γ_c is usually related to a single crack surface. Referred to the crack area it is $\mathcal{G}_c = 2\gamma_c$.

required for the corresponding creation of crack surface. $\mathcal{G}_c$ is also referred to as critical energy release rate. However, the energetic balance (2.150) must be satisfied at any time. Consequently, in case of crack growth $\dot{A} > 0$ the relation (2.151) must hold as well. In addition, since the fracture process is irreversible the rate of fracture energy must be non-negative, i.e. $\dot{\Gamma} = \mathrm{d}\Gamma/\mathrm{d}t \geq 0$. By using (2.144) this is equivalent with the non-negative extension of crack area $\dot{A} \geq 0$.

For linear elastic fracture mechanics the relation between stress intensity factors and the energy release rate is

$$\mathcal{G} = \frac{1}{E'} \left(K_{\mathrm{I}}^2 + K_{\mathrm{II}}^2 \right) + \frac{K_{\mathrm{III}}^2}{2\mu} \tag{2.152}$$

The parameter E' takes into account the difference between the plane stress state on one hand with $E' := E$ and the plane strain state or three-dimensional assumption on the other hand with $E' := E/(1 - \nu^2)$, see App. A.6. Due to relation (2.152) for pure mode I loading the K-concept (2.139) and the energetic criterion (2.151) are equivalent.

J-integral

Another approach to compute the energy release rate was introduced independently by Cherepanov (1967) and Rice (1968a) and is based on the evaluation of a conservation integral. The so-called J-integral vector is defined as

$$J_k = \int_{\partial P} \underbrace{\left(\psi(\boldsymbol{\varepsilon})\, \delta_{jk} - \sigma_{ij} \frac{\partial u_i}{\partial x_k} \right)}_{=: \Sigma_{kj}} n_j \, \mathrm{d}A \tag{2.153}$$

and evaluated by integrating over an arbitrary closed surface ∂P (with the outward unit normal vector $\boldsymbol{n}$) inside a homogeneous elastic body (of volume $\Omega \supseteq P$) by neglecting volume forces, see Fig. 2.9a. The second order tensor Σ_{kj} in (2.153) is the Eshelby stress tensor (Eshelby, 1951). Since under the presumed circumstances its divergence vanishes, application of the divergence theorem (A.44) gives the conservation law

$$\begin{bmatrix} J_1 \\ J_2 \\ J_3 \end{bmatrix} = \int_{\partial P} \boldsymbol{\Sigma} \cdot \boldsymbol{n} \, \mathrm{d}A = \int_P \mathrm{div}(\boldsymbol{\Sigma}) \, \mathrm{d}V = \boldsymbol{0} \quad , \tag{2.154}$$

which can be utilized in the following. If the volume P contains any inhomogeneities like defects or cracks the Eshelby stress tensor Σ_{kj} becomes discontinuous or singular and $J_k \neq 0$.

In a plane problem the integral in (2.153) can be evaluated by integrating along an arbitrary closed curve C_{P} with the arc length parameter s. To evaluate the situation at a crack tip, a straight crack in x-direction with traction-free crack faces is considered, as depicted in

Fig. 2.9b. The path of the contour integral C_P encloses the crack tip and starts and ends at the opposite crack faces. As discussed e.g. by Gross and Seelig (2011), the quantity J_k (2.153) characterizes the energy change of an elastic system if any kind of enclosed discontinuity (surface, line or point defect) undergoes translational movement in k-direction. Since the only kinematically admissible displacement corresponds to crack advance in x-direction (indicated by da in Fig. 2.9b) the corresponding J-integral is

$$J := J_1 = \int_{C_\mathrm{P}} \left(\psi(\boldsymbol{\varepsilon})\, n_1 - \sigma_{ij}\, n_j\, \frac{\partial u_i}{\partial x} \right) \mathrm{d}s \quad . \tag{2.155}$$

By using the Cauchy theorem (2.30) and the relation $\mathrm{d}y = n_1\, \mathrm{d}s$, the J-integral (2.155) can alternatively be written as

$$J = \int_{C_\mathrm{P}} \left(\psi(\boldsymbol{\varepsilon})\, \mathrm{d}y - t_i\, \frac{\partial u_i}{\partial x}\, \mathrm{d}s \right) \quad . \tag{2.156}$$

Since, as discussed by Rice (1968b), the J-integral corresponds to the energy release for an infinitesimal crack advance da in crack direction, i.e.

$$J = \lim_{\Delta a \to 0} -\frac{\Delta \Pi}{\Delta a} = -\frac{\mathrm{d}\Pi}{\mathrm{d}a} = \mathcal{G} \quad , \tag{2.157}$$

it can be directly applied as a fracture parameter and the corresponding fracture criterion $J = J_\mathrm{c}$ is equivalent to the energy criterion given in (2.151). Furthermore, for linear elastic material

$$J = \frac{1}{E'} \left(K_\mathrm{I}^2 + K_\mathrm{II}^2 \right) + \frac{1}{2\mu} K_\mathrm{III}^2 \quad , \tag{2.158}$$

Subjected to a pure mode I, mode II or mode III loading the fracture criterion corresponds to the K-concept.

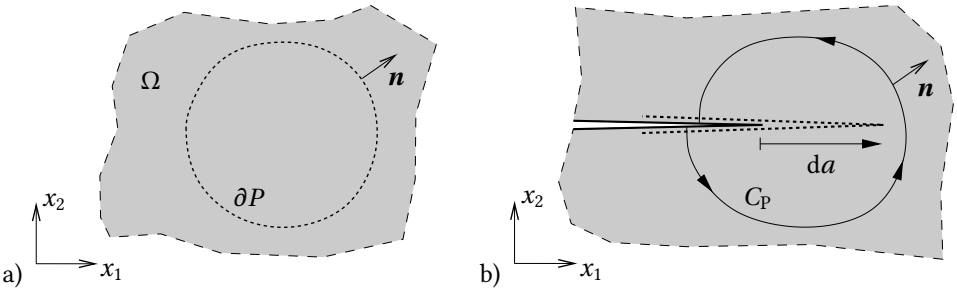


Figure 2.9: a) Evaluation of the J-integral by integrating over an arbitrary closed surface ∂P. b) Contour path C_P enclosing the crack tip with outward pointing normal $\boldsymbol{n}$ to compute the J-integral. Hypothetical crack movement in x_1-direction by da.

A key feature of the J-integral, besides the fact that it is not limited to linear elasticity as the K-concept, is its path independence for unloaded and straight cracks and in absence

of volume forces. This becomes relevant for numerical approaches, where the appropriate approximation of the crack tip singularity and evaluation of strong gradients in the vicinity of the crack tip can be avoided by choosing the integration path at a sufficiently large distance. If one or several of the requirements are not fulfilled (e.g. in case of a loaded or curved crack) the J-integral is not path independent any more. In this case the parameter can only be evaluated if the integration contour C_P is shrunk directly to the crack tip

$$J = \lim_{r \to 0} \int_{C_\mathrm{P}} \left(\psi(\boldsymbol{\varepsilon}) \, \mathrm{d}y - t_i \, u_{i,x} \, \mathrm{d}s \right) \quad . \tag{2.159}$$

2.2.5 Fracture criteria for mixed-mode loading

While for straight crack growth (pure mode I loading) only the critical load must be determined, e.g. by applying one of the concepts presented above, for mixed loading also the direction of crack progress is unknown. Apart from some very specific situations, crack advance in case of mixed-mode loading does not take place in tangential direction. Therefore, fracture criteria to determine both the critical load and the angle (defined relative to straight crack progress) of crack growth have been proposed. The most relevant criteria for mixed mode I-mode II fracture are the maximum hoop stress criterion, the strain-energy-density criterion, the maximum energy release rate criterion, and the principle of local symmetry. It should be noted, that these criteria are based on classical fracture mechanics. Thus, they do not account for the micro-mechanisms of fracture, which vary between different materials.

The maximum hoop/tangential stress criterion (MTS) proposed by Erdogan and Sih (1963), who first studied fracture under non-symmetric loading conditions, is based on the assumption that the crack grows perpendicular to the maximum circumferential stress σ_ϕ of the near-field solution[16]. In addition, crack growth is assumed to start as soon as a critical stress some distance in front of the crack tip reaches the critical value of a pure mode I loading. One outcome is, that for pure mode II loading, the critical load is $K_\mathrm{II} = \sqrt{3/4}\, K_\mathrm{Ic} \approx 0.866\, K_\mathrm{Ic}$ and the deflection angle is $\phi = -70.6^\circ$.

The strain-energy-density criterion (SED), also known as S-criterion, was proposed by Sih (1973). It assumes that crack growth takes place in the direction where the (singular) strain energy density has a minimum. Crack growth starts if a critical energy density is reached.

In contrast to the criteria mentioned before, the maximum energy release rate criterion (ERR) does not depend on the near-field solution prior to crack propagation. Instead, it is based on Griffith's concept of energy release rate. Crack growth takes place in the direction with the maximum energy release. Numerical methods must be used to evaluate

[16] According to the comments by Paris (2014), the experiments did not accurately capture a brittle elastic response, so the conclusions are in dispute.

the corresponding relations. The idea was already mentioned by Erdogan and Sih (1963), but analyzed in more detail by Hussain et al. (1974) and Wu (1978).

The principle of local symmetry (PLS) according to Gol'dstein and Salganik (1974) states that crack growth takes place in such a way that in-plane shear stresses vanish in the vicinity of the crack tip, i.e. the mode II stress intensity factor ahead of the propagated crack is $K_{\mathrm{II}} = 0$. Thus, crack growth starts when $K_{\mathrm{I}} = K_{\mathrm{Ic}}$ is reached.

2.2.6 Hybrid fracture initiation criterion

Crack initiation from sharp notches with finite opening angle (V-notches, see Fig. 2.8a), blunt notches, cavities and holes belongs to a class of problems not accessible to classical fracture mechanics discussed above where the pre-existence of a well established crack (with a singular $r^{-1/2}$ stress concentration) is required. As a remedy and an alternative to the classical idea of infinitesimal crack growth controlled by a stress intensity factor (Sect. 2.2.3) or, equivalently, by the energy release rate (Sect. 2.2.4), several researchers have suggested to instead consider the instantaneous formation of a crack of small, yet finite, length (e.g. Hashin, 1996; Leguillon, 2001; Taylor et al., 2005). As introduced in the pioneering work of Hashin (1996), this concept is referred to as *finite fracture mechanics* (FFM) in the following. Bringing this approach into a rigorous form Leguillon (2002) proposed a two-part fracture initiation criterion that involves two well-defined material parameters, the (tensile) strength σ_{t} and the specific (i.e. per area) fracture energy $\mathcal{G}_{\mathrm{c}}$ (toughness). This *hybrid fracture initiation criterion* states that over the prospective finite crack increment (with the area ΔA_{c}) a certain (pointwise) stress measure $F(\boldsymbol{\sigma})$ as well as the released energy (on average) $\bar{\mathcal{G}} := -\Delta\Pi/\Delta A_{\mathrm{c}}$, with Π being the total potential energy, must exceed critical values, i.e.

$$F(\boldsymbol{\sigma}) \geq \sigma_{\mathrm{t}} \qquad \underline{\text{and}} \qquad \bar{\mathcal{G}} \geq \mathcal{G}_{\mathrm{c}} \quad . \tag{2.160}$$

The quantity $\bar{\mathcal{G}}$ in the latter part of (2.160) is referred to as the *average energy release rate* in the following. For brittle materials typical stress measures are the maximum principal stress

$$F(\boldsymbol{\sigma}) := \sigma_1 \tag{2.161}$$

or the normal stress perpendicular to the finite crack increment

$$F(\boldsymbol{\sigma}) := \boldsymbol{n} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{n} \quad , \tag{2.162}$$

with $\boldsymbol{n}$ being the normal vector on the surface of the finite crack increment. Alternatively, instead to demand the pointwise fulfillment of the stress condition its fulfillment might only be required in the mean along ΔA_{c} according to (Cornetti et al., 2006)

$$F(\boldsymbol{\sigma}) := \frac{1}{\Delta A_{\mathrm{c}}} \int_0^{\Delta A_{\mathrm{c}}} \sigma_1 \, \mathrm{d}A \quad , \tag{2.163}$$

which is sometimes referred to as Line Method (Taylor, 2007). According to Gross and Seelig (2011), the evaluation of (2.163) is more practical in many cases.

As a consequence, the hybrid criterion encompasses the two limiting cases of crack initiation as elaborated by, e.g., Hebel (2010). In case of a homogeneous stress field failure is controlled essentially by the stress part of the hybrid criterion (2.160). Considering a linear elastic 1D bar of length L with the cross-sectional area A and Young's modulus E subjected to uniaxial tension σ the elastic energy is $\Pi = \sigma^2 \, L/(2E)$. By assuming the formation of a transverse crack (with the area A) the energy part of the hybrid criterion (2.160) is fulfilled for

$$L \geq 2\frac{\mathcal{G}_\mathrm{c} \, E}{\sigma^2} \quad . \tag{2.164}$$

Thus, if the body is sufficiently large, i.e. the 1D bar is longer than $L = 2\mathcal{G}_\mathrm{c} \, E/\sigma_\mathrm{t}^2$, fracture is controlled solely by the stress part which corresponds to a classical strength hypothesis (2.134). In the other limiting case, a sharp crack, the stress part of the hybrid criterion (2.160) is always fulfilled due to the singular stress field at the crack tip. Fracture is then controlled solely by the energy part of the hybrid criterion (2.160) which corresponds to the established relation of an incremental energy release rate, see Sect. 2.2.4.

By using a hybrid fracture criterion it is then possible to determine crack initiation from V-notched specimens (e.g. Leguillon, 2002; Yosibash et al., 2006), from circular holes (e.g. Li and Zhang, 2006), for indentation fracture (Dowgiałło, 2016), and in many other situations as thoroughly discussed by Weißgraeber et al. (2016). The criterion is, moreover, capable to predict size effects, see e.g. Leguillon et al. (2007). Hertzian indentation fracture where crack initiation is assumed to take place from a defect-free surface solely due to the presence of a (nonsingular) stress concentration is analyzed in detail in Chapter 4.

2.2.7 Cohesive zone model

Instead of assuming complete and instantaneous separation of the material at the crack tip in a process zone of vanishing size, a cohesive zone model treats fracture as a gradual separation of the crack surfaces in a zone of finite size ahead of the crack tip. The so-called *cohesive zone* is characterized by material bridges between the crack flanks, which are able to transfer forces depending on the opening state as illustrated in Fig. 2.10. For their numerical treatment, e.g. in finite element simulations, cohesive zones are discretized by cohesive elements of vanishing (initial) thickness. Cohesive behavior is characteristic, e.g., for ductile materials, where a large amount of plastic/inelastic deformation appears and the crack tip is blunted. The state of the plastically deformed crack tip is commonly characterized either by the crack-tip opening displacement (CTOD), independently proposed by Cottrell (1960) and Wells (1961), or the J-integral (Rice, 1968a). The first cohesive zone approaches date back to the models of Barenblatt (1959) and Dugdale (1960).

Dugdale (1960) introduced a model to analyze the influence of a thin, but elongated plastic zone ahead of the crack tip, which appears in thin plates of ductile material. In this

case, the plastic zone is idealized as line, but its length d can be of the order of the crack length or the structure size. With the assumption of perfect plasticity, the stress in the plastic zone is constant (equals the yield stress σ_y) and independent of the separation δ of the crack flanks, the crack-tip opening displacement. The size d of the plastic zone increases with the applied load σ and can be determined from the conditions, that the stress cannot exceed the yield stress and that the singularity at the tip of the plastic zone must vanish. While for the limit load ($\sigma \to \sigma_y$) plastic collapse occurs with $d \to \infty$, for small loads ($\sigma \ll \sigma_y$) the size of the plastic zone is $d = \pi K_I^2/(8\,\sigma_y^2)$, similar to the zone size Irwin (1960) obtained for small scale yielding $d_{ssy} = K_I^2/(\pi\,\sigma_y^2)$. In the latter case, the fracture criterion reduces to that of linear elastic fracture mechanics (2.139).

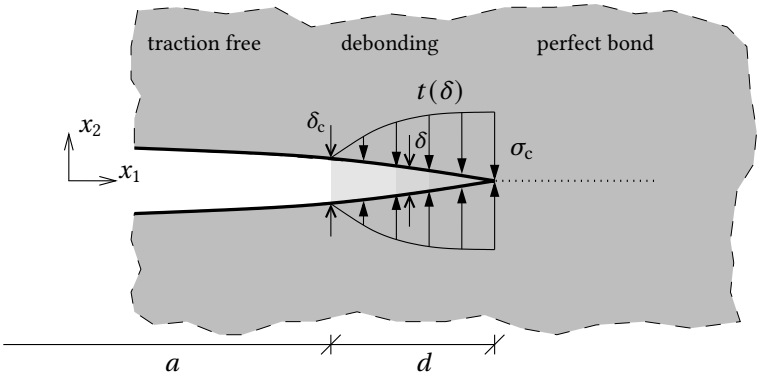


Figure 2.10: Schematic representation of a crack with cohesive zone ahead of the crack tip.

The model of Barenblatt (1959) is based on a cohesive stress $\sigma_{coh}(\delta)$ to describe brittle fracture. Thus, the cohesive zone is assumed to be small. The surface energy increases with crack opening up to a finite amount of crack opening (large compared to atomic length) where traction forces vanish and the Griffith energy $\mathcal{G}_c$ is reached. With the additional assumption, that the stress distribution and crack opening in the cohesive zone are material dependent and by assuming a vanishing singularity at the tip of the plastic zone, the concept finally yields a fracture criterion which is equivalent to linear elastic fracture mechanics (2.139).

In general, the size of the cohesive zone is not small and depends on the load. A traction separation or cohesive law of the format $t(\delta) = f(\delta)$ (here only normal crack opening is regarded) should represent the material behavior and usually depends on a small number of material specific parameters such as the ultimate stress σ_c with the associated separation δ_0 or the maximum separation δ_c for which traction exists. Common choices are, e.g., an exponential law, a function with trapezoidal shape or a bilinear function. The

energy release rate is equivalent to the area below the traction separation curve. Thus, the specific work of separation is

$$\mathcal{G}_\mathrm{c} = \int_0^{\delta_\mathrm{c}} t(\delta)\, \mathrm{d}\delta \quad . \tag{2.165}$$

In the case that the traction separation relation is characterized by two parameters, e.g. the strength σ_c and the specific work of separation $\mathcal{G}_\mathrm{c}$, the cohesive laws can be regarded as hybrid criterion (Sect. 2.2.6).

2.2.8 Continuum damage mechanics

Most materials exhibit microstructural defects like cracks and voids. Subjected to mechanical load, these defects might grow, coalesce to larger defects, and even new defects can nucleate. The material undergoes microstructural changes, which are considered to be irreversible in most cases. Different mechanisms on the micro- and mesoscale change material properties on the continuum scale, which are said to be caused by damage. Continuum damage theories are either formulated based on micromechanical or phenomenological approaches. Different forms of damage exist, e.g. creep, fatigue or ductile damage, with each of them caused by one or several different mechanisms. In the following, only phenomenological approaches for brittle damage are considered. Brittle damage is mainly governed by initiation and growth of microcracks, eventually leading to macroscopic cracks as discussed in Sect. 2.2.2 and failure of the structure.

To describe the state of damage in a material phenomenologically, at least one internal variable, denoted as damage parameter $D \in [0,\, 1]$, has to be introduced. It characterizes the gradual deterioration of material between two limit states, $D = 0$ for intact material and $D = 1$ at complete loss of integrity which is often interpreted as "fracture" (e.g. Kachanov, 1986). This concept goes back to the work of Kachanov (1958)[17]. Since "the direct measuring of damage is impossible" (Kachanov, 1986), the current state of damage and its evolution must be estimated from changed material properties. From Fig. 2.11, it is quite obvious to correlate damage (e.g. microcracks) in the cross sectional plane with the "loss" of area $\mathrm{d}A_\mathrm{D}$ caused by defects. The remaining area is

$$\mathrm{d}A = \mathrm{d}A_0 - \mathrm{d}A_\mathrm{D} \quad , \tag{2.166}$$

with A_0 being the initial cross sectional area. Accordingly, damage might be introduced by

$$D := \frac{\mathrm{d}A_\mathrm{D}}{\mathrm{d}A_0} = 1 - \frac{\mathrm{d}A}{\mathrm{d}A_0} \tag{2.167}$$

[17] In Kachanov's work dealing with rupture after ductile creep (Kachanov, 1958), damage, in fact, was characterized by a scalar variable denoted as "continuity" ψ_D with $\psi_\mathrm{D} = 1 - D$. In the original application, continuity was assumed to be a positive, monotonically decreasing function, i.e. $\dot{\psi}_\mathrm{D} < 0$. The widely-used definition of D seems to emanate from the work of Rabotnov (1959).

in relation to the reduced area and the current area[18], respectively. Starting from the definition of normal stress $\sigma = F/A_0$ without taking notice of damage, the application of a force F as depicted in Fig. 2.11 yields the *effective stress*

$$\tilde{\sigma} := \frac{F}{A} = \frac{1}{(1-D)}\,\frac{F}{A_0} = \frac{\sigma}{1-D} \quad , \tag{2.168}$$

with $\tilde{\sigma} \to \infty$ as $D \to 1$. Note, that the definitions so far are only valid for the considered plane. By assuming damage to be isotropic, its behavior and influence is equal in all directions (independent of $\boldsymbol{n}$). Accordingly, the effective stress tensor (related to the specific definition of D) is

$$\tilde{\boldsymbol{\sigma}} := \frac{\boldsymbol{\sigma}}{1-D} \quad . \tag{2.169}$$

Since the definition of cross sectional area in multiple dimensions is cumbersome and, more important, because the effect on stress components in other directions is not clear, alternative concepts are commonly used.

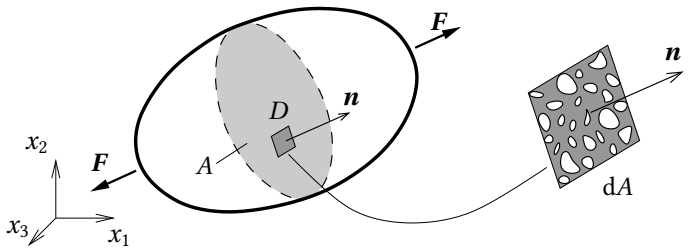


Figure 2.11: Solid body with reduced cross-sectional area (in the plane specified by its normal vector $\boldsymbol{n}$) due to damage, subjected to a force $\boldsymbol{F}$.

Damage equivalence principles

To further identify the notion of damage, the actual stress-strain relation of a solid subjected to damage

$$\boldsymbol{\sigma} = \mathbb{C}(D) : \boldsymbol{\varepsilon} = (1-D)\,\tilde{\boldsymbol{\sigma}} \quad , \tag{2.170}$$

is considered in conjunction with the effective stress definition (2.169). Here, the stiffness tensor $\mathbb{C}(D)$ is assumed to be influenced by damage. A second, hypothetical undamaged configuration is introduced, in which the effective stress is given in relation to the hypothetical strain $\tilde{\boldsymbol{\varepsilon}}$ by

$$\tilde{\boldsymbol{\sigma}} := \mathbb{C}(D{=}0) : \tilde{\boldsymbol{\varepsilon}} = \mathbb{C}_0 : \tilde{\boldsymbol{\varepsilon}} \quad . \tag{2.171}$$

[18] The quantification of damage in ductile materials is based on the same idea, but the void volume fraction $f := \mathrm{d}V_\mathrm{D}/\mathrm{d}V$ is used instead (e.g. Gurson, 1977).

Since the hypothetical undamaged state must correspond with the physical state of no damage ($D = 0$), the stiffness tensor of the undamaged state must be equal to the original (undamaged) material stiffness $\mathbb{C}(D = 0) = \mathbb{C}_0$ with the subscript "0" indicating the undamaged constitutive law, e.g. the linear elastic law (2.85).

By assuming that strains in both configurations (2.170) and (2.171) are equal, i.e. $\tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}$, the *strain equivalence principle* (Rabotnov, 1969; Lemaitre, 1971) identifies the degradation of stiffness

$$\mathbb{C}(D) := (1 - D)\, \mathbb{C}_0 \tag{2.172}$$

with increasing damage. In this context the "effective stress" is identified to be

$$\tilde{\boldsymbol{\sigma}} := \mathbb{C}_0 : \boldsymbol{\varepsilon} := \frac{1}{1 - D}\, \mathbb{C} : \boldsymbol{\varepsilon} \quad . \tag{2.173}$$

The concept implies, that the "strain behavior is modified by damage only through the effective stress" (Lemaitre, 1985). Alternatively, the equivalence of strain energies in both configurations $\psi(\boldsymbol{\varepsilon}) = \tilde{\psi}(\tilde{\boldsymbol{\varepsilon}})$ can be postulated, with (by assuming linear elastic behavior prior to damage)

$$\psi(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\sigma} : \boldsymbol{\varepsilon} \qquad \text{and} \qquad \tilde{\psi}(\tilde{\boldsymbol{\varepsilon}}) = \frac{1}{2}\tilde{\boldsymbol{\sigma}} : \tilde{\boldsymbol{\varepsilon}} \quad . \tag{2.174}$$

With (2.169), the effective elastic strain is $\tilde{\boldsymbol{\varepsilon}} = (1 - D)\, \boldsymbol{\varepsilon}$ and stiffness degradation due to damage according to the *energy equivalence principle* is

$$\mathbb{C}(D) := (1 - D)^2\, \mathbb{C}_0 \quad . \tag{2.175}$$

Note, that both concepts do not make any statement about the evolution of D.

Thermodynamic consideration of damage

To investigate dissipation associated with the appearance of damage, for further details see Holzapfel (2000), the Helmholtz free energy function

$$\psi(\boldsymbol{\varepsilon}, D) = (1 - D)\, \psi_0(\boldsymbol{\varepsilon}) \tag{2.176}$$

is introduced, with the (scalar) internal damage variable D and the effective strain-energy function ψ_0 being a function depending only on strains. Differentiation with respect to time yields

$$\dot{\psi} = (1 - D)\, \frac{\partial \psi_0(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} : \dot{\boldsymbol{\varepsilon}} - \psi_0(\boldsymbol{\varepsilon})\, \dot{D} \quad . \tag{2.177}$$

Application of the Clausius-Planck inequality (2.79) yields

$$\mathscr{D}_{\text{int}} = \left(\boldsymbol{\sigma} - (1 - D)\, \frac{\partial \psi_0(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \right) : \dot{\boldsymbol{\varepsilon}} + \psi_0(\boldsymbol{\varepsilon})\, \dot{D} \geq 0 \quad . \tag{2.178}$$

While the actual stress is given by

$$\boldsymbol{\sigma} = (1 - D)\, \frac{\partial \psi_0(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} = (1 - D)\, \boldsymbol{\sigma}_0 \quad , \tag{2.179}$$

the non-negative energy dissipation attributed due to damage is

$$\mathscr{D}_{\text{int}} = \psi_0(\boldsymbol{\varepsilon})\, \dot{D} \quad . \tag{2.180}$$

The introduction of the thermodynamic force f_{D} which is work conjugate to the damage variable in (2.180) yields, by using (2.176),

$$\mathscr{D}_{\text{int}} = f_{\text{D}}\, \dot{D} \qquad \text{with} \qquad f_{\text{D}} = \psi_0(\boldsymbol{\varepsilon}) = -\frac{\partial \psi(\boldsymbol{\varepsilon}, D)}{\partial D} \quad . \tag{2.181}$$

The thermodynamic force f_{D} quantifies the amount of energy release per unit volume (Holzapfel, 2000), which is here the effective strain energy.

Damage evolution should represent the irreversible structural changes in the material and is formulated, e.g. in terms of strains or strain energy. A simple assumption to illustrate the thermodynamic properties would be

$$D(\alpha) = \frac{\alpha}{1 + \alpha} \tag{2.182}$$

with the phenomenological variable

$$\alpha = \max_{\tau \in [0,\, t]} \psi_0(\tau) \tag{2.183}$$

related to the maximum strain energy (density) obtained in history (the closed time interval $\tau \in [0,\, t]$) which acts as threshold value for further damage. The damage function in the strain space is

$$F(\boldsymbol{\varepsilon}, \alpha) = f_{\text{D}}(\boldsymbol{\varepsilon}) - \alpha \leq 0 \tag{2.184}$$

and is the damage criterion. While for $F < 0$ no damage occurs, for $F = 0$ damage occurs with three possible cases $\dot{f}_{\text{D}} < 0$ (unloading), $\dot{f}_{\text{D}} = 0$ (neutral loading), and $\dot{f}_{\text{D}} > 0$ (loading). Only the latter case is relevant for the evolution of the phenomenological parameter α, i.e.

$$\dot{\alpha} = \begin{cases} \dot{f}_{\text{D}} = \boldsymbol{\sigma} : \boldsymbol{\varepsilon} & \text{if } F = 0 \text{ and } \dot{f}_{\text{D}} > 0 \\ 0 & \text{otherwise.} \end{cases} \quad , \tag{2.185}$$

From (2.179), the stress evolution takes the form

$$\dot{\boldsymbol{\sigma}} = \begin{cases} \left[(1 - D)\, \mathbb{C}_0 : \dot{\boldsymbol{\varepsilon}} - \dfrac{\partial D}{\partial \alpha}\, \boldsymbol{\sigma}_0 \otimes \boldsymbol{\sigma}_0 \right] : \dot{\boldsymbol{\varepsilon}} & \text{if } F = 0 \text{ and } \dot{f}_{\text{D}} > 0 \\ (1 - D)\, \mathbb{C}_0 : \dot{\boldsymbol{\varepsilon}} & \text{otherwise,} \end{cases} \quad , \tag{2.186}$$

with the effective elasticity tensor $\mathbb{C}_0 := \partial^2 \psi_0 / \partial \boldsymbol{\varepsilon}^2$ as already introduced. In the first case of (2.186), the second term represents stress softening due to damage evolution, while in the second case this term is missing due to a constant amount of damage ($\dot{D} = 0$) during unloading or elastic loading below the damage threshold.

Tension-compression asymmetry of damage

Many materials show a different damaging behavior when subjected to tensile or compressive loads. Brittle materials like rocks (Shao and Rudnicki, 2000), concrete (Berthaud et al., 1990; Mazars et al., 1990), ceramics (Chaboche et al., 1995) or glass typically sustain much larger compressive forces than tensile forces. In addition, as brittle damage is governed by microcracks, the elastic behavior is different for closed cracks under compression, i.e. contact between closed crack flanks allows partial or even full stiffness recovery so that damage-induced area reduction (2.167) has no effect. While the first effect can be incorporated by relating damage evolution only to, e.g. positive principle strains or a combination thereof (Mazars, 1982), the second effect is accounted for by modifying the stress-strain relation (2.179). In the literature, this effect is often referred to as "quasi-unilateral crack condition" (Ladeveze and Lemaitre, 1984; Lemaitre, 1996), "crack closure effect" (Murakami, 1988), "activation" (Ju, 1989) and "deactivation" (Skrzypek, 2015) of cracks, "active" or "passive" crack states (Murakami, 2012) or — much more general — *tension-compression asymmetry*. Typically, the stiffness is not fully degraded by the factor $1 - D$ but only some portion related to (micro-) crack opening. Since brittle fracture is known to be dominated by normal stresses (Sect. 2.2.1) Ladeveze and Lemaitre (1984) and Mazars (1986), for example, independently from each other introduced a continuum damage model with a split[19] based on the spectral decomposition[20] of the Cauchy stress tensor

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^+ + \boldsymbol{\sigma}^- \tag{2.187}$$

with

$$\boldsymbol{\sigma}^+ := \sum_{\alpha=1}^{3} \langle \sigma_\alpha \rangle \; \overset{\sigma}{\boldsymbol{n}}_\alpha \otimes \overset{\sigma}{\boldsymbol{n}}_\alpha \quad \text{and} \quad \boldsymbol{\sigma}^- := -\sum_{\alpha=1}^{3} \langle -\sigma_\alpha \rangle \; \overset{\sigma}{\boldsymbol{n}}_\alpha \otimes \overset{\sigma}{\boldsymbol{n}}_\alpha \quad . \tag{2.188}$$

Not much later, Ortiz (1985) proposed a model with a similar split in the strain space, i.e.

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^+ + \boldsymbol{\varepsilon}^- \tag{2.189}$$

with

$$\boldsymbol{\varepsilon}^+ := \sum_{\alpha=1}^{3} \langle \varepsilon_\alpha \rangle \; \overset{\varepsilon}{\boldsymbol{n}}_\alpha \otimes \overset{\varepsilon}{\boldsymbol{n}}_\alpha \quad \text{and} \quad \boldsymbol{\varepsilon}^- := -\sum_{\alpha=1}^{3} \langle -\varepsilon_\alpha \rangle \; \overset{\varepsilon}{\boldsymbol{n}}_\alpha \otimes \overset{\varepsilon}{\boldsymbol{n}}_\alpha \quad , \tag{2.190}$$

see also the notes by Ju (1989). Consequently, only the portion of the strain energy density related to "crack opening" is affected, i.e. degraded. In the latter case, the stress is

$$\boldsymbol{\sigma} = (1 - D) \, \frac{\partial \psi_0^+(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} + \frac{\partial \psi_0^-(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \tag{2.191}$$

[19] For this purpose, the bracket operator $\langle x \rangle := \big(|x| + x\big)/2$ to provide the (unique) additive decomposition of a scalar quantity x into its positive and negative portions $x = \langle x \rangle + \big(- \langle -x \rangle \big)$ is used, for more details see App. A.9.

[20] The application of the bracket operator requires scalar quantities, e.g. principal values or invariants of the strain or stress tensor. For further details on the spectral decomposition, see App. A.5.

based on the strain energy density portions

$$\psi_0^+ := \frac{\lambda}{2} \langle \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 + \mu \left(\langle \varepsilon_1 \rangle^2 + \langle \varepsilon_2 \rangle^2 + \langle \varepsilon_3 \rangle^2 \right) \qquad \text{and} \tag{2.192}$$

$$\psi_0^- := \frac{\lambda}{2} \langle - \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 + \mu \left(\langle -\varepsilon_1 \rangle^2 + \langle -\varepsilon_2 \rangle^2 + \langle -\varepsilon_3 \rangle^2 \right) \quad , \tag{2.193}$$

while the first model yields similar, but more lengthy expressions. Alternatively, by following the continuum damage models proposed and discussed by, e.g., Comi and Perego (2000) and Lemaitre and Desmorat (2005) crack closure can also be modeled based on the volumetric-deviatoric decomposition. In this case, the stress (2.191) is derived from the strain energy density portions (here somewhat simplified compared to the original models)

$$\psi_0^+ := \frac{K}{2} \langle \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 + \mu \left(\boldsymbol{\varepsilon}_{\mathrm{dev}} : \boldsymbol{\varepsilon}_{\mathrm{dev}} \right) \qquad \text{and} \tag{2.194}$$

$$\psi_0^- := \frac{K}{2} \langle - \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 \quad , \tag{2.195}$$

with the deviatoric strain $\boldsymbol{\varepsilon}_{\mathrm{dev}}$ according to (2.19).

2.3 Variational formulation of brittle fracture

The analysis of initial boundary value problems with numerical techniques like the finite element method has become a standard approach in engineering. With the appearance of cracks, however, internal surfaces are created. They introduce additional boundaries which can lead to discontinuous and singular mechanical fields. Thus, the problem does not correspond with the continuity requirements of classical finite element approaches. It has long been a challenge to simulate nucleation and growth of cracks numerically, and it still is. Instead of modeling cracks in a discrete manner (e.g. with the extended finite element method) or in a continuous manner (with non-local or gradient enhanced continuum damage formulations) a variational treatment is introduced in the following.

The application of Griffith's theory as discussed in Sect. 2.2.4 to simulate crack growth requires both an initial crack (yields incremental crack growth) and some a-priori knowledge about the crack path, e.g. by additionally using one of the criteria introduced in Sect. 2.2.5. Ambrosio and Braides (1995) and, later on, Francfort and Marigo (1998) proposed a method based on the minimization of the sum of stored elastic energy and surface energy of displacement discontinuity sets. While the first work was assigned to stationary cracks, the latter approach enables crack evolution under quasi-static conditions. The main feature of the so called *variational approach to fracture* is that the crack path/topology does neither have to be known a priori nor crack extension is predicted by classical fracture criteria. Crack growth is a result of energetic considerations only. In the limit case of straight crack growth of a pre-crack, the approach corresponds to Griffith's theory. In addition, it can

also handle cases like $\mathcal{G} > \mathcal{G}_c$ or crack initiation from different types of singular points or non-singular settings. This becomes beneficial for far more complex three-dimensional situations which are naturally included in the formulation without additional effort.

Here, Ω denotes a smooth bounded connected open domain of $\mathbb{R}^n$ with the dimension $n \in [1, 3]$ to represent the crack-free configuration with a smooth boundary $\partial\Omega$. While in classical fracture mechanics only a single pre-crack or a small number of pre-cracks is considered, a set with a large number of defects is possible, represented by closed subdomains of $\bar{\Omega} = \Omega \cup \partial\Omega$. The shape of the defects is not restricted, but their dimension is not greater than $n - 1$. In the following, only cracks with $n - 1$ are regarded, which are comprised in the crack set $\mathscr{S} \subset \bar{\Omega}$. Due to physical reasons, it is reasonable to limit the geometry and size of a crack at time t by its predecessors (at all $s < t$). Thus, the current crack set $\mathscr{S}(t) \in \bar{\Omega}$ contains previous crack sets

$$\mathscr{S}(s) \subseteq \mathscr{S}(t) \quad , \tag{2.196}$$

which is an irreversibility constraint. According to Griffith's theory, the surface energy of a crack is proportional to its surface area (2.144). With the proportionality factor $\mathcal{G}_c$, the surface energy Γ associated to the crack set $\mathscr{S}$ is

$$\Gamma(\mathscr{S}) = \mathcal{G}_c \, \mathscr{H}^{n-1}(\mathscr{S}) \quad , \tag{2.197}$$

with $\mathscr{H}^{n-1}$ being the $n-1$-dimensional Hausdorff measure. The Hausdorff measure[21] is a surface measure for smooth hypersurfaces (e.g. Evans and Gariepy, 1992). In the remaining part of the domain the body is assumed to behave elastically. Due to absence of smoothness of $\Omega \setminus \mathscr{S}$, a smooth domain Ω_u is constructed such that $\Omega_u \cap \partial\Omega = \partial\Omega_u$. Only displacement boundary conditions $\bar{\boldsymbol{u}}(\boldsymbol{x}, t)$ on $\partial\Omega_u$ are admissible, while the remaining part of the boundary $\partial\Omega_f = \partial\Omega \setminus \partial\Omega_u$ and the crack lips are assumed to be traction free. Since cracks can also occur along the displacement boundary, the displacement field in the body is subjected to

$$\boldsymbol{u}(\boldsymbol{x}, t) = \bar{\boldsymbol{u}}(\boldsymbol{x}, t) \quad \text{on} \quad \partial\Omega_u \setminus \mathscr{S}(t) \quad . \tag{2.198}$$

The set of kinematically admissible displacement fields depends on the imposed loading (2.198) and the crack set $\mathscr{S}$ and is introduced, according to Francfort and Marigo (1998), as

$$\mathscr{C}(\mathscr{S}, \bar{\boldsymbol{u}}) = \{\boldsymbol{u} \in H^1(\Omega_u \setminus \mathscr{S}; \mathbb{R}^n) \mid \boldsymbol{u} = \bar{\boldsymbol{u}}(\boldsymbol{x}, t) \text{ in } \Omega_u \setminus \bar{\Omega}\} \quad , \tag{2.199}$$

[21] The German mathematician Felix Hausdorff (1868-1942) introduced fundamental concepts to assign, e.g., volume and area measures to non-integer dimensions (Hausdorff, 1918). In particular, the Hausdorff measure assigns a number in $[0, \infty)$ to each set in, e.g., $\mathbb{R}^n$ (Rogers, 1998). The length of a simple curve in $\mathbb{R}^n$, for instance, is given by the one-dimensional Hausdorff measure.

with the Sobolev space H^1. With the assumption of small deformations (2.16), which, however, is no prerequisite for the approach (Dal Maso et al., 2005), the elastic energy in the bulk for any kinematically admissible field $\boldsymbol{u} \in \mathscr{C}$ is

$$E_{\mathrm{el}}(\mathscr{S}, \boldsymbol{u}) = \int_{\Omega \setminus \mathscr{S}} \frac{1}{2} \boldsymbol{\varepsilon}(\boldsymbol{u}) : \mathbb{C} : \boldsymbol{\varepsilon}(\boldsymbol{u}) \, \mathrm{d}V \quad . \tag{2.200}$$

Finally, the key idea of the variational approach to fracture is to find the pair of a kinematically admissible displacement field $\boldsymbol{u}$ and an irreversible crack set $\mathscr{S}$ for which the total energy

$$E_{\mathrm{tot}}(\mathscr{S}, \bar{\boldsymbol{u}}) = E_{\mathrm{el}}(\mathscr{S}, \bar{\boldsymbol{u}}) + \Gamma(\mathscr{S}) \tag{2.201}$$

is a global minimum.

A significant restriction of the original approach is the unavailability of surface tractions and body forces. In the example outlined by Francfort and Marigo (1998; 2005), an elastic body of some size Ω is subjected to a traction force $\boldsymbol{t}$ at $\partial\Omega_{\mathrm{t}}$. If, as one possible (and admissible) solution, the loaded part of the boundary is cut-off by a crack from the rest of the domain, i.e. $\mathscr{S} = \partial\Omega_{\mathrm{t}}$, the surface energy is $\Gamma = \mathcal{G}_{\mathrm{c}} \, \mathscr{H}^{\mathrm{n}-1}(\partial\Omega_{\mathrm{t}})$, while the elastic energy in the bulk is $E_{\mathrm{el}} = 0$. The movement of the loaded boundary is parametrized by $\boldsymbol{u} \equiv \lambda \boldsymbol{t}$, so that the corresponding work of external forces $-\int_{\partial\Omega_{\mathrm{t}}} \boldsymbol{t} \cdot \boldsymbol{u} \, \mathrm{d}\mathscr{H}^{\mathrm{n}-1}$ can be added to the total energy

$$E_{\mathrm{tot}} = -c \, \lambda + \Gamma(\mathscr{S}) \quad , \tag{2.202}$$

with a constant positive c. It is obvious, that $E_{\mathrm{tot}} \to -\infty$ as $\lambda \to \infty$. Thus, the energy is not bounded from below and no global minimizer exists.

Global or local minimization?

Global minimization as driving principle is advertised as one of the key merits of the approach Francfort and Marigo (1998). The example above, however, might motivate the usage of local minimization instead. Definitions of globally and locally stable equilibrium states are given by Charlotte et al. (2000) as follows. By considering only displacements, an admissible field $\boldsymbol{u}$ corresponds to the globally stable equilibrium state if the total energy $E_{\mathrm{tot}}(\boldsymbol{u}, \mathscr{S})$ is less or equal than the total energy $E_{\mathrm{tot}}(\boldsymbol{v}, \mathscr{S})$ of the domain in any admissible state, i.e.

$$E_{\mathrm{tot}}(\boldsymbol{u}, \mathscr{S}) \leq E_{\mathrm{tot}}(\boldsymbol{v}, \mathscr{S}) \quad \forall \boldsymbol{v} \text{ admissible} \quad . \tag{2.203}$$

Here, admissible means, that $\boldsymbol{u}$ and $\boldsymbol{v}$ are piecewise smooth and satisfy the kinematic boundary conditions. By contrast, an admissible field $\boldsymbol{u}$ corresponds to the locally stable equilibrium state if the total energy $E_{\mathrm{tot}}(\boldsymbol{u}, \mathscr{S})$ is less than the total energy $E_{\mathrm{tot}}(\boldsymbol{v}, \mathscr{S})$ of the domain in any admissible state in a neighborhood of $\boldsymbol{u}$, i.e.

$$\begin{aligned} E_{\mathrm{tot}}(\boldsymbol{u}, \mathscr{S}) < E_{\mathrm{tot}}(\boldsymbol{v}, \mathscr{S}) \quad & \forall \boldsymbol{v} \text{ admissible}, \ \exists \delta(\boldsymbol{u}) > 0, \\ & ||\boldsymbol{v} - \boldsymbol{u}|| \leq \delta(\boldsymbol{u}), \ \boldsymbol{u} \neq \boldsymbol{v} \quad . \end{aligned} \tag{2.204}$$

The solution obtained by local minimization might be influenced by the "choice of the distance between admissible test fields" (Francfort and Marigo, 2005). To elaborate physical implications of global and local minimization, two boundary value problems are analyzed briefly.

A one-dimensional bar subjected to an increasing displacement $\bar{u} > 0$, see Fig. 2.12a, was already analyzed by Charlotte et al. (2000). The globally stable equilibrium state predicts elastic response up to $u_{\text{crit}} = \sqrt{2\mathcal{G}_{\text{c}}\, L/E}$. For $\bar{u} = u_{\text{crit}}$ fracture occurs at an arbitrary point $x_{\text{crack}} \in L$. The bar is then unloaded, i.e. $u(x) = 0$ for $x < x_{\text{crack}}$ and $u(x) = \bar{u}$ for $x > x_{\text{crack}}$. Thus, global minimization predicts fracture, as soon as the elastic energy in the bar suffices to create crack surfaces according to Griffith's theory. With increasing bar length ($L \to \infty$), the critical stress at which failure occurs, i.e.

$$\sigma_{\text{crit}} = \sqrt{\frac{2\mathcal{G}_{\text{c}}\, E}{L}} \quad , \tag{2.205}$$

approaches zero ($\sigma_{\text{crit}} \to 0$), which contradicts the concept of ultimate stress/strength emanating from experimental observations, see Sect. 2.2.1. By contrast, the locally stable equilibrium state yields the load, at which fracture occurs, to be $u_{\text{crit}} = \sigma_{\text{c}}\, L/E$.

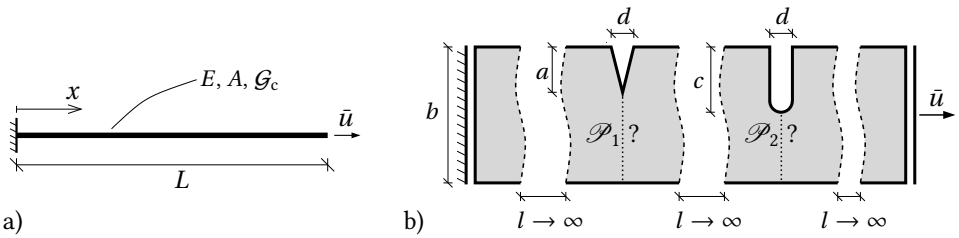


Figure 2.12: Global versus local energy minimization examples: a) crack formation in elastic bar, b) crack progress in elastic specimen with V-notch and rounded-off notch.

A second example is used to analyze the impact of different stress concentrations in an infinitely long elastic strip of finite width b, as depicted in Fig. 2.12b. Far away from the boundary and from each other, a V-notch of length a and a rounded-off notch of length $c > a$ are introduced. While the V-notch introduces a stress singularity, the stress concentration at the rounded-off notch is expected to be finite. For a large enough displacement load ($\bar{u} \to \infty$), a through-cracked state (no bulk energy), either with the crack set $\mathscr{S} = \mathscr{P}_1$ or $\mathscr{S} = \mathscr{P}_2$ is obtained. Since the surface energy related to $\Gamma(\mathscr{P}_2) \sim b - c$ is obviously smaller than the surface energy related to $\Gamma(\mathscr{P}_1) \sim b - a$, the globally stable equilibrium state is obtained for a crack initialized at the rounded-off notch. By contrast, the locally stable equilibrium state takes into account each loading increment separately to search a stable solution in a neighborhood to its prior configuration. Thus, minimization of the total energy predicts crack growth starting at the "more critical" tip of the V-notch, which finally results in the crack set $\mathscr{S} = \mathscr{P}_1$, although the energy consumption is larger compared to a crack emanating from the rounded-off notch.

As demonstrated by an one-dimensional analysis (Charlotte et al., 2000), the elastic solution is always a local minimizer of the total energy in absence of a singularity, e.g. caused by an initial crack. This can be circumvented by replacing Griffith's surface energy in (2.197) by a cohesive surface energy, in particular Barenblatt's model as introduced in Sect. 2.2.7, to allow crack nucleation in absence of singularities also for local minimization. Such a model is analyzed, e.g., by Francfort (2006).

Results obtained by the application of the method on different boundary value problems have been published in the literature. Francfort and Marigo (1998) compared the results of two examples, uniaxial tension of a cylinder and the tearing of a reinforcement cylinder, with the exact solution. The problem of antiplane shear was analyzed by Dal Maso and Toader (2002) and Francfort and Larsen (2003). Chambolle (2003) analyzed a planar elastic problem. It should be noted, that above investigations are made with some restrictive assumptions, e.g., the assumption of monotonically increasing loads. In addition, only fracture problems are analyzed for which the "no-contact condition" $\boldsymbol{\sigma}^{\mathrm{T}} \cdot \boldsymbol{n} = \mathbf{0}$ (with the crack surface normal vector $\boldsymbol{n}$) holds a priori.

Types of crack growth

As in Griffith's theory (Sect. 2.2.4), the type of displacement/stress singularity at, e.g., a V-shaped notch, plays an essential role also in the variational approach of fracture. Thus, Francfort and Marigo (1998) provided an analysis of the crack growth behavior in planar problems, which is summarized as follows. For an initial crack set $\mathscr{S}_0$, which might be empty, and the external load $\bar{\boldsymbol{u}}_0$ the elastic equilibrium in terms of the displacement is

$$\boldsymbol{u}_0 \equiv \boldsymbol{u}(\mathscr{S}_0, \bar{\boldsymbol{u}}_0) \qquad \text{on } \Omega \setminus \mathscr{S}_0 \quad . \tag{2.206}$$

By assuming that there is a single singular point $\boldsymbol{x}_\mathrm{s}$, at least in the neighborhood, it is useful to represent the displacement field (as sum of its eigenfunctions $\hat{\boldsymbol{u}}_{(\alpha)}$)

$$\boldsymbol{u}_0(r, \theta) = \boldsymbol{u}_0(\boldsymbol{x}_\mathrm{s}) + r^\alpha \, \hat{\boldsymbol{u}}_{(\alpha)}(\theta) + \ldots \tag{2.207}$$

in polar coordinates (r, θ) with the origin at $\boldsymbol{x}_\mathrm{s}$. The first exponent α denotes the type of singularity in the displacement field (by being dominant for $r \to 0$) and is restricted to $\alpha \in (0,\, 1)$. While the upper bound ensures the singular character of the displacement field at $\boldsymbol{x}_\mathrm{s}$, the lower bound ensures the finiteness of bulk energy. It is assumed, that crack growth starts from the singularity. Leguillon (1989) provided the asymptotic expansion of the bulk energy

$$E_\mathrm{el}(\mathscr{S}_0 \bigcup \mathscr{S}_\mathrm{l}, \bar{\boldsymbol{u}}) = E_\mathrm{el}(\mathscr{S}_0, \bar{\boldsymbol{u}}_0) - \left(K \, l^{2\alpha} + o(l^{2\alpha})\right) \tag{2.208}$$

for a small added crack of length l, which constitutes the crack set $\mathscr{S}_\mathrm{l}$. K is a positive factor depending on the defect shape and singularity strength, but not on its length l. With the assumption that crack growth is caused by a monotonically increasing displacement load $\bar{\boldsymbol{u}}$, Francfort and Marigo (1998) provides the following cases (equipped with proofs):

- Crack growth at a strong singularity $\alpha \in (0, 1/2)$ is progressive and starts as soon as load is applied (zero initiation time). As noted by Kuhn (2013), this type of singularity is typical for isolated points and becomes weaker, as soon as crack growth takes place.
- From the tip of a perfect pre-crack ($\alpha = 1/2$), with $\sqrt{r}$-singularity in the displacement field, crack growth is also progressive and starts at a finite load $\bar{\boldsymbol{u}}_{\text{crit}\,1}$ (non-zero initiation time). In this particular case, the variational approach of fracture coincides with Griffith's theory in Sect. 2.2.4.
- Crack growth from a weak singularity $\alpha \in (1/2, 1)$ takes place with a crack increment of finite length denoted as "brutal crack growth" (Francfort and Marigo, 1998) at a finite load $\bar{\boldsymbol{u}}_{\text{crit}\,2}$ (non-zero initiation time). According to classical Griffith's theory, the energy release rate at a weak singularity is zero. Thus, no crack growth is possible for energetic reasons, which is contrary to the variational approach of fracture.

By also considering the case of a non-singular/regular displacement field ($\alpha = 1$), either no crack growth occurs or the crack grows with a crack increment of finite length at a finite load $\bar{\boldsymbol{u}}_{\text{crit}\,3}$ (non-zero initiation time).

According to this analysis, progressive crack growth takes place only at $\sqrt{r}$-singularities or stronger singularities. The energy release of an infinite crack increment in non-singular problems or at weak singularities is not sufficient to create corresponding crack surfaces. Thus, crack initiation always starts with a crack length of finite length. This is in accordance with results obtained by finite fracture mechanics (Sect. 2.2.6), see also Leguillon (2002). In addition, crack initiation always starts in finite time.

Incorporation of non-interpenetration condition

To avoid possible interpenetration of crack lips, Francfort and Marigo (1998) illustrated ways to incorporate "unilateral contact" without friction. For convenience, the additional boundary conditions on the crack set $\mathscr{S}$ are given here in the form of Karush Kuhn-Tucker (KKT) conditions

$$\delta_{\text{n}} := [\![\boldsymbol{u}]\!] \cdot \boldsymbol{n} \geq 0 \quad , \qquad \sigma_{\text{nn}} := \boldsymbol{\sigma}^{\text{T}} : (\boldsymbol{n} \otimes \boldsymbol{n}) \leq 0 \quad , \qquad \delta_{\text{n}}\, \sigma_{\text{nn}} = 0 \quad , \tag{2.209}$$

with $[\![\boldsymbol{u}(\boldsymbol{x}_{\text{c}})]\!] := \boldsymbol{u}^{+}(\boldsymbol{x}_{\text{c}}) - \boldsymbol{u}^{-}(\boldsymbol{x}_{\text{c}})$ being the displacement jump across the crack at a point $\boldsymbol{x}_{\text{c}} \in \mathscr{S}$ and $\boldsymbol{n}(\boldsymbol{x}_{\text{c}}^{+}) = -\boldsymbol{n}(\boldsymbol{x}_{\text{c}}^{-})$ being the outward normal vector. In contact mechanics, (2.209) is the classical complementary condition. The first requirement limits crack opening δ_{n} to non-negative values, the second condition restricts the normal stress σ_{nn} to non-positive values (compressive stress in case of crack closure), and the third condition enforces either crack opening ($\sigma_{\text{nn}} = 0$) or crack closure ($\delta_{\text{n}} = 0$).

In Francfort and Marigo (1998), two approaches to incorporate (2.209) in the variational approach of fracture are proposed. Either the surface energy associated to the crack set $\mathscr{S}$ (2.200) is extended according to

$$\Gamma(\mathscr{S}, \boldsymbol{u}) = \int_{\mathscr{S}} \mathcal{G}_{\mathrm{c}}(\boldsymbol{x}_{\mathrm{c}}, \boldsymbol{n}, \delta_{\mathrm{n}}) \, \mathrm{d}\mathscr{H}^{\mathrm{n}-1}(\boldsymbol{x}_{\mathrm{c}}) \tag{2.210}$$

with $\mathcal{G}_{\mathrm{c}}(\boldsymbol{x}_{\mathrm{c}}, \boldsymbol{n}, \delta_{\mathrm{n}} < 0) = \infty$ or the set of kinematically admissible displacements (2.199) is modified to

$$\mathscr{C}(\mathscr{S}, \bar{\boldsymbol{u}}) = \{\boldsymbol{u} \in H^1(\Omega_{\mathrm{u}} \setminus \mathscr{S}; \mathbb{R}^n) \mid \boldsymbol{u} = \bar{\boldsymbol{u}}(\boldsymbol{x}, t) \text{ in } \Omega_{\mathrm{u}} \setminus \bar{\Omega} \mid \delta_{\mathrm{n}} \geq 0 \text{ on } \mathscr{S}\} \quad . \tag{2.211}$$

In the first case, the surface energy term becomes energetically unfavorable for interpenetration of crack lips while the latter case excludes displacement fields with crack lip interpenetration from the solution. The implementation of both approaches would require special attention since all involved parameters have to be adequately defined (Evans and Gariepy, 1992) and the finding of solutions is not trivial since, e.g., bounded variation type spaces come into play (Ambrosio, 1989).

3 Phase field approach to brittle fracture

In the recent one and a half decades, the phase field approach to fracture has enjoyed great popularity in academic research. This can be traced back to the method's clear concept, its simple numerical implementation, robust solution algorithms, and increasing computing power. Thus, phase field fracture models enjoy a good reputation and a wide variety of extensions has been developed, e.g., to compute fracture in the context of large deformations, plasticity, fatigue, anisotropy and inhomogeneous problems. This chapter introduces a phase field description of fracture in brittle materials. A key issue is the coupling of elastic fields with the phase field. Constitutive ingredients of common phase field approaches to fracture are carefully examined to identify and discuss severe consequences of the models presently available. A novel requirement for the continuous representation of macroscopic cracks is formulated and incorporated in the phase field approach. Furthermore, constitutive aspects of the crack's regularization as well as issues of the discrete numerical treatment are discussed. All issues are demonstrated by means of representative analytical analyses or numerical examples.

3.1 Regularization of free discontinuity problems

The variational formulation of fracture (Sect. 2.3) provides very promising features to investigate non-standard problems of fracture mechanics, such as crack formation in absence of pre-cracks. Especially compared to methods of discrete crack representation, which often rely on existing cracks, and methods of smeared crack representation, which usually are not primarily intended to represent the state of complete material separation, a variational approach to find the minimum of the sum of volume and surface energies appears superior. Such problems form the wide class of *free discontinuity problems* (e.g. Braides, 1998; De Giorgi, 1991) and are present in many mathematical and physical problems such as image segmentation (Mumford and Shah, 1985) or more general, signal processing, harmonic maps (Brezis et al., 1986), and statics of liquid crystals Ericksen (1976). Due to the unknown set of crack surfaces, which is part of the solution and might cause discontinuities in the displacement field, the variational formulation of brittle fracture cannot be treated numerically by a direct implementation into, e.g., the

finite element framework. Thus, at least two fundamentally different numerical methods, both based on the key procedure of energy minimization, have been developed in recent years. In the first approach, the discretization is adapted to take (potential) crack surfaces into account (Bourdin et al., 2000; Fraternali, 2007). Instead of discretizing a continuous functional the discretization itself is part of the minimization procedure. Besides simple discretization techniques, e.g. triangulation, the implementation of Bourdin et al. (2000) relies on mesh optimization realized by introducing an additional variable and using anisotropic mesh generation. Crack evolution does not suffer from local crack tip singularities due to the overall energy minimization procedure. In the second approach, the (potential) crack surfaces and corresponding displacement discontinuities are regularized, i.e. "smeared", in an appropriate manner to provide a continuous functional (Bourdin et al., 2000), which can then be discretized. Both approaches are strongly inspired by the solution of similar problems in the well-developed field of image segmentation. While the first approach (Chambolle and Dal Maso, 1999; Bourdin and Chambolle, 2000) is based on the Mumford-Shah functional, the second approach (Bourdin, 1999) also uses the elliptic approximation introduced by Ambrosio and Tortorelli (1990; 1992), as summarized below. Due to its implementation simplicity, only the second approach is used in the following.

Image segmentation: extraction of edges

Image segmentation is a typical problem of image analysis. Usually, an image on a domain Ω consists of pixels with different gray levels, described by $g \in [0, 1]$. The image might consist of several objects/regions separated by boundaries/edges, which occur in terms of gray level discontinuities. However, also shadows, meaningless details, noise and other small irregularities can be the source of discontinuities. To partition images the object boundaries (combined in the hypersurface or, in two dimensions, the edge set K) have to be extracted by ignoring these irregularities. This is achieved by using piecewise smooth functions u, which represent the original image gray level g on $\Omega \setminus K$, are thus separated by the edges, and give much better results than global smoothing (see e.g. Rosenfeld, 1982). For this purpose, Mumford and Shah (1985; 1989) developed a minimization procedure for image segmentation based on the functional

$$F(u, K) = \int_{\Omega \setminus K} \beta\,(u - g)^2 + |\nabla u|^2 \,\mathrm{d}x + \alpha\,\mathscr{H}^{n-1}(K) \quad , \tag{3.1}$$

with $\Omega \subset \mathbb{R}^n$ being an open bounded set, the image domain. K is a sufficiently closed subset of Ω representing the union of all edges, and $\mathscr{H}^{n-1}$ is the $n-1$-dimensional Hausdorff measure as introduced in Sect. 2.3, which, in the two-dimensional case, gives the edge length of all edges in K. The fixed parameters $\alpha > 0$ and $\beta > 0$ weight the contributions of the gray level approximation (first term) and the edge set (third term) relative to the gradient term. The Mumford-Shah functional (3.1) penalizes large edge sets K as well as large gradients of u outside the edge set. Thus, u varies smoothly and, in addition, must be a good approximation of g on $\Omega \setminus K$.

De Giorgi et al. (1989) proposed the functional

$$F(u, S_u) = \int_\Omega \beta\,(u-g)^2 + |\nabla u|^2 \,\mathrm{d}x + \alpha\,\mathscr{H}^{n-1}(S_u) \tag{3.2}$$

and proved it to be equivalent to (3.1), see also Dal Maso et al. (1992). The functional (3.2), however, is based on the space of special bounded variation functions, i.e. $u \in \mathrm{SBV}(\Omega)$ as introduced by De Giorgi and Ambrosio (1988) and Ambrosio (1989). Thus, u is allowed to be discontinuous along a set of codimension one, i.e. the discontinuity set S_u.

To approximate (3.2) by an elliptic functional defined on the Sobolev space, the discontinuity must be regularized appropriately. For this purpose, an additional variable $v \in [0,\,1]$ is introduced to replace the jump set S_u by a continuous approximation. The regularization is controlled by a small parameter ϵ. Different functionals have been proposed, e.g., by Ambrosio and Tortorelli (1990; 1992), Bellettini and Coscia (1994), and Richardson and Mitter (1997). Their equivalence with the original problem is often discussed in terms of the so-called *Γ-convergence* as elaborated, e.g., by (Dal Maso, 1993). Usually, the proof contains steps to show the convergence relative to the lower and upper bounds as $\epsilon \to 0$, i.e.

$$F(u,v) \le \liminf_{\epsilon\to 0} F_\epsilon(u,v) \qquad \text{and} \qquad \limsup_{\epsilon\to 0} F_\epsilon(u,v) \le F(u,v) \quad , \tag{3.3}$$

respectively. Accordingly, minimum points of the regularized functional $F_\epsilon(u,v)$ converge to the minimum points of the limit functional $F(u,v)$, see also Braides (1998; 2002). In order to prove Γ-convergence of the approximation Ambrosio and Tortorelli (1990), for example, formally introduced the elliptic functional

$$F_{\tilde\epsilon}(u,\tilde v) = \int_\Omega \beta\,(u-g)^2 + (|\nabla u|^2 + |\nabla \tilde v|^2)\,(1-\tilde v^2)^{2/\tilde\epsilon} + \frac{\alpha^2}{4\tilde\epsilon^2}\,\tilde v^2 \,\mathrm{d}x \quad , \tag{3.4}$$

which has been the first approximating functional of this type, equipped with the "control variable" $\tilde v := 1 - v$ acting on the gradient term of u and $\tilde\epsilon = \alpha\,\epsilon$. Because of the term $(1-\tilde v^2)^{2/\epsilon}$ the functional cannot be discretized as $\epsilon \to 0$ (Bourdin, 1999). Alternatively, Ambrosio and Tortorelli (1990) also proposed the functional

$$F_{\tilde\epsilon}(u,v) = \int_\Omega \beta\,(u-g)^2 + v^2\,|\nabla u|^2 + \frac{\alpha^2}{4\tilde\epsilon}\,(1-v)^2 + \tilde\epsilon\,|\nabla v|^2 \,\mathrm{d}x \quad , \tag{3.5}$$

which is "formally very similar" to the functionals used by Ericksen (1976) in the context of the equilibrium theory of liquid crystals. After some minor modifications, Ambrosio and Tortorelli (1992) finally introduced the functional

$$F_\epsilon(u,v) = \int_\Omega \beta\,(u-g)^2 + (v^2 + k_\epsilon)\,|\nabla u|^2 + \alpha\left(\frac{(1-v)^2}{4\epsilon} + \epsilon\,|\nabla v|^2\right)\mathrm{d}x \quad , \tag{3.6}$$

which contains the small positive parameter $k_\epsilon \ll 1$ to circumvent numerical difficulties in case of $v \to 0$.

A key feature of (3.6) is its simple numerical treatment. Variation, discretization and its (straightforward) finite element implementation are provided in App. C.1. In contrast to, e.g., piecewise linear functionals as the one proposed by Bellettini and Coscia (1994), which require constrained minimization techniques, the quadratic terms in (3.6) impose natural constraints on v at $v = 0$ and $v = 1$. However, strict convexity cannot be guaranteed for the mixed term $v^2 |\nabla u|^2$ in (3.6). Thus, Bourdin (1999) proposed the usage of an *alternate minimization scheme*. Minimization of $F_\epsilon(u, v)$ is carried out separately for both fields, i.e. $\min_{\{u\}} F_\epsilon(u, v \text{ is fixed})$ and then $\min_{\{v\}} F_\epsilon(u \text{ is fixed}, v)$, with both subproblems being strictly convex.

To illustrate the capabilities of the method, the gray level image given in Fig. 3.1a, with 384×512 pixels of gray levels ranging from $g = 0$ (black) to $g = 1$ (white), is segmented. Here, the mesh is chosen to coincide with the pixels to simplify the input of gray level values, but the approach in general allows non-regular/unstructured meshes. The numerical parameters are $\alpha = 0.02/l_{\text{img}}$, $\beta = 0.5/l_{\text{img}}^2$, $\epsilon = 0.1\, l_{\text{img}}$, $k_\epsilon = 0.01$ in relation to the image width of $384\, l_{\text{img}}$. Figure 3.1b shows the smooth approximation u, which is close to the original image some distance away from edges, and Fig. 3.1c depicts the extracted edge set in terms of the scalar field v (with $v = 0$ indicated by black).

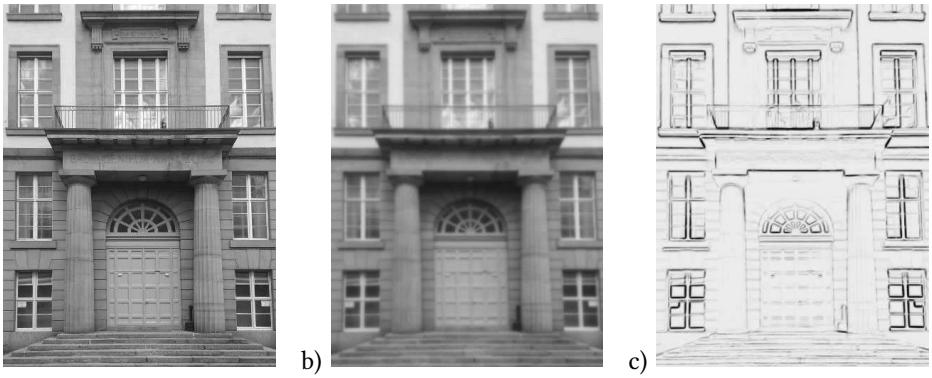


Figure 3.1: Image segmentation by using a discrete formulation of the Mumford-Shah functional with $\alpha = 0.02/l_{\text{img}}$, $\beta = 0.5/l_{\text{img}}^2$, $\epsilon = 0.1\, l_{\text{img}}$, $k_\epsilon = 0.01$; a) original gray level image with 384×512 pixels and a width of $384\, l_{\text{img}}$, b) approximated (smoothed) gray level u, c) final edge set represented by v.

Fracture mechanics: physical incorporation of crack surfaces

Inspired by the promising results of the image segmentation problem based on a standard finite element implementation with the so-called alternate minimization algorithm (Bourdin, 1999), Bourdin et al. (2000) used a straightforward adaption of the algorithms to solve

problems of the variational approach to fracture as introduced in Sect. 2.3. For this purpose, a sharp crack $\mathscr{S}$ (see Fig. 3.2a) is approximated in a smooth manner by an additional scalar-valued field variable, here denoted by $s(\boldsymbol{x}) \in [0,\, 1]$. It describes the current state of the material between $s=0$ directly at the position of the crack and $s=1$ some distance away, where the material is in its virgin, fully intact state.

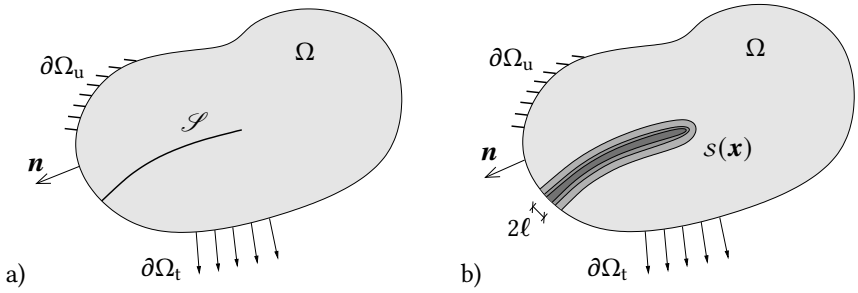


Figure 3.2: Elastic body with a) sharp crack described by the crack set $\mathscr{S}$ and b) corresponding phase field approximation $s(\boldsymbol{x})$ of the crack with smooth transition between broken and intact material state.

The regularized description of the crack set $\mathscr{S}$ by the "crack field" $s(\boldsymbol{x})$ (e.g. Bourdin et al., 2008) and discontinuity approximation of the (vector-valued) displacement field $\boldsymbol{u}$ across the crack resembles the approximation of the edge set and the gray level by the scalar fields u and v, respectively. While edges are extracted from given (and thus fixed) image pixels, a crack set typically results from previous and current mechanical loads. According to this, the strain energy (2.200) is approximated by

$$E_{\mathrm{e}}(\boldsymbol{u}, s) = \int_{\Omega} \frac{1}{2}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \left[(s^2 + \eta_\epsilon)\, \mathbb{C}_0\right] : \boldsymbol{\varepsilon}(\boldsymbol{u})\, \mathrm{d}V \quad , \tag{3.7}$$

while the approximation of the crack surface energy (2.197) is given by

$$\Gamma(s) = \mathcal{G}_{\mathrm{c}} \int_{\Omega} \frac{(1-s)^2}{2\ell} + \frac{\ell}{2}|\nabla s|^2\, \mathrm{d}V \quad . \tag{3.8}$$

The small parameter ℓ has the dimension of a length and controls the width of the transition zone between the two limit states (see Fig. 3.2b). Finally, the solution of the regularized approach is obtained by minimizing the total free energy

$$E_{\mathrm{tot}}(\boldsymbol{u}, s) = \int_{\Omega} \left[\frac{1}{2}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \left[(s^2 + \eta_\epsilon)\, \mathbb{C}_0\right] : \boldsymbol{\varepsilon}(\boldsymbol{u}) + \mathcal{G}_{\mathrm{c}} \left(\frac{(1-s)^2}{2\ell} + \frac{\ell}{2}|\nabla s|^2\right)\right] \mathrm{d}V \tag{3.9}$$

with the alternate minimization algorithm (Bourdin et al., 2000) and the backtracking algorithm, which has been introduced later on (Bourdin et al., 2008). For functionals of the type of (3.9), it has been discussed in terms of Γ-convergence, that the approximated

solution converges to that of the sharp crack topology as the parameter $\ell \to 0$, see e.g. Bourdin et al. (2000). Accordingly, the zero set of $s(\boldsymbol{x})$, which is defined by

$$\mathscr{S}_{\mathrm{s}} = \{\boldsymbol{x} \in \Omega : \; s(\boldsymbol{x}) = 0\} \quad , \tag{3.10}$$

approaches the crack set $\mathscr{S}$. It is noted that the subscript s is used to refer to the approximated solution. In addition, the approach is said to be able to solve a wide variety of fracture problems like crack initiation, crack growth and crack branching by common finite element techniques without additional assumptions (Bourdin et al., 2000).

3.2 General framework and governing equations

As already stated by Bourdin et al. (2000), the straightforward adaptation of the regularization technique of the Mumford-Shah problem to the variational formulation of brittle fracture resembles the approach of regularized phase transition problems. In this spirit, Kuhn and Müller (2008) reinterpreted the approach introduced above as phase transition problem and the state variable s as the corresponding "phase field order parameter", see also Kuhn (2013). Later on, Miehe et al. (2010a); Miehe et al. (2010c) used the term "crack phase field" for a similar approach.

To avoid confusion, it should be noted that there is a number of alternative fracture modeling approaches which stem directly from phase transition physics, see e.g. Aranson et al. (2000), Karma et al. (2001), Henry and Levine (2004), Karma and Lobkovsky (2004), Gugenberger (2009), Hakim and Karma (2009), Spatschek et al. (2011). They include similar terms, however, the underlying models and approximated physical quantities deviate significantly. Most significantly, crack propagation is controlled by an auxiliary energetic barrier between the intact and broken state instead of relating the (smeared) surface energy with the energy release, i.e. according to the Griffith criterion. Thus, in the remainder of the work, the regularized approach as introduced in Sect. 3.1 is used, but in conjunction with the terminology "phase field".

For a broader discussion of the approach, the energy functional of the regularized version of the variational approach to fracture (3.9) is rewritten in a more general manner. It can be seen as a general framework of phase field modeling of (brittle) fracture. Inspired by the previous derivation, a more general form of (3.9) is given by the total energy

$$E_{\mathrm{tot}}(\boldsymbol{\varepsilon}, s) = \int_{\Omega} \psi_{\mathrm{e}}(\boldsymbol{\varepsilon}, s) + \overbrace{\underbrace{\mathcal{G}_{\mathrm{c}} \frac{1}{2\ell \, c_{\mathrm{w}}} \left(w(s) + \ell^2 \, |\nabla s|^2 \right)}_{=: \, \gamma_\ell(s, \nabla s)}}^{=: \, \psi_{\mathrm{s}}(s, \nabla s)} \mathrm{d}V \quad , \tag{3.11}$$

where $\psi_{\mathrm{e}}(\boldsymbol{\varepsilon}, s)$ is the strain energy density degraded by the phase field and $\psi_{\mathrm{s}}(s, \nabla s)$ is the crack surface energy density. Typically, the strain energy density is isotropically

degraded (the stiffness in all directions is degraded in the same manner). This is achieved by introducing a degradation function $g(s)$, which relates the change of the phase field parameter to the corresponding loss of stiffness. A simple choice, in accordance with (3.9), is the full degradation (the whole strain energy is degraded) by a quadratic degradation function $g(s) := s^2 + \eta_\epsilon$ with the small remaining stiffness parameter $0 < \eta_\epsilon \ll 1$, i.e.

$$\psi_e(\boldsymbol{\varepsilon}, s) := g(s)\, \psi_0(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon} : \left[g(s)\, \mathbb{C}_0 \right] : \boldsymbol{\varepsilon} \quad . \tag{3.12}$$

Following the terminology introduced in the field of continuum damage mechanics (see Sect. 2.2.8) the effective stress tensor according to (2.169), i.e. the Cauchy stress tensor for intact, not degraded material, is identified to be

$$\tilde{\boldsymbol{\sigma}} := \mathbb{C}_0 : \boldsymbol{\varepsilon} = \frac{\boldsymbol{\sigma}}{g(s)} \tag{3.13}$$

in phase field models with full and isotropic degradation.

Meanwhile, the regularized crack is represented by the crack surface density function $\gamma_\ell(s, \nabla s)$, which consists of a local part described by the function $w(s)$ and a non-local part based on the phase field gradient ∇s. To incorporate the non-local part, an internal length parameter ℓ is required for dimensional reasons, which acts as regularization parameter. The scaling parameter c_w depends on the choice of the local crack surface function $w(s)$ and ensures the fundamental property

$$\int_{-\infty}^{\infty} \gamma_\ell(s, \nabla s)\, \mathrm{d}x_n = 1 \quad , \tag{3.14}$$

where x_n denotes the coordinate normal to the fracture surface. In a domain of finite size, e.g. in (3.11), the property (3.14) is fulfilled for a vanishing length of the regularization parameter $\ell \to 0$. The surface energy of a discrete crack is represented by a volume integral

$$\int_\Gamma \mathcal{G}_c\, \mathrm{d}A \approx \int_\Omega \mathcal{G}_c\, \gamma_\ell(s, \nabla s) \mathrm{d}V \tag{3.15}$$

and defined in the whole domain Ω. In accordance with (3.9), a common choice of the local crack surface function is the quadratic function $w(s) := (1 - s)^2$ with $c_w = 1$.

3.2.1 Application of Hamilton's principle

The governing equations of phase field approaches to fracture are typically derived from variational principles, see, e.g., Kuhn and Müller (2010a) and Miehe et al. (2010a; 2010c). The application of Hamilton's principle of stationary action, as introduced in Sect. 2.1.4, yields the coupled set of *Euler-Lagrange-equations*, which provides the strong form of the governing equations. Dynamic phase field fracture models have been introduced

structurally similar to one another by Borden et al. (2012), Hofacker and Miehe (2012) and Schlüter et al. (2014).

By following Borden et al. (2012), we start from the Lagrangian

$$\begin{aligned} L(\dot{\boldsymbol{u}}, \boldsymbol{\varepsilon}, s, \nabla s) &= K - E_{\text{tot}} \\ &= \int_{\Omega} \frac{1}{2} \rho\, \dot{\boldsymbol{u}} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}V - \int_{\Omega} \psi_{\text{e}}(\boldsymbol{\varepsilon}, s) \, \mathrm{d}V - \int_{\Omega} \mathcal{G}_{\text{c}} \, \gamma_{\ell}(s, \nabla s) \, \mathrm{d}V \quad , \end{aligned} \tag{3.16}$$

with the kinetic energy (2.56) and the total free energy, which is the sum of strain energy and crack surface energy according to (3.11). To conserve the mass of the body, the kinetic energy K is not affected by the phase field, see Hofacker and Miehe (2012) and Borden et al. (2012). In addition, by treating the critical energy release rate $\mathcal{G}_{\text{c}}$ as constant, it is assumed to be independent of the crack speed. Correspondingly, the Lagrangian density is

$$\mathcal{L}(\dot{\boldsymbol{u}}, \boldsymbol{\varepsilon}, s, \nabla s) = \frac{1}{2} \rho\, \dot{\boldsymbol{u}} \cdot \dot{\boldsymbol{u}} - \psi_{\text{e}}(\boldsymbol{\varepsilon}, s) - \mathcal{G}_{\text{c}} \, \gamma_{\ell}(s, \nabla s) \quad . \tag{3.17}$$

Applied body forces $\bar{\boldsymbol{f}}$ and boundary tractions $\bar{\boldsymbol{t}}$ can be non-conservative. In addition, there might be a viscous regularization of the crack advance, controlled by the viscosity parameter η_{vc}. These contributions are included in terms of generalized forces Q. Their virtual work (2.114) is

$$\delta W_{\text{f}} = \int_{\Omega} \bar{\boldsymbol{f}} \cdot \delta \boldsymbol{u} \, \mathrm{d}V \, , \quad \delta W_{\text{t}} = \int_{\partial \Omega_t} \bar{\boldsymbol{t}} \cdot \delta \boldsymbol{u} \, \mathrm{d}A \qquad \text{and} \qquad \delta W_{\eta} = \int_{\Omega} \left(\eta_{\text{vc}} \, \dot{s} \right) \delta s \, \mathrm{d}V \quad , \tag{3.18}$$

respectively. According to the principle of stationary action, the actual displacement field $\boldsymbol{u}$ and phase field s must fulfill

$$\int_{t_1}^{t_2} \left(\delta L + \int_{\Omega} \bar{\boldsymbol{f}} \cdot \delta \boldsymbol{u} \, \mathrm{d}V + \int_{\partial \Omega_t} \bar{\boldsymbol{t}} \cdot \delta \boldsymbol{u} \, \mathrm{d}A + \int_{\Omega} \left(\eta_{\text{vc}} \, \dot{s} \right) \delta s \, \mathrm{d}V \right) \mathrm{d}t = 0 \quad . \tag{3.19}$$

Further steps are based on the detailed derivation of the governing equations provided by Schlüter et al. (2014). With the exchange of integration over the domain and variation, the first term in (3.19) is rewritten

$$\int_{t_1}^{t_2} \left(\int_{\Omega} \delta \mathcal{L} \, \mathrm{d}V + \int_{\Omega} \bar{\boldsymbol{f}} \cdot \delta \boldsymbol{u} \, \mathrm{d}V + \int_{\partial \Omega_t} \bar{\boldsymbol{t}} \cdot \delta \boldsymbol{u} \, \mathrm{d}A + \int_{\Omega} \left(\eta_{\text{vc}} \, \dot{s} \right) \delta s \, \mathrm{d}V \right) \mathrm{d}t = 0 \quad . \tag{3.20}$$

The variation of the Lagrangian density with respect to the unknown fields $\boldsymbol{u}$ and s results in, by applying the product rule (A.47),

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}}\cdot\delta\dot{\boldsymbol{u}} + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}} : \delta\boldsymbol{\varepsilon} + \frac{\partial\mathcal{L}}{\partial s}\,\delta s + \frac{\partial\mathcal{L}}{\partial\nabla s}\cdot\nabla\delta s \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}}\cdot\delta\boldsymbol{u}\right) - \frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}}\cdot\delta\boldsymbol{u} - \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}}\right)\cdot\delta\boldsymbol{u} + \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}}\cdot\delta\boldsymbol{u}\right) \\ &\quad + \frac{\partial\mathcal{L}}{\partial s}\,\delta s - \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\nabla s}\right)\delta s + \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\nabla s}\,\delta s\right) \quad . \end{aligned} \tag{3.21}$$

Integration over time (in the interval $[t_1,\ t_2]$) and domain and employment of the divergence theorem (A.40) finally yields

$$\begin{aligned}\int_{t_1}^{t_2}\int_{\Omega}\delta\mathcal{L}\,\mathrm{d}V\,\mathrm{d}t &= \int_{t_1}^{t_2}\int_{\Omega}\left[-\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}} - \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}}\right)\right]\cdot\delta\boldsymbol{u}\,\mathrm{d}V\,\mathrm{d}t \\ &\quad + \int_{t_1}^{t_2}\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial s} - \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\nabla s}\right)\right]\delta s\,\mathrm{d}V\,\mathrm{d}t \\ &\quad + \int_{t_1}^{t_2}\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}}\cdot\boldsymbol{n}\right]\cdot\delta\boldsymbol{u}\,\mathrm{d}A\,\mathrm{d}t + \int_{t_1}^{t_2}\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\nabla s}\cdot\boldsymbol{n}\right]\delta s\,\mathrm{d}A\,\mathrm{d}t \\ &\quad + \int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}}\cdot\delta\boldsymbol{u}\right]\mathrm{d}V\Big|_{t_1}^{t_2}\end{aligned} \tag{3.22}$$

With the domain Ω being independent of time, the order of integration of time and domain can be exchanged according to

$$\int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}}\cdot\delta\boldsymbol{u}\right]\mathrm{d}V\Big|_{t_1}^{t_2} = \int_{\Omega}\left[\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}}\cdot\delta\boldsymbol{u}\right]\Big|_{t_1}^{t_2}\mathrm{d}V = 0 \quad , \tag{3.23}$$

which vanishes, if $\delta\boldsymbol{u}(\boldsymbol{x},t_1) = \delta\boldsymbol{u}(\boldsymbol{x},t_2) = \boldsymbol{0}$ is chosen. By employing the fundamental Lemmas of variational calculus, see (2.117) and (2.118), the square bracket terms in (3.22) must vanish independently, i.e. for each admissible variation $\delta\boldsymbol{u}$ and δs. Thus, the coupled set of Euler-Lagrange equations of the variational expression (3.18) is given by

$$\bar{\boldsymbol{f}} - \frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{u}}} - \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}}\right) = \boldsymbol{0} \tag{3.24}$$

$$\eta_{\mathrm{vc}}\,\dot{s} + \frac{\partial\mathcal{L}}{\partial s} - \operatorname{div}\left(\frac{\partial\mathcal{L}}{\partial\nabla s}\right) = 0 \tag{3.25}$$

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}}\cdot\boldsymbol{n} + \bar{\boldsymbol{t}} = \boldsymbol{0} \quad \text{on } \partial\Omega_t \tag{3.26}$$

and

$$\frac{\partial \mathcal{L}}{\partial \nabla s} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega_s \quad . \tag{3.27}$$

By using the actual expressions of the Lagrangian density $\mathcal{L}(\dot{\boldsymbol{u}}, \boldsymbol{\varepsilon}, s, \nabla s)$, see (3.17), one finally obtains from (3.24-3.25) the momentum balance

$$\rho\, \ddot{\boldsymbol{u}} = \operatorname{div} \boldsymbol{\sigma} + \bar{\boldsymbol{f}} \tag{3.28}$$

with the Cauchy stress tensor

$$\boldsymbol{\sigma} = \frac{\partial \psi_e(\boldsymbol{\varepsilon}, s)}{\partial \boldsymbol{\varepsilon}} = g(s)\, \frac{\partial \psi_0(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \tag{3.29}$$

and the phase field equation

$$\begin{aligned} &\eta_{vc}\, \dot{s} + \frac{\mathrm{d}\psi_e(\boldsymbol{\varepsilon}, s)}{\mathrm{d}s} + \frac{\mathrm{d}\psi_s(s, \nabla s)}{\mathrm{d}s} - \mathcal{G}_c\, \frac{\ell}{c_w}\, \Delta s \\ &= \eta_{vc}\, \dot{s} + \underbrace{\frac{\mathrm{d}g(s)}{\mathrm{d}s}\, \psi_0(\boldsymbol{\varepsilon})}_{=:\, f_s(\boldsymbol{\varepsilon}, s)} + \frac{\mathcal{G}_c}{c_w} \left(\frac{1}{2\ell}\, \frac{\mathrm{d}w(s)}{\mathrm{d}s} - \ell\, \Delta s \right) = 0 \end{aligned} \tag{3.30}$$

where Δs denotes the Laplacian of the phase field. Equation (3.30) governs the evolution of the phase field driven by the elastic energy. In absence of any phase field constraints the phase field evolution is driven by the strain energy only. Consequently, in terms of thermodynamic forces the second term in (3.30) can be identified as *crack driving force* $f_s(\boldsymbol{\varepsilon}, s)$ while the third term is the *crack resistance force* (Miehe et al., 2015). While displacement boundary conditions $\boldsymbol{u} = \bar{\boldsymbol{u}}$ are prescribed on $\partial\Omega_u$ (see Fig. 3.2), the traction boundary conditions

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \bar{\boldsymbol{t}} \quad \text{on } \partial\Omega_t \quad , \tag{3.31}$$

with $\boldsymbol{n}$ being the outward normal vector on $\partial\Omega = \partial\Omega_u \cup \partial\Omega_t$ with $\partial\Omega_t \cap \partial\Omega_u = \emptyset$, as well as the boundary condition for the phase field

$$\nabla s \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega \tag{3.32}$$

are obtained from the variational principle (3.26-3.27). In addition, initial cracks can be modeled by the phase field, e.g., by imposing phase field boundary conditions $s = 0$ along $\mathscr{S}_s$, see App. C.4.

3.2.2 Leaving the variational scheme

A variational principle has been used to obtain the coupled set of partial differential equations and the corresponding boundary conditions (3.28-3.32). Foremost, using a variational principle enforces the energy transfer from the elastic field to the phase field to be fully consistent, i.e. there is no artificial loss of energy besides the viscous regularization.

However, there are several reasons to break the variational derivation of the governing equations of the phase field approaches to fracture, some of them are discussed in the remainder of this work.

The coupling of both fields, the displacement field and the phase field, is clearly visible from the momentum balance (3.28) in conjunction with stress degradation (3.29) and the evolution equation (3.30), where the elastic energy drives the phase field evolution. While the momentum balance (3.28) typically remains unchanged, the phase field evolution equation is reformulated as

$$\eta_{\mathrm{vc}}\,\dot{s} + \frac{\mathrm{d}g(s)}{\mathrm{d}s}\,D_{\mathrm{s}}(\boldsymbol{\varepsilon}) + \frac{\mathcal{G}_{\mathrm{c}}}{c_{\mathrm{w}}}\left(\frac{1}{2\ell}\,\frac{\mathrm{d}w(s)}{\mathrm{d}s} - \ell\,\Delta s\right) = 0 \quad , \tag{3.33}$$

with the modified *crack driving energy* $D_{\mathrm{s}}(\boldsymbol{\varepsilon})$. In comparison with (3.30), the specific choice of

$$D_{\mathrm{s}}(\boldsymbol{\varepsilon}) := \psi_0(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon} : \mathbb{C}_0 : \boldsymbol{\varepsilon}/2 \tag{3.34}$$

corresponds to the variationally "consistent" version as obtained in Sect. 3.2.1. For all other possible choices of $D_{\mathrm{s}}(\boldsymbol{\varepsilon})$, the present set of equations (3.28) and (3.33) cannot be derived from a variational principle since the underlying energy functional does not exist. Note, that the crack driving energy can also be formulated in terms of the effective stress tensor (3.13), which is an isotropic function of the strain tensor (Sect. 2.1.3). For the sake of clarity the modified crack driving energy formulated in the effective stress space is denoted by $D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}})$. A variational structure — though mathematically pleasant — is neither necessary nor sufficient for the physical correctness of a model and its governing equations. It should also be emphasized that the basic structure of the evolution equation (3.30) is still retained in (3.33) by keeping the derivative $\mathrm{d}g/\mathrm{d}s$ as a scaling factor in front of the crack driving energy $D_{\mathrm{s}}(\boldsymbol{\varepsilon})$ in order to render the energy transfer from the elastic field to the phase field consistent. A more detailed discussion is provided in Sect. 3.5.4.

3.2.3 Incorporation of crack irreversibility

As introduced in Sect. 2.2, macroscopic fracture describes material separation which is assumed to be of irreversible nature, so effects of crack healing are not considered[1]. The equations obtained from the variational principle (Sect. 3.2.1) or their somewhat modified version (Sect. 3.2.2), however, do not take the irreversible character of fracture into account. The phase field variable s can evolve freely in both directions (between 0 and 1) only depending on the current loading conditions. If, for example, external loads are reduced after a crack (described by the phase field) has been formed, the phase

[1] Crack healing, which means at least partial reestablishment of bonding between crack surfaces, can take place on a microscopic level in many different materials, e.g. in vulcanized rubber due to the rearrangement of inherent sulfur cross-linked networks (Xiang et al., 2015) or in soda-lime glass Wiederhorn and Townsend (1970). Since in this work only macroscopic cracks are considered, these effects are out of the scope.

field might regenerate towards its initial value $s = 1$. By regarding the phase field evolution equation (3.30) or (3.33) it is evident that the phase field is driven only by strain energy and that strain energy and surface energy can be converted in both direction. From the conceptual perspective, there are two different approaches to incorporate crack irreversibility, which are presented in the following.

Crack-like irreversibility

The "crack-like" irreversibility constraint

$$\mathscr{S}_{\mathrm{s}}(t) \subseteq \mathscr{S}_{\mathrm{s}}(t + \mathrm{d}t) \qquad \forall \mathrm{d}t > 0 \tag{3.35}$$

is close to the original idea, that the current crack set contains all previous cracks (2.196), and has been introduced by Bourdin et al. (2000). Therefore, the phase field is fixed by applying suitable constraints as soon as a crack is formed, i.e. the cracked "state" $s = 0$ is reached. To allow further crack growth, indicated in the model by further evolution of $s(\boldsymbol{x})$, the phase field is only constrained directly at the crack, e.g. by imposing Dirichlet-type boundary conditions

$$s(t + \mathrm{d}t) \stackrel{!}{=} 0 \qquad \text{if } s(t) \leq s_{\mathrm{tol}} \qquad \forall \mathrm{d}t > 0 \quad . \tag{3.36}$$

As introduced in Section 3.1, the crack in the phase field approach is indicated by the zero set of the phase field (3.10). For numerical reasons, usually a small threshold value

$$0 < s_{\mathrm{tol}} \ll 1 \tag{3.37}$$

is used, which defines the phase field crack set

$$\mathscr{S}_{\mathrm{s}} = \{\boldsymbol{x} \in \Omega : \; s(\boldsymbol{x}) \leq s_{\mathrm{tol}}\} \quad . \tag{3.38}$$

A reasonable choice is, for example, $s_{\mathrm{tol}} = 10^{-8}$ (Schlüter, 2018). Dirichlet boundary conditions can be imposed by reformulating the system of algebraic equations. The corresponding procedure for the "crack-like" irreversibility in the phase field approach has been proposed by Kuhn (2013).

After the release of all mechanical loads (and by neglecting viscosity) the local and non-local term of the phase field are in equilibrium with each other, i.e.

$$\frac{\mathrm{d}w(s)}{\mathrm{d}s} = 2\ell^2 \, \Delta s \quad . \tag{3.39}$$

Since the phase field on $\Omega \setminus \mathscr{S}_{\mathrm{s}}$ is still governed by its evolution equation (3.33) it can recover with $\dot{s} > 0$. This is also the case before s_{tol} is reached. Thus, by applying "crack-like" irreversibility constraints, the phase field acts as an "auxiliary indicator field for cracks" (everywhere where it does not regularize a crack), see Kuhn (2013).

Damage-like irreversibility

Conceptually different is the "damage-like" irreversibility, which has been introduced in different versions by Kuhn and Müller (2008), and Miehe et al. (2010a; 2010c). As the name indicates, the phase field variable $s(\boldsymbol{x})$, to be more precise, its counterpart $d := 1 - s$ is interpreted as damage parameter. It quantifies locally the amount of damage and via the degradation function $g(s)$ the related permanent loss of stiffness, see Sect. 2.2.8. Thus, the phase field is non-increasing with

$$\dot{s} \leq 0 \quad . \tag{3.40}$$

The conversion of kinetic and elastic energy into crack surface energy is a unidirectional process for every, also partial, phase field evolution. Instead of directly constraining the direction of phase field evolution, Miehe et al. (2010a) proposed the modification of the crack driving energy based on the non-variational approach given in Sect. 3.2.2. By introducing a "local history field of maximum positive reference energy" $\mathcal{H}(\psi_0)$ the approach is, in general, given by

$$D_\mathrm{s} := \mathcal{H}(\psi_0) = \max_{\tau \in [0,\, t]} \psi_0(\boldsymbol{\varepsilon}(\boldsymbol{x}, \tau)) \quad . \tag{3.41}$$

Imposing irreversibility by artificial viscosity

As discussed by Kuhn (2013), the phase field evolution equation (3.30) can also be derived from considering the balance law in a microforce system[2]. For this purpose, Kuhn and Müller (2008) introduced the kinetic "mobility parameter" M with

$$M := \frac{1}{\eta_\mathrm{vc}} \quad . \tag{3.42}$$

While the evolution equation with a constant non-negative mobility corresponds to the classical Ginzburg-Landau equation, a more general form allows the mobility itself to be non-constant (Gurtin, 1996), i.e. $\tilde{M}(s, \dot{s}, \dots)$, to form the evolution equation

$$\dot{s} = -\tilde{M} \left[\frac{\mathrm{d}g(s)}{\mathrm{d}s}\, \psi_0 + \frac{\mathcal{G}_\mathrm{c}}{c_\mathrm{w}} \left(\frac{1}{2\ell}\, \frac{\mathrm{d}w(s)}{\mathrm{d}s} - \ell\, \Delta s \right) \right] \quad . \tag{3.43}$$

The rate-independent limit case of phase field evolution is reached for the mobility being a constant with $\tilde{M} = M \to \infty$, which is equivalent to $\eta_\mathrm{vc} = 0$ in (3.30). To enforce a "crack-like" irreversibility constraint, Kuhn (2013) proposes to apply

$$\tilde{M}(s) = \begin{cases} M & \text{if} \quad s > 0 \\ 0 & \text{else} \end{cases} \quad , \tag{3.44}$$

[2] Gurtin (1996) uses a microforce balance law to set up a framework for Cahn-Hilliard and Ginzburg-Landau type equations. Both equations describe the evolution of order parameters in two-phase systems. While the first describes the transport of atoms between unit cells the latter describes the ordering of atoms within unit cells on a lattice.

while the "damage-like" irreversibility is obtained by

$$\tilde{M}(\dot{s}) = \begin{cases} M & \text{if} \quad \dot{s} \leq 0 \\ 0 & \text{else} \end{cases} \quad . \tag{3.45}$$

In both cases, the unintended behavior, either the release of the phase field from its minimum value $s = 0$ or its positive evolution $\dot{s} > 0$ is suppressed by imposing locally a zero mobility, which corresponds to an (almost) infinite amount of artificial viscosity.

Similar approaches to enforce a "damage-like" irreversibility constraint have been proposed by Miehe et al. (2010c). One of them is based on a three-field approach, where a separate auxiliary field equation handles phase field-related forces. Another one is a common two-field approach with an "artificial viscous hardening" term that penalizes the unintended increase of the phase field $\dot{s} > 0$. The phase field evolution equation in this context is then

$$\epsilon_{\text{ir}} \langle \dot{s} \rangle + \frac{\mathrm{d}g(s)}{\mathrm{d}s} \psi_0 + \frac{\mathcal{G}_{\text{c}}}{c_{\text{w}}} \left(\frac{1}{2\ell} \frac{\mathrm{d}w(s)}{\mathrm{d}s} - \ell \, \Delta s \right) = 0 \quad , \tag{3.46}$$

with $\epsilon_{\text{ir}} \gg 0$ being a penalty parameter, which has to be chosen as large as possible in relation to the time step size the conditioning number of the system (Miehe et al., 2010c).

It should be noted, that in absence of the enforcement of irreversibility, additional viscous damping is not required in dynamic models as stated by Borden et al. (2012) and Hofacker and Miehe (2012).

3.2.4 Weak form and numerical solution

The set of coupled partial differential equations (3.28) and (3.30) or its modified version (3.33) for the displacement field $\boldsymbol{u}(\boldsymbol{x})$ and the phase field $s(\boldsymbol{x})$ is solved by the finite element method. By introducing test functions $\delta\boldsymbol{u}$, $\delta\boldsymbol{\varepsilon} := \nabla_{\text{s}} \delta\boldsymbol{u}$ and δs, the weak forms of the momentum balance (3.28) and phase field evolution equation (3.33) are

$$- \int_{\Omega} \left[\delta\boldsymbol{u} \cdot \left(\rho \, \ddot{\boldsymbol{u}} \right) + \delta\boldsymbol{\varepsilon} : \boldsymbol{\sigma}(\boldsymbol{\varepsilon}, s) \right] \, \mathrm{d}V + \int_{\Omega} \delta\boldsymbol{u} \cdot \bar{\boldsymbol{f}} \, \mathrm{d}V + \int_{\partial\Omega_{\text{t}}} \delta\boldsymbol{u} \cdot \bar{\boldsymbol{t}} \, \mathrm{d}A = 0 \tag{3.47}$$

and

$$\int_{\Omega} \left[\delta s \left(\eta_{\text{vc}} \, \dot{s} \right) + \delta s \, g'(s) \, D_{\text{s}}(\boldsymbol{\varepsilon}) + \delta s \left(\frac{\mathcal{G}_{\text{c}}}{\ell \, c_{\text{w}}} \, w'(s) \right) + \nabla \delta s \cdot \left(\mathcal{G}_{\text{c}} \, \ell \, \nabla s \right) \right] \, \mathrm{d}V = 0 \, , \tag{3.48}$$

respectively. In case of a variationally consistent approach the crack driving energy in (3.48) is given by $D_{\text{s}}(\boldsymbol{\varepsilon}) = \psi_0(\boldsymbol{\varepsilon})$, if a local history field is used the crack driving energy is $D_{\text{s}}(\boldsymbol{\varepsilon}) = \max_{\tau \in [0, \, t]} \psi_0(\boldsymbol{\varepsilon}(\boldsymbol{x}, \tau))$, see (3.41). After time integration and spatial discretization

(more details are provided in App. C.2) the global system of coupled equations in terms of internal forces $\boldsymbol{P}$ and external forces $\boldsymbol{F}$ is

$$\boldsymbol{R} = \boldsymbol{F} - \boldsymbol{P}(\boldsymbol{d}, \dot{\boldsymbol{d}}, \ddot{\boldsymbol{d}}) = \boldsymbol{0} \quad , \tag{3.49}$$

with $\boldsymbol{d}$ being the vector of unknown displacements $\boldsymbol{u}$ and phase field values s.

In order to solve the resulting coupled set of algebraic equations two solution procedures are available: a monolithic approach and an alternate minimization procedure, also referred to as staggered scheme (Miehe et al., 2010b). To achieve reliable solutions small time/load steps are necessary in most situations. In a monolithic solution scheme the governing equations for both unknown fields are solved simultaneously. Since in the quasi-static case convexity with respect to both unknown fields cannot be guaranteed (Bourdin, 2007a; Bourdin et al., 2000), this can be accomplished by either introducing an artificial viscosity (e.g. Kuhn and Müller, 2010a), a special type of line search algorithm (Gerasimov and De Lorenzis, 2016) or by including inertia (Borden et al., 2012; Schlüter et al., 2014; Steinke et al., 2017). By contrast, by using an alternate minimization procedure the displacements $\boldsymbol{u}$ and the phase field s are computed separately in an alternating manner based on the respective other results (Bourdin et al., 2000). The main advantage is that two, yet much smaller systems of equations have to be solved, where each problem is convex. Moreover, any modification of the phase field evolution equation (Sect. 3.2.2) would cause unsymmetric entries in the coupling terms of the tangent matrix in a monolithic approach resulting in higher computational costs (Zienkiewicz et al., 2013). In addition to small time/load steps, several iterations in each load step between the solutions of both equation systems might be necessary until convergence for the coupled problem is reached. According to Bourdin et al. (2000), convergence can be assumed when the change of the phase field between consecutive iterations is below a tolerance δ_s as illustrated in Fig. 3.3, see also Borden (2012) and Wheeler et al. (2014). Keeping track of only the phase field is sufficient for a converged solution, since without further change in the phase field solution also the displacement solution does not change anymore. Moreover, the phase field is a dimensionless quantity and therefore eligible to check convergence. Alternative criteria to control convergence of alternate minimization have been investigated by Ambati et al. (2015a). The number of iterations strongly depends on the choice of the specific phase field formulation, the load step size, the state of crack propagation and also the finite element mesh[3]. Miehe et al. (2010a) proposed a "robust"[4] alternate minimization procedure (there called staggered scheme) in which both field solutions are not reiterated. Also used by Borden et al. (2012), Hofacker and Miehe (2012), Hesch and Weinberg (2014), Streich (2015), Nguyen et al. (2016), Molnár and Gravouil (2017), to name a few, solutions obtained by this scheme might not be in equilibrium even for small load steps, especially at the

[3] Without providing a systematic investigation, a too coarse finite element mesh in the zone of crack growth (thus the spatial and temporal approximation are of low quality) seems to converge much better/faster.

[4] In fact, the scheme is extremely simple and robust and almost always yields a solution. However, one cannot make sure it is the correct one. E.g. phase field evolution at rapid failure might take place over many time steps, while unphysical loading can still be applied.

onset of crack growth. Thus, crack formation and propagation is often delayed, which can be easily avoided by allowing for reiterations.

Alternate minimization for load increment n:
→ initialization: $k = 1$, set $\boldsymbol{u}_n^{(0)} = \boldsymbol{u}_{n-1}$ and $\boldsymbol{s}_n^{(0)} = \boldsymbol{s}_{n-1}$
→ alternate minimization cycle $k \geq 1$
→ solve $\boldsymbol{s}_n^{(k)}$ with constraint $\boldsymbol{u}_n = \boldsymbol{u}_n^{(k-1)}$
→ solve $\boldsymbol{u}_n^{(k)}$ with constraint $\boldsymbol{s}_n = \boldsymbol{s}_n^{(k)}$
repeat cycle with $k + 1 \rightarrow k$ until $\|\boldsymbol{s}_n^{(k)} - \boldsymbol{s}_n^{(k-1)}\|_\infty < \delta_s = 10^{-5}$

Figure 3.3: Alternate minimization solution procedure; $\boldsymbol{u}$ and $\boldsymbol{s}$ denote sets of all nodal displacements and phase field values as introduced in App. C.2.

3.3 Distribution of crack surface and stiffness degradation

The phase field evolution equation (3.30) or (3.33) governs the energy transfer between the strain field and the phase field. In this context, the exact form of the (stiffness) degradation function $g(s)$, which is also present in the momentum balance equation (3.28), and the function $w(s)$, in the following referred to as (crack) *surface function*, play key roles, which are analyzed in this section.

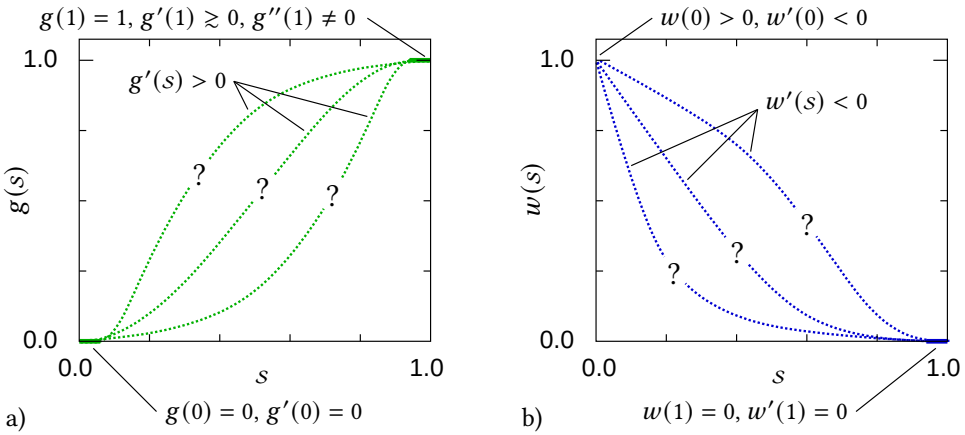


Figure 3.4: Illustration of a) degradation functions $g(s)$ and b) surface functions $w(s)$ in combination with typical requirements.

In fact, the functions describe the local distributions of stiffness degradation and crack surface, respectively. Figure 3.4 illustrates typical degradation and surface functions in conjunction with some of their requirements. Since the regularization of the crack is meaningful without any elastic contribution, the discussion starts with the surface function. Since most of the following discussion is regarded in the context of one-dimensional equations, spatial derivatives are rather indicated by comma notation, i.e. $(\cdot)_{,x}$ for the first and $(\cdot)_{,xx}$ for the second spatial derivative, than by symbolic notation. By contrast, derivatives with respect to the phase field s are given by

$$g'(s) := \frac{\mathrm{d}g(s)}{\mathrm{d}s} \qquad \text{and} \qquad w'(s) := \frac{\mathrm{d}w(s)}{\mathrm{d}s} \quad . \tag{3.50}$$

3.3.1 Surface function

The key ingredient of the phase field approach to fracture is the smeared representation of the crack surface, see Sect. 3.1. This feature is based on the crack surface density

$$\gamma_{\tilde{\ell}}(s, \nabla s, \Delta s, \dots) = \frac{1}{2\tilde{\ell}\, \tilde{c}_{\mathrm{w}}} \left(a_0\, w(s) + a_1\, \tilde{\ell}^{\,2}\, |\nabla s|^2 + a_2\, \tilde{\ell}^{\,4}\, (\Delta s)^2 + \dots \right) \quad , \tag{3.51}$$

which consists of a local and a non-local part. The local part is typically formulated in terms of the surface function $w(s)$, while the non-local part is based on the first until n-th spatial derivative of the phase field $s(\boldsymbol{x})$. There is a scaling constant $\tilde{c}_{\mathrm{w}}$ and, for dimensional reasons, the internal length parameter $\tilde{\ell}$. In the most general context, each summand in (3.51) is weighted by a factor $a_0, a_1, \dots, a_n$. This work is focused on crack surface densities including only the first spatial derivative of the phase field. Accordingly, the crack surface density (3.51) is

$$\gamma_{\ell}(s, \nabla s) = \frac{1}{2\ell\, c_{\mathrm{w}}} \left(w(s) + \ell^2\, |\nabla s|^2 \right) \quad . \tag{3.52}$$

Note, that due to the scaling by $\ell = \tilde{\ell}\, \sqrt{a_1/a_0}$ and $c_{\mathrm{w}} = \tilde{c}_{\mathrm{w}}/\sqrt{a_0\, a_1}$ the weight factors cannot be chosen independently in this specific case. The scaling parameter can then be obtained from

$$c_{\mathrm{w}} = 2 \int_0^1 \sqrt{w(s)}\, \mathrm{d}s \quad , \tag{3.53}$$

see App. B.3. A crack surface density which also includes higher-order spatial derivatives, more specifically the Laplacian Δs, has been introduced by Borden (2012) and is further discussed by Borden et al. (2014) and Weinberg and Hesch (2015).

In absence of external and internal mechanical forces the current phase field solution $s(\boldsymbol{x}, t)$ is governed by the equilibrium between the local, described by the function $w(s)$, and non-local forces (3.39) and related first- and second-type boundary conditions $s = 0$ on $\mathscr{S}_{\mathrm{s}}$ and (3.32), respectively. The crack surface density $\gamma_{\ell}(s, \nabla s)$ approximates the crack

surface A in terms of a Dirac delta distribution along the crack $\mathscr{S}$ as outlined by Verhoosel and Borst (2013). Thus, obtained phase field solutions resemble Dirac distributions. Based on a point-wise definition, the dimensionless function $w(s)$ relates the fractured surface area to the total cross sectional area. For this purpose, fundamental requirements on the local part of the surface description $w(s)$ are

$$w(s=1) = 0 \qquad \text{and} \qquad w(s=0) = 1 \quad . \tag{3.54}$$

In addition, it must be ensured that the energy release is non-positive as the phase field decreases. This in typically ensured by requesting the function $w(s)$ to be strictly monotonically decreasing between the two limit states, i.e.

$$w'(s) < 0 \quad \forall s \in (0,\, 1) \quad . \tag{3.55}$$

In order to give the natural solution $s = 1$ if neither mechanical forces nor non-natural boundary conditions are present, it is required that

$$w'(s=1) = 0 \quad . \tag{3.56}$$

This is no physically motivated restriction, but from the modeling standpoint of view, a pleasant property and simplifies the implementation and solution as discussed below.

Quadratic surface function

A common basis for the surface function $w(s)$ in (3.52) is one of the quadratic functions

$$w_\beta(s) := \beta\, s^2 - (\beta + 1)\, s + 1 \quad . \tag{3.57}$$

By considering all essential requirements from above, i.e. (3.54) and (3.55), the coefficient is limited to $\beta \in [-1, 1]$, as derived in App. B.2. Moreover, the quadratic function with (B.8_1) fulfills the optional requirement (3.56) only for $\beta = 1$. Not only because it was the first approach Bourdin (1998), but also since the choice

$$w_{\beta=1}(s) = (1 - s)^2 \tag{3.58}$$

is beneficial for implementation reasons (see below), it is the most widely used phase field approach to fracture in the literature. The function is sometimes denoted as "single-well potential" (e.g. Ambati et al., 2015a; Levitas et al., 2018), loosely referring to the terminology of multi-phase physics and to distinguish it from the double-well potential used by Karma et al. (2001). Since it is the only quadratic function with a single well (in the range of definition), this specific choice is here referred to as *quadratic single-well function.*

A special case of (3.57) is the linear surface function ($\beta = 0$)

$$w_{\beta=0}(s) = 1 - s \quad , \tag{3.59}$$

which has been analyzed in the context of phase field fracture by Pham et al. (2011). Another quadratic function is the concave quadratic function with $\beta = -1$

$$w_{\beta=-1}(s) = 1 - s^2 \quad , \tag{3.60}$$

which has been proposed and analyzed by Wu (2017). The scaling parameter for the quadratic function (3.57) can then be calculated as, see (B.23) and (B.24),

$$c_{\mathrm{w}} = \begin{cases} \dfrac{1}{4\beta}\left[\dfrac{(1-\beta)^2}{\sqrt{-\beta}}\left(\arcsin\left(\dfrac{1+\beta}{1-\beta}\right) - \dfrac{\pi}{2}\right) + 2(1+\beta)\right] & \forall \beta \in [-1,\, 0) \\ \dfrac{4}{3} & \forall \beta = 0 \\ \dfrac{1}{4\beta}\left[-\dfrac{(1-\beta)^2}{\sqrt{\beta}}\ \operatorname{arcosh}\left(\dfrac{1+\beta}{1-\beta}\right) + 2(1+\beta)\right] & \forall \beta \in (0,\, 1) \\ 1 & \forall \beta = 1 \end{cases} . \tag{3.61}$$

Treatment of unbounded solutions

For all choices of $\beta \neq 1$ the resulting problem becomes a bound-constrained minimization problem, where the phase field is not guaranteed to be smaller or equal to one if not treated adequately. Thus, special solution techniques such as bound-constrained solvers are required, see, e.g., the discussions by Wu (2017). Bleyer et al. (2017) for example, use a bound-constrained optimization solver integrated in the PETSc library, while Pham et al. (2011) use the large-scale bound-constrained quadratic optimization solver included in the Matlab Optimization Toolbox. As an alternative, Gerasimov and De Lorenzis (2019) applied a penalization technique to enforce $s \leq 1$. For this purpose, the additional penalty term

$$P_\epsilon = \frac{\epsilon_{\mathrm{s}}}{2} \int_\Omega \langle s - 1 \rangle^2 \, \mathrm{d}V \tag{3.62}$$

has to be added to the energy functional (3.11). The choice of the penalty parameter $\epsilon_{\mathrm{s}} \gg 0$ is critical, since a too small penalty parameter is not able to keep the phase field in a meaningful range. A too large penalty parameter, however, results in ill-conditioning of the system. Thus, Gerasimov and De Lorenzis (2019) also provide a procedure to calculate the lower bound of the penalty parameter to obtain a sufficiently accurate crack approximation (in terms of crack surface energy).

To avoid both special solution techniques and the global treatment of sufficiently local problems, a further alternative is to incorporate a penalization term in the formulation

itself by superimposing the desired surface function, e.g. w_β with $\beta \neq 1$, and a penalizing function $p_\epsilon(s)$, see App. B.4. In case of the quadratic surface function, this yields

$$\begin{aligned} w_\epsilon(s) &:= w_\beta(s) + p_\mathrm{w}(s) \\ &= \beta\, s^2 - (\beta+1)\, s + 1 + (1-\beta)\left[\, \exp\left(\epsilon_\mathrm{w}\,(s-1)\right) - 1\right]/\epsilon_\mathrm{w} \quad . \end{aligned} \tag{3.63}$$

Figure 3.5 shows the function graphs for two specific choices of β depending on the choice of the penalty parameter ϵ_w. It should be noted, that the approximation in the range $s < 1$ is quite accurate for $\epsilon_\mathrm{w} = 10^2$. The scaling parameter c_w can be calculated by numerical integration of (3.53). All corresponding values and their deviation due to penalization are provided in Tab. B.1 in App. B.4.

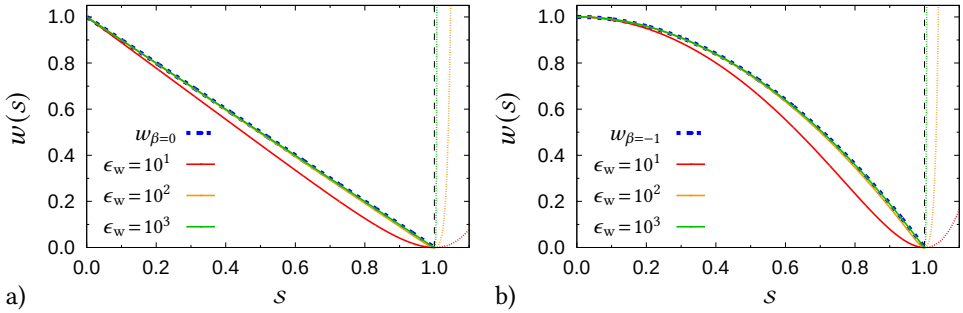


Figure 3.5: Surface function $w_\epsilon(s)$ depending on the choice of the penalty parameter ϵ_w to approximate a) the linear surface function $w(s) \approx 1 - s$ and b) the concave surface function $w(s) \approx 1 - s^2$.

3.3.2 Degradation function

The degradation function $g(s)$ introduced in (3.12) establishes the coupling between the strain field and the phase field. Its role is twofold: On the one hand it describes the decay of the elastic stiffness in (3.29) towards a vanishing stiffness with the phase field evolution $s \to 0$, on the other hand its first derivative $g'(s)$ quantifies the elastic energy contribution to the phase field evolution in (3.33).

To be consistent with the original material behavior in the intact state ($s=1$) and the (almost) complete loss of stiffness at the crack ($s=0$), the degradation function is subject to the conditions

$$g(s=1) = 1 \qquad \text{and} \qquad g(s=0) = \eta_\epsilon \quad , \tag{3.64}$$

with $0 < \eta_\epsilon \ll 1$ being a small parameter to avoid an ill-posed (zero) stiffness. From a spatially homogeneous consideration, η_ϵ must be large enough so that the conditioning number of the resulting algebraic system remains well-posed. However, since the fully degraded state does not affect large regions, it is omitted here for brevity (e.g. Borden,

2012; Kuhn, 2013). To guarantee positive energy release from the strain energy density at every instant of phase field evolution ($\dot{s} < 0$) between the two limit states, i.e.

$$- g'(s) \underbrace{D_s(\boldsymbol{\varepsilon})}_{>0} \dot{s} > 0 \quad , \tag{3.65}$$

with a positive definite crack driving energy $D_s(\boldsymbol{\varepsilon})$, the degradation function must be strictly monotonically increasing, i.e. its first derivative must be

$$g'(s) > 0 \quad \forall s \in (0,\, 1) \quad . \tag{3.66}$$

Note, that the combination of (3.64) and (3.66) already contains the requirement

$$g(s) > 0 \quad \forall s \in (0,\, 1] \quad , \tag{3.67}$$

which ensures the positive definiteness of the degraded strain energy. Aforementioned conditions are similar to those introduced in continuum damage mechanics formulations, see Sect. 2.2.8. Moreover, the phase field evolution has to seize in the fully broken state to keep s non-negative, which is enforced by the additional condition

$$g'(s{=}0) = 0 \quad . \tag{3.68}$$

Accordingly, the phase field approaches zero as $D_s(\boldsymbol{\varepsilon}) \to \infty$ because the crack driving strain energy term $g'(s)\, D_s(\varepsilon)$ has to be eliminated. Note that the other limit state (keeping $s \leq 1$), which is important e.g. if no strain energy is present, has to be either ensured by an appropriate choice of the local crack surface function $w(s)$ or by special solution techniques, see Sect. 3.3.1. In conjunction with the quadratic single-well function (3.58), Borden et al. (2012) additionally postulates

$$g'(s{=}1) \gtrapprox 0 \tag{3.69}$$

to perturb the homogeneous response $s = 1$, see also Kuhn et al. (2015). In the following, different choices of $g(s)$ are discussed.

Polynomial degradation function

The simplest polynomial degradation function which satisfies all essential requirements from above is the quadratic function

$$g(s) = s^2 \quad . \tag{3.70}$$

This function has already been used in (3.9) as introduced by Bourdin et al. (2000) and is extensively used in the literature, see e.g. Miehe et al. (2010c), Kuhn and Müller (2010a) and Steinke et al. (2017). In conjunction with the quadratic single-well function (3.58) its major drawback is that phase field evolution (and hence the stiffness degradation) starts already

at minimal strains, see also the analysis in Sect. 3.4. This gives rise to a significantly non-linear stress-strain response prior to fracture, which is unrealistic for brittle materials. To some extent, this behavior can be avoided by using the cubic degradation function

$$g(s) = (3 - a_s)\, s^2 + (a_s - 2)\, s^3 \tag{3.71}$$

as suggested by Borden et al. (2012). With

$$g'(s) = 2(3 - a_s)\, s + 3(a_s - 2)\, s^2 = s\left[6(1 - s) + (3s - 2)\, a_s\right] \tag{3.72}$$

the parameter a_s with

$$0 < a_s := g'(s{=}1) \ll 1 \tag{3.73}$$

takes the requirement (3.69) into account and serves to circumvent the numerical difficulties in detecting the onset of material degradation mentioned in Kuhn et al. (2015). Even though a_s has been introduced for numerical reasons, it can also be used as a modeling parameter (Schlüter, 2018) in its valid range $a_s \in (0,\, 3]$. With $a_s = 2$ the degradation function corresponds to the quadratic choice (3.70).

Karma et al. (2001) used a quartic degradation function, but in conjunction with the fundamentally different crack surface regularization by a double-well potential. Kuhn et al. (2015) analyzed the quartic degradation function in conjunction with the quadratic single-well function. By using the parameter a_s, the quartic degradation function is

$$g(s) = (4 - a_s)\, s^3 + (a_s - 3)\, s^4 \quad . \tag{3.74}$$

Inspired by the specific choice of cubic and quartic degradation functions, Strobl and Seelig (2018) proposed a specific family of higher-order polynomial degradation functions

$$g(s) = (n - a_s)\, s^{n-1} + (a_s - n + 1)\, s^n \tag{3.75}$$

with the exponent $n \in \mathbb{N}$, $n > 2$. Its construction is discussed in detail in App. B.5. Evidently, the function satisfies all requirements discussed above and the choice of the exponent n is, in fact, not restricted to natural numbers. With $a_s = 0$ the degradation function (3.75) can be simplified to

$$g(s) = n\, s^{n-1} - (n - 1)\, s^n = s^{n-1}\left[n - (n - 1)\, s\right] \quad . \tag{3.76}$$

One motivation to increase the polynomial order of the degradation function in this way is related to the role of the internal length parameter ℓ which governs the material strength as discussed in detail in Sect. 3.4.

Other types of degradation functions

To provide a wide overview of possible degradation functions and their underlying motivation, some different types are presented in the following. To model brittle fracture in piezoelectric ceramics[5], Wilson et al. (2013) introduced the degradation function

$$g(s) = \alpha \left[1 - \left(\frac{a-1}{a} \right)^{(1-s)^2} \right] \tag{3.77}$$

with $\alpha \in (1, \infty)$ being a modeling parameter. In contrast to the quadratic degradation function (3.70), which results for $\alpha \ll 1$ in the region of interest ($s \in [0, 1]$), degradation with (3.77) massively suppresses the crack driving term in the unphysical region $s > 1$. As discussed by Wilson et al. (2013), this is necessary for the proposed piezoelectric model and seems to be an issue of the coupling with the electric field, denoted as "anti-damaging" effect.

Sargado et al. (2018) introduced and discussed a family of degradation functions based on the exponential function. The most general form is

$$g(s) = (1 - w_{\text{deg}}) \, \frac{1 - \exp(-k \, s^n)}{1 - \exp(-k)} + w_{\text{deg}} \, f_{\text{corr}}(s) \tag{3.78}$$

with the modeling parameter $k > 0$, the exponent $n \geq 2$, a weighting factor $w_{\text{deg}} \in [0, 1]$ and the "corrector function" $f_{\text{corr}}(s)$. While increasing the modeling parameter k reduces the effect of non-linear behavior prior to fracture, it also increases some kind of snap-back behavior of the spatially homogeneous solution. Thus, Sargado et al. (2018) propose to chose k in accordance with the order n, i.e.

$$k(n) = \frac{(n-2)\,(1-s^*) + 1}{n\,(1-s^*)\,(s^*)^n} \tag{3.79}$$

with

$$s^* = \begin{cases} \dfrac{2}{3} & \text{for } n = 2 \\ 1 + \dfrac{(n+1) - \sqrt{5n^2 - 6n + 1}}{2(n^2 - 2n)} & \text{for } n > 2 \end{cases} . \tag{3.80}$$

Since for large values of $k(n)$ the crack driving force becomes almost zero long before reaching the broken state, Sargado et al. (2018) propose to add the corrector function

$$f_{\text{corr}}(s) = a_2 \, s^2 + a_3 \, s^3 \tag{3.81}$$

with

$$a_2 = \frac{3(1-s^*)^2 - 3}{3(1-s^*)^2 - 1} \qquad \text{and} \qquad a_3 = \frac{2}{3(1-s^*)^2 - 1} \tag{3.82}$$

[5] To model fracture in piezoelectric ceramics, the phase field has to be coupled with both the displacement field and an additional electric field.

to increase the slope of $g(s)$ in the post-failure range. Since the function $f_{\text{corr}}(s)$ does not satisfy all requirement of the degradation function, i.e. $f_{\text{corr}}(s) > 0$ for some states $s \in (0,\, 1)$, the weighting factor w_{deg} and thus the impact of $f_{\text{corr}}(s)$ should be "kept small" (Sargado et al., 2018).

Another degradation function based on the exponential function has been proposed by Steinke and Kaliske (2019). It is given by

$$g(s) = \frac{\exp[\alpha\,(1-s)] - \exp(\alpha)\,[1+\alpha(s)]}{(\alpha-1)\,\exp(\alpha)+1} \tag{3.83}$$

with $\alpha > 0$ and constructed in such a way, that it approaches for $\alpha \to \infty$ the proportional behavior between $g(s)$ and s, i.e. $g(s) \approx s$, however, by taking the fundamental requirement (3.68) into account.

Other types of degradation functions are the rational function

$$g(s) = \frac{s^2}{1+\alpha\,(1-s)^2} \tag{3.84}$$

used by Lorentz and Godard (2011) (in combination with a different surface function) with a positive parameter $\alpha > 0$ and the trigonometric function

$$g(s) = 1 - \cos\left(\frac{\pi}{2}\,s\right) \quad . \tag{3.85}$$

introduced and used by Yin and Kaliske (2019) and Yin et al. (2020).

3.3.3 Threshold function

Based on the non-variational treatment of the phase field equations introduced in Sect. 3.2.2, Miehe et al. (2015) proposed alternative "crack driving criteria". For this purpose, the (non-variational) phase field evolution equation (3.33) in conjunction with the quadratic single-well function (3.57) and the quadratic degradation function (3.70) is rescaled according to

$$\underbrace{-\tilde{\eta}_{\text{vc}}\,\dot{s}}_{\text{“evolution”}} \;=\; \underbrace{s \max_{\tau\in[0,\,t]} \mathcal{D}(\tau)}_{\text{“driving force”}} \;-\; \underbrace{\left[(1-s) + \ell^2\,\Delta s\right]}_{\text{“geometric resistance”}} \quad . \tag{3.86}$$

with the scaled viscosity $\tilde{\eta}_{\text{vc}} := \eta_{\text{vc}}/(\mathcal{G}_{\text{c}}/\ell)$. The crack driving energy is also rescaled to reveal the dimensionless "crack driving state function" $\mathcal{D} := 2D_{\text{s}}(\boldsymbol{\varepsilon})/(\mathcal{G}_{\text{c}}/\ell)$. In this format, the phase field evolution equation is a dimensionless function of s only and reveals the "driving force" to be the decisive constitutive choice. Different proposals have been made, most guided by the introduction of a critical threshold, which must be exceeded

to obtain phase field evolution (Karma et al., 2001). While the "strain criterion without threshold" resembles the original formulation with

$$\mathcal{D} = \frac{2\psi_0(\boldsymbol{\varepsilon})}{\mathcal{G}_c/\ell} \tag{3.87}$$

also the "strain criterion with threshold"

$$\mathcal{D} = \left\langle \frac{\psi_0(\boldsymbol{\varepsilon})}{\psi_c} - 1 \right\rangle \tag{3.88}$$

is introduced, with ψ_c being a critical energy density. Based on the Legendre transformation $\tilde{\psi}_0(\tilde{\boldsymbol{\sigma}}) = \sup_{\boldsymbol{\varepsilon}}[\tilde{\boldsymbol{\sigma}} : \boldsymbol{\varepsilon} - \psi_0(\boldsymbol{\varepsilon})]$ in terms of the effective stress tensor (3.13), i.e. $\tilde{\boldsymbol{\sigma}} := \boldsymbol{\sigma}/s^2$, the "stress criteria with and without threshold"

$$\mathcal{D} = \frac{2\tilde{\psi}_0(\tilde{\boldsymbol{\sigma}})}{\mathcal{G}_c/\ell} \qquad \text{and} \qquad \mathcal{D} = \left\langle \frac{\tilde{\psi}_0(\tilde{\boldsymbol{\sigma}})}{\psi_c} - 1 \right\rangle \tag{3.89}$$

and a "principal tensile stress criterion with threshold"

$$\mathcal{D} = \zeta \left\langle \sum_{\alpha=1}^{3} \left(\frac{\langle \tilde{\sigma}_\alpha \rangle}{\sigma_c} \right)^2 - 1 \right\rangle \quad , \tag{3.90}$$

have been introduced by Miehe et al. (2015). The additional parameter $\zeta > 0$ controls the slope in the post-critical range while σ_c is a critical amount of stress. It is also noted, that a "pure Rankine-type criterion" can be constructed by

$$\mathcal{D} = \zeta \left\langle \frac{1}{\sigma_c^2} \max\left(\langle \tilde{\sigma}_1 \rangle^2,\ \langle \tilde{\sigma}_2 \rangle^2,\ \langle \tilde{\sigma}_3 \rangle^2 \right) - 1 \right\rangle \quad , \tag{3.91}$$

which is close to classical failure hypotheses, see Sect. 2.2.1. According to Miehe et al. (2015), aforementioned criteria can be understood as generalization of the phase field approach and resemble damage models, either by introducing a specific fracture energy per unit volume ψ_c or by defining a failure stress σ_c.

While the regularization controlled by the internal length parameter ℓ is still present due to the nonlocal term in (3.86), the threshold based criteria as proposed by Miehe et al. (2015) do not take the critical fracture energy $\mathcal{G}_c$ into account. To overcome this, the strain and stress criteria with threshold are reformulated by reintroducing $\psi_c = \mathcal{G}_c/(2\ell)$. In this case, the strain energy based criterion (3.88) can be rewritten as

$$\mathcal{D} = \frac{2}{\mathcal{G}_c/\ell} \left\langle \psi_0(\boldsymbol{\varepsilon}) - \frac{1}{2}\mathcal{G}_c/\ell \right\rangle \tag{3.92}$$

to fit into the original structure

$$\eta_{vc}\,\dot{s} + \frac{\mathrm{d}g(s)}{\mathrm{d}s} \left\langle \psi_0(\boldsymbol{\varepsilon}) - \frac{1}{2}\mathcal{G}_c/\ell \right\rangle - \mathcal{G}_c \left(\frac{1-s}{\ell} + \ell\,\Delta s \right) = 0 \quad . \tag{3.93}$$

By comparing the evolution equations (3.33) and (3.93) it is evident that by introducing the threshold value, the early phase field evolution $\mathrm{d}s < 0$, i.e. at small strains, is suppressed until "sufficient" energy is available to form the cracked state. In a more general version the phase field evolution equation can be constructed according to

$$\eta_{\mathrm{vc}}\,\dot{s} + \frac{\mathrm{d}g(s)}{\mathrm{d}s}\,\langle \psi_0(\boldsymbol{\varepsilon}) - \psi_{\mathrm{c}}\rangle + \frac{\mathcal{G}_{\mathrm{c}}}{c_{\mathrm{w}}}\left(\frac{1}{2\ell}\,\frac{\mathrm{d}w(s)}{\mathrm{d}s} - \ell\,\Delta s\right) = 0 \quad , \tag{3.94}$$

also used by Morand (2015) and Strobl and Seelig (2016).

3.4 Analysis of phase field transition

In order to discuss the impact of different degradation functions $g(s)$ and surface functions $w(s)$ on the behavior of the coupled problem, a symmetric, homogeneous, initially elastic one-dimensional problem of length $2L$ with applied displacements

$$u(x = \pm L) = u_0 \geq 0 \tag{3.95}$$

as depicted in Fig. 3.6 is analyzed thoroughly. In addition, basic properties of the phase field approach to fracture can be shown.

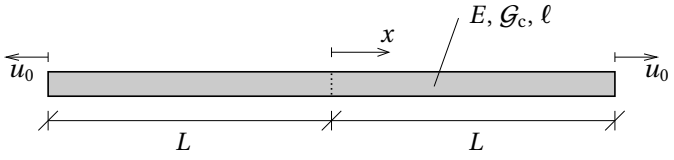


Figure 3.6: Uniaxial tension of 1D bar with symmetric boundary conditions.

The problem is kept simple by only regarding the quasi-static limit state (inertia and viscous damping are out of the scope of the present investigation) and neglecting traction and volume forces. The set of governing equations consists of the momentum balance, the constitutive relation and the phase field evolution equation

$$\sigma_{,x} = 0 \quad , \tag{3.96}$$

$$\sigma = g(s)\,E\,\varepsilon \quad , \tag{3.97}$$

$$0 = \frac{1}{2}\,g'(s)\,E\,\varepsilon^2 + \frac{\mathcal{G}_{\mathrm{c}}}{c_{\mathrm{w}}}\left(\frac{w'(s)}{2\ell} - \ell\,s_{,xx}\right) \quad , \tag{3.98}$$

the kinematic relation $\varepsilon(x) = u_{,x}$ and the natural boundary condition $s_{,x}\,(x = \pm L) = 0$. Here, $u_{,x}$, $\sigma_{,x}$, and $s_{,x}$ denote the first spatial derivatives of the displacement u, the stress σ the phase field s (with respect to the coordinate x) and $s_{,xx}$ denotes the second

spatial derivative of the phase field s. Note, that in general, neither the displacement field $u(x)$, nor the strain field $\varepsilon(x)$, nor the phase field $s(x)$ are homogeneous, i.e. they do not possess a constant solution independent of the location x. By contrast, the stress is always homogeneous due to the fulfillment of the momentum balance (3.96). The strain in the domain is related to the displacement imposed at the boundary $u(x=\pm L) = \pm u_0$ by

$$\pm\, u_0 = \int_0^{\pm L} \varepsilon(x)\, \mathrm{d}x \quad . \tag{3.99}$$

With the constitutive relation (3.97) and the introduction of $h(s) := 1/g(s) - 1$ (Wu, 2017) it is possible to rewrite (3.99) as

$$\pm\, u_0 = \int_0^{\pm L} \frac{1}{g(s)}\, \frac{\sigma}{E}\, \mathrm{d}x = \int_0^{\pm L} \big(1 + h(s)\big)\, \frac{\sigma}{E}\, \mathrm{d}x = \underbrace{\pm \frac{\sigma}{E}\, L}_{=:\, u_{\mathrm{elast}}} + \underbrace{\frac{\sigma}{E} \int_0^{\pm L} h(s)\, \mathrm{d}x}_{=:\, \pm \delta_{\mathrm{s}}/2} \quad . \tag{3.100}$$

The first summand can be identified to be the elastic response, while the remaining part depends on the phase field solution $s(x)$ and might be interpreted in terms of a *crack opening displacement* (cf. Sect. 2.2.7)

$$\delta_{\mathrm{s}} = [\![u]\!] = 2 \frac{\sigma}{E} \int_0^{L} h(s)\, \mathrm{d}x = 2 \frac{\sigma}{E} \int_0^{L} \left(\frac{1}{g(s)} - 1 \right) \mathrm{d}x \quad , \tag{3.101}$$

which corresponds to the displacement jump between crack flanks $[\![u]\!] := u^+ - u^-$. For fracture (causing $\sigma \to 0$) crack opening is determined by the applied displacements, i.e. $\delta_{\mathrm{s}} \to 2u_0$. In combination with the current stress σ it yields a "traction separation" behavior similar to that of cohesive zone models, see Sect. 2.2.7. Its specific behavior, however, depends on the choice of the degradation function $g(s)$ and the spatial formation of $s(x)$ during phase field evolution, also governed by the surface function $w(s)$.

The application of (3.96) on (3.97) reveals

$$\sigma_{,x} = \big(g(s)\, E\, \varepsilon\big)_{,x} = g'(s)\, s_{,x}\, E\, \varepsilon + g(s)\, E\, \varepsilon_{,x} = 0 \quad . \tag{3.102}$$

For $\varepsilon > 0$, it yields the relation between the spatial derivatives of strain and phase field

$$\varepsilon_{,x} = - \underbrace{\frac{g'(s)\, \varepsilon}{g(s)}}_{>\, 0}\, s_{,x} \quad . \tag{3.103}$$

Thus, a spatially decreasing phase field, e.g. when approaching a crack, is directly linked to a spatially increasing strain, independent of the choice of the degradation function $g(s)$.

3.4.1 Phase field in absence of mechanical forces

In absence of mechanical loads, i.e. $u(x=\pm L) = 0$, the phase field solution is governed by the evolution equation (3.98) only, which is

$$w'(s) - 2\ell^2 s_{,xx}(x) = 0 \quad . \tag{3.104}$$

Evidently, the solution does not depend on the choice of degradation function $g(s)$ and is independent of the scaling constant c_w. The solution of the perhaps nonlinear second-order ordinary differential equation (3.104) strongly depends on the choice of the local crack surface function $w(s)$. As already mentioned in the preceding section, typical choices of the local crack surface function are quadratic, so that the evolution equation (3.104) is linear.

Solution on an infinite domain

The "cracked" solution, usually referred to as *phase field profile*, represents the ultimate state of fracture. In an infinite domain ($L \to \infty$) the solution is obtained by solving (3.104) with the specific boundary conditions (assuming a crack at $x = 0$)

$$s(x=0) = 0 \qquad \text{and} \qquad s_{,x}(x=\pm\infty) = 0 \tag{3.105}$$

as done in App. B.6. The phase field "profile" for the quadratic single-well function (3.58) is given by

$$s(x) = 1 - \exp\left(-\frac{|x|}{\ell}\right) \tag{3.106}$$

and depicted in Fig. 3.7 for different choices of ℓ. For some other choices of β of the general quadratic surface function (3.57) solutions are given in the literature. As discussed, e.g., by Pham et al. (2011), the solution for the linear surface function consists of several pieces, i.e.

$$s(x) = \begin{cases} 1 & \text{for } x \le -2\ell \\ 1 - \left(1 - \dfrac{|x|}{2\ell}\right)^2 & \text{for } -2\ell \le x \le 2\ell \\ 1 & \text{for } x \ge 2\ell \quad . \end{cases} \tag{3.107}$$

The concave surface function presented by Wu (2017) also yields a piecewise solution given by

$$s(x) = \begin{cases} 1 & \text{for } x \le -(\pi/2)\,\ell \\ \sin\left(\dfrac{|x|}{\ell}\right) & \text{for } -(\pi/2)\,\ell \le x \le (\pi/2)\,\ell \\ 1 & \text{for } x \ge (\pi/2)\,\ell \quad . \end{cases} \tag{3.108}$$

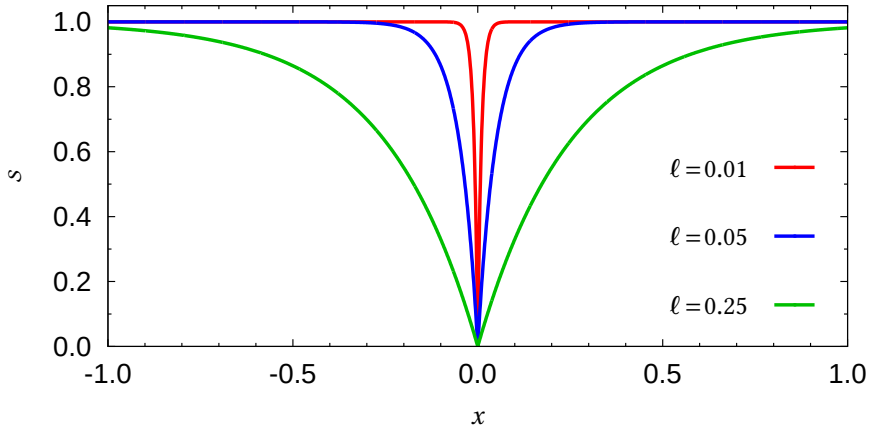


Figure 3.7: Phase field profile $s(x) = 1 - \exp(-|x|/\ell)$ valid in $x \in (-\infty, \infty)$ for varying ℓ.

Influence of finite domain size

To analyze the specific influence of a finite domain (i.e. free boundaries at $x = \pm L$), the phase field formulation based on the quadratic single-well function (3.58) is investigated. By applying the phase field boundary conditions $s(x=0) = 0$ and $s'(x=\pm L) = 0$, the specific solution for a crack at $x = 0$ is

$$s(x) = 1 - \cosh\left(\frac{|x|}{\ell}\right) \pm \tanh\left(\frac{L}{\ell}\right) \sinh\left(\frac{|x|}{\ell}\right) \tag{3.109}$$

as derived in App. B.6. In this context, the phase field value at the free boundary is, by using the relation $\cosh^2(x) - \sinh^2(x) = 1$,

$$s(x=\pm L) = 1 - \cosh\left(\frac{L}{\ell}\right) + \tanh\left(\frac{L}{\ell}\right) \sinh\left(\frac{L}{\ell}\right) = \frac{\cosh(L/\ell) - 1}{\cosh(L/\ell)} \quad . \tag{3.110}$$

For $\ell/L \to 0$, it approaches the undamaged state $s(x=\pm L) = 1$ and (3.109) approaches the well-known solution (3.106).

3.4.2 Homogeneous degradation of stiffness

For a number of specific approaches the spatially homogeneous problem has already been investigated in the literature. In conjunction with the single-well surface function, Pham et al. (2011) and Kuhn (2013) analyzed the quadratic degradation function. In addition, both Borden (2012) and Schlüter (2018) discussed a cubic degradation function with transition to the quadratic one. Kuhn et al. (2015) investigated the cubic and quartic degradation function, which often is used in models inspired by phase transition problems

(e.g. Aranson et al., 2000; Henry and Levine, 2004; Karma et al., 2001). The behavior based on other surface functions, e.g. the linear function, has been analyzed by Pham et al. (2011) and Wu (2017). In the following, general relations in the phase field approach are pointed out and used to discuss qualitative and quantitative features of different, also new approaches.

General discussion

Although not relevant from a physical point of view, the investigation of the spatially homogeneous phase field solution and the related degradation of the stress/stiffness is of practical importance. If the analysis starts from a homogeneous problem, e.g. without pre-cracks as it is the case in many initiation problems, the homogeneous behavior and succeeding localization are relevant.

Since the phase field is assumed to be spatially homogeneous, i.e. $s_{,x} = s_{,xx} = 0$, also the strain field must be homogeneous as a direct consequence of (3.103), i.e. $\varepsilon_{,x} = 0$. Relation (3.99) can then be simplified to $u_0 = L\,\varepsilon$. From (3.98) one obtains the strain based relation

$$g'(s)\,\frac{E\,\ell\,c_\mathrm{w}}{\mathcal{G}_\mathrm{c}}\,\varepsilon^2 + w'(s) = 0 \tag{3.111}$$

or, by inserting (3.97) in (3.98), the stress based relation

$$\frac{g'(s)}{\big(g(s)\big)^2}\,\frac{\ell\,c_\mathrm{w}}{\mathcal{G}_\mathrm{c}\,E}\,\sigma^2 + w'(s) = 0 \tag{3.112}$$

for the spatially homogeneous problem. The strain as function of the phase field value is

$$\varepsilon(s) = \sqrt{-\frac{w'(s)}{g'(s)}\,\frac{\mathcal{G}_\mathrm{c}}{E\,\ell\,c_\mathrm{w}}} = \sqrt{-\frac{w'(s)}{\bar{E}'(s)}\,\frac{\mathcal{G}_\mathrm{c}}{\ell\,c_\mathrm{w}}} \quad , \tag{3.113}$$

where the material stiffness function $\bar{E}(s) := g(s)\,E$ and the related change caused by the phase field $\bar{E}'(s) = g'(s)\,E$ are used. By using (B.2) the stress can be written as a function of the phase field value

$$\sigma(s) = \sqrt{-\big(g(s)\big)^2\,\frac{w'(s)}{g'(s)}\,\frac{\mathcal{G}_\mathrm{c}\,E}{\ell\,c_\mathrm{w}}} = \sqrt{\frac{w'(s)}{\big(1/g(s)\big)'}\,\frac{\mathcal{G}_\mathrm{c}\,E}{\ell\,c_\mathrm{w}}} = \sqrt{\frac{w'(s)}{\bar{S}'(s)}\,\frac{\mathcal{G}_\mathrm{c}}{\ell\,c_\mathrm{w}}} \quad . \tag{3.114}$$

Inspired by Pham and Marigo (2009), the material compliance function $\bar{S}(s) := 1/\big(g(s)\,E\big)$ is introduced. Its change due to phase field evolution is

$$\bar{S}'(s) := \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{\bar{E}(s)}\right) = \frac{1}{E}\,\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{g(s)}\right) = \frac{1}{E}\left(\frac{1}{g(s)}\right)' \quad , \tag{3.115}$$

as already used in (3.114). Evidently, (3.113) and (3.114) depend on the specific choice of the degradation function $g(s)$ and the surface function $w(s)$. More specifically, both relations reveal their functional dependency with the dimensionless expressions

$$f_\varepsilon(s) := -\frac{w'(s)}{g'(s)} \quad , \tag{3.116}$$

$$f_\sigma(s) := \frac{w'(s)}{\big(1/g(s)\big)'} = -\big(g(s)\big)^2 \, \frac{w'(s)}{g'(s)} \quad , \tag{3.117}$$

which are relevant for the further investigation. In addition, by using (3.116), it is

$$f_\sigma(s) = \big(g(s)\big)^2 \, f_\varepsilon(s) \quad . \tag{3.118}$$

With

$$f'_\varepsilon(s) := \frac{\mathrm{d}f_\varepsilon(s)}{\mathrm{d}s} = \frac{w'(s)\, g''(s) - w''(s)\, g'(s)}{\big(g'(s)\big)^2} \quad , \tag{3.119}$$

the derivative of (3.113) with respect to s is

$$\begin{aligned} \varepsilon'(s) &:= \frac{\mathrm{d}\varepsilon(s)}{\mathrm{d}s} = \frac{1}{2} \frac{f'_\varepsilon(s)}{\sqrt{f_\varepsilon(s)}} \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell\,c_\mathrm{w}}} \\ &= \frac{1}{2} \frac{w'(s)\, g''(s) - w''(s)\, g'(s)}{\big(g'(s)\big)^2} \sqrt{-\frac{g'(s)}{w'(s)}} \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell\,c_\mathrm{w}}} \quad . \end{aligned} \tag{3.120}$$

Similarly, the derivative of (3.114) is

$$\sigma'(s) := \frac{\mathrm{d}\sigma(s)}{\mathrm{d}s} = \frac{1}{2} \frac{f'_\sigma(s)}{\sqrt{f_\sigma(s)}} \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell\, c_\mathrm{w}}} = \frac{1}{2} \frac{2\, f_\varepsilon(s) + g(s)\, f'_\varepsilon(s)}{\sqrt{f_\varepsilon(s)}} \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell\, c_\mathrm{w}}} \quad , \tag{3.121}$$

with

$$\begin{aligned} f'_\sigma(s) :=& \frac{\mathrm{d}f_\sigma(s)}{\mathrm{d}s} = -2g(s)\, \frac{w'(s)}{g'(s)} + \big(g(s)\big)^2 \, \frac{w'(s)\, g''(s) - w''(s)\, g'(s)}{\big(g'(s)\big)^2} \\ =& -2g(s)\, \frac{w'(s)}{g'(s)} + \big(g(s)\big)^2 \, \frac{w'(s)\, g''(s)}{\big(g'(s)\big)^2} - \big(g(s)\big)^2 \, \frac{w''(s)}{g'(s)} \quad . \end{aligned} \tag{3.122}$$

Alternatively, by using relation (3.118), it is

$$\sigma'(s) = 2g\, g'(s)\, f_\varepsilon(s) + \big(g(s)\big)^2 \, f'_\varepsilon(s) \quad . \tag{3.123}$$

By directly using the constitutive relation (3.97), it is

$$\sigma'(s) := \frac{\mathrm{d}\sigma(s)}{\mathrm{d}s} = g'(s)\, E\, \varepsilon(s) + g(s)\, E\, \varepsilon'(s) \tag{3.124}$$

and thus, the stress-strain relation is governed by

$$\frac{\mathrm{d}\sigma(s)}{\mathrm{d}\varepsilon} = \frac{\mathrm{d}\sigma(s)}{\mathrm{d}s}\,\frac{\mathrm{d}s}{\mathrm{d}\varepsilon} = \left[\frac{g'(s)\,\varepsilon(s)}{\varepsilon'(s)} + g(s)\right] E \quad . \tag{3.125}$$

With the above relations at hand, it is possible to discuss some general requirements on $g(s)$ and $w(s)$ and the impact of their specific choice on the behavior of the modeling approach. All points are discussed only for the range of interest, i.e. $s \in (0,\,1)$, and both limiting states $s=1$ and $s=0$. Since some function values are undefined or indeterminate, l'Hôspital's rule[6] is utilized to evaluate the function. In particular, if $g(s=1) \to 0$, the limit state of $w'(s)/g'(s)$ is determined by the limit state of $w''(s)/g''(s)$ as $s \to 0$. The most important facts are:

- If $g(s)$ and $w(s)$ are polynomial functions, (3.116) and (3.117) are rational functions with $g'(s)$ being the denominator. To obtain a continuous transition from the limit state $s=1$ to $s \to 0$, the denominator, i.e. the slope of the degradation function, must not vanish. In fact, $g'(s) > 0$ was already a requirement according to (3.66) for exactly this reason.
- In the case that a quadratic single-well function is combined with a degradation function with $g'(s=1) \neq 0$, the phase field starts to decrease from the limit strain and stress zero.
- By contrast, for a combination of the single-well function with a degradation function with $g'(s=1) = 0$, the strain and stress limits of the intact state ($s=1$) are determined by $w''(s)/g''(s)$, which yields a finite value if $g''(s=1) \neq 0$. This is an additional requirement not given in Sect. 3.3.2 (but already added in Fig. 3.4a).
- While the dimensionless stress function (3.117) approaches zero as $s \to 0$, the dimensionless strain function continuously increases to infinity if $w'(s=0) \neq 0$.
- Depending on the choice of $g(s)$ and $w(s)$, the stress function either possess a maximum at $s < 1$ or reaches its maximum at the limit state $s = 1$. In the first case, which is relevant if $f'_\sigma(s \to 1^-) < 0$, the maximum stress in the homogeneous problem is obtained from

$$\sigma_\mathrm{t} = \sup_{s\in(0,\,1]} \sqrt{\frac{w'(s)}{\big(1/g(s)\big)'}\,\frac{\mathcal{G}_\mathrm{c}\,E}{\ell\,c_\mathrm{w}}} \quad , \tag{3.126}$$

  see, e.g., Pham et al. (2011).

[6] Although the rule was published in Guillaume de l'Hôspital's book about differential calculus (De L'Hôpital, 1696) the rule was introduced to him by Johann Bernoulli in 1694 (Sonar, 2016).

In the following, some relevant combinations of degradation functions and surface functions are analyzed by using the above relations. Figure 3.8 shows the functions $f_\varepsilon(s)$ and $f_\sigma(s)$ for specific combinations of the polynomial degradation function $g(s)$ with the surface function $w(s)$.

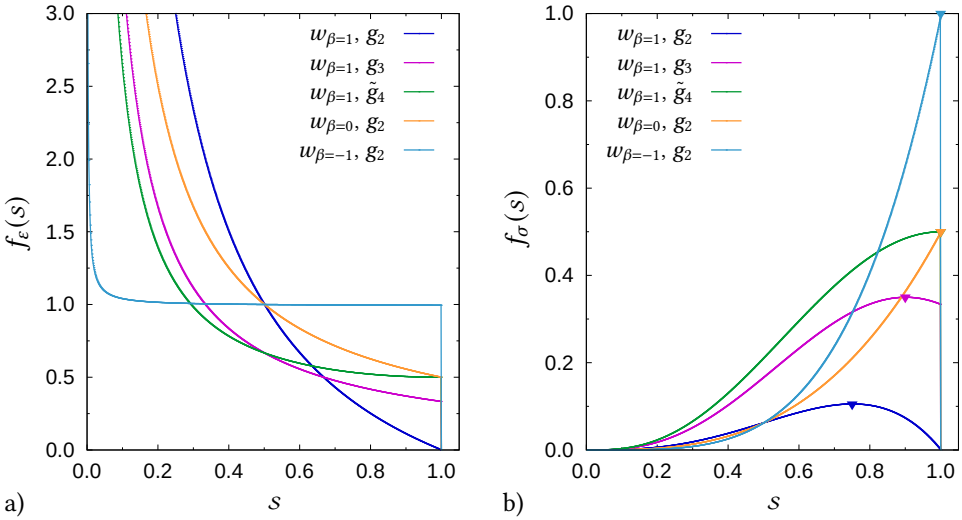


Figure 3.8: Qualitatively different relations for different combinations of $g(s)$ and $w(s)$ between a) strain and phase field evolution, represented by the function $f_\varepsilon(s)$ and b) stress and phase field evolution, represented by the function $f_\sigma(s)$.

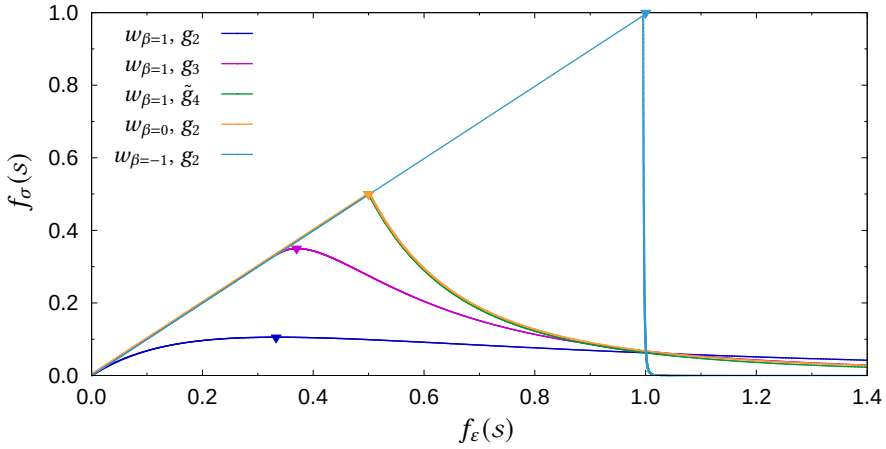


Figure 3.9: Qualitatively different stress-strain relations for different combinations of $g(s)$ and $w(s)$.

By combining $f_\varepsilon(s)$ and $f_\sigma(s)$ the stress-strain relation of the homogeneous phase field problem can be obtained as depicted in Fig. 3.9 for the same combinations of $g(s)$ and $w(s)$. All combinations exhibit two loading regimes: Starting from the unloaded state $f_\sigma = f_\varepsilon = 0$, both stress and strain functions increase while, after exceeding a specific state strain softening takes place, i.e. the stress decreases for increasing strain. By regarding the stress, strain and phase field graphs in Figs. 3.8 and 3.9, it is evident that the choice of $g(s)$ has a strong impact on the approach and features qualitatively different behavior, but must always be discussed in conjunction with the choice of $w(s)$.

Single-well surface function in combination with quadratic degradation

At first, the combination of the single-well surface function (3.58) with the quadratic degradation function (3.70) is investigated. By using the strain and stress relations (3.116) and (3.117) and their derivatives (3.119) and (3.122), respectively, the corresponding limit states are

$$\lim_{s\to 0^+} f_\varepsilon(s) = \infty\,, \quad \lim_{s\to 0^+} f'_\varepsilon(s) = -\infty\,, \quad \lim_{s\to 1^-} f_\varepsilon(s) = 0\,, \quad \lim_{s\to 1^-} f'_\varepsilon(s) = -1\,, \tag{3.127}$$

$$\lim_{s\to 0^+} f_\sigma(s) = 0\,, \quad \lim_{s\to 0^+} f'_\sigma(s) = 0\,, \quad \lim_{s\to 1^-} f_\sigma(s) = 0\,, \quad \lim_{s\to 1^-} f'_\sigma(s) = -1 \quad . \tag{3.128}$$

Due to (3.127_3), phase field evolution $\mathrm{d}s < 0$ starts already at a minimum strain, as also depicted in Fig. 3.8a (indicated by "$w_{\beta=1}$, g_2"). As a consequence, the stress-strain response is not linear, see Fig. 3.9, as it is supposed to be in brittle materials. This is also indicated by (3.128_3), i.e. phase field evolution starts from a minimum stress. As depicted in Fig. 3.8b, the stress reaches a maximum at the specific value $0 < s_\mathrm{t} < 1$, which is obtained by solving (3.126). In particular, either by setting (3.122) or (3.125) equal to zero, the stress maximum is obtained at

$$s_\mathrm{t} = \frac{3}{4} \quad . \tag{3.129}$$

Inserting this result into (3.114) and (3.113) one obtains the maximum stress

$$\sigma_\mathrm{t} := \sigma(s_\mathrm{t}) = \underbrace{\sqrt{\frac{27}{256}}}_{:= f_\mathrm{gw}} \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell}} \approx 0.325 \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell}} \quad . \tag{3.130}$$

In this regard, f_gw quantifies the relation between σ_t and ℓ, and the strain at maximum stress

$$\varepsilon_\mathrm{t} := \varepsilon(s_\mathrm{t}) = \sqrt{\frac{1}{3}} \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \approx 0.577 \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \quad . \tag{3.131}$$

Note, that (3.131) is significantly (> 75 %) larger than the (linear) elastic strain would be. As shown in Fig. 3.9 the stress slowly decreases in the post-peak range.

To return to the original question of the 1D problem (Fig. 3.6) the homogeneous phase field response for a given displacement $u_0 > 0$ is

$$s = \frac{\mathcal{G}_c}{\mathcal{G}_c + E\,\ell\,\varepsilon^2} = \frac{\mathcal{G}_c\,L^2}{\mathcal{G}_c\,L^2 + E\,\ell\,u_0^2} \quad . \tag{3.132}$$

and obtained from (3.113). It decreases strictly monotonically with increasing displacement. Combination of (3.132) and (3.97) yields the corresponding stress

$$\sigma = \left(\frac{\mathcal{G}_c}{\mathcal{G}_c + E\,\ell\,\varepsilon^2}\right)^2 E\,\varepsilon = \left(\frac{\mathcal{G}_c\,L^2}{\mathcal{G}_c\,L^2 + E\,\ell\,u_0^2}\right)^2 \frac{E}{L}\,u_0 \quad , \tag{3.133}$$

which increases up to the maximum stress (3.130) at (3.131), i.e.

$$u_t := u_0(\sigma_t) = \sqrt{\frac{1}{3}}\,\sqrt{\frac{\mathcal{G}_c}{E\,\ell}}\,L \quad . \tag{3.134}$$

Single-well surface function in combination with higher-order degradation

The limit states of (3.116), (3.117), (3.119) and (3.122) for the combination of the single-well surface function (3.58) with a higher order degradation function (3.76), for simplicity $a_s = 0$ is assumed, are

$$\lim_{s\to 0^+} f_\varepsilon(s) = \infty\,, \quad \lim_{s\to 0^+} f'_\varepsilon(s) = -\infty\,, \quad \lim_{s\to 0^+} f_\sigma(s) = 0\,, \quad \lim_{s\to 0^+} f'_\sigma(s) = 0\,, \tag{3.135}$$

$$\lim_{s\to 1^-} f_\varepsilon(s) = \frac{2}{n\,(n-1)}\,, \qquad \lim_{s\to 1^-} f'_\varepsilon(s) = -\frac{2(n-2)}{n\,(n-1)}\,, \tag{3.136}$$

$$\lim_{s\to 1^-} f_\sigma(s) = \frac{2}{n\,(n-1)}\,, \qquad \lim_{s\to 1^-} f'_\sigma(s) = -\frac{2(n-2)}{n\,(n-1)} \quad . \tag{3.137}$$

By contrast to the quadratic degradation function, the strain and stress limit states (3.136_1) and (3.137_1) are of finite size. Consequently, phase field evolution $\mathrm{d}s < 0$ does not start before a specific strain $\varepsilon_{\text{lin}} > 0$ and stress $\sigma_{\text{lin}} > 0$ are exceeded. This is also visible in Figs. 3.8a and b (indicated by "$w_{\beta=1},\ g_3$"), where, for illustration purposes, a cubic degradation function ($n = 3$) has been used.

Due to the specific choice of the surface function and the way how the degradation function is constructed (see App. B.5), the strain relation (3.114) has a quite simple structure

$$\varepsilon(s) = \sqrt{\frac{2(1-s)}{s^{n-2}\,(1-s)\,n\,(n-1)}}\,\sqrt{\frac{\mathcal{G}_c}{E\,\ell}} = \sqrt{\frac{1}{s^{n-2}}\,\frac{2}{n\,(n-1)}}\,\sqrt{\frac{\mathcal{G}_c}{E\,\ell}} \tag{3.138}$$

and thus, is easily invertible. The strain increases monotonically as the phase field decreases, see also Fig. 3.8a. Obviously, the original function (3.138) has multiple roots. Thus, the homogeneous phase field solution for a given displacement $u_0 > 0$ is

$$s = \begin{cases} 1 & \text{for} \quad u_0 \leq \sqrt{\dfrac{2}{n\,(n-1)}}\,\sqrt{\dfrac{\mathcal{G}_\mathrm{c}}{E\,\ell}}\,L \\ \left[\dfrac{2}{n\,(n-1)}\,\dfrac{1}{u_0^2}\,\dfrac{\mathcal{G}_\mathrm{c}\,L^2}{E\,\ell}\right]^{1/(n-2)} & \text{for} \quad u_0 > \sqrt{\dfrac{2}{n\,(n-1)}}\,\sqrt{\dfrac{\mathcal{G}_\mathrm{c}}{E\,\ell}}\,L \end{cases} \quad . \tag{3.139}$$

The first solution ($s = 1$) governs the elastic range for $\varepsilon < \varepsilon_\mathrm{lin}$, see the linear stress-strain dependency in Fig. 3.9 starting at zero. While this solution is always valid from a mathematical point of view, the second solution is invalid for small strains ($\varepsilon < \varepsilon_\mathrm{lin}$), since it would result in $s >1$, but becomes preferable for $\varepsilon > \varepsilon_\mathrm{lin}$ from an energetic point of view. The specific strain and stress, where phase field evolution starts and the stress-strain relation becomes non-linear is given by the intersection point of the two branches (3.139), i.e.

$$\varepsilon_\mathrm{lin} = \sqrt{\frac{2}{n\,(n-1)}}\,\sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \qquad \text{and} \qquad \sigma_\mathrm{lin} = E\,\varepsilon_\mathrm{lin} = \sqrt{\frac{2}{n\,(n-1)}}\,\sqrt{\frac{\mathcal{G}_\mathrm{c}\,E}{\ell}} \quad . \tag{3.140}$$

Since the limit (3.137_2) is positive, the stress increases as the phase field starts to evolve from $s = 1$ until the stress reaches its maximum (Fig. 3.8b). In this state, the stress-strain response is also non-linear. By solving (3.126), either by setting (3.122) or (3.125) equal to zero, the stress maximum is obtained at

$$s_\mathrm{t} = \frac{n^2}{n^2 + n - 2} \quad . \tag{3.141}$$

Inserting this result into (3.114) and (3.113) one obtains the maximum stress

$$\sigma_\mathrm{t} = \underbrace{\sqrt{\frac{8n}{(n-1)\,(n+2)^2}\left(\frac{n^2}{n^2+n-2}\right)^n}}_{= f_\mathrm{gw}}\,\sqrt{\frac{\mathcal{G}_\mathrm{c}\,E}{\ell}} \tag{3.142}$$

and the strain at maximum stress

$$\varepsilon_\mathrm{t} = \sqrt{\frac{2n^3}{n^{2n}}\,\frac{(n^2+n-2)^{n-2}}{(n-1)}}\,\sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \quad . \tag{3.143}$$

For the cubic (third order) degradation function (3.71), for example, the phase field value at maximum stress, the maximum stress and corresponding strain are

$$s_\mathrm{t} = \frac{9}{10} \tag{3.144}$$

$$\sigma_\mathrm{t} = \sqrt{\frac{2187}{6250}} \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell}} \approx 0.592 \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell}} \tag{3.145}$$

$$\varepsilon_\mathrm{t} = \sqrt{\frac{10}{27}} \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \approx 0.609 \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \quad , \tag{3.146}$$

while for the quartic (fourth order) degradation function (3.74) the phase field value at maximum stress, the maximum stress and corresponding strain are

$$s_\mathrm{t} = \frac{8}{9} \tag{3.147}$$

$$\sigma_\mathrm{t} = \sqrt{\frac{32768}{177147}} \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell}} \approx 0.430 \sqrt{\frac{\mathcal{G}_\mathrm{c}\, E}{\ell}} \tag{3.148}$$

$$\varepsilon_\mathrm{t} = \sqrt{\frac{27}{128}} \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \approx 0.459 \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,\ell}} \quad . \tag{3.149}$$

As a measure of non-linear behavior prior to the peak stress, the strain at the peak stress is compared with the (linear) elastic solution. In case of the cubic degradation function (3.146), the strain is about 3 % larger and in case of the quartic degradation function (3.149) it is 7 % larger. In both cases, this is far below the > 75 % obtained for the quadratic degradation function. It should be noted, that it is possible to construct other polynomial degradation functions so that $f'_\sigma(s) \geq 0$. In this case, the maximum stress is obtained for $s_\mathrm{t} = 1$. One specific example with $f'_\sigma(s) = 0$ is the 4th-order degradation function

$$g(s) = s^4 - 4s^3 + 4s^2 \quad , \tag{3.150}$$

for further details, see App. B.5. The impact of this degradation function on strain and stress relations is shown in Figs. 3.8 and 3.9 (indicated by "$w_{\beta=1}$, $\tilde{g}_4$").

The function $\sigma(s)$ is illustrated in Fig. 3.10a for different orders n. In addition, the peak stress (3.142) is illustrated for various orders of the degradation function in Fig. 3.10b. For increasing order n the maximum stress decreases. As shown in Fig. 3.9 the stress slowly decreases in the post-peak range. By taking (3.135–3.137) into account, the limit states of (3.125) at the branching point are equal, i.e.

$$\lim_{s\to 1^+} \frac{\mathrm{d}\sigma(s)}{\mathrm{d}\varepsilon} = E \quad , \qquad \lim_{s\to 1^-} \frac{\mathrm{d}\sigma(s)}{\mathrm{d}\varepsilon} = \lim_{s\to 1^-} \Big[\underbrace{\frac{1}{\varepsilon'(s)}}_{<0} \underbrace{g'(s)}_{\to 0} \underbrace{\varepsilon(s)}_{>0} + \underbrace{g(s)}_{=1} \Big] E = E \quad , \tag{3.151}$$

and the transition is smooth. From a numerical point of view, this might be a drawback since it is not possible to distinguish both solutions at the bifurcation point. This can

result in delayed or even suppressed bifurcation. For this purpose, the parameter a_s (3.73) and the requirement (3.69) have been introduced by Borden et al. (2012). As an alternative, a small perturbation on the numerical solution can be applied as discussed by Kuhn et al. (2015).

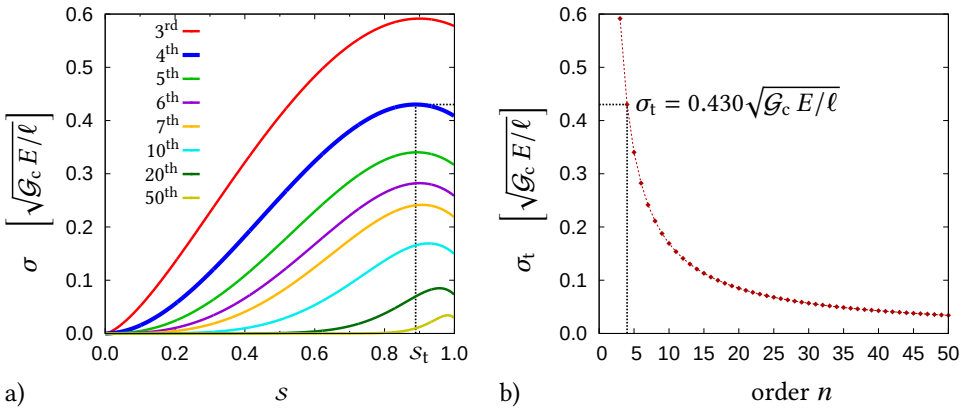


Figure 3.10: a) Function $\sigma(s) := \sqrt{f_\sigma(s)}\,\sqrt{\mathcal{G}_c\, E/\ell}$ for polynomial degradation functions of different order n, b) maximum stress of the spatially homogeneous solution σ_t depending on order $n > 3$ of degradation function $g(s)$. The particular solution for the 4th order degradation function utilized in Sect. 4.3 is indicated.

Linear surface function in combination with quadratic degradation

In the following, the combination of the linear surface function (3.58) with the quadratic degradation function (3.70) is briefly discussed in order to highlight qualitatively similar and distinct features. Due to the choice of the linear surface function the scaling constant according to (3.61) is $c_w = 4/3$. The relevant limit states of (3.116), (3.117), (3.119) and (3.122) are

$$\lim_{s\to 0^+} f_\varepsilon(s) = \infty\,, \quad \lim_{s\to 0^+} f'_\varepsilon(s) = -\infty\,, \quad \lim_{s\to 1^-} f_\varepsilon(s) = 1/2\,, \quad \lim_{s\to 1^-} f'_\varepsilon(s) = -1/2\,, \tag{3.152}$$

$$\lim_{s\to 0^+} f_\sigma(s) = 0\,, \quad \lim_{s\to 0^+} f'_\sigma(s) = 0\,, \quad \lim_{s\to 1^-} f_\sigma(s) = 1/2\,, \quad \lim_{s\to 1^-} f'_\sigma(s) = 3/2\quad. \tag{3.153}$$

Qualitatively similar to the previous case, phase field evolution does not start before exceeding a specific strain $\varepsilon_{lin} > 0$ or stress $\sigma_{lin} > 0$ as indicated by (3.152_3) and (3.153_3), respectively. This is also visible in Figs. 3.8a and b (indicated by "$w_{\beta=0}$, g_2"). However, with (3.153_4) being positive, the maximum stress is obtained for $s = 1$. As a consequence, the transition from the linear range is not smooth any more (Fig. 3.9). From the strain relation (3.114)

$$\varepsilon(s) = \sqrt{\frac{1}{2s}}\,\sqrt{\frac{\mathcal{G}_c}{E\,\ell\,c_w}} \tag{3.154}$$

the solution

$$s = \begin{cases} 1 & \text{for} \quad u_0 \le \sqrt{\frac{3}{8}} \sqrt{\frac{\mathcal{G}_c}{E\,\ell}}\, L \\ \frac{3}{8}\,\frac{1}{u_0^2}\,\frac{\mathcal{G}_c\, L^2}{E\,\ell} & \text{for} \quad u_0 > \sqrt{\frac{3}{8}} \sqrt{\frac{\mathcal{G}_c}{E\,\ell}}\, L \end{cases} \tag{3.155}$$

is obtained. Note, that the linear solution $s = 1$ is not obtained by the phase field approach but has to be enforced by appropriate solution techniques, see the corresponding remarks in Sect. 3.3.1. The characteristic parameters of this approach are $s_t = 1$,

$$\varepsilon_t = \varepsilon_{lin} = \sqrt{\frac{3}{8}} \sqrt{\frac{\mathcal{G}_c}{E\,\ell}} \approx 0.4330 \sqrt{\frac{\mathcal{G}_c}{E\,\ell}} \tag{3.156}$$

and

$$\sigma_t = \sigma_{lin} = \underbrace{\sqrt{\frac{3}{8}}}_{= f_{gw}} \sqrt{\frac{\mathcal{G}_c\, E}{\ell}} \approx 0.4330 \sqrt{\frac{\mathcal{G}_c\, E}{\ell}} \quad . \tag{3.157}$$

For comparison, the (almost) linear surface function (3.63), where $w'(s{=}1) = 0$ is incorporated by a penalization term, is investigated in the same way. Due to the penalization with an exponential function, the solution can only be expressed in terms of the Lambert W function[7]. Because of this, and to compare the resulting critical values, a penalty factor $\epsilon_w = 100$ is chosen. Note, that due to the slight deviation of the phase field approximation the scaling parameter $c_w = 1.3211$ (see Tab. B.1) must be considered. The limit values are

$$\lim_{s\to 0^+} f_\varepsilon(s) = \infty\,, \quad \lim_{s\to 0^+} f'_\varepsilon(s) = -\infty\,, \quad \lim_{s\to 1^-} f_\varepsilon(s) = 0\,, \quad \lim_{s\to 1^-} f'_\varepsilon(s) = -\epsilon_w/2\,, \tag{3.158}$$

$$\lim_{s\to 0^+} f_\sigma(s) = 0\,, \quad \lim_{s\to 0^+} f'_\sigma(s) = 0\,, \quad \lim_{s\to 1^-} f_\sigma(s) = 0\,, \quad \lim_{s\to 1^-} f'_\sigma(s) = -\epsilon_w/2\,. \tag{3.159}$$

The phase field value of the peak stress is

$$s_t = \frac{1}{\epsilon_w}\left[W\big(3\exp(3)\exp(\epsilon_w)\big) - 3\right] \overset{\epsilon_w=100}{\approx} 0.965 \quad . \tag{3.160}$$

With this value, both the strain at the peak stress

$$\varepsilon_t = \frac{\epsilon_w}{2W\big(3\exp(\epsilon_w + 3)\big)} \sqrt{\frac{\mathcal{G}_c}{1.3211\, E\,\ell}} \overset{\epsilon_w=100}{\approx} 0.4376 \sqrt{\frac{\mathcal{G}_c}{E\,\ell}} \tag{3.161}$$

[7] The function, named after Johann Heinrich Lambert, cannot be expressed in terms of elementary functions and possess the important property $z = W(z)\,\exp\big(W(z)\big) = W\big(z\,\exp(z)\big)$ with $z \in \mathbb{C}$.

and the peak stress

$$\sigma_{\mathrm{t}} = -\frac{\left[W\left(3\exp(\epsilon_{\mathrm{w}}+3)\right)-3\right]^3}{2\epsilon_{\mathrm{w}}^3}\left(\frac{3}{W\left(3\exp(\epsilon_{\mathrm{w}}+3)\right)}-1\right)\sqrt{\frac{\mathcal{G}_{\mathrm{c}}\,E}{1.3211\,\ell}}$$
$$\overset{\epsilon_{\mathrm{w}}=100}{\approx} 0.4028\sqrt{\frac{\mathcal{G}_{\mathrm{c}}\,E}{\ell}} \tag{3.162}$$

are obtained. Note the slight deviation from the original formulation (3.156) and (3.157).

Concave surface function in combination with quadratic degradation

In a similar manner, the combination of the concave surface function (3.60), here with $\beta \to -1$, with the quadratic degradation function (3.70) is investigated. With the specific choice of the concave surface function the scaling constant is $c_{\mathrm{w}} = \pi/2$. From (3.116), (3.117), (3.119) and (3.122) the relevant limit states are

$$\lim_{s\to 0^+} f_\varepsilon(s) = 1\,, \quad \lim_{s\to 0^+} f_\varepsilon'(s) = 0\,, \quad \lim_{s\to 1^-} f_\varepsilon(s) = 1\,, \quad \lim_{s\to 1^-} f_\varepsilon'(s) = 0\,, \tag{3.163}$$

$$\lim_{s\to 0^+} f_\sigma(s) = 0\,, \quad \lim_{s\to 0^+} f_\sigma'(s) = 0\,, \quad \lim_{s\to 1^-} f_\sigma(s) = 1\,, \quad \lim_{s\to 1^-} f_\sigma'(s) = 4 \quad . \tag{3.164}$$

The model's overall behavior regarding the linear behavior up to a specific strain $\varepsilon_{\mathrm{lin}} > 0$ or stress $\sigma_{\mathrm{lin}} > 0$ is qualitatively similar to the one based on the linear surface function, see also Figs. 3.8a, b and 3.9 (indicated by "$w_{\beta=0}$, g_2"). It also exhibits a non-smooth transition to the post-peak range. However, for $\beta = -1$ the concave surface function yields a constant strain

$$\varepsilon(s) = \sqrt{\frac{-2s}{2s}}\sqrt{\frac{\mathcal{G}_{\mathrm{c}}}{E\,\ell\,c_{\mathrm{w}}}} \tag{3.165}$$

as depicted in Fig. 3.8a. The critical quantities in this approach are $s_{\mathrm{t}} = 1$,

$$\varepsilon_{\mathrm{t}} = \varepsilon_{\mathrm{lin}} = \sqrt{\frac{\mathcal{G}_{\mathrm{c}}}{\pi/2\,E\,\ell}} \tag{3.166}$$

and

$$\sigma_{\mathrm{t}} = \sigma_{\mathrm{lin}} = \sqrt{\frac{\mathcal{G}_{\mathrm{c}}\,E}{\pi/2\,\ell}} = \underbrace{\sqrt{\frac{2}{\pi}}}_{=f_{\mathrm{gw}}}\sqrt{\frac{\mathcal{G}_{\mathrm{c}}\,E}{\ell}} \approx 0.7979\sqrt{\frac{\mathcal{G}_{\mathrm{c}}\,E}{\ell}} \quad . \tag{3.167}$$

In conjunction with the smoothly decreasing stress $\sigma(s)$ as the phase field value decreases, this yields an almost vertical drop in the stress-strain curve (already for the spatially homogeneous problem!), see Fig. 3.9.

Quadratic single-well function in combination with quadratic degradation and threshold

The phase field evolution equation (3.94) including an energetic threshold ψ_c is investigated in combination with the single-well surface function (3.58) and the quadratic degradation function (3.70). It is assumed, that the threshold value scales with

$$\psi_c := \alpha_c \frac{\mathcal{G}_c}{\ell} \quad , \tag{3.168}$$

with $\alpha_c \geq 0$ being a dimensionless parameter. The homogeneous phase field response of the 1D problem (Fig. 3.6) for a given displacement $u_0 > 0$ is

$$s = \begin{cases} 1 & \text{for} \quad u_0 \leq \sqrt{2\alpha_c \dfrac{\mathcal{G}_c}{E\,\ell}}\, L \\ \dfrac{\mathcal{G}_c\, L^2}{\mathcal{G}_c(1-2\,\alpha_c)\, L^2 + E\,\ell\, u_0^2} & \text{for} \quad u_0 > \sqrt{2\alpha_c \dfrac{\mathcal{G}_c}{E\,\ell}}\, L \end{cases} \quad . \tag{3.169}$$

After the threshold is exceeded, i.e. $E\,\varepsilon^2/2 > \psi_c$, the phase field decreases strictly monotonically with increasing displacement. Thus, the stress-strain relation is linear up to

$$\varepsilon_{\text{lin}} = \sqrt{2\alpha_c \frac{\mathcal{G}_c}{E\,\ell}} \quad . \tag{3.170}$$

The stress for the homogeneous phase field response is

$$\sigma(\varepsilon) = \left(\frac{\mathcal{G}_c}{\mathcal{G}_c + \langle E\,\ell\,\varepsilon^2 - 2\alpha_c\,\mathcal{G}_c \rangle} \right)^2 E\,\varepsilon \quad . \tag{3.171}$$

For $\alpha_c \geq 1/8$, which is equal to

$$\psi_c \geq \frac{1}{8} \frac{\mathcal{G}_c}{\ell} \quad , \tag{3.172}$$

the stress-strain relation is linear up to the maximum at (3.170), for $0 < \alpha_c < 1/8$ there is a linear and non-linear part. The stress maximum is

$$\sigma_t = \begin{cases} \dfrac{\sqrt{27 \dfrac{\mathcal{G}_c\, E}{\ell}\,(1-2\alpha_c)}}{16(1-2\alpha_c)^2} & \text{at } \varepsilon_t = \sqrt{\dfrac{\mathcal{G}_c}{3E\,\ell}\,(1-2\alpha_c)} \quad \text{for } 0 \leq \alpha_c < 1/8 \\ \sqrt{2\alpha_c \dfrac{\mathcal{G}_c E}{\ell}} & \text{at } \varepsilon_t = \sqrt{2\alpha_c \dfrac{\mathcal{G}_c}{E\,\ell}} \quad \text{for } \alpha_c \geq 1/8 \end{cases} \quad . \tag{3.173}$$

To obtain a linear response it is

$$f_{\text{gw}} = \sqrt{2\alpha_c} \geq 0.5 \quad . \tag{3.174}$$

From an energetic point of view, it should be noted that the strain energy in an elastic bar of length L and cross sectional area A at failure (which here is assumed at maximum strain) is

$$E_\mathrm{e} = L\,A \int_0^{\varepsilon_\mathrm{t}} \sigma(\varepsilon)\,\mathrm{d}\varepsilon \qquad \text{and for } \alpha_\mathrm{c} \geq \frac{1}{8} \text{ it is} \qquad E_\mathrm{e} = L\,A\,\frac{\sigma_\mathrm{t}^2}{2E} \quad , \tag{3.175}$$

while the surface energy is $E_\mathrm{s} = \mathcal{G}_\mathrm{c}\,A$. For $\alpha_\mathrm{c} > 1/8$, the (quasi-static) energetic requirement $E_\mathrm{e} \geq E_\mathrm{s}$ is fulfilled if $\ell \leq \alpha_\mathrm{c}\,L$, which is always the case due to the requirement $\ell \ll L$ in phase field approaches to fracture. Thus, phase field evolution according to (3.94) in homogeneous problems is clearly controlled by the stress state (motivated by a 1D equivalent).

3.4.3 Inhomogeneous solution after exceeding stability limit

As discussed by Benallal and Marigo (2007), Pham and Marigo (2008) and Kuhn and Müller (2013) among others, the spatially homogeneous solution is stable before exceeding the stability limit $\varepsilon_\mathrm{stab}$. The stability limit is slightly above the strain at maximum stress, i.e. $\varepsilon_\mathrm{stab} > \varepsilon_\mathrm{t}$ with ε_t serving as lower bound. After passing this limit, the phase field starts to localize and becomes inhomogeneous, eventually it forms a crack. In fact, the difference between both states is negligible in the common case that $\ell \ll L$ (Kuhn and Müller, 2013). For this reason, the maximum stress of the spatially homogeneous solution can be interpreted as material strength. Following the analyses of Hakim and Karma (2009), Borden et al. (2012) and Kuhn et al. (2015), the spatially inhomogeneous phase field solution is investigated.

The combination of the constitutive relation (3.97) and the phase field evolution equation (3.98) yields the non-linear ordinary differential equation

$$\frac{g'(s)}{\big(g(s)\big)^2}\,\frac{\ell\,c_\mathrm{w}}{2\mathcal{G}_\mathrm{c}\,E}\,\sigma^2 + \frac{w'(s)}{2} - \ell^2\,s_{,xx} = 0 \quad . \tag{3.176}$$

In the following, it is assumed that the problem is symmetric as depicted in Fig. 3.6. Thus, localization takes place at the center $x = 0$ and is quantified by the phase field value $s_\mathrm{loc} := s(x{=}0)$. Moreover, the solution of (3.176) must be differentiable at every point except the center. The differential equation (3.176) is supplemented with the far-field boundary conditions $\lim_{x\to\pm\infty} s_{,x} = 0$. In this specific case, there is a smooth transition of $s(x)$ to the homogeneous response $\lim_{x\to\pm\infty} s = s_\mathrm{hom}(\sigma)$, which is a function of the stress as worked out in the preceding section. Figure 3.11 shows a function which contains all requirements specified above. To find the solution, the differential equation (3.176) is integrated with respect to s over the interval $[s_\mathrm{loc},\ s(x)]$, i.e.

$$\int_{s_\mathrm{loc}}^{s(x)} \left(\frac{g'(\bar{s})}{\big(g(\bar{s})\big)^2}\,\frac{\ell\,c_\mathrm{w}}{2\mathcal{G}_\mathrm{c}\,E}\,\sigma^2 + \frac{w'(\bar{s})}{2} - \bar{\ell}^{\,2}\,\bar{s}_{,xx} \right) \mathrm{d}\bar{s} = 0 \quad . \tag{3.177}$$

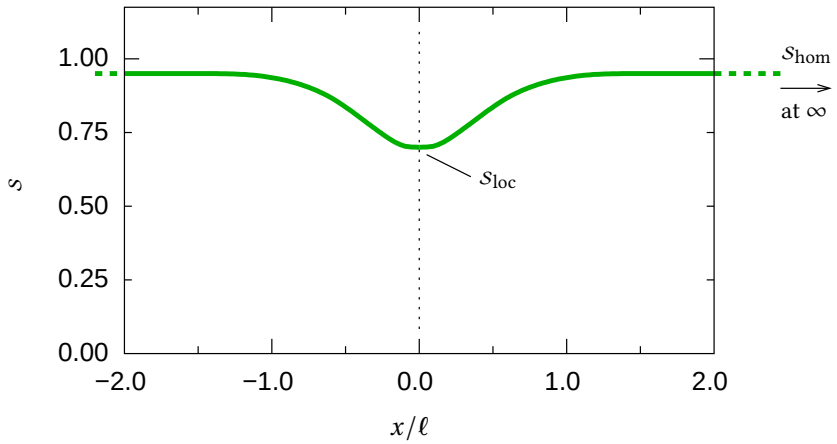


Figure 3.11: Sketch of non-homogeneous phase field solution $s(x)$ displaying minimal value s_{loc} at $x = 0$ and s_{hom} at $x \to \infty$. s_{hom} is obtained from the homogeneous phase field solution and only depends on the stress, i.e. $s_{\mathrm{hom}}(\sigma)$.

By making use of the fact that the stress is homogeneous (3.96), i.e. independent of $s(x)$, and by utilizing (B.1) and (B.2) one obtains

$$\underbrace{\frac{1}{g(s)}\,\frac{\ell\, c_{\mathrm{w}}}{2\mathcal{G}_{\mathrm{c}}\, E}\,\sigma^2 - \frac{w(s)}{2}}_{=:\, V_{\mathrm{eff}}(s)} + \underbrace{\frac{\ell^2}{2}\,(s_{,x})^2}_{=:\, V_{\mathrm{kin}}(s_{,x})} = \underbrace{\frac{1}{g(s_{\mathrm{loc}})}\,\frac{\ell\, c_{\mathrm{w}}}{2\mathcal{G}_{\mathrm{c}}\, E}\,\sigma^2 - \frac{w(s_{\mathrm{loc}})}{2}}_{=:\, V_{\mathrm{eff}}(s_{\mathrm{loc}})} \quad . \tag{3.178}$$

Following the interpretation of Hakim and Karma (2009) that (3.178) is a conservation law which yields an "equation of motion" (Karma et al., 2001), terms of "effective potential" energy $V_{\mathrm{eff}}(s)$ and "kinetic" energy $V_{\mathrm{kin}}(s_{,x})$ can be identified. Note, that since $s_{,x}\,(x \to \pm\infty) = 0$, evaluation of (3.178) for $x \to \pm\infty$ returns the homogeneous state, i.e. $V_{\mathrm{kin}}(s_{,x}) = 0$ and $V_{\mathrm{eff}}\big(s(x)\big) = V_{\mathrm{eff}}\big(s_{\mathrm{hom}}\big)$. Since the minimum value of the phase field s_{loc} is unknown for a specific stress $\sigma(s_{\mathrm{hom}})$, it must be determined from

$$V_\sigma(s) := V_{\mathrm{eff}}(s) - V_{\mathrm{eff}}(s_{\mathrm{hom}}) \quad . \tag{3.179}$$

In general, the function $V_\sigma(s)$ has positive function values at the limit states $s \to 0$ and $s \to 1$ and is negative in a specific interval in between. Thus, the function has two roots. Figure 3.12a shows $V_\sigma(s)$ for different stress levels for the combination of the single-well function with the quartic degradation function. The root closer to one (marked by upward pointing triangles) corresponds to the homogeneous solution, i.e. the phase field limit far way from $x = 0$, and increases as σ decreases. The other root (marked by downward pointing triangles) is located closer to zero. It corresponds to the minimum

value of the inhomogeneous solution s_{loc} and decreases as σ decreases. At the stability limit σ_{t} both solutions coincide. Integration of (3.178) finally yields the inverse relation

$$x(s) = \operatorname{sgn}(x) \int_{s_{\mathrm{loc}}}^{s} \frac{\ell}{\sqrt{2V_{\mathrm{eff}}(s_{\mathrm{loc}}) - 2V_{\mathrm{eff}}(\bar{s})}} \, \mathrm{d}\bar{s} \tag{3.180}$$

By evaluating (3.180) for (many) values $s \in (s_{\mathrm{loc}}, s_{\mathrm{hom}})$, phase field profiles for each load level can be constructed. The result obtained for the single-well surface in conjunction with the quartic degradation function is shown in Fig. 3.12b. In case of the cracked state (with $\sigma = 0$), the solution reduces to the well-known solution (3.106).

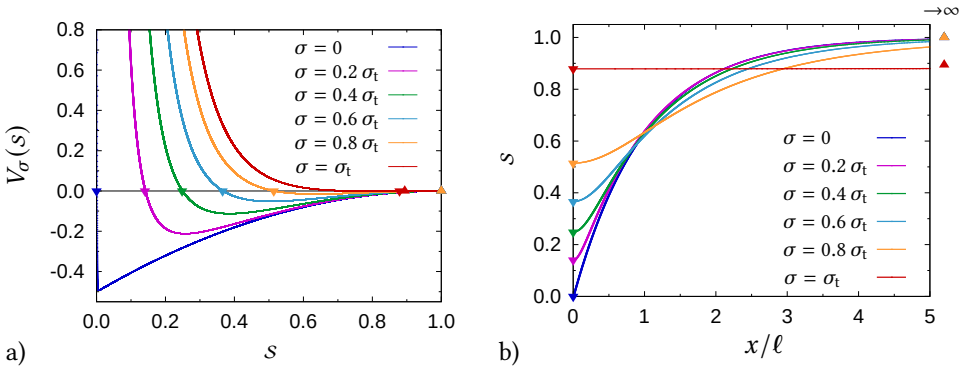


Figure 3.12: Evaluation of the inhomogeneous problem for different load levels $\sigma \in [0,\ \sigma_{\mathrm{t}}]$ for the single-well surface in conjunction with the quartic degradation function, a) evaluation of $V_\sigma(s)$ with roots indicated, b) inhomogeneous phase field solution $s(x)$.

By comparing the graphs in Fig. 3.12b for different evolution states, the phase field value some distance away from the center increases as $s_{\mathrm{loc}} \to 0$. The maximum increase of the phase field can be evaluated by the difference between the final phase field profile and the phase field value during evolution from $s_{\mathrm{loc}} = s_{\mathrm{hom}}$ to $s_{\mathrm{loc}} = 0$ at each location, i.e.

$$s_{\mathrm{rel}}(x) := \max_{s_{\mathrm{evol}}} \left\{ s(x, s_{\mathrm{loc}}{=}0) - s(x, s_{\mathrm{loc}}{=}s_{\mathrm{evol}}) \right\} \qquad \text{with} \quad s_{\mathrm{evol}} \in [0,\ s_{\mathrm{hom}}] \quad , \tag{3.181}$$

The corresponding results obtained for the single-well function in conjunction with the quadratic (see App. B.7) and the quartic degradation function are shown in Fig. 3.13. Interestingly, in case of the quadratic degradation function, the phase field increases at every point but the center, i.e.

$$s_{\mathrm{rel}}(x) > 0 \quad \forall x \setminus \{0\} \quad . \tag{3.182}$$

For the quartic degradation function, there is a finite location around the center where the phase field only decreases during localization.

In this context, the application of a “damage-like” irreversibility constraint as introduced in Sect. 3.2.3 seems to be questionable, since it hinders the phase field to approximate

the intended solution. One severe consequence is, that the stiffness away from the crack is erroneously reduced by $g(s_\mathrm{t})$, for the quartic degradation function $g(s_\mathrm{t}=3/4) \approx 0.56$. In addition, the surface energy is overestimated by $\approx 3L/\ell$ %, which, in fact, is a drastic deviation as $\ell/L \to 0$.

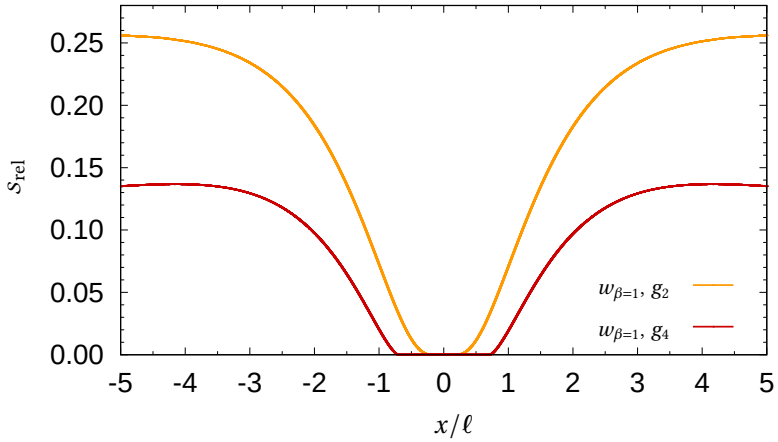


Figure 3.13: Maximum release of phase field s during evolution from the unloaded intact state $\sigma = 0$, $s = 1$ via the maximum load $\sigma = \sigma_\mathrm{t}$ to the broken state $\sigma = 0$, $s(x_0) = 0$.

3.4.4 Numerical assessment

In addition to the analytical considerations above, the effect of different degradation functions on the numerical solution, obtained by the finite element method (see App. C.2), is investigated. For this purpose, the 1D problem given in Fig. 3.6 is slightly modified as shown in Fig. 3.14a. Instead of applying the displacement at both ends, one end is fixed $u(-L) = 0$, while the other end is loaded by an linearly increasing displacement $u(t) = 2u\,t$. Since the analysis is meant to be static, i.e. neither inertia ($\rho=0$) nor viscosity ($\eta_\mathrm{vc}=0$) exist, this modification does not have an effect on the behavior. The mesh consists of either 200, 500 or 1000 regular quadrilateral elements.

Due to its importance for further studies, the focus is on the phase field approach based on the single-well surface function (3.58) in conjunction with the polynomial degradation functions (3.70) and (3.76). In a first step, both the monolithic solution scheme and the alternate minimization procedure are used. As long as no perturbation is present, only the homogeneous solutions are obtained for each $u \to \infty$, even above the stability limit. All simulation results are in accordance with the analytical results obtained above. Once there is a small (accidental) perturbation after the stability point is exceeded the solution is expected to become inhomogeneous. In this case, the static monolithic scheme usually fails in finding a solution while the alternate minimization procedure is still able to find a

solution (the inhomogeneous one), although many iterations might be necessary. To also obtain inhomogeneous solution by a monolithic solution scheme, an additional viscosity, as small as possible, must be applied (Kuhn, 2013; Schlüter, 2018). For this purpose, the alternate minimization procedure is preferred in the following.

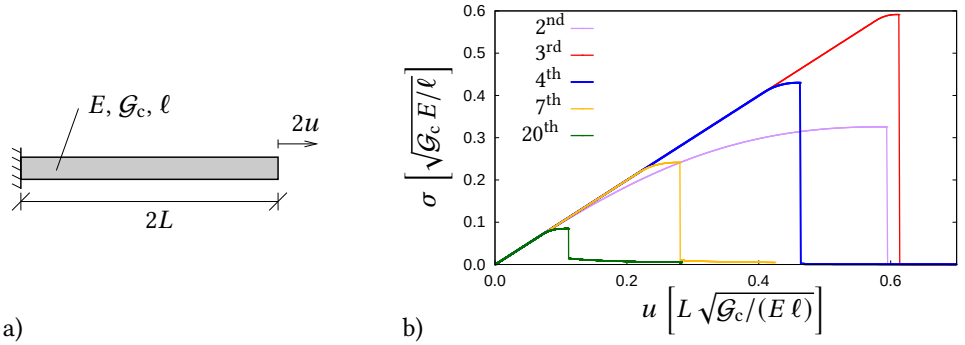


Figure 3.14: Uniaxial tension of 1D bar: a) boundary value problem, b) normalized load-displacement curves for polynomial degradation functions of different ($2^{\text{nd}} \ldots 20^{\text{th}}$) order.

From an energetic point of view, localization at one of the bar ends is preferred, since only one half of the fracture energy is required while the energy release is independent of the location of the crack. To investigate fracture at the center of the problem the location of the localization is prescribed, typically by reducing the fracture toughness $\mathcal{G}_c$ at the center by about 10^{-2} %. This also avoids the problem of incomplete surface energy approximation if localization starts too close to the free/fixed boundary. Figure 3.14b shows the load-displacement curves for polynomial degradation functions of different order. Since neither in the homogeneous nor the inhomogeneous state there is a noticeable difference between the three different meshes, only results of the fine mesh are shown. Before the peak stress is reached, all curves show the result of the homogeneous response, while shortly after exceeding the maximum stress, the results are governed by the inhomogeneous (localized) solution. The approach based on the quadratic degradation function (3.70) behaves strongly nonlinear prior to the peak stress. By contrast, an almost perfectly linear response is obtained up to a point σ_{lin} somewhat below the peak stress by using higher-order degradation functions (with $a_s = 10^{-4}$). This represents the behavior of brittle materials much better than the formulation based on the quadratic degradation function does. As indicated by Fig. 3.14b, the peak stress decreases with increasing order of the degradation function, as already analyzed before.

3.4.5 Discussion of consequences

Ambiguous role of the internal length parameter

As discussed in the previous section, in the framework of phase field modeling, the peak stress value of the homogeneous solution can be interpreted as the material's tensile strength σ_t, which is immediately followed by strain localization controlled by the phase field equation. The peak stress depends on the choice of the degradation function $g(s)$, the surface function $w(s)$, the internal length ℓ and the material parameters E and $\mathcal{G}_\mathrm{c}$. With σ_t being a given material parameter, along with E and $\mathcal{G}_\mathrm{c}$, and a specific choice of $g(s)$ and $w(s)$, the value of the internal length parameter

$$\ell = {f_\mathrm{gw}}^2 \frac{\mathcal{G}_\mathrm{c}\, E}{\sigma_\mathrm{t}^2} \tag{3.183}$$

is given via

$$f_\mathrm{gw} = \begin{cases} \dfrac{1}{\sqrt{c_\mathrm{w}}} \sup\limits_{s\in(0,1]} \sqrt{f_\sigma(s)} & \text{for } \lim\limits_{s\to 1^-} f_\sigma'(s) < 0 \\ \dfrac{1}{\sqrt{c_\mathrm{w}}} \sqrt{\lim\limits_{s\to 1^-} f_\sigma(s)} & \text{else} \end{cases} . \tag{3.184}$$

Since it is directly related to the material strength, it can be regarded as material parameter. For this reason, it is important to note that for phase field evolution in absence of pre-cracks, the internal length ℓ parameter is not only a regularization parameter of the crack and thus cannot be chosen from purely geometric considerations.

If the internal length ℓ is determined by the material strength and controls the width of the regularization zone, two potential difficulties arise:

- ℓ is small: The regularized zone must be adequately resolved by the mesh. If the problem of interest is comparatively large (compared to ℓ), the numerical solution is extremely costly, if not impossible. In the case that the location of crack onset is known a-priori, local mesh refinement can be used. However, if a location of potential crack onset/progress is not refined adequately, fracture is suppressed by the discretization and the correct solution cannot be obtained. As an alternative, Kuhn and Müller (2010b) proposed to use exponential shape functions (instead of "expensive" mesh refinement) to better approximate the phase field in the vicinity of cracks for small values of the regularization parameters ℓ. The phase field approximation of stationary cracks in absence of mechanical forces turned out to be very accurate even for coarse meshes (Kuhn and Müller, 2011; Kuhn, 2013). Challenges are the construction of a valid element approximation, especially for non-regular meshes, the numerical integration and that the elements must be adjusted to the crack location and orientation. Besides these technical issues, the requirement of vanishing stiffness across open cracks (typically causing

broadening of the broken state in the phase field over one finite element as discussed in Sect. 3.7.3) does not seem to be fulfilled for coarse meshes[8].

- ℓ is large: Regularization sets a lower bound for the correct representation of physical effects. If the effect of interest is comparatively small (compared to ℓ), e.g. the length of a crack, it is not possible to use the model. In particular, this becomes relevant in Sect. 4.3.

The first point is rather a technical issue. In addition, by an appropriate choice of the degradation function, there are possibilities to increase the critical stress to infinity, so ℓ increases equivalently. To circumvent the second point, there are only limited possibilities. As indicated above, degradation functions of higher polynomial order may give rise to smaller values of the internal length. However, a major drawback of such higher-order degradation functions is the vanishing crack-driving energy in the phase field evolution equation long before reaching the broken state, thus hampering a further decay of the phase field as observed by Strobl and Seelig (2018). This does not much influence the stiffness degradation (which is almost completed), but leads to a wrong approximation of the crack surface energy (3.15).

By introducing a threshold function in the phase field evolution equation, e.g. according to (3.94), it is possible to decouple the internal length ℓ from the material strength σ_t to some extent as analyzed in Sect. 3.4.2. However, as $\ell \longrightarrow 0$ for a given set of material parameters (E, $\mathcal{G}_\mathrm{c}$, σ_t) the parameter α_c must decrease, but is limited to $\alpha_\mathrm{c} \geq 1/8$. Thus, this helps in the first, but not in the second case (from above).

Width of the crack zone

As already discussed in Sect. 3.4.2, the choice of the quadratic single-well function (3.58) to approximate the crack surface density (3.52) requires the application of a higher-order degradation function to prevent an unrealistic non-linear stress-strain response prior to fracture. An alternative for this might be seen in linear surface function (3.59), which in conjunction with the quadratic degradation function (3.70) yields a linear stress-strain response prior to fracture (Sect. 3.4.2). This formulation has recently received great attention and is sometimes claimed to provide a more narrow transition zone between the intact material state and the crack than the quadratic single-well function, e.g., Schneider et al. (2016), Bleyer et al. (2017), Gerasimov and De Lorenzis (2019).

In the following, the quantitative effects of these formulations are briefly analyzed and compared to each other. At this point, one mainly focuses on the size of the internal length parameter which is related to the material strength (see the discussion above) and the formation of the regularized zone in the vicinity of a crack. Since the internal

[8] In fact, Kuhn (2013) introduced a priori "one extra layer of thin elements" at the location of the pre-existing and along the propagating crack.

length parameters of the different formulations cannot be directly compared to each other, they are related in a first step to the corresponding peak stress (Sect. 3.4.2). Using the linear surface function (3.59) with $c_w = 4/3$ in conjunction with the quadratic degradation function (3.70) the internal length is

$$\ell_{\beta=0} = \frac{3}{8}\,\frac{\mathcal{G}_c\,E}{\sigma_t^2} = 0.375\,\frac{\mathcal{G}_c\,E}{\sigma_t^2} \quad , \tag{3.185}$$

whereas the analysis in Sect. 3.4.2 yields

$$\ell_{\beta=1} \approx 0.185\,\frac{\mathcal{G}_c\,E}{\sigma_t^2} \tag{3.186}$$

for the single-well surface function (3.58) in combination with the quartic degradation function (3.74).

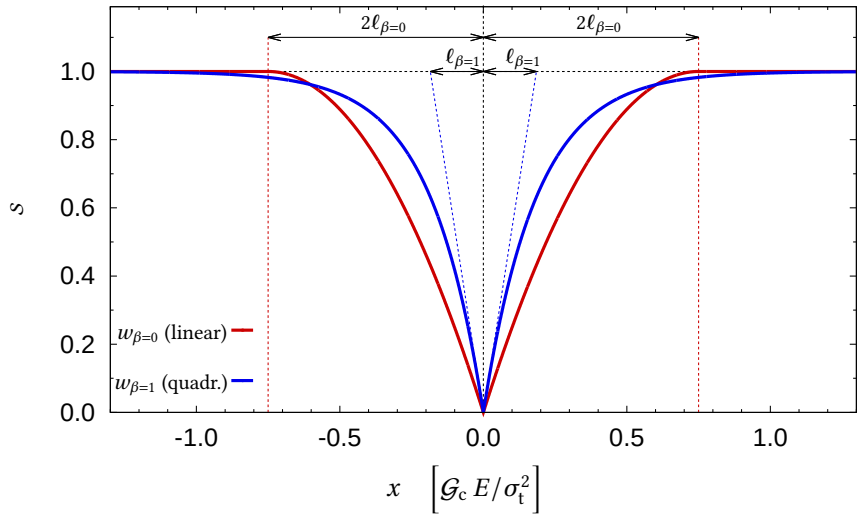


Figure 3.15: Localized 1D phase field solution $s(x)$ in the vicinity of a crack (at $x=0$) in absence of mechanical load; comparison of the linear surface function $w_{\beta=0}=(1-s)$ with regularization parameter $\ell_{\beta=0}=0.375\,\mathcal{G}_c\,E/\sigma_t^2$ and the quadratic single-well surface function $w_{\beta=1}=(1-s)^2$ with regularization parameter $\ell_{\beta=1}\approx 0.185\,\mathcal{G}_c\,E/\sigma_t^2$.

In a second step, the inhomogeneous phase field solutions $s(x)$ of both formulations in the ultimate ("cracked") state, see Sect. 3.4.1, are compared. Figure 3.15 shows both phase field solutions (3.106 and 3.107). Note, that the two formulations do not only give rise to the different values (3.185) and (3.186), but that $\ell_{\beta=0}$ and $\ell_{\beta=1}$ occur in different ways in the phase field profiles. While the width of the transition zone for the quadratic single-well function is infinite and the length parameter $\ell_{\beta=1}$ quantifies the (inverse) slope of $s(x)$ at the crack, the length parameter $\ell_{\beta=0}$ is equal to one fourth of the width of the regularized

zone resulting from the linear surface function. Hence, with the ratio $\ell_{\beta=0} \approx 2\ell_{\beta=1}$ from (3.185) and (3.186) accounted for in Fig. 3.15, it is obvious that both formulations lead to very similar crack profiles and, correspondingly, crack surface energy approximations. In a similar fashion, the application of the concave surface function (3.60) in conjunction with the quadratic degradation function (3.70) yields with (3.167)

$$\ell_{\beta=-1} \approx 0.637 \frac{\mathcal{G}_c E}{\sigma_t^2} \quad . \tag{3.187}$$

The resulting width of the transition zone $\pi\ell_{\beta=-1} = 2$ is even larger than the width of the transition zone of the linear surface function with $4\ell_{\beta=0} = 1.5$.

3.5 Tension-compression asymmetric response

Degradation of the complete strain energy, or in other words, full stiffness degradation at a crack (i.e. for $s=0$ depicted in Fig. 3.16a) as introduced without further discussion in (3.7) and (3.12) has two major flaws. First, crack closure (e.g. caused by compression perpendicular to a crack) does not cause any elastic response, which results in unphysical interpenetration of the crack flanks as illustrated in Fig. 3.16b. Second, positive (tensile) and negative (compressive) mechanical forces contribute to phase field evolution in the same manner, see the phase field evolution equation (3.30). In case of pure shear (mainly mode II loading) the phase field model predicts symmetric crack growth in the tensile and compressive regime, see Fig. 3.16c, which clearly contradicts experimental evidence and theoretical predictions (Sect. 2.2.5). Both issues have already been reported by Bourdin et al. (2000).

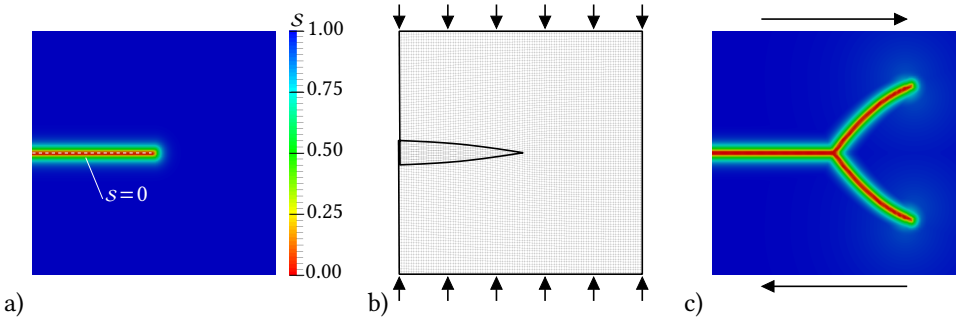


Figure 3.16: Illustration of negative consequences of full stiffness degradation: a) pre-crack modeled by the phase field with $s \to 0$, b) interpenetration of crack surfaces under compressive loading, c) symmetric phase field evolution in the tensile (lower half) and compressive regime (upper half).

Strain energy decomposition

To account for both problems while keeping the variational structure of the method (as introduced in Sect. 3.2) different modifications of the strain energy density have been proposed (Amor et al., 2009; Freddi and Royer-Carfagni, 2010; Miehe et al., 2010a). All approaches have in common, that the strain energy density (2.82)

$$\psi_0(\boldsymbol{\varepsilon}) = \psi_{(1)0}(\boldsymbol{\varepsilon}) + \psi_{(2)0}(\boldsymbol{\varepsilon}) + \cdots + \psi_{(n)0}(\boldsymbol{\varepsilon}) = \frac{1}{2}\,\boldsymbol{\varepsilon} : \mathbb{C}_0 : \boldsymbol{\varepsilon} \tag{3.188}$$

is additively split into several portions $\psi_{(1)0} \ldots \psi_{(n)0}$, where each portion has to be non-negative for any possible strain state. This entails that none of the energy portions in (3.188) exceeds the strain energy density (2.82), i.e.

$$0 \le \psi_{(i)0}(\boldsymbol{\varepsilon}) \le \psi_0(\boldsymbol{\varepsilon}) \tag{3.189}$$

with $i \in [0,\ n]$. Note, that the subscript "0" here indicates the undegraded (linear) elastic case. By degrading (or not degrading) each portion by a separate degradation function (following the restrictions provided in Sect. 3.3.2) it is guaranteed in advance that the strain energy dissipation is positive in case of phase field evolution ($\mathrm{d}s < 0$). This results in the strain energy density

$$\psi_{\mathrm{e}}(\boldsymbol{\varepsilon}, s) = g_{(1)}(s)\,\psi_{(1)0}(\boldsymbol{\varepsilon}) + g_{(2)}(s)\,\psi_{(2)0}(\boldsymbol{\varepsilon}) + \cdots + g_{(n-1)}(s)\,\psi_{(n-1)0}(\boldsymbol{\varepsilon}) + \psi_{(n)\,0}(\boldsymbol{\varepsilon}) \quad , \tag{3.190}$$

where one (here the last) portion $\psi_{(n)0}(\boldsymbol{\varepsilon})$ is assumed to be not degraded. Variation with respect to both unknowns (according to Sect. 3.2.1) finally yields the Cauchy stress

$$\begin{aligned} \boldsymbol{\sigma}(\boldsymbol{\varepsilon}, s) &:= \frac{\partial\, \psi_{\mathrm{e}}(\boldsymbol{\varepsilon}, s)}{\partial\, \boldsymbol{\varepsilon}} \\ &= g_{(1)}(s)\,\frac{\partial \psi_{(1)0}(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} + \cdots + g_{(\mathrm{n}-1)}(s)\,\frac{\partial \psi_{(n-1)0}(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} + \frac{\partial \psi_{(n)0}(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \quad , \end{aligned} \tag{3.191}$$

the stiffness

$$\begin{aligned} \mathbb{C}(\boldsymbol{\varepsilon}, s) &:= \frac{\partial^2\, \psi_{\mathrm{e}}(\boldsymbol{\varepsilon}, s)}{\partial\, \boldsymbol{\varepsilon}^2} \\ &= g_{(1)}(s)\,\frac{\partial \psi_{(1)0}(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} + \cdots + g_{(\mathrm{n}-1)}(s)\,\frac{\partial \psi_{(n-1)0}(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} + \frac{\partial \psi_{(n)0}(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \end{aligned} \tag{3.192}$$

and the (variationally consistent) crack driving force

$$\begin{aligned} f_{\mathrm{s}}(\boldsymbol{\varepsilon}, s) &:= \frac{\partial\, \psi_{\mathrm{e}}(\boldsymbol{\varepsilon}, s)}{\partial\, s} \\ &= \frac{\mathrm{d}g_{(1)}(s)}{\mathrm{d}s}\,\psi_{(1)0}(\boldsymbol{\varepsilon}) + \cdots + \frac{\mathrm{d}g_{(\mathrm{n}-1)}(s)}{\mathrm{d}s}\,\psi_{(\mathrm{n}-1)0}(\boldsymbol{\varepsilon}) \quad . \end{aligned} \tag{3.193}$$

Note, that the undegraded strain energy portion $\psi_{(n)0}(\boldsymbol{\varepsilon})$ does not appear in the crack driving force (3.193). Consequently, the related strain portions are decoupled from the phase field equation.

The energetic portions $\psi_{(1)0} \dots \psi_{(n)0}$ (and the degradation functions $g_{(1)} \dots g_{(n-1)}$) must then be specified in order to both realize a non-interpenetration condition by (3.191) or (3.192) and prevent phase field evolution under compression in (3.193). Typically, this is done by a split in only two strain energy density portions, one related to tensile states which is degraded by the degradation function $g(s)$, and one related to compressive states which is not degraded, i.e

$$\psi_e(\boldsymbol{\varepsilon}, s) = g(s)\, \psi_0^+(\boldsymbol{\varepsilon}) + \psi_0^-(\boldsymbol{\varepsilon}) \quad . \tag{3.194}$$

In the case of only two portions, the choice of one portion directly yields the other portion, either

$$\psi_0^- = \frac{1}{2}\, \boldsymbol{\varepsilon} : \mathbb{C}_0 : \boldsymbol{\varepsilon} - \psi_0^+ \qquad \text{or} \qquad \psi_0^+ = \frac{1}{2}\, \boldsymbol{\varepsilon} : \mathbb{C}_0 : \boldsymbol{\varepsilon} - \psi_0^- \quad . \tag{3.195}$$

In the literature, they are referred to as "positive" or "negative" (Li et al., 2016; Miehe et al., 2010a), "active" or "passive" (Strobl and Seelig, 2016), "active" or "non-active" (Cajuhi, 2019), "crack driving" or "persistent" (Steinke and Kaliske, 2019), "tensile" (Ambati et al., 2015a), "tension sensitive" (Hesch et al., 2017) or "fracture insensitive" (Hesch and Weinberg, 2014), to name a few.

Clear situation in 1D

As a continuation of the one-dimensional problem analyzed in Sect. 3.4 (see Fig. 3.6) the constitutive choice of the energetic portions in one dimension is discussed. To separate positive and negative portions of a scalar quantity $x \in \mathbb{R}$ the bracket operator[9]

$$\langle x \rangle := \frac{1}{2}\big(|x| + x\big) = \begin{cases} 0 & \forall x \leq 0 \\ x & \forall x > 0 \end{cases} \quad , \tag{3.196}$$

is used. The decomposition reads

$$x \equiv \underbrace{\langle x \rangle}_{\geq 0} + \underbrace{\big(- \langle -x \rangle \big)}_{\leq 0} \quad , \tag{3.197}$$

where the first term is related to the positive portion of x and the latter term is related to the negative portion of x. For more details see App. A.9.

[9] Most often named after the British mathematician William Herrick Macauley, the bracket operator notation was first introduced by Clebsch (1862) to calculate beam deflections. At this time, Rudolf Friedrich Alfred Clebsch held the chair of theoretical mechanics at the Carlsruhe Polytechnicum (Ben-Menahem, 2009). In German-speaking countries the mathematical notation is named after August Föppl, who extended Clebsch's idea, see e.g. Weissenburger (1964).

The application of the additive decomposition (3.197) on the strain $\varepsilon = \langle\varepsilon\rangle - \langle -\varepsilon\rangle$ or, equivalently, the undegraded stress $\sigma_0 = \langle E\,\varepsilon\rangle - \langle -E\,\varepsilon\rangle$ allows to (uniquely) identify the elastic energy density portions related to tension and compression

$$\psi_0^+(\boldsymbol{\varepsilon}) := \frac{1}{2} E \langle\varepsilon\rangle^2 \qquad \text{and} \qquad \psi_0^-(\boldsymbol{\varepsilon}) := \frac{1}{2} E \langle -\varepsilon\rangle^2 \quad , \tag{3.198}$$

respectively. Substitution in (3.194) yields the strain energy density

$$\psi_\mathrm{e}(\varepsilon, s) = g(s)\,\frac{1}{2} E \langle\varepsilon\rangle^2 + \frac{1}{2} E \langle -\varepsilon\rangle^2 \quad , \tag{3.199}$$

with the degradation function being the only constitutive choice. The corresponding Cauchy stress is

$$\sigma(\varepsilon, s) = g(s)\, E \langle\varepsilon\rangle - E \langle -\varepsilon\rangle = \begin{cases} E\,\varepsilon & \forall \varepsilon \leq 0 \\ g(s)\, E\,\varepsilon & \forall \varepsilon > 0 \end{cases} \tag{3.200}$$

while the crack driving force is

$$f_\mathrm{s}(\varepsilon, s) = g'(s)\,\frac{1}{2} E \langle\varepsilon\rangle^2 = \begin{cases} 0 & \forall \varepsilon \leq 0 \\ g'(s)\,\dfrac{1}{2} E\,\varepsilon^2 & \forall \varepsilon > 0 \end{cases} . \tag{3.201}$$

In case of compression ($\varepsilon < 0$) the mechanical response (3.200) is independent of the phase field value. Thus, the original stiffness is recovered and no interpenetration of the crack flanks occurs. The latter can be seen by regarding the crack opening displacement which is, by using (3.100) and (3.101),

$$\delta_\mathrm{s} = \begin{cases} 0 & \forall \varepsilon \leq 0 \\ 2\dfrac{\sigma}{E} \displaystyle\int_0^L \frac{1}{g(s)} - 1 \,\mathrm{d}x & \forall \varepsilon > 0 \end{cases} , \tag{3.202}$$

and, consequently, non-negative. In addition, according to (3.201) no phase field evolution takes place under compression. For tension ($\varepsilon > 0$) both (3.200) and (3.201) yield the behavior analyzed in Sect. 3.4. Note, that for $s \neq 1$ the different treatment of the positive and negative part in (3.200) introduces non-linearity.

Ambiguous situation in 3D

The bracket operator (3.196) can be applied on scalar quantities only. Thus, a straightforward extension of the approach discussed above to two or three dimensions is not possible. Besides this technical reason, the states of "tension" and "compression" are unclear when applied to more than one dimension, at least they are not uniquely defined.

To illustrate some conceptual possibilities, different scalar projections of the strain tensor (see Sect. 2.1.1) are regarded, i.e. the volumetric strain and principal strains

$$\varepsilon_{\mathrm{vol}} = \boldsymbol{I} : \boldsymbol{\varepsilon} \,, \quad \varepsilon_1 = \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_1 : \boldsymbol{\varepsilon} \,, \quad \varepsilon_2 = \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_2 : \boldsymbol{\varepsilon} \,, \quad \varepsilon_3 = \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_3 : \boldsymbol{\varepsilon} \tag{3.203}$$

with $\overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_\alpha := \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_\alpha \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_\alpha$. Each of the four projections and combinations thereof (in conjunction with elastic parameters, e.g., λ and μ) can be separately decomposed by (3.197) into its positive and negative portion as long as each resulting strain energy density portion is non-negative for any possible strain state to fulfill (3.189). In present phase field models to fracture this is, to some extent, a constitutive choice (Strobl and Seelig, 2016). In the following, a simple test to validate the mechanical response of a phase field induced crack is presented, before two specific and widely-used phase field approaches based on strain energy decomposition are analyzed in detail.

Test to validate the mechanical response at a crack

In contrast to damage models, where the notion of "damage" — quite often — is rather unspecific (see Sect. 2.2.8), the phase field s in the present approach is dedicated to approximate cracks. This concerns both the energetic balance of the surface energy according to the idea of Griffith, see Sect. 3.1, as well as the mechanical behavior at cracks. The latter comprises not only non-interpenetration of the crack flanks. In particular, the state of the material for $s = 0$ should reproduce the behavior of a macroscopic crack, i.e. crack opening and (frictionless) sliding of the crack surfaces should not cause any traction while there are compressive surface forces in case of crack closure. To validate the mechanical behavior of phase field models at cracks a simple numerical test was introduced by Strobl and Seelig (2015). An elastic block with initial through-crack is subjected to different loads in terms of displacements relative to the crack orientation. The key aspect is that the initial crack has to be consistently modeled by the phase field[10] as shown in Fig. 3.17a, either by imposing $s = 0$ on a double row of nodes or, alternatively, by applying an initial crack driving force as discussed in App. C.4. Of course, initial cracks can also be modeled without using the phase field by free surfaces/double nodes[11], however, newly grown cracks (modeled by the phase field) should also behave correctly. In particular, positive and negative relative displacements normal to the crack plane are imposed to test homogeneous tension (Fig. 3.17b) and compression states (Fig. 3.17c). The tensile loading should cause crack opening without any elastic response while compression should yield the well-known elastic behavior (also including Poisson's effect). The left and

[10] At the time when the test was proposed, the single-edged notched test was one of the standard "benchmarks", typically with the initial crack modeled by a discrete crack, i.e. with double nodes. Thus, the specific (perhaps wrong) model response at phase field induced cracks was not immediately apparent.

[11] In fact, Klinsmann et al. (2015) compared the results for pure mode I loading in the case where the initial crack is prescribed by free surfaces and the case where the initial case is modeled by the phase field. While the results for the phase field modeled initial crack show convergence to analytical results (as $\ell \to 0$) results for the mesh induced initial crack deviate significantly.

right boundaries of the tension and compression test are traction-free to capture lateral elastic effects.

Furthermore, as proposed by Strobl and Seelig (2016), the through-cracked block can be tested by applying opposite displacements parallel to the crack as shown in Fig. 3.18a (shear loading). Note, that this test is carried out with $\nu = 0$ and vertical displacement constraints are prescribed at the left and right boundary (Fig. 3.18a) to apply a homogeneous shear loading. In case of a frictionless crack surface one would expect both parts sliding on each other without causing any elastic response as illustrated in Fig. 3.18b.

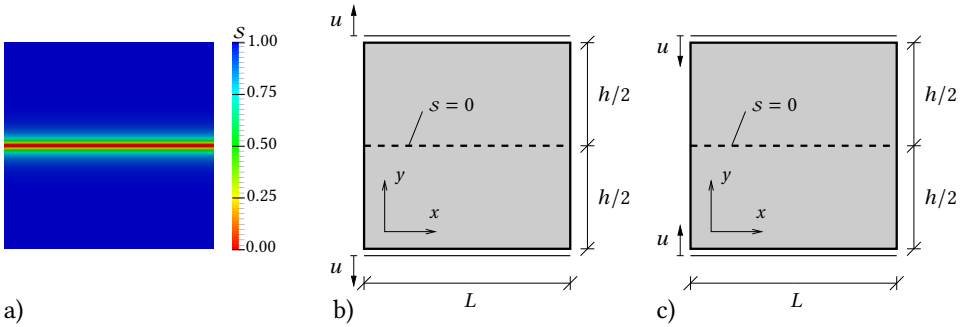


Figure 3.17: Elastic body with initial through-crack modeled by the phase field $s \approx 0$ a) phase field solution of the initially unloaded specimen, b) crack subjected to tension, c) crack subjected to compression.

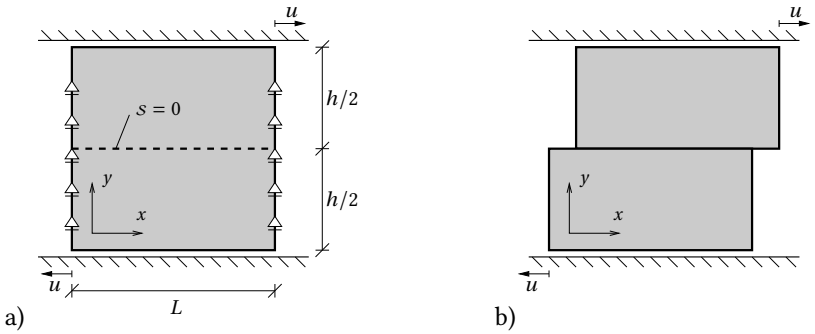


Figure 3.18: Elastic body with initial through-crack modeled by the phase field $s \approx 0$ a) crack subjected to shear deformation, b) expected deformation.

3.5.1 Models based on spectral decomposition

Brittle fracture is dominated by normal stresses, i.e. the fracture surface is typically oriented perpendicular to the largest principal positive stress. Thus, a common assumption

in models of continuum damage mechanics is that (micro) crack opening and growth is induced by positive principal strains or stresses, see Sect. 2.2.8.

Spectral strain decomposition

In terms of the phase field approach to fracture, Hofacker et al. (2009) introduced a tension-compression asymmetric response based on the spectral decomposition of strains (2.23). For simplicity, this approach is referred to as *spectral strain decomposition* in the following[12]. In particular, the elastic strain energy density is decomposed into portions related positive and negative principal strains and a portion related to the positive or negative first strain invariant $\mathrm{I}_\varepsilon := \operatorname{tr} \boldsymbol{\varepsilon}$ (cf. App. A.4), i.e.

$$\psi_0^+(\boldsymbol{\varepsilon}) := \frac{\lambda}{2}\langle \operatorname{tr} \boldsymbol{\varepsilon}\rangle^2 + \mu \left(\langle \varepsilon_1\rangle^2 + \langle \varepsilon_2\rangle^2 + \langle \varepsilon_3\rangle^2\right) \quad , \tag{3.204}$$

$$\psi_0^-(\boldsymbol{\varepsilon}) := \frac{\lambda}{2}\langle -\operatorname{tr} \boldsymbol{\varepsilon}\rangle^2 + \mu \left(\langle -\varepsilon_1\rangle^2 + \langle -\varepsilon_2\rangle^2 + \langle -\varepsilon_3\rangle^2\right) \quad . \tag{3.205}$$

Obviously, the approach is guided by the idea, that crack closure is (partially) detected by negative strain eigenvalues related to three mutually perpendicular planes in combination with volume reduction. It, however, should be noted that the orientation of the planes is determined by the principal strain directions and not by a crack surface. This will result in serious problems as analyzed below.

Although the approach is sometimes referred to as being anisotropic (e.g. Hofacker, 2014; Kiendl et al., 2016; Miehe et al., 2010c), it is in fact non-linear but isotropic. The choice resembles the approach (2.192) used in continuum damage mechanics. Substitution of (3.204) into (3.194) and variation provides both the Cauchy stress tensor

$$\begin{aligned}\boldsymbol{\sigma}(\boldsymbol{\varepsilon}, s) = g(s) \Big[&\lambda \langle \operatorname{tr} \boldsymbol{\varepsilon}\rangle \boldsymbol{I} + 2\mu \left(\langle \varepsilon_1\rangle \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_1 + \langle \varepsilon_2\rangle \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_2 + \langle \varepsilon_3\rangle \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_3 \right) \Big] \\ &- \lambda \langle -\operatorname{tr}(\boldsymbol{\varepsilon})\rangle \boldsymbol{I} - 2\mu \left(\langle -\varepsilon_1\rangle \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_1 + \langle -\varepsilon_2\rangle \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_2 + \langle -\varepsilon_3\rangle \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_3 \right) \quad ,\end{aligned} \tag{3.206}$$

already including the relation $\partial \varepsilon_\alpha / \partial \boldsymbol{\varepsilon} = \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_\alpha$, and the crack driving force

$$f_\mathrm{s}(\boldsymbol{\varepsilon}, s) = g'(s) \left[\frac{\lambda}{2}\langle \operatorname{tr} \boldsymbol{\varepsilon}\rangle^2 + \mu \left(\langle \varepsilon_1\rangle^2 + \langle \varepsilon_2\rangle^2 + \langle \varepsilon_3\rangle^2\right) \right] \quad . \tag{3.207}$$

Due to its motivation related to brittle fracture, the approach was over years (the perhaps most) widely used, e.g. by Hofacker and Miehe (2012), Borden et al. (2012), Klinsmann et al. (2015), May et al. (2015), among others and extended to finite deformations by Hesch and Weinberg (2014).

[12] To be precise, it should be noted that the split/decomposition of strain, stress or energy itself is at first only a reformulation. Separate treatment of each portion, in the specific case by different degradation, constitutes the core of each formulation. Nevertheless, naming the formulation after the used split is quite common in the phase field community and thus used in the following.

Inaccurate model behavior at cracks subjected to compression

In a first step, the model response caused by uniaxial compression perpendicular to a phase field modeled initial crack as depicted in Fig. 3.17a and c is investigated. As already discussed by Strobl and Seelig (2015), the approach based on the spectral strain decomposition is not able to reproduce a homogeneous stress response for $\nu \neq 0$. This is caused by the fact that positive lateral strains cause erroneous stiffness degradation in the plane of the crack where $s < 1$. The corresponding results showing the inhomogeneous stress response and the resulting inhomogeneous deformations are depicted in Fig. 3.19a and b. Compared to the effects discussed next, this inaccuracy, however, appears to be secondary.

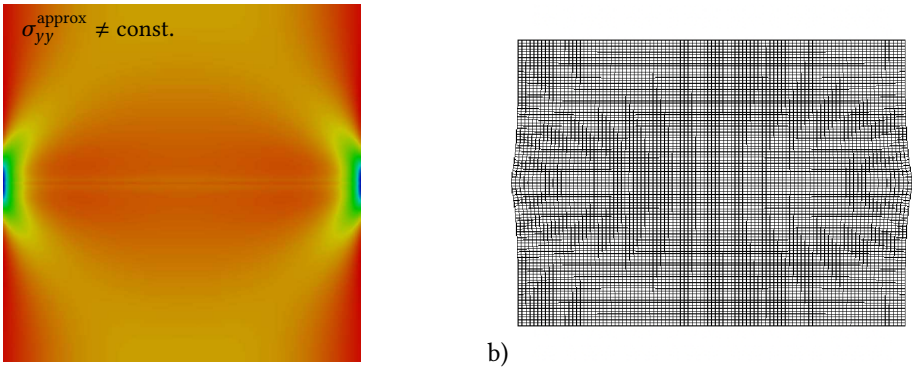


Figure 3.19: Validation of mechanical behavior of the phase field approach based on spectral strain decomposition at a crack subjected to compression (with $\nu = 0.3$), a) inhomogeneous stress response as consequence of erroneous degradation due to (positive) lateral strains at the crack, b) resulting inhomogeneous deformations at the degraded zone.

Inadequate model behavior at cracks subjected to shear

One of the main requirements on the phase field approach to fracture is the correct, i.e. traction-free, separation of crack flanks. By contrast, the correct modeling of crack closure is an additional feature which is only required in situations where existing cracks are closed. The purpose of the strain energy modification discussed in this section is to provide this additional feature (Miehe et al., 2010c). In addition to the tensile (not shown) and the inaccurate compressive state (see above), the model response caused by shear as depicted in Fig. 3.18a investigated. The initial crack is consistently modeled by the phase field as depicted in Fig. 3.17a. As shown by the deformed mesh in Fig. 3.20a the result does not correspond to the expected deformation in Fig. 3.18. The upper and lower part in Fig. 3.20a are sticking on each other along the phase field modeled crack ($s \approx 0$). Obviously, the crack surfaces are transmitting significant forces across the crack surface, which is depicted in Fig. 3.20b. As a consequence, elastic deformations are present in the

whole domain. Thus, the phase field approach based on the spectral strain decomposition is not able to reproduce the correct behavior. With increased shear displacement the phase field in the vicinity of the crack decreases due to non-zero traction forces across the crack, making the crack to "grow" in thickness direction. In addition, unphysical crack growth also occurs along the outer boundaries, see Fig. 3.20c.

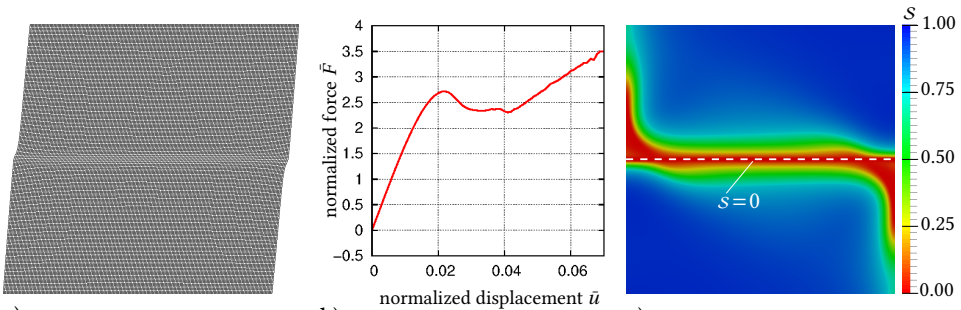


Figure 3.20: Validation of mechanical behavior of the phase field approach based on spectral strain decomposition at a crack subjected to shear: a) erroneous stiff elastic response, b) load-displacement curve showing stiff "elastic" response, c) unphysical expansion of the phase field caused by not traction-free crack surfaces.

As shown above, the phase field approach based on the spectral decomposition does not provide traction-free crack surfaces if the crack direction does not coincidentally coincide with one of the principal strain directions. Klinsmann et al. (2015) investigated mode I dominated fracture from an initial phase field modeled crack where the spectral strain decomposition model did not fail. The failure of the model is further demonstrated by the single-edge notched specimen shown in Fig. 3.21a subjected to shear (Strobl and Seelig, 2016). With increased load, the crack does not start do expand at the crack tip. At first, phase field evolution takes place at the boundary. With increased displacement, an additional crack forms with an inclination of $\approx 45°$ at both sides of the crack tip as depicted in Fig. 3.21b. It should be noted, that some problems have also been made public by May et al. (2015). They compared the test with a discrete initial crack and a phase field modeled initial crack. In the latter case, the results are similar to Fig. 3.21b. However, May et al. (2015) concluded that the "phase field model for brittle fracture is sensitive to the exact form of the applied boundary conditions", i.e. the erroneous behavior was related to the application of phase field boundary conditions at the crack and not related to the spectral strain decomposition.

Taking into account these results it must be concluded that the phase field approach based on the spectral strain decomposition completely fails in reproducing traction-free crack surfaces for arbitrary loading scenarios. Therefore, this approach cannot be used to model cracks. In this context, however, Ambati et al. (2015a), Kiendl et al. (2016) and Pham et al. (2017) published "single-edge notched shear test" and "mixed-mode I + II loading" simulations, with the initial crack modeled by the phase field. Although subjected to shear the phase field model based on the spectral strain decomposition did not fail. This is in

contradiction to the findings above, thus the applied initial crack conditions or modeling techniques in the aforementioned publications raise some questions.

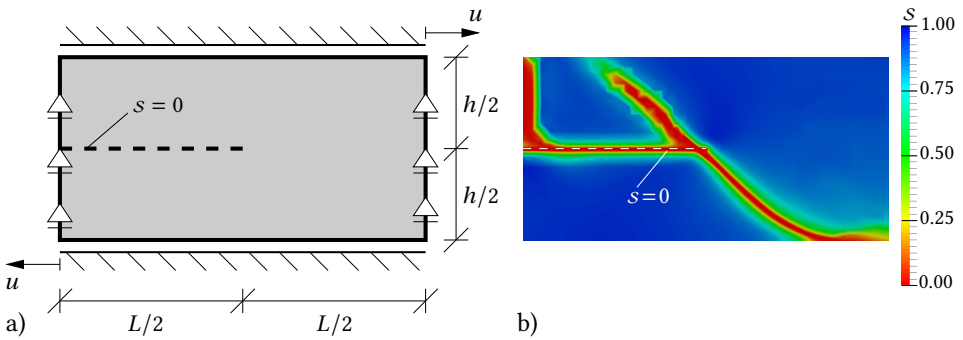


Figure 3.21: Single-edge notched shear test with phase field modeled initial crack: a) boundary value problem, b) erroneous phase field evolution caused by not traction-free crack surfaces as predicted by the phase field model based on the spectral strain decomposition.

Spectral (effective) stress decomposition

As an alternative to the spectral strain decomposition, the elastic energy density can be decomposed in the effective stress-space. For this purpose, the linear elastic (undegraded) relation (3.13) is used to calculate the principal values of the effective stress tensor

$$\tilde{\sigma}_\alpha = \lambda \ \mathrm{tr}\,\boldsymbol{\varepsilon} + 2\mu\, \varepsilon_\alpha \quad , \tag{3.208}$$

with $\alpha \in [1, 3]$. The elastic strain energy density is then decomposed into a portion related to the positive or negative first effective stress invariant

$$\mathrm{I}_{\tilde{\sigma}} := \mathrm{tr}\,\tilde{\boldsymbol{\sigma}} = (3\lambda + 2\mu)\ \mathrm{tr}\,\boldsymbol{\varepsilon} \tag{3.209}$$

and portions related positive and negative principal effective stresses, i.e.

$$\tilde{\psi}_0^+(\tilde{\boldsymbol{\sigma}}) := \frac{1}{4\mu}\left(\langle\tilde{\sigma}_1\rangle^2 + \langle\tilde{\sigma}_2\rangle^2 + \langle\tilde{\sigma}_3\rangle^2\right) - \frac{\lambda}{4\mu\,(3\lambda + 2\mu)}\langle \mathrm{tr}\,\tilde{\boldsymbol{\sigma}}\rangle^2 \quad , \tag{3.210}$$

$$\tilde{\psi}_0^-(\tilde{\boldsymbol{\sigma}}) := \frac{1}{4\mu}\left(\langle-\tilde{\sigma}_1\rangle^2 + \langle-\tilde{\sigma}_2\rangle^2 + \langle-\tilde{\sigma}_3\rangle^2\right) - \frac{\lambda}{4\mu\,(3\lambda + 2\mu)}\langle -\mathrm{tr}\,\tilde{\boldsymbol{\sigma}}\rangle^2 \quad . \tag{3.211}$$

Substitution of (3.208) into (3.194) and variation provides both the Cauchy stress tensor

$$\begin{aligned}\boldsymbol{\sigma}(\tilde{\boldsymbol{\sigma}}, s) = g(s)\left[\frac{1}{2\mu}\left(\langle\tilde{\sigma}_1\rangle\,\hat{\boldsymbol{M}}_1 + \langle\tilde{\sigma}_2\rangle\,\hat{\boldsymbol{M}}_2 + \langle\tilde{\sigma}_3\rangle\,\hat{\boldsymbol{M}}_3\right) - \frac{\lambda}{2\mu}\,\langle\mathrm{tr}\,\tilde{\boldsymbol{\sigma}}\rangle\,\boldsymbol{I}\right] \\ - \frac{1}{2\mu}\left(\langle\tilde{\sigma}_1\rangle\,\hat{\boldsymbol{M}}_1 + \langle\tilde{\sigma}_2\rangle\,\hat{\boldsymbol{M}}_2 + \langle\tilde{\sigma}_3\rangle\,\hat{\boldsymbol{M}}_3\right) + \frac{\lambda}{2\mu}\,\langle\mathrm{tr}\,\tilde{\boldsymbol{\sigma}}\rangle\,\boldsymbol{I} \quad ,\end{aligned} \tag{3.212}$$

and the crack driving force

$$f_s(\tilde{\boldsymbol{\sigma}}, s) = g'(s) \left[\frac{1}{4\mu} \left(\langle \tilde{\sigma}_1 \rangle^2 + \langle \tilde{\sigma}_2 \rangle^2 + \langle \tilde{\sigma}_3 \rangle^2 \right) - \frac{\lambda}{4\mu\,(3\lambda + 2\mu)} \langle \operatorname{tr} \tilde{\boldsymbol{\sigma}} \rangle^2 \right] \quad . \tag{3.213}$$

It should be noted, that with

$$\frac{\partial \mathrm{I}_{\tilde{\sigma}}}{\partial \boldsymbol{\varepsilon}} = (3\lambda + 2\mu)\, \boldsymbol{I} \qquad \text{and} \qquad \hat{\boldsymbol{M}}_\alpha := \frac{\partial \tilde{\sigma}_\alpha}{\partial \boldsymbol{\varepsilon}} = \lambda\, \boldsymbol{I} + 2\mu\, \hat{\boldsymbol{N}}_\alpha \quad , \tag{3.214}$$

the stress (3.212) still depends on the principal strain directions, not on the orientation of the crack. Thus, the spectral effective stress decomposition does not improve the phase field approach to model existing cracks. For $\nu = 0$ (which means $\lambda = 0$), the spectral effective stress decomposition and the spectral strain decomposition are equal.

Generalized spectral decomposition

Inspired by the decomposition in the (effective) stress space, which is a linear transformation from the strain space up to the split and separate treatment of each portion, one might think of other spectral decompositions. For this purpose, the auxiliary projected tensor

$$\tilde{\boldsymbol{\eta}} := \left(\mathbb{C}_0 \right)^{\tilde{\eta}} : \boldsymbol{\varepsilon} \tag{3.215}$$

with $\tilde{\eta} \in \mathbb{R}$ is introduced. By using the principal values of the stiffness tensor (2.87), the principal values of the projected tensor are

$$\tilde{\eta}_\alpha = \frac{1}{3} \left((3\lambda + 2\mu)^{\tilde{\eta}} - (2\mu)^{\tilde{\eta}} \right) \operatorname{tr} \boldsymbol{\varepsilon} + (2\mu)^{\tilde{\eta}}\, \varepsilon_\alpha \quad , \tag{3.216}$$

with $\alpha \in [1,\, 3]$. Its first invariant is

$$\mathrm{I}_{\tilde{\eta}} := \operatorname{tr} \tilde{\boldsymbol{\eta}} = (3\lambda + 2\mu)^{\tilde{\eta}}\, \operatorname{tr} \boldsymbol{\varepsilon} \quad . \tag{3.217}$$

The (undegraded) strain energy (2.82) can be rewritten in terms of the projected tensor (3.215) according to

$$\psi_0 = \frac{1}{2} \tilde{\boldsymbol{\eta}} : \left(\mathbb{C}_0 \right)^{(1-2\tilde{\eta})} : \tilde{\boldsymbol{\eta}} \quad . \tag{3.218}$$

Accordingly, the generalized spectral decomposition is

$$\begin{aligned} \tilde{\psi}_0^+(\tilde{\boldsymbol{\eta}}) := {} & \frac{1}{6} \left[(3\lambda + 2\mu)^{1-2\tilde{\eta}} - (2\mu)^{1-2\tilde{\eta}} \right] \langle \operatorname{tr} \tilde{\boldsymbol{\eta}} \rangle^2 \\ & + \frac{1}{2} (2\mu)^{1-2\tilde{\eta}} \left(\langle \tilde{\eta}_1 \rangle^2 + \langle \tilde{\eta}_2 \rangle^2 + \langle \tilde{\eta}_3 \rangle^2 \right) \quad , \end{aligned} \tag{3.219}$$

$$\begin{aligned} \tilde{\psi}_0^-(\tilde{\boldsymbol{\eta}}) := {} & \frac{1}{6} \left[(3\lambda + 2\mu)^{1-2\tilde{\eta}} - (2\mu)^{1-2\tilde{\eta}} \right] \langle - \operatorname{tr} \tilde{\boldsymbol{\eta}} \rangle^2 \\ & + \frac{1}{2} (2\mu)^{1-2\tilde{\eta}} \left(\langle -\tilde{\eta}_1 \rangle^2 + \langle -\tilde{\eta}_2 \rangle^2 + \langle -\tilde{\eta}_3 \rangle^2 \right) \quad . \end{aligned} \tag{3.220}$$

By substituting (3.219) into (3.194), variation provides with

$$\frac{\partial \mathrm{I}_{\tilde{\eta}}}{\partial \boldsymbol{\varepsilon}} = (3\lambda + 2\mu)^{\tilde{\eta}}\, \boldsymbol{I} \tag{3.221}$$

and the projection tensor

$$\boldsymbol{M}_\alpha := \frac{\partial \tilde{\eta}_\alpha}{\partial \boldsymbol{\varepsilon}} = \frac{1}{3}\left((3\lambda + 2\mu)^{\tilde{\eta}} - (2\mu)^{\tilde{\eta}}\right) \boldsymbol{I} + (2\mu)^{\tilde{\eta}}\, \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_\alpha \tag{3.222}$$

both the Cauchy stress tensor

$$\begin{aligned}\boldsymbol{\sigma}(\tilde{\boldsymbol{\eta}}, s) = g(s) \Big[& \frac{1}{3}\left((3\lambda + 2\mu)^{1-\tilde{\eta}} - (3\lambda + 2\mu)^{\tilde{\eta}}\, (2\mu)^{1-2\tilde{\eta}}\right) \langle \operatorname{tr} \tilde{\boldsymbol{\eta}} \rangle\, \boldsymbol{I} \\ & + (2\mu)^{1-2\tilde{\eta}} \left(\langle \tilde{\eta}_1 \rangle\, \boldsymbol{M}_1 + \langle \tilde{\eta}_2 \rangle\, \boldsymbol{M}_2 + \langle \tilde{\eta}_3 \rangle\, \boldsymbol{M}_3 \right) \Big] \\ & - \frac{1}{3}\left((3\lambda + 2\mu)^{1-\tilde{\eta}} - (3\lambda + 2\mu)^{\tilde{\eta}}\, (2\mu)^{1-2\tilde{\eta}}\right) \langle - \operatorname{tr} \tilde{\boldsymbol{\eta}} \rangle\, \boldsymbol{I} \\ & - (2\mu)^{1-2\tilde{\eta}} \left(\langle -\tilde{\eta}_1 \rangle\, \boldsymbol{M}_1 + \langle -\tilde{\eta}_2 \rangle\, \boldsymbol{M}_2 + \langle -\tilde{\eta}_3 \rangle\, \boldsymbol{M}_3 \right) \quad , \end{aligned} \tag{3.223}$$

and the crack driving force

$$\begin{aligned} f_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}, s) = g'(s) \Big[& \frac{1}{6}\left[(3\lambda + 2\mu)^{1-2\tilde{\eta}} - (2\mu)^{1-2\tilde{\eta}} \right] \langle \operatorname{tr} \tilde{\boldsymbol{\eta}} \rangle^2 \\ & + \frac{1}{2}(2\mu)^{1-2\tilde{\eta}} \left(\langle \tilde{\eta}_1 \rangle^2 + \langle \tilde{\eta}_2 \rangle^2 + \langle \tilde{\eta}_3 \rangle^2 \right) \Big] \quad . \end{aligned} \tag{3.224}$$

Since the stress of the generalized spectral decomposition (3.223) depends via (3.222) on the principal strain directions instead of the crack orientation it does not improve the phase field approach to model existing cracks. It should be further noted, that also in the special case $\tilde{\eta} = 0.5$ the crack driving force

$$f_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}, s) = g'(s)\, \frac{1}{2}\left(\langle \tilde{\eta}_1 \rangle^2 + \langle \tilde{\eta}_2 \rangle^2 + \langle \tilde{\eta}_3 \rangle^2 \right) \tag{3.225}$$

including

$$\tilde{\eta}_\alpha = \frac{1}{3}\left(\sqrt{3\lambda + 2\mu} - \sqrt{2\mu} \right) \operatorname{tr} \boldsymbol{\varepsilon} + \sqrt{2\mu}\, \varepsilon_\alpha \tag{3.226}$$

and the stress tensor

$$\begin{aligned} \boldsymbol{\sigma}(\tilde{\boldsymbol{\eta}}, s) = & g(s) \sqrt{2\mu} \left(\langle \tilde{\eta}_1 \rangle\, \boldsymbol{M}_1 + \langle \tilde{\eta}_2 \rangle\, \boldsymbol{M}_2 + \langle \tilde{\eta}_3 \rangle\, \boldsymbol{M}_3 \right) \\ & - \sqrt{(2\mu} \left(\langle -\tilde{\eta}_1 \rangle\, \boldsymbol{M}_1 + \langle -\tilde{\eta}_2 \rangle\, \boldsymbol{M}_2 + \langle -\tilde{\eta}_3 \rangle\, \boldsymbol{M}_3 \right) \quad , \end{aligned} \tag{3.227}$$

with

$$\boldsymbol{M}_\alpha = \frac{1}{3}\left(\sqrt{3\lambda + 2\mu} - \sqrt{2\mu} \right) \boldsymbol{I} + \sqrt{2\mu}\, \overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_\alpha \quad , \tag{3.228}$$

still contain λ. Thus, the stress in the crack plane is erroneously degraded.

Further spectral decompositions

By Freddi and Royer-Carfagni (2010) the tension-compression asymmetry is also introduced based on the spectral decomposition. However, they propose a slightly different treatment of the principal strain values (by also taking Poisson's ratio into account) so that only negative principal stresses exists. This treatment was inspired by the "no-tension" masonry-like material which has been formulated by Del Piero (1989).

Dally and Weinberg (2017), and later on, Bilgen et al. (2018) and Bilgen and Weinberg (2019) proposed a spectral strain decomposition according to

$$\psi_0 = \psi_0^+ + \psi_0^- = \frac{1}{2}\boldsymbol{\varepsilon}^+ : \mathbb{C}_0 : \boldsymbol{\varepsilon}^+ + \frac{1}{2}\boldsymbol{\varepsilon}^- : \mathbb{C}_0 : \boldsymbol{\varepsilon}^- \tag{3.229}$$

with the positive and negative strain tensors according to (2.190). However, as discussed in App. B.8, the decomposition does not comply with (3.189) if not all principal strains have the same sign or $\nu = 0$. Since the positive and negative energy portions are not related to the energy of the sound material, (3.229) is no valid decomposition.

3.5.2 Models based on volumetric-deviatoric decomposition

The second widely-used approach is based on the *volumetric-deviatoric decomposition*. In terms of phase field modeling to fracture[13]. Amor et al. (2009) introduced their so-called "unilateral contact" model with the clear intention to "avoid crack interpenetration in compression". The formulation is guided by the idea that crack closure can be detected by a negative volume change $\varepsilon_{\text{vol}} = \operatorname{tr}\boldsymbol{\varepsilon} < 0$. Therefore, the volumetric strain energy is decomposed into positive and negative portions. Since crack opening is governed by volume expansion and distortion both portions must be degraded by the phase field to provide a traction-free crack surface. Accordingly, the strain energy portions for (3.194) are

$$\psi_0^+(\boldsymbol{\varepsilon}) := \frac{K}{2}\langle \operatorname{tr}\boldsymbol{\varepsilon}\rangle^2 + \mu\,(\boldsymbol{\varepsilon}_{\text{dev}} : \boldsymbol{\varepsilon}_{\text{dev}}) \quad , \tag{3.230}$$

$$\psi_0^-(\boldsymbol{\varepsilon}) := \frac{K}{2}\langle -\operatorname{tr}\boldsymbol{\varepsilon}\rangle^2 \quad . \tag{3.231}$$

Approaches based on the volumetric-deviatoric decomposition of strain energy have already been used in continuum damage mechanics for similar reasons, see, e.g., (2.194) in Sect. 2.2.8. Further, it should be noted, that the original decomposition (3.230) is based on a dimension-dependent bulk modulus $K_{\text{n}} := \lambda + 2\mu/n_{\text{dim}}$ with $n_{\text{dim}} \in [1, 3]$ (Amor et al., 2009) and has further been used by Schlüter et al. (2014) and Ambati et al. (2015a; 2015b). Since most related fracture problems (Sect. 2.2) are, however, discussed in terms of plane strain, plane stress or axisymmetric settings (which in fact are three-dimensional)

[13] In fact, Amor et al. (2009) called their approach "regularized formulation of the variational brittle fracture".

only the bulk modulus K according to (2.91) is regarded here. Variation provides both the Cauchy stress tensor

$$\boldsymbol{\sigma}(\boldsymbol{\varepsilon}, s) = g(s) \left[K \langle \operatorname{tr} \boldsymbol{\varepsilon} \rangle \boldsymbol{I} + 2\mu \boldsymbol{\varepsilon}_{\mathrm{dev}} \right] - K \langle - \operatorname{tr} \boldsymbol{\varepsilon} \rangle \boldsymbol{I} \quad , \tag{3.232}$$

and the crack driving force

$$f_{\mathrm{s}}(\boldsymbol{\varepsilon}, s) = g'(s) \left[\frac{K}{2} \langle \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 + \mu \left(\boldsymbol{\varepsilon}_{\mathrm{dev}} : \boldsymbol{\varepsilon}_{\mathrm{dev}} \right) \right] \quad . \tag{3.233}$$

Since the last term in (3.232) is not degraded the non-interpenetration of crack lips (at $g = 0$) is to be ensured by a hydrostatic stress state. In addition, because of the partial coupling of strain energy and the phase field in (3.233), phase field evolution is only governed by volume expansion and distortion.

Inadequate model behavior for uniaxial compression

By continuing the one-dimensional analysis of Sect. 3.4 also for compression, e.g. for the quadratic single-well function in combination with the quadratic degradation function, the maximum compressive stress (absolute value) is

$$\sigma_{\mathrm{c}} = \sqrt{\frac{81}{512\,(1+\nu)}} \sqrt{\frac{\mathcal{G}_{\mathrm{c}}\, E}{\ell}} \quad , \tag{3.234}$$

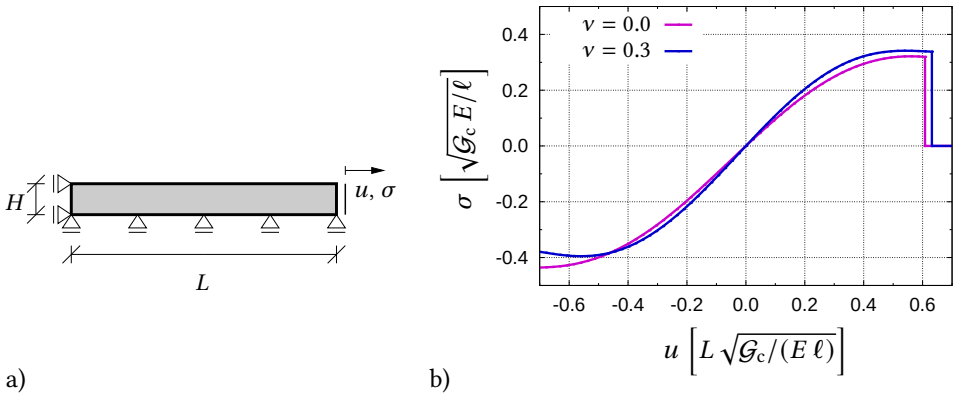
which in fact can be interpreted as strength. By utilizing (3.130) the ratio of compressive strength to tensile strength is

$$\frac{\sigma_{\mathrm{c}}}{\sigma_{\mathrm{t}}} = \sqrt{\frac{3}{2+2\nu}} \approx \begin{cases} 1.225 & \text{for} \quad \nu = 0.0 \\ 1.074 & \text{for} \quad \nu = 0.3 \\ 1 & \text{for} \quad \nu \to 0.5 \end{cases} . \tag{3.235}$$

As already reported by Amor et al. (2009) and further analyzed by Li et al. (2016), the comparatively small difference between the compressive and tensile strength does not correspond to experimental results, especially for brittle materials. Thus, phase field evolution and consequently stiffness degradation take place even under compression if the load is not hydrostatic.

In order to further investigate the non-local asymmetric response a bar with plane strain conditions as depicted in Fig. 3.22a is separately subjected to tension and compression. For monotonically increasing loading $u > 0$ the phase field homogeneously decreases until $s \approx s_{\mathrm{t}} = 0.75$ is reached, see (3.129). Then the phase field solution becomes inhomogeneous and the strain localizes, leading to a sudden drop of the stress response to zero as shown in Fig. 3.22b. Applying a monotonically decreasing displacement $u < 0$ also leads to a homogeneously decreasing phase field in a quite similar manner as for

tension. The maximum negative stress $|\sigma_c|$ obviously is of similar size which corresponds with (3.235). However the solution does not localize, so no crack is formed and the stiffness continuously degrades. This might be explained by the quite complex hydrostatic stress state obtained under compression.



a) b)

Figure 3.22: Numerical evaluation of tension and compression response: a) boundary value problem, b) global stress response.

Inadequate model behavior at cracks subjected to compression

The capability of the phase field model based on the volumetric-deviatoric split to approximate crack closure and thus prevent interpenetration of the crack flanks is tested by compressing a through-cracked block (Fig. 3.17c). The initial crack is modeled by the phase field as proposed by Strobl and Seelig (2015), see Fig. 3.17a. Since the deviatoric stress is degraded to zero at a crack ($g = 0$) uniaxial compression perpendicular to the crack results in a hydrostatic stress state at the crack. Therefore the stress response is inhomogeneous as depicted in Fig. 3.23a and does not correspond with the expected (homogeneous) uniaxial stress state. Moreover, the unphysical hydrostatic stress state at the crack requires compressive forces parallel to the crack, which in fact are equal to the normal stress. In the case that the phase field modeled crack crosses a traction-free outer boundary there are no forces from outside. In addition, shear forces are almost completely degraded in the vicinity of the crack. Thus extremely large displacements in the crack plane occur as shown in Fig. 3.23b. They are causing extremely large shear forces in the undegraded configuration which then are degraded by a factor of almost zero and provide the required normal forces. However, if the fully broken state $s \approx 0$ is wider than a single row of elements, even infinite lateral displacements do not cause a significant hydrostatic stress any more and thus the hydrostatic stress and corresponding mechanical contact forces across the crack surface tend to zero. Thus, if the crack crosses a free outer boundary the model completely looses its ability to cause elastic response or causes elastic response at the expense of almost infinite displacements. This problem does not occur

if the crack is completely enclosed by material or (potentially artificial) displacement boundary conditions are imposed where the crack crosses the free outer boundary as done, e.g., by Schlüter et al. (2014). Due to the dominating compressive hydrostatic stress state in the vicinity of a phase field modeled crack the model's mechanical response can be characterized as "fluid-like" behavior. In the compressed block test, a layer of liquid is compressed between the upper and lower part of the solid. As long as the boundary is "sealed off" and the liquid cannot run out the contact force is equal to the liquid's pressure, but as soon as the boundary is not adequately prepared the liquid leaks out and the pressure (and with it the contact force) drops to zero.

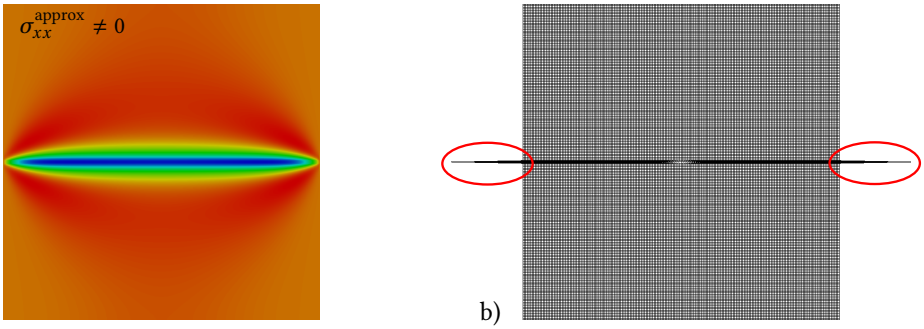


Figure 3.23: Validation of mechanical behavior of the phase field approach based on volumetric-deviatoric decomposition at a crack subjected to compression: a) inhomogeneous stress response due to (unphysical) hydrostatic stress state at the crack, b) resulting deformations with extremely large displacements in the degraded zone.

Further volumetric-deviatoric decompositions

The volumetric-deviatoric decomposition can also be applied on the effective stress tensor. Since the traces of effective stress and strain are proportional to each other, i.e.

$$\operatorname{tr} \tilde{\boldsymbol{\sigma}} = \operatorname{tr}\left(K \operatorname{tr} \boldsymbol{\varepsilon}\, \boldsymbol{I} + 2\mu\, \boldsymbol{\varepsilon}_{\text{dev}}\right) = 3K \operatorname{tr} \boldsymbol{\varepsilon} \quad , \tag{3.236}$$

their sign is equal. Thus, with $\tilde{\boldsymbol{\sigma}}_{\text{dev}} = 2\mu\, \tilde{\boldsymbol{\varepsilon}}_{\text{dev}}$ the volumetric-deviatoric split of the effective stress tensor results in

$$\psi_0^+(\boldsymbol{\varepsilon}) := \frac{1}{18K} \langle 3K\, \boldsymbol{\varepsilon} \rangle^2 + \frac{1}{4\mu} \left(4\mu^2\, \boldsymbol{\varepsilon}_{\text{dev}} : \boldsymbol{\varepsilon}_{\text{dev}}\right) = \frac{K}{2} \langle \boldsymbol{\varepsilon} \rangle^2 + \mu \left(\boldsymbol{\varepsilon}_{\text{dev}} : \boldsymbol{\varepsilon}_{\text{dev}}\right) \quad , \tag{3.237}$$

$$\psi_0^-(\boldsymbol{\varepsilon}) := \frac{1}{18K} \langle -3K\, \boldsymbol{\varepsilon} \rangle^2 = \frac{K}{2} \langle -\boldsymbol{\varepsilon} \rangle^2 \tag{3.238}$$

and is identical with the volumetric-deviatoric decomposition in the strain space (3.232).

Lancioni and Royer-Carfagni (2009) also introduced a phase field approach to fracture based on the volumetric-deviatoric decomposition. Their focus, however, was on the modeling of shear-like fracture as it appears in the ashlar masonry work of the French

Panthéon. Thus, the intention rather lies on the modification of the crack driving force to promote "mode II fracture" caused by coalescence of micro-cracked shear bands. This is done by coupling only the deviatoric strain energy portion via a degradation function $g(s)$ with the phase field. The corresponding strain energy density portions are

$$\psi_0^+(\boldsymbol{\varepsilon}) \coloneqq \mu \left(\boldsymbol{\varepsilon}_{\mathrm{dev}} : \boldsymbol{\varepsilon}_{\mathrm{dev}}\right) \quad , \tag{3.239}$$

$$\psi_0^-(\boldsymbol{\varepsilon}) \coloneqq \frac{K}{2} \left(\operatorname{tr} \boldsymbol{\varepsilon}\right)^2 \quad . \tag{3.240}$$

3.5.3 First conclusion on isotropic strain energy decompositions

The incorporation of a tension-compression asymmetry by decomposing the elastic energy density and applying only a partial coupling with the phase field is an elegant but, regarding the outcome (see above sections), also limited approach. By regarding only the correct mechanical behavior at a (macroscopic) crack, both widely-used approaches, the volumetric-deviatoric and the spectral decomposition, fail in some situations. With both approaches, it is not possible to obtain a homogeneous stress response for compression due to erroneous degradation of the lateral stress components (except the spectral strain decomposition if $\nu = 0$). More severely, the spectral split fails in reproducing traction-free sliding of the crack surfaces while the volumetric-deviatoric split causes unphysical hydrostatic stress states in the vicinity of the crack, which make the model to fail if the crack crosses the free boundary. Thus, the question arises, how to model the response of a macroscopic crack correctly, e.g. by finding an appropriate split?

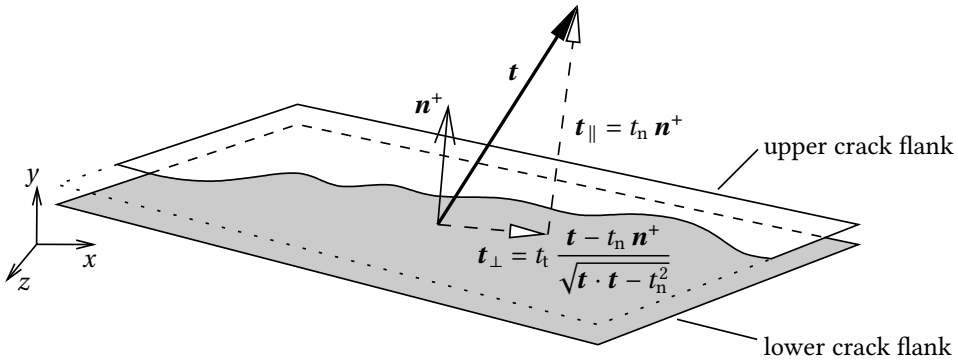


Figure 3.24: Regarding a discrete crack with a hypothetical traction vector $\boldsymbol{t}$ and its in-plane component t_{t} and out-of-plane component t_{n}.

For this purpose a hypothetical discrete crack with its flanks as depicted in Fig. 3.24 is regarded. The orientation of the crack surface is described by its outward normal vector $\boldsymbol{n}^+$. Since in terms of infinitesimal deformations the orientation of the opposite surface does not change, the orientation of the opposite crack surface (partially shown in Fig. 3.24) is given by $\boldsymbol{n}^- = -\boldsymbol{n}^+$. To obtain the orientation independent of the positive or negative

crack flank (the decomposition is independent of the direction of the vector) the second order tensor

$$\boldsymbol{N}_{\mathrm{c}} := \boldsymbol{n}^{+} \otimes \boldsymbol{n}^{+} = \boldsymbol{n}^{-} \otimes \boldsymbol{n}^{-} \tag{3.241}$$

is used. The mechanical response of the crack is characterized by a non-positive normal stress and, in case of a frictionless crack surface, by a zero tangential stress, i.e. $t_{\mathrm{n}} \leq 0$ and $t_{\mathrm{t}} = 0$. In a homogeneous setting without the hypothetical discrete crack, the stress tensor can be decomposed according to

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_{\mathrm{nn}} + \boldsymbol{\sigma}_{\mathrm{nt}} + \boldsymbol{\sigma}_{\mathrm{tt}} \quad . \tag{3.242}$$

It consists of the normal stress normal to the crack plane

$$\boldsymbol{\sigma}_{\mathrm{nn}} := \underbrace{(\boldsymbol{\sigma} : \boldsymbol{N}_{\mathrm{c}})}_{\leq 0} \boldsymbol{N}_{\mathrm{c}} \tag{3.243}$$

and a stress tensor related to the tangential stress on the crack plane

$$\boldsymbol{\sigma}_{\mathrm{nt}} := \boldsymbol{\sigma} \cdot \boldsymbol{N}_{\mathrm{c}} + \boldsymbol{N}_{\mathrm{c}} \cdot \boldsymbol{\sigma} - 2(\boldsymbol{\sigma} : \boldsymbol{N}_{\mathrm{c}})\, \boldsymbol{N}_{\mathrm{c}} \quad . \tag{3.244}$$

The third stress tensor $\boldsymbol{\sigma}_{\mathrm{tt}}$ contains stresses in the plane and is for now irrelevant. The focus lies on the shear stress tensor which should vanish for tension and compression, i.e.

$$\boldsymbol{\sigma}_{\mathrm{nt}} \overset{!}{=} \boldsymbol{0} \quad . \tag{3.245}$$

Since the crack direction is — so far — not available in the phase field approach and thus a priori unknown the relation (3.245) must be fulfilled for any direction. By using the principal stresses $\sigma_1 \dots \sigma_3$ and related directions $\boldsymbol{N}_1 \dots \boldsymbol{N}_3$ (aligned with the principal strain directions) and making use of $\boldsymbol{N}_3 = \boldsymbol{I} - \boldsymbol{N}_1 - \boldsymbol{N}_2$ one obtains

$$\begin{aligned} \boldsymbol{\sigma}_{\mathrm{nt}} = {} & \big(\sigma_1 - \sigma_3\big)\, \boldsymbol{N}_1 \cdot \boldsymbol{N}_{\mathrm{c}} + \big(\sigma_1 - \sigma_3\big)\, \boldsymbol{N}_{\mathrm{c}} \cdot \boldsymbol{N}_1 - 2\big(\sigma_1 - \sigma_3\big)\, \big(\boldsymbol{N}_1 : \boldsymbol{N}_{\mathrm{c}}\big)\, \boldsymbol{N}_{\mathrm{c}} \\ & + \big(\sigma_2 - \sigma_3\big)\, \boldsymbol{N}_2 \cdot \boldsymbol{N}_{\mathrm{c}} + \big(\sigma_2 - \sigma_3\big)\, \boldsymbol{N}_{\mathrm{c}} \cdot \boldsymbol{N}_2 - 2\big(\sigma_2 - \sigma_3\big)\, \big(\boldsymbol{N}_2 : \boldsymbol{N}_{\mathrm{c}}\big)\, \boldsymbol{N}_{\mathrm{c}} = \boldsymbol{0} \quad . \end{aligned} \tag{3.246}$$

Equation (3.246) is either satisfied if one principal stress/strain direction coincidentally corresponds to the crack direction, e.g. $\boldsymbol{N}_1 = \boldsymbol{N}_{\mathrm{c}}$, or if all principal stresses are equal, i.e.

$$\sigma_1 = \sigma_2 = \sigma_3 \quad . \tag{3.247}$$

Since the crack direction is fixed while principal strain directions depend on the applied loading/deformation the first case is not obtained in general. Thus, the shear stress on a crack surface vanishes only in the case of a hydrostatic stress state. As discussed in Section 3.5.2, severe problems related to a non-zero hydrostatic stress state occur. Thus, it can be concluded, that in the present (isotropic) phase field framework no decomposition of the strain energy density exists, which is able to reproduce the mechanical behavior of a macroscopic crack appropriately.

Apart from the physical situation, phase field approaches using tension-compression splits increase the computational costs. Since the spectral strain decomposition contains

six different cases (in three dimensions), corresponding phase field approaches converge significantly slower to the solution compared to models based on the volumetric-deviatoric decomposition, which includes only two different cases (see also Ambati et al., 2015a). In addition, the volumetric-deviatoric decomposition is more robust. Since the decomposition is carried out at each integration point separately and the cases can alternate in every iteration, there might be the situation that no convergence is obtained. Thus, it is a common practice to increase the convergence tolerance when using the spectral strain decomposition. Another solution can be the regularization of the non-smooth stress-strain relation (Schmitt et al., 2015). For example, the one-dimensional stiffness in (3.200) can be approximated by the function

$$\tilde{E}(\epsilon_k) = \frac{E}{2}\left[\left(1 - g(s)\right)\,\tanh(-\varepsilon/\epsilon_k) + 1 + g(s)\right] \tag{3.248}$$

with the additional regularization parameter $0 < \epsilon_k \ll 1$ to compute the stress $\sigma = \tilde{E}(\epsilon_k)\,\varepsilon$. Figures 3.25 shows the negligible difference between the original non-smooth and the smooth regularized function as $\epsilon_k \to 0$. While in 1D problems convergence is usually obtained, regularizations of the type of (3.248) might improve the situation in two- and three-dimensional problems. Since the present work is focused on the physical aspects, improvements on the numerical robustness are out of scope.

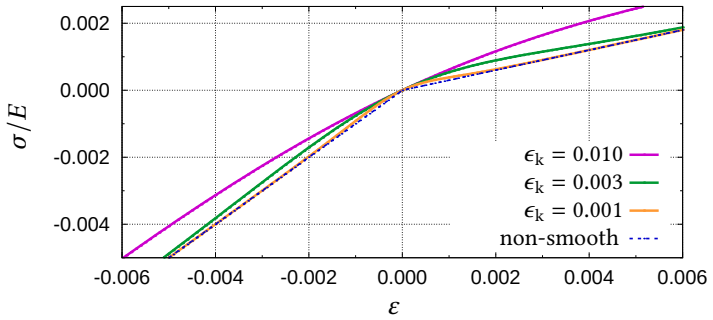


Figure 3.25: Comparison of stress calculations for the original non-smooth tension-compression split in 1D (dashed line) and the smooth regularized function as $\epsilon_k \to 0$. Here, the tensile stiffness is degraded by $g(s) = 0.3$.

3.5.4 Application in evolution equation only

The changing physical conditions caused by opened or closed cracks and the related stiffness degradation in continuous approaches to fracture are clear, thus they can be assigned as boundary conditions. By contrast, the crack driving forces and related crack formation and propagation can vary between different materials, e.g. to account for differences between ductile and brittle behavior, see Sect. 2.2. Thus, their modeling has to be adaptable (to some extent), i.e. they have to be constitutively chosen. This

contradicts the variational approach (at least for the current structure of the approach), which is typically used to derive the governing equations in phase field models to fracture (Sect. 3.2).

As already introduced in Sect. 3.2.2, it is possible — at least to some extent — to chose the crack driving energy independent from the strain energy density degradation. If cracks remain open once they have formed, the stiffness recovery to model closed cracks is not required. Accordingly, the stress tensor is not decomposed and all components are degraded according to

$$\boldsymbol{\sigma} := g(s)\, \mathbb{C}_0 : \boldsymbol{\varepsilon} \quad . \tag{3.249}$$

Mainly to suppress phase field evolution caused by compression, but also to be flexible regarding the modes of crack driving forces, the crack driving energy is constitutively chosen. Since cracks in brittle materials tend to propagate in mode I, an obvious assumption is that they are driven by positive normal strains Miehe et al. (2010c) or stresses Miehe et al. (2015). Accordingly, the crack driving energy in phase field models can be formulated as

$$D_{\mathrm{s}}(\boldsymbol{\varepsilon}) := \frac{\lambda}{2} \langle \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 + \mu \left(\langle \varepsilon_1 \rangle^2 + \langle \varepsilon_2 \rangle^2 + \langle \varepsilon_3 \rangle^2 \right) \tag{3.250}$$

or, in terms of the effective stress tensor (3.13), more precisely, its spectral decomposition

$$\tilde{\boldsymbol{\sigma}} = \sum_{\alpha=1}^{3} \tilde{\sigma}_\alpha \, \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha \quad , \tag{3.251}$$

as

$$D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}) := \frac{1+\nu}{2E} \left(\langle \tilde{\sigma}_1 \rangle^2 + \langle \tilde{\sigma}_2 \rangle^2 + \langle \tilde{\sigma}_3 \rangle^2 \right) - \frac{\nu}{2E} \langle \operatorname{tr}(\tilde{\boldsymbol{\sigma}}) \rangle^2 \quad . \tag{3.252}$$

In conjunction with (3.249), these choices neglect possible crack closure effects, i.e. crack flanks can interpenetrate in each other, but automatically guarantees a traction-free crack for $s \to 0$. It should be noted, that the effective stress based driving energy (3.252) contains positive addends (in fact energetic contributions) with negative sign. For this reason a separate treatment (i.e. degradation) of each addend might artificially increase[14] the internal energy. By contrast, strain based crack driving energies with separate treatment of each strain energy portion, e.g.

$$D_{\mathrm{s}}(\boldsymbol{\varepsilon}) := \mu \langle \varepsilon_1 \rangle^2 \qquad \text{or} \qquad D_{\mathrm{s}}(\boldsymbol{\varepsilon}) := \frac{\lambda}{2} \langle \operatorname{tr} \boldsymbol{\varepsilon} \rangle^2 + \mu \langle \varepsilon_1 \rangle^2 \quad , \tag{3.253}$$

[14] The formulation of arbitrary crack driving forces f_{s} neither depending on the current phase field value nor based on portions of strain energy, see e.g. Bilgen and Weinberg (2019), entails the risk of being thermodynamically inconsistent. While additional dissipation of strain energy in the event of damage or fracture might be motivated by physics the arbitrary creation of energy, i.e. the creation of surface energy without supplying other forms of energy, does not seem to be a reasonable approach.

can be formulated. By using the generalized spectral decomposition (3.219) formulated in terms of the auxiliary projected tensor (3.215) a generalized crack driving energy might be

$$D_{\mathrm{s}}(\tilde{\boldsymbol{\eta}}) := \frac{p_0}{6}\left[(3\lambda+2\mu)^{1-2\tilde{\eta}} - (2\mu)^{1-2\tilde{\eta}}\right]\langle \operatorname{tr}\tilde{\boldsymbol{\eta}}\rangle^2 + \frac{1}{2}(2\mu)^{1-2\tilde{\eta}}\left(p_1\langle\tilde{\eta}_1\rangle^2 + p_2\langle\tilde{\eta}_2\rangle^2 + p_3\langle\tilde{\eta}_3\rangle^2\right) \quad , \tag{3.254}$$

with $p_0, p_1, p_2, p_3 \in [0,\,1]$ if $\tilde{\eta} \geq 0.5$ and $p_0 = p_1 = p_2 = p_3$ with $p_0 \in [0,\,1]$ if $\tilde{\eta} < 0.5$.

Primarily to improve the computational efficiency[15], Ambati et al. (2015a) proposed the use of (3.249) in combination with (3.250) and an additional criterion to prevent the interpenetration of crack faces as so-called "hybrid" phase field formulation (one should not be confused with the "hybrid fracture initiation criterion" introduced in Sect. 2.2.6). They also worked out the link between non-variational phase field approaches to fracture and (non-local) gradient damage models. It should be noted, that in absence of the variational structure only some portion of the dissipated energy is transferred to the surface energy. To show its thermodynamic consistency, a short discussion on the dissipated energy, which must be non-negative, is provided in the following.

Discussion of non-negative energy dissipation

The following discussion serves to clarify the amount of artificial energy dissipation introduced by giving up the variational structure in phase field modeling of fracture. We start from the balance of internal and external power (at a given state $\boldsymbol{u}$ and s)

$$\underbrace{\int_\Omega \rho\,\ddot{\boldsymbol{u}}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}V}_{=\dot{K}} + \int_\Omega \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V = \underbrace{\int_{\partial\Omega_{\mathrm{t}}} \bar{\boldsymbol{t}}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}A}_{=P_{\mathrm{ext}}} \tag{3.255}$$

and the rate of change of elastic energy

$$\dot{E}_{\mathrm{el}} = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \psi_{\mathrm{el}}(\boldsymbol{\varepsilon}, s)\,\mathrm{d}V = \int_\Omega \left(\underbrace{\frac{\partial\psi_{\mathrm{el}}}{\partial\boldsymbol{\varepsilon}}}_{=\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}} + \frac{\partial\psi_{\mathrm{el}}}{\partial s}\,\dot{s}\right)\mathrm{d}V \quad . \tag{3.256}$$

On account of the choice of a full strain energy degradation (3.12) the corresponding elastic energy density can be written as

$$\psi_{\mathrm{el}}(\boldsymbol{\varepsilon}, s) = g(s)\,\psi_{0\,\mathrm{el}}(\boldsymbol{\varepsilon}) \tag{3.257}$$

[15] After comparing a phase field model based on the spectral strain decomposition (there called "anisotropic") and their "hybrid" formulation, Ambati et al. (2015a) concluded that "this formulation was shown to lead to results very similar to those of the available anisotropic models, at a small fraction of their computational cost", without discovering the severe shortcomings of the spectral strain decomposition discussed in Sect. 3.5.1.

where $\psi_{0\,\mathrm{el}}(\boldsymbol{\varepsilon})$ denotes the elastic energy density of the intact material. The latter can, equivalently, be expressed in terms of the effective stress tensor $\tilde{\boldsymbol{\sigma}}$ introduced in (3.13) as

$$\tilde{\psi}_{0\,\mathrm{el}}(\tilde{\boldsymbol{\sigma}}) = \psi_{0\,\mathrm{el}}(\boldsymbol{\varepsilon}) \quad . \tag{3.258}$$

Combining (3.255) and (3.256) and inserting (3.257) and (3.258) yields

$$P_{\mathrm{ext}} - \dot{K} - \dot{E}_{\mathrm{el}} = -\int_{\Omega} g'(s)\, \tilde{\psi}_{0\,\mathrm{el}}(\tilde{\boldsymbol{\sigma}})\, \dot{s}\, \mathrm{d}V \quad . \tag{3.259}$$

On the other hand, the rate of change of crack surface energy (3.15) is

$$\dot{E}_{\mathrm{s}} = \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega} \mathcal{G}_{\mathrm{c}} \underbrace{\left[\frac{(1-s)^2}{2\ell} + \frac{\ell}{2} |\nabla s|^2 \right]}_{= \gamma_\ell(s, \nabla s)} \mathrm{d}V = -\int_{\Omega} \mathcal{G}_{\mathrm{c}} \left[\frac{1-s}{\ell} + \ell\, \Delta s \right] \dot{s}\, \mathrm{d}V \tag{3.260}$$

and inserting the phase field evolution equation (3.33) yields

$$\dot{E}_{\mathrm{s}} = -\int_{\Omega} g'(s)\, D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}})\, \dot{s}\, \mathrm{d}V \quad . \tag{3.261}$$

Combining (3.259) and (3.261) one gets

$$P_{\mathrm{ext}} - \dot{K} - \dot{E}_{\mathrm{el}} = \dot{E}_{\mathrm{s}} \underbrace{- \int_{\Omega} g'(s) \left(\tilde{\psi}_{0\,\mathrm{el}}(\tilde{\boldsymbol{\sigma}}) - D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}) \right) \dot{s}\, \mathrm{d}V}_{=: \mathcal{D}_{\mathrm{art}}} \tag{3.262}$$

with the artificial dissipation $\mathcal{D}_{\mathrm{art}}$ inspired by (2.74). The choice of the crack driving energy $D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}})$ in (3.252) is guided by the property

$$0 \leq D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}) \leq \tilde{\psi}_{0\,\mathrm{el}}(\tilde{\boldsymbol{\sigma}}) \tag{3.263}$$

which can easily be shown by using the Cauchy-Schwarz inequality. Hence, the artificial dissipation $\mathcal{D}_{\mathrm{art}}$ introduced in (3.262) by which the rate of energy release $P_{\mathrm{ext}} - \dot{K} - \dot{E}_{\mathrm{el}}$ exceeds the rate of crack surface energy $\dot{E}_{\mathrm{s}}$ is non-negative for crack growth $\dot{s} < 0$.

The amount of artificially dissipated energy depends on the presence of negative effective stresses in the region where the phase field evolves with $\dot{s} < 0$. For the constitutive choice $D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}) = \tilde{\psi}_{0\,\mathrm{el}}(\tilde{\boldsymbol{\sigma}})$ the phase field model would be identical with the variationally derived model of Bourdin et al. (2000) and corresponds to the classical Griffith theory (with $\mathcal{D}_{\mathrm{art}} = 0$) discussed in Sect. 2.2.4. In the other limiting case $D_{\mathrm{s}}(\tilde{\boldsymbol{\sigma}}) = 0$ the phase field solution is only governed by the equality between the local and non-local part of the surface energy (3.15) with $\dot{s} = 0$ and consequently $\mathcal{D}_{\mathrm{art}} = 0$.

If a phase field model is used in conjunction with the "crack-like" irreversibility constraint (described in App. C.3) the phase field can recover ($\dot{s} > 0$) where the broken state is not reached and artificially dissipated energy $\mathcal{D}_{\text{art}}$ can thus recover in regions where no crack is formed. As explored later on (Sect. 4.3.5), this behavior is beneficial to reduce the amount of erroneously degraded elastic energy.

3.6 Incorporation of crack boundary conditions

All phase field approaches discussed above have in common that the decomposition of the strain energy density in portions related to crack opening and closing are isotropic functions of the strain tensor. As a result, the mechanical response at the crack, also including the determination of an open or closed crack state, depends on the particular constitutive choice and current strain state, which usually does not correspond to the behavior of a macroscopic crack. In particular, the appearance of a discrete ("real") crack in a solid introduces two surfaces with specific boundary conditions on the crack plane, i.e. frictionless crack faces should be free of shear forces and positive normal forces. In a continuous setting the positive normal stress perpendicular to the crack surface and shear stresses on the crack plane should vanish, i.e. must be degraded to zero by the degradation function $g(s)$. As already indicated in Sect. 3.5.3, it appears to be necessary to take the crack orientation into account. In the following, the focus lies on the incorporation of a crack boundary conditions by a specific constitutive choice of the strain energy decomposition.

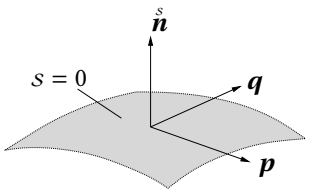


Figure 3.26: Schematic of isosurface with $s = 0$ representing the (hypothetical) crack surfaces. The direction of the crack surface is represented by its outward normal vector $\overset{s}{\boldsymbol{n}}$ and corresponding Cartesian coordinates with tangential directions $\boldsymbol{p}$ and $\boldsymbol{q}$.

The decomposition of strains/stresses with respect to the crack surface obviously requires the knowledge of the crack orientation. Initially, for the construction of equations, the orientation of the crack plane in the phase field approach is taken for granted. Further details about the determination of the crack direction are provided later in Sect. 3.6.1. The orientation of a single crack surface is specified by its outward normal vector $\overset{s}{\boldsymbol{n}}$. For simplicity a local Cartesian coordinate system with $\overset{s}{\boldsymbol{n}}$, $\boldsymbol{p}$ and $\boldsymbol{q}$ is introduced, aligned with the crack surface as depicted in Fig. 3.26. Note, that the other two directions are not

uniquely determined and thus two arbitrary unit vectors $\boldsymbol{p}$ and $\boldsymbol{q}$, which are perpendicular to each other and form a right-handed system are introduced, i.e.

$$|\boldsymbol{p}| = |\boldsymbol{q}| = 1 \quad , \qquad \overset{s}{\boldsymbol{n}} \cdot \boldsymbol{p} = \overset{s}{\boldsymbol{n}} \cdot \boldsymbol{q} = \boldsymbol{p} \cdot \boldsymbol{q} = 0 \quad , \qquad \boldsymbol{q} = \overset{s}{\boldsymbol{n}} \times \boldsymbol{p} \quad . \tag{3.264}$$

The direction of $\overset{s}{\boldsymbol{n}}$ changes with the surface. Thus, in a continuous setting it is more appropriate to use the second order tensor $\overset{s}{\boldsymbol{N}} = \overset{s}{\boldsymbol{n}} \otimes \overset{s}{\boldsymbol{n}}$, see (3.241), which is independent of the individual crack surface.

3.6.1 Crack orientation dependent degradation

Identification of relevant stress and strain portions

In order to decompose elastic portions related to the kinematic states of a crack (opened and closed) the equivalence between a discrete crack and its "smeared" representation is regarded. In particular, the deformations of two through-cracked infinitesimal elements and their equivalent in the phase field approach as depicted in Fig. 3.27 are analyzed.

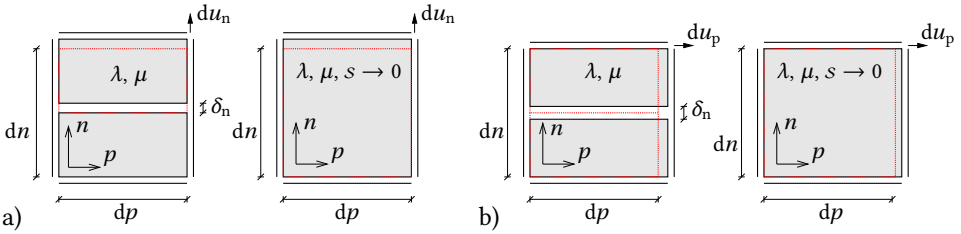


Figure 3.27: Identification of elastic and inelastic portions by comparing discrete crack opening and its "smeared" representation. Initial configurations are indicated by red color. Crack opening (under plane strain conditions) can be caused by a) (positive) displacement normal to crack plane, b) normal displacement in the crack plane.

In the first case, crack opening is caused by a displacement normal to the crack plane (Fig. 3.27a) with $\mathrm{d}u_\mathrm{n} = \delta_\mathrm{n} + \varepsilon_\mathrm{nn}\,\mathrm{d}n$. Since no elastic response is expected in the discrete setting the (normal) distance of crack opening is $\delta_\mathrm{n} = \mathrm{d}u_\mathrm{n}$. Its counterpart in the continuous setting (no discrete crack opening exists) is represented by

$$\tilde{\delta}_\mathrm{n} = \mathrm{d}u_\mathrm{n} = \cancel{\delta_\mathrm{n}} + \varepsilon_\mathrm{nn}\,\mathrm{d}n \quad . \tag{3.265}$$

To be consistent, the normal stress in the continuous setting must vanish, i.e.

$$\sigma_\mathrm{nn} = C_\mathrm{nn}\,\varepsilon_\mathrm{nn} = 0 \quad , \tag{3.266}$$

$$\sigma_\mathrm{pp} = C_\mathrm{pn}\,\varepsilon_\mathrm{nn} = 0 \quad , \tag{3.267}$$

$$\sigma_\mathrm{qq} = C_\mathrm{qn}\,\varepsilon_\mathrm{nn} = 0 \quad . \tag{3.268}$$

Obviously, the three elastic components C_nn, C_pn and C_qn must be fully degraded as $s \to 0$ at an open crack. In the second case, crack opening is caused by a displacement in the

crack plane (in the p-direction) with $\mathrm{d}u_\mathrm{p} = \varepsilon_\mathrm{pp}\,\mathrm{d}p > 0$ (see Fig. 3.27b). For $\nu > 0$ the crack opens due to the Poisson effect some distance δ_n. The total displacement in the crack normal direction is

$$\mathrm{d}u_\mathrm{n} = \delta_\mathrm{n} + \varepsilon_\mathrm{nn}\,\mathrm{d}n \overset{!}{=} 0 \quad . \tag{3.269}$$

With the opened crack being traction-free, the normal strain is

$$\varepsilon_\mathrm{nn} = -\frac{\lambda}{\lambda + 2\mu}\,\varepsilon_\mathrm{pp} \quad . \tag{3.270}$$

Thus, the (normal) distance of crack opening and its representation in the continuous setting are

$$\delta_\mathrm{n} = \frac{\lambda}{\lambda + 2\mu}\,\varepsilon_\mathrm{pp}\,\mathrm{d}n \overset{!}{=} \tilde{\delta}_\mathrm{n} \quad . \tag{3.271}$$

The non-zero stress components are

$$\sigma_\mathrm{pp} = \left[\lambda + 2\mu - \frac{\lambda^2}{\lambda + 2\mu}\right]\varepsilon_\mathrm{pp} \qquad \text{and} \qquad \sigma_\mathrm{qq} = \left[\lambda - \frac{\lambda^2}{\lambda + 2\mu}\right]\varepsilon_\mathrm{pp} \quad . \tag{3.272}$$

Since the response of the continuous setting must be equal, the stress-strain relations are

$$\sigma_\mathrm{nn} = \underbrace{C_\mathrm{nn}}_{\to 0}\,\varepsilon_\mathrm{nn} + C_\mathrm{np}\,\varepsilon_\mathrm{pp} = 0 \quad , \tag{3.273}$$

$$\sigma_\mathrm{pp} = \underbrace{C_\mathrm{pn}}_{\to 0}\,\varepsilon_\mathrm{nn} + C_\mathrm{pp}\,\varepsilon_\mathrm{pp} = \left[\lambda + 2\mu - \frac{\lambda^2}{\lambda + 2\mu}\right]\varepsilon_\mathrm{pp} \quad , \tag{3.274}$$

$$\sigma_\mathrm{qq} = \underbrace{C_\mathrm{qn}}_{\to 0}\,\varepsilon_\mathrm{nn} + C_\mathrm{qp}\,\varepsilon_\mathrm{pp} = \left[\lambda - \frac{\lambda^2}{\lambda + 2\mu}\right]\varepsilon_\mathrm{pp} \quad . \tag{3.275}$$

While the elastic component C_np must be fully degraded as $s \to 0$ at an open crack, the components $C_\mathrm{pp} = \lambda + 2\mu - \lambda^2/(\lambda + 2\mu)$ and $C_\mathrm{qp} = \lambda - \lambda^2/(\lambda + 2\mu)$ are identified. Accordingly, if crack opening is caused by a displacement in the q-direction the components $C_\mathrm{nq} \to 0$, $C_\mathrm{pq} = \lambda - \lambda^2/(\lambda + 2\mu)$, and $C_\mathrm{qq} = \lambda + 2\mu - \lambda^2/(\lambda + 2\mu)$ are obtained. Superposition of the crack opening distances (3.265), (3.271) and its counterpart caused by a displacement in the q-direction reveals the non-negative crack opening equivalent in the continuous setting

$$\begin{aligned}\tilde{\delta}_\mathrm{n} &= \varepsilon_\mathrm{nn}\,\mathrm{d}n + \frac{\lambda}{\lambda + 2\mu}\,\varepsilon_\mathrm{pp}\,\mathrm{d}n + \frac{\lambda}{\lambda + 2\mu}\,\varepsilon_\mathrm{qq}\,\mathrm{d}n \\ &= \Bigg[\underbrace{\varepsilon_\mathrm{nn} + \frac{\lambda}{\lambda + 2\mu}\,\varepsilon_\mathrm{pp} + \frac{\lambda}{\lambda + 2\mu}\,\varepsilon_\mathrm{qq}}_{=:\,-\tilde{\varepsilon}}\Bigg]\mathrm{d}n \geq 0\end{aligned} \tag{3.276}$$

with $\varepsilon_{nn} - \tilde{\varepsilon} \geq 0$ indicating an opened crack.

Strobl and Seelig (2015) started from the requirement of a non-positive normal stress on the crack surface, which also leads to an inequality condition. For isotropic linear elastic material this restriction can be written in terms of strain including the material parameters and the normal strains in lateral directions

$$\sigma_{nn} \leq 0 \qquad \Longleftrightarrow \qquad \varepsilon_{nn} \leq -\frac{\lambda}{\lambda + 2\mu} \left(\varepsilon_{pp} + \varepsilon_{qq}\right) =: \tilde{\varepsilon} \quad , \tag{3.277}$$

or equivalently, $\varepsilon_{nn} - \tilde{\varepsilon} \leq 0$, which indicates a closed crack. In order to enforce this condition and prevent a positive normal stress some parts of the stiffness must be degraded by the phase field. By means of the bracket operator (3.196) the normal strain ε_{nn} can be rearranged by using the identity (3.197) as

$$\varepsilon_{nn} = \underbrace{\langle \varepsilon_{nn} - \tilde{\varepsilon} \rangle}_{\text{crack opening}} \underbrace{- \langle \tilde{\varepsilon} - \varepsilon_{nn} \rangle + \tilde{\varepsilon}}_{\text{crack closure}} . \tag{3.278}$$

The first bracket term is identified to be related to crack opening and the rest corresponds to crack closure. Due to the fact that the constitutive behavior in case of compression should remain unaffected, so-called passive, only the stiffness which is responsible for a positive normal stress is degraded by the phase field via a degradation function $g(s)$, see Sect. 3.3.2. The preceding split of the normal strain permits the identification of the individual parts related to the opened and closed crack state in (3.278). The normal stress on the crack surface is

$$\begin{aligned} \sigma_{nn} &= (\lambda + 2\mu) \left[g(s) \langle \varepsilon_{nn} - \tilde{\varepsilon} \rangle - \langle \tilde{\varepsilon} - \varepsilon_{nn} \rangle + \tilde{\varepsilon} \right] + \lambda \left(\varepsilon_{pp} + \varepsilon_{qq}\right) \\ &= \begin{cases} g(s) \left[(\lambda + 2\mu) \, \varepsilon_{nn} + \lambda \left(\varepsilon_{pp} + \varepsilon_{qq}\right) \right] & \text{for } \varepsilon_{nn} - \tilde{\varepsilon} \geq 0 \text{ (open crack)} \\ (\lambda + 2\mu) \, \varepsilon_{nn} + \lambda \left(\varepsilon_{pp} + \varepsilon_{qq}\right) & \text{for } \varepsilon_{nn} - \tilde{\varepsilon} < 0 \text{ (closed crack)} \end{cases} . \end{aligned} \tag{3.279}$$

Applying this approach to the normal stresses in the crack plane results in

$$\begin{aligned} \sigma_{pp} &= \lambda \left[g(s) \langle \varepsilon_{nn} - \tilde{\varepsilon} \rangle - \langle \tilde{\varepsilon} - \varepsilon_{nn} \rangle + \tilde{\varepsilon} \right] + (\lambda + 2\mu) \, \varepsilon_{pp} + \lambda \, \varepsilon_{qq} \\ &= \begin{cases} g(s) \, \lambda \, \varepsilon_{nn} + \dfrac{g(s) \, \lambda^2 + 4\lambda \, \mu + 4\mu^2}{\lambda + 2\mu} \, \varepsilon_{pp} + \dfrac{g(s) \, \lambda^2 + 2\lambda \, \mu}{\lambda + 2\mu} \, \varepsilon_{qq} & \text{for } \varepsilon_{nn} - \tilde{\varepsilon} \geq 0 \\ (\lambda + 2\mu) \, \varepsilon_{pp} + \lambda \left(\varepsilon_{nn} + \varepsilon_{qq}\right) & \text{for } \varepsilon_{nn} - \tilde{\varepsilon} < 0 \end{cases} \end{aligned} \tag{3.280}$$

and

$$\sigma_{qq} = \lambda \left[g(s) \langle \varepsilon_{nn} - \tilde{\varepsilon} \rangle - \langle \tilde{\varepsilon} - \varepsilon_{nn} \rangle + \tilde{\varepsilon} \right] + (\lambda + 2\mu)\, \varepsilon_{pp} + \lambda\, \varepsilon_{qq}$$
$$= \begin{cases} g(s)\, \lambda\, \varepsilon_{nn} + \dfrac{g(s)\, \lambda^2 + 2\lambda\, \mu}{\lambda + 2\mu}\, \varepsilon_{pp} + \dfrac{g(s)\, \lambda^2 + 4\lambda\, \mu + 4\mu^2}{\lambda + 2\mu}\, \varepsilon_{qq} & \text{for } \varepsilon_{nn} - \tilde{\varepsilon} \geq 0 \\ (\lambda + 2\mu)\, \varepsilon_{qq} + \lambda\, (\varepsilon_{nn} + \varepsilon_{pp}) & \text{for } \varepsilon_{nn} - \tilde{\varepsilon} < 0 \end{cases} , \tag{3.281}$$

respectively. Interestingly, the normal stress response in the crack plane given by (3.280) and (3.281) is affected by the tension-compression split. In case of an open crack the normal stresses parallel to the crack are partially degraded. Assuming a frictionless crack surface the shear stresses on the crack plane vanish independent of an open or closed crack

$$\sigma_{np} = g(s)\, 2\mu\, \varepsilon_{np} \tag{3.282}$$

and

$$\sigma_{nq} = g(s)\, 2\mu\, \varepsilon_{nq} \quad . \tag{3.283}$$

The shear stress component in the crack plane

$$\sigma_{pq} = 2\mu\, \varepsilon_{pq} \tag{3.284}$$

remains unaffected by the crack. Alternatively, the crack orientation dependent degradation can also be carried out in terms of undegraded (effective) stresses according to

$$\boldsymbol{\sigma} = g(s)\, \tilde{\boldsymbol{\sigma}}^{+} + \tilde{\boldsymbol{\sigma}}^{-} \quad . \tag{3.285}$$

This results in an identical approach as elucidated in App. B.9.

The resulting elasticity tensors are

$$\mathbb{C}_{\text{open}}(s) = \frac{g(s)\, \lambda^2 + 2\lambda\mu}{\lambda + 2\mu}\, \boldsymbol{I} \otimes \boldsymbol{I} + 2\mu\, \mathbb{I}_{\text{sym}} + \big(g(s) - 1\big)\, \frac{2\lambda\, \mu}{\lambda + 2\mu} \left[\overset{s}{\boldsymbol{N}} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \overset{s}{\boldsymbol{N}} \right]$$
$$+ \big(1 - g(s)\big)\, \frac{4\mu\, (\lambda + \mu)}{\lambda + 2\mu}\, \overset{s}{\boldsymbol{N}} \otimes \overset{s}{\boldsymbol{N}} + \big(g(s) - 1\big)\, 2\mu\, \mathbb{N}_{\text{s}} \tag{3.286}$$

for an opened and

$$\mathbb{C}_{\text{closed}}(s) = \lambda\, \boldsymbol{I} \otimes \boldsymbol{I} + \mu\, \mathbb{I}_{\text{sym}} + \big(1 - g(s)\big)\, 2\mu\, \overset{s}{\boldsymbol{N}} \otimes \overset{s}{\boldsymbol{N}} + \big(g(s) - 1\big)\, 2\mu\, \mathbb{N}_{\text{s}} \tag{3.287}$$

for a closed crack for a fixed crack direction and contains the fourth order symmetrized tensor $\mathbb{N}_{\text{s}}$ defined by (B.90). The Voigt notations of (3.286) and (3.287) provided in App. C.2 by (C.80) and (C.81), respectively, reveal the correction term $\tilde{\Lambda}_{\text{g}}(s) := (g(s) - 1)\, \Lambda$ in the lateral directions with the elastic portion

$$\Lambda := \frac{\lambda^2}{\lambda + 2\mu} \quad . \tag{3.288}$$

It is required to obtain plane stress conditions in the plane of the opened crack.

In both cases the elasticity tensors (3.286) and (3.287) describing the crack-induced anisotropy display a special form of transverse isotropy with symmetry about the crack normal. The five parameters required to quantify transverse isotropic behavior are combinations of the elastic material parameters (e.g. λ and μ) and the degradation function $g(s)$. All details are provided in App. B.10.

Approximation of the crack orientation

The crack orientation dependent degradation requires a reliable approximation of the crack normal direction. Since in the phase field approach neither the crack position nor its direction are explicitely tracked a reliable determination becomes a difficult task. However, the smeared crack representation by the phase field offers some possibilities.

From a theoretical point of view, the crack normal direction in a continuous setting can be computed from the gradient of the phase field according to

$$\overset{s}{\boldsymbol{n}} := -\frac{\nabla s}{|\nabla s|} \tag{3.289}$$

as proposed by Strobl and Seelig (2015). The negative sign is caused by the fact that the crack normal points outward and the phase field value slopes towards the crack. Accordingly, the dyadic product of the crack normal direction which describes the orientation independent of the sign is

$$\overset{s}{\boldsymbol{N}} := \frac{\nabla s \otimes \nabla s}{|\nabla s|^2} \quad . \tag{3.290}$$

Since in the phase field approach to fracture the crack is approximated in a "smeared" sense, all related equations and decompositions must be applicable in the whole domain, not only at the location of the limit state $s = 0$. There are situation where the computation of the crack normal direction appears to be difficult. Directly at the crack a finite region can reach $s = 0$ or far away from the crack it is $s = 1$ with vanishing gradient in both cases. While the latter case is irrelevant since the decomposition plays no role, the first case is very decisive. Close to the crack tip, however, no unique phase field gradient is available. The situation is depicted in Fig. 3.28. While the crack normal $\overset{s}{\boldsymbol{n}}$ at A computed from the phase field gradient clearly indicates the crack direction, the normal vector computed at B and C are incoherent. Moreover, the physical meaning of the crack direction at or in front of the crack tip is unclear. Under no circumstances, the phase field gradient in front of the crack tip does represent the original idea of strain energy decomposition related to open and closed cracks. Thus, Strobl and Seelig (2015) proposed to use the crack orientation dependent degradation only some distance ($\approx \ell$) behind the crack tip. This, however, requires either the decomposition of the problem into two subdomains, one "behind the crack tip" and one "ahead of the crack tip", or a smooth transition between two modeling approaches, one based on the crack direction and one which is isotropic (not taking the

phase field and its gradient into account). Both options are discussed in Sect. 3.6.2. Up to this point, this remains an open issue.

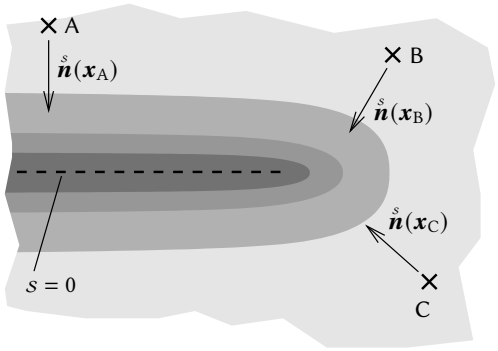


Figure 3.28: Schematic of the crack tip with the crack normal direction computed from the phase field gradient pointing in different directions.

From a technical point of view, the approximation of the crack orientation from the phase field according to (3.289) or (3.290) is sensitive to the orientation of the finite element mesh. Since the phase field is approximated by finite elements of Lagrangian type (see App. C.2) the discretized solution s_h is only piecewise differentiable (C^0-continuous). Thus, the approximated phase field gradient ∇s_h is discontinuous at finite element edges and does not necessarily represent the intended direction at every (integration) point. In fact, it turns out, that the approximation of the orientation by the phase field gradient inside the finite element is very sensitive. Deviations between integration points can be large, so that a local (element-wise) smoothing does not work. Note, that the phase field gradient is a key ingredient of the phase field approach (3.11), but only its scalar product $\nabla s_\mathrm{h} \cdot \nabla s_\mathrm{h}$ which represents a norm is used which is much more accurate. Thus, the original procedure (3.289) works only in situations with a regular finite element mesh which is aligned with the crack.

There are a number of references, where the phase field approach for fracture is used in conjunction with the normal direction of the crack. Usually, the methods are also based on the phase field gradient (3.289) or are only applicable to very special situations (Verhoosel and Borst, 2013; Mauthe and Miehe, 2015; Kienle and Keip, 2019). Singh et al. (2018) propose the bi-sectioning-based segmentation scheme and Verhoosel et al. (2015) use a α-contour to compute the normal direction of the crack surface for pressure loads. Other, very promising options to compute the crack direction are:

- Higher-order approximation: By using an approximation with a higher order continuity the phase field gradient can be computed smoothly in the whole domain. This can be done by the framework of isogeometric analysis (IGA), see, e.g., Cottrell (2009). In the context of phase field models to fracture Borden (2012) and Borden et al. (2014) were the first who used isogeometric analysis. However, they only used

the formulation based on the spectral strain decomposition (Sect. 3.5.1) without any attempts to approximate the crack direction.

- Least-square projection: The approximated phase field gradient can be smoothened over the whole domain by using a least square projection, e.g. a lumped L_2 projection (Govindjee et al., 2012). Typically, such projections are used to post-process element variables. Stresses or other variables computed at the integration points can be projected (component-wise) to nodes to obtain a continuous representation (Zienkiewicz et al., 2013). The normal vector $\overset{s}{\boldsymbol{n}}$ represents a direction with one angle in two dimensions and three angles in three directions. Since averaging angles (with large differences) is ambiguous it is more practical to project the components of the second order tensor (3.290).
- Separate approximation of crack direction: Yoshioka and Naumov (2019) presented and compared two methods to quantify the crack opening displacement in phase field approaches, see also Yoshioka et al. (2020). In case of hydraulic fracturing crack propagation is driven by the fluid pressure, which depends on the crack opening displacement. To compute the crack opening displacement also the crack normal direction is required. One approximation is based on the line integral method (Chukwudozie et al., 2019), the other one is based on a level-set function (Lee et al., 2017).

Regarding the effort, the least square projection is by far the simplest approach.

Accurate model behavior at cracks

In order to show the correct behavior of the crack orientation dependent degradation, the test of a through-cracked block as introduced in Sect. 3.5 is applied. The initial crack is modeled by the phase field as depicted in (Fig. 3.17a). Due to the regular mesh which is perfectly aligned with the crack orientation, the crack normal is obtained from the phase field gradient (3.289). In case of tension (Fig. 3.17b) the model shows crack opening without any contractions in lateral direction and elastic response (not shown). For compression according to Fig. 3.17c the model is able to approximate crack closure. The elastic response is homogeneous with the expected uniaxial stress (Fig. 3.29a) and with a homogeneous lateral expansion. Lastly, the model's ability to reproduce frictionless sliding (Fig. 3.18) is validated. As shown by the deformed mesh in Fig. 3.29b the upper and lower parts are sliding on each other without showing elastic deformations. This correspond perfectly to the expected deformation shown in Fig. 3.18. So, in contrast to the phase field models based on the spectral strain decomposition (Sect. 3.5.1) or the volumetric-deviatoric decomposition (Sect. 3.5.2) the phase field model based on the crack orientation dependent degradation is able to treat opened and closed cracks consistently.

It should be noted, that no (strain or stress) inhomogeneities occur in the vicinity of the crack. In isotropic formulations (with and without split), they are either caused by the Poisson effect or erroneous degradation of lateral components, see Figs. 3.19 and 3.23.

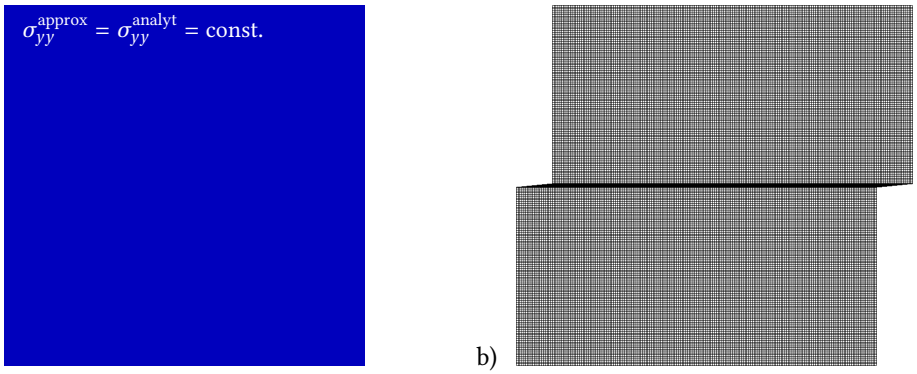


Figure 3.29: Validation of mechanical behavior of the phase field approach based on crack orientation dependent decomposition, a) homogeneous stress response in case of compression (with $\nu = 0.3$), b) perfect sliding of both parts without elastic response.

For a closed crack, the crack orientation dependent degradation degrades only two shear components while for an opened crack the model corrects the Poisson effect by (3.288) Regarding transversely isotropic behavior, the elastic component $\lambda + 2\mu - \lambda^2/(\lambda + 2\mu)$ sometimes appears as one of the Young's moduli (depending on the given parameters), see, e.g. Torquato (2002). The treatment of crack-induced anisotropy and the correction of the Poisson effect, both introduced by Strobl and Seelig (2015), are new features in phase field approaches to fracture and — to the best of the author's knowledge — haven't been used in (continuum) damage mechanics of isotropic materials either.

3.6.2 Further conceptual improvements

The approach introduced in Sect. 3.6.1 raises the question how to determine crack progress. While the approach was constructed so that it treats the situation at the opened and closed crack correctly (at least if the crack normal direction can be computed) it is so far unclear how to incorporate phase field evolution. A variationally consistent treatment would require, besides the variation with respect to $\boldsymbol{u}$ and s, also the variation with respect to the crack direction. This appears to be questionable from a physical perspective. Especially possible reorientation in the vicinity of existing cracks would demand additional irreversibility criteria. Moreover, as already mentioned, the current normal direction ahead of the crack tip (in particular relevant for the crack progress) has no physical meaning and should thus not be of interest for crack growth. Figure 3.30 shows the situation for possible crack progress in case of mode I and mode II loading. While the crack normal computed at A is obvious and does not change, the crack directions computed at B and C change with crack progress. More relevant is the fact, that it appears to be difficult to decide at which state of crack growth the reorientation is completed. It does not seem to be possible to make it dependent on local (i.e. point-wise) values like the phase field value.

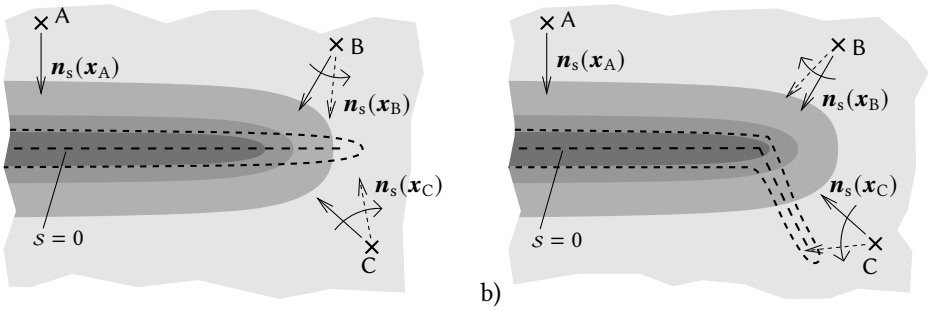


Figure 3.30: Schematic of the crack tip with the crack normal direction computed from the phase field gradient pointing in different directions. In addition, they change their direction with crack advance, for example, caused by a) mode I loading, b) mode II loading.

A conceptually simple approach would be to combine the crack orientation dependent degradation (Sect. 3.6.1) with ψ^{open} and ψ^{closed} at small phase field values (to approximate the crack) with one of the isotropic decompositions, e.g. the spectral strain decomposition (Sect. 3.5.1), with ψ^{act} and ψ^{pas} at phase field values close to one (to compute phase field evolution at crack initiation and crack progress). Such a formulation would look like

$$\psi(\boldsymbol{\varepsilon}, s) = f_{s0}(s)\left[g(s)\,\psi^{\text{open}}(\boldsymbol{\varepsilon}) + \psi^{\text{closed}}(\boldsymbol{\varepsilon}, s)\right] + f_{s1}(s)\left[g(s)\,\psi^{\text{act}}(\boldsymbol{\varepsilon}, s) + \psi^{\text{pas}}\right] \quad , \tag{3.291}$$

with a function $f_{s0}(s)$ which fades in the first part as $s \to 0$ and a second function f_{s1} which fades out the second part as $s \to 0$. There are, in general, at least two reasons against this: Close to the crack the decomposition is based on the crack direction while far away the decomposition is based on the principal strain direction. If the strain is not aligned with the crack normal, there is an inhomogeneous transition zone between both states (due to the change of orientation and different degradation). The second reason is that artificial reorientation due to the transition between both parts occurs at crack progress. Close to the crack front, the crack orientation dependent degradation would be still dominating due to the small phase field value while some distance away, the other driving mechanism starts to degrade the phase field.

Based on the crack orientation dependent degradation, Steinke and Kaliske (2019) enforced the approach to predict crack progress by using additional assumptions and constraints. A key assumption is, that the crack normal direction is aligned with the maximum principle strain direction during phase field evolution $s \to 0$. This direction is locally stored in combination with the maximum positive (crack driving) strain energy obtained at every integration point by using history variables, similar to those introduced in Sect. 3.2.3. A reorientation of the crack direction is only possible if the current crack driving energy exceeds the previously stored one. While storing the direction seems to be beneficial in situations where the phase field gradient is not available (e.g. in a fully degraded finite element), there are a number of drawbacks of this rather technical approach. Some of them are:

1. In this approach, the orientation of the crack (given by the localized phase field solution) is, in general, independent of the locally stored direction given by the mechanical response shortly before the finally broken state has been reached. Phase field evolution might lead to equal directions, but conformity is not guaranteed.
2. In this context, a local change of the principal (strain) direction during localization results in a wrong crack direction. This can only be prevented by applying monotonic loading normal to the crack direction (proportional loading) during crack growth.
3. The possible reorientation of the crack direction in the vicinity of the crack tip (especially in front of the crack tip) strongly depends on the local criterion to exceed a critical energy. The orientation of a kinking crack, for example, might be updated on its "tensile" side and remain unchanged on the other side, see Steinke and Kaliske (2019, Figure 42).
4. Initial cracks must be initialized by imposing information on direction and a critical energy for reorientation at all integration points.

To further investigate these effects in a structured manner, Storm et al. (2019; 2020) started to develop the concept of *representative crack elements*, which should provide a framework to consistently derive crack kinematics and crack driving forces.

3.7 Further numerical studies

One of the unbeatable advantages of the phase field approach to fracture is its simple implementation to existing finite element codes. For this work, the phase field approach as provided in App. C.2 has been implemented in different versions into FEAP (Finite Element Analysis Program), Version 8.4 (Taylor, 2014). With some exceptions, most of the implementation can be done on element level.

3.7.1 Mesh orientation dependency

In order to account for the asymmetric tension-compression response of a crack, the elastic energy density is decomposed into at least two different portions ψ_0^+ and ψ_0^-, where only one is degraded at a crack. As already discussed in Sect. 3.5, isotropic splits which do not account for the crack direction like the volumetric-deviatoric or spectral strain decomposition generally violate crack boundary conditions and thus lead to an unphysical behavior. For this reason, the crack orientation dependent degradation has been introduced (Sect. 3.6). However, it has been observed that all kinds of tension-compression decompositions reintroduce a mesh orientation dependency. This behavior results from the only partial (split-dependent) degradation of stiffness components in conjunction with the limited kinematics of finite elements as analyzed in the following.

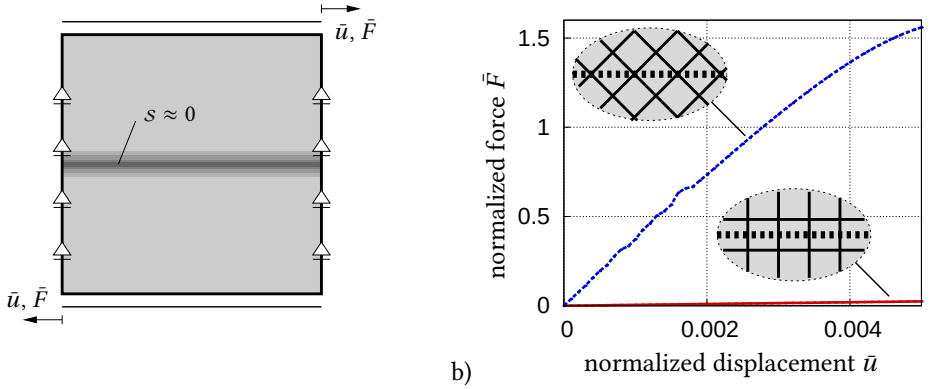


Figure 3.31: Through-cracked block subjected to simple shear: a) initial crack modeled by imposing $s \approx 0$ on at least one continuous row of finite elements, b) non-vanishing reaction force for mesh not aligned with crack orientation.

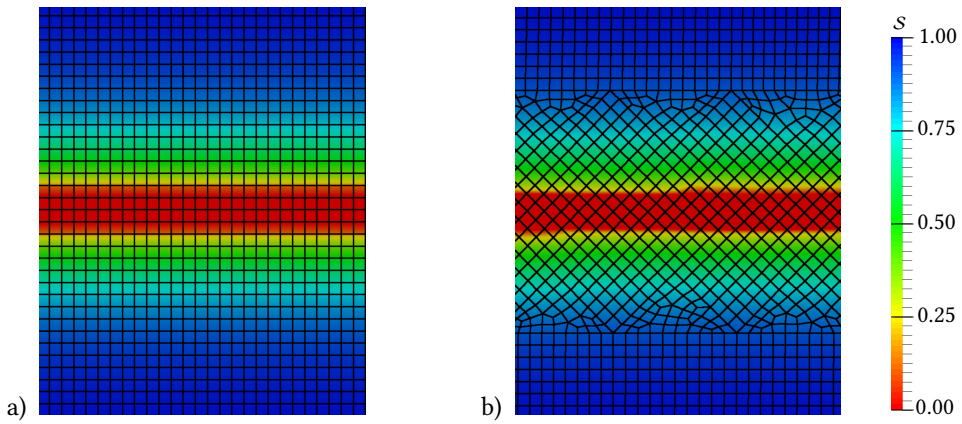


Figure 3.32: Phase field modeled crack for a) structured quadrilateral mesh aligned with crack direction, b) unstructured mesh rotated by about 45° at the crack.

Strobl and Seelig (2018) used the already known test of a through-cracked block subjected to simple shear as depicted in Fig. 3.31a. Both a regular mesh aligned with the through-crack and a mesh not aligned with the crack (quadrilateral finite elements rotated by about 45° at the crack) are compared. To minimize erroneous elastic effects, the through-crack is modeled by imposing $s \approx 0$ on at least one continuous row of elements (App. C.4), i.e. on three rows of nodes along the initial crack due to the rotated element configuration. Figure 3.32 shows the phase field approximation for the structured and unstructured mesh. Due to its simplicity, only the volumetric-deviatoric split (Sect. 3.5.2) is regarded in the following. In case of the finite mesh being aligned with the crack direction, the volumetric-deviatoric split correctly reproduces the behavior of the through-cracked block with almost no reaction force (see Fig. 3.31b). By contrast, using the same split, a mesh

which is not aligned with the crack yields a stiff elastic response as shown by the load-displacement diagram in Fig. 3.31b. A qualitatively similar result is obtained when using the crack orientation dependent degradation, in addition with the difficulty to approximate the crack direction correctly. The reintroduction of mesh orientation dependency is only avoided when all stiffness components are degraded according to (3.249).

Investigation of strain approximation in finite elements

The situation of finite element deformations and approximations at the phase field modeled crack is investigated in more detail. For this purpose, the through-cracked block is evaluated analytically. Figures 3.33–3.36 illustrate the situation of finite elements along a (horizontal) phase field modeled crack, in particular for homogeneous meshes consisting of (linear) triangular and quadrilateral finite elements. To reduce the stiffness across the crack, scenarios where a vanishing phase field ($s = 0$) is imposed at two rows of nodes (Fig. 3.33), three rows of nodes (Fig. 3.34) and four rows of nodes (Fig. 3.35) are regarded and full stiffness degradation is assumed. In the latter two cases, it should be noted that all nodes within the fully degraded zone (indicated by white background) can move freely without causing any elastic response.

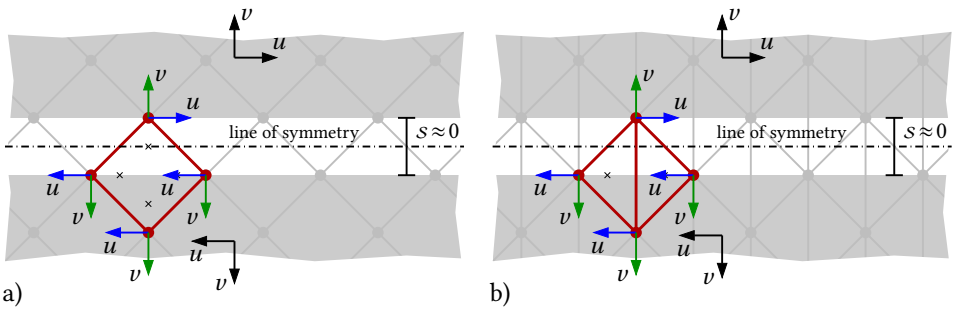


Figure 3.33: Element deformation at phase field crack (2 rows of nodes) with a) quadrilateral elements, b) triangular elements not aligned with crack.

This can be avoided by applying a remaining stiffness $0 < \eta_\epsilon \ll 1$ as introduced in (3.7), which however is neglected for simplicity. By applying horizontal displacements $u \geq 0$ and vertical displacements $v \geq 0$ on the upper and lower part, frictionless sliding and crack opening are studied. It is assumed, that all (localized) deformations take place in the zone of fully degraded stiffness without showing any elastic response. Thus, all nodes in the upper and lower part of the block are subjected to homogeneous displacements u and v, indicated by blue and green arrows, respectively. The antisymmetry and symmetry about the crack plane is used in the following to discuss possible nodal displacements, i.e. $u = v = 0$ along the line of symmetry. Due to the continuity of deformations nodes between the line of symmetry and the stiff upper and lower part move by $\tilde{u} := \alpha_\mathrm{u}\, u$ and $\tilde{v} := \alpha_\mathrm{v}\, v$ in horizontal and vertical direction with $\alpha_\mathrm{u}, \alpha_\mathrm{v} \in [0,\, 1]$. The quadrilateral mesh is rotated by 45° (Fig. 3.33a), while triangular elements are either placed with one

edge being vertical (not aligned with the crack, see Fig. 3.33b) or with one edge being horizontal (aligned with the crack, see Fig. 3.36). In each row of elements (which contribute to sliding and crack opening displacements) one element is representative for the whole row of elements and thus highlighted by red color, see Figs. 3.33–3.36.

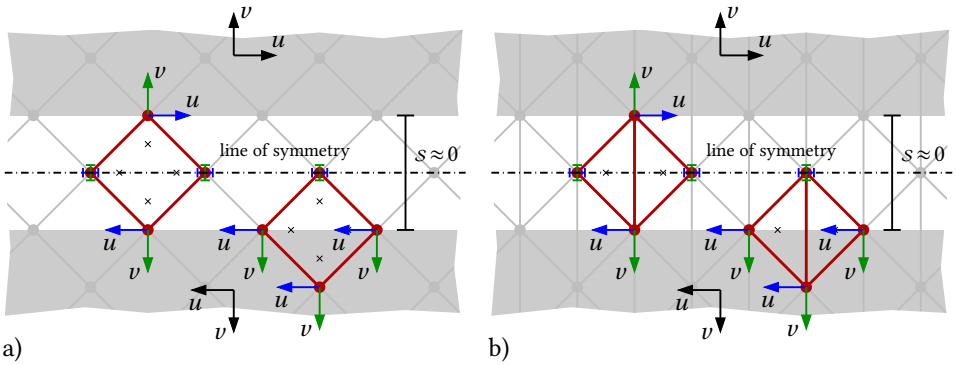


Figure 3.34: Element deformation at phase field crack (3 rows of nodes) with a) quadrilateral elements, b) triangular elements not aligned with crack.

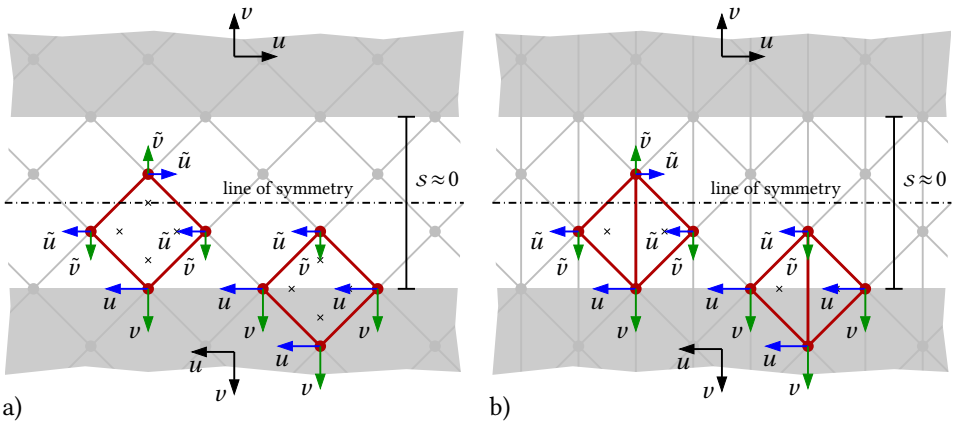


Figure 3.35: Element deformation at phase field crack (4 rows of nodes) with a) quadrilateral elements, b) triangular elements not aligned with crack.

After subtracting homogeneous displacements, all highlighted elements are captured in the test depicted in Fig. 3.37. In particular, the situations depicted in Fig. 3.36 are captured by Fig. 3.37a, the situations depicted in Fig. 3.33b, Fig. 3.34b and Fig. 3.35b are captured by Fig. 3.37b, and the situations depicted in Fig. 3.33a, Fig. 3.34a and Fig. 3.35a are captured by Fig. 3.37c. The outer length of the square geometry is $\sqrt{2}h$. Two nodes are kept at fixed positions while two are moved by u and v or $\tilde{u}$ and $\tilde{v}$. In the following, the approximation of strains (C.31) and their implications on the tension-compression asymmetric response (see Sect. 3.5 and 3.6) are analyzed, based on the discretized quantities provided in App. C.

All elements are evaluated at their integration points indicated by ① and ② for the triangular mesh and ①, ②, ③ and ④ for the fully integrated quadrilateral element.

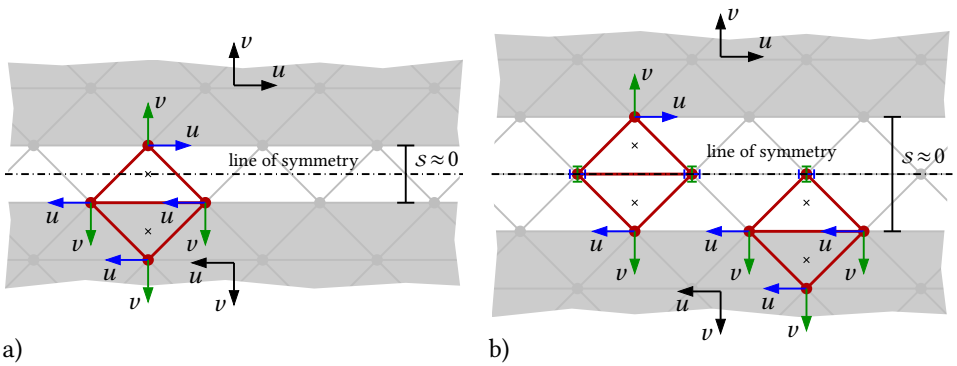


Figure 3.36: Element deformation at phase field crack with triangular element edges aligned with crack, phase field crack modeled by a) 2 rows of nodes, b) 3 rows of nodes.

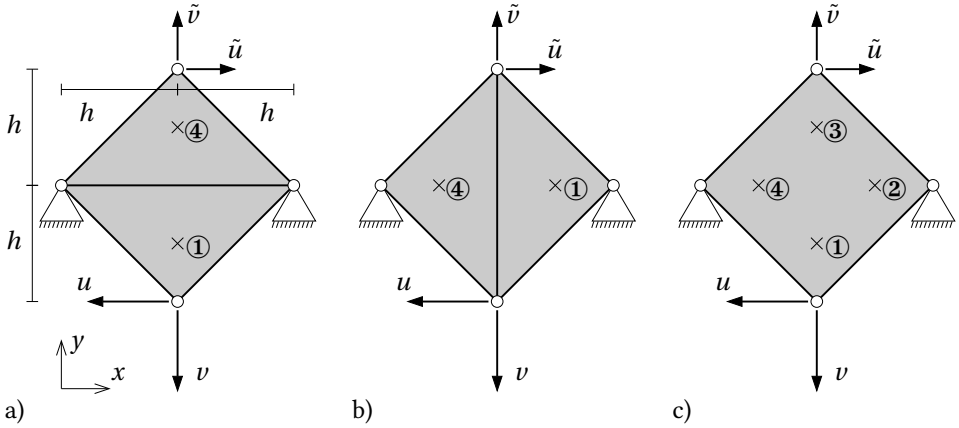


Figure 3.37: Evaluation of strain and stress approximation for typical deformations at the phase field modeled crack (assumed to be horizontal), a) triangular elements aligned with crack, b) triangular elements not aligned with crack, c) quadrilateral element not aligned with crack.

For the discretization with two triangular elements as depicted in Fig. 3.37a (their horizontal edges correspond to the direction of the crack plane) the strain approximations at the integration points are

$$\boldsymbol{\varepsilon}_{\mathrm{h}}^{①} = \begin{bmatrix} 0 \\ v/h \\ u/h \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\varepsilon}_{\mathrm{h}}^{②} = \begin{bmatrix} 0 \\ \alpha_{\mathrm{v}}\, v/h \\ \alpha_{\mathrm{u}}\, u/h \end{bmatrix} \tag{3.292}$$

yielding

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{①} = v/h > 0 \qquad \text{and} \qquad \operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{②} = \alpha_{\mathrm{v}}\, v/h > 0 \quad . \tag{3.293}$$

Since neither for horizontal nor for vertical displacements $\operatorname{tr} \boldsymbol{\varepsilon}_\mathrm{h} < 0$ the stress response of the volumetric-deviatoric decomposition (3.232) is zero as expected.

For the discretization with two triangular elements as depicted in Fig. 3.37b (their vertical edges do not correspond to the direction of the crack plane) the strain approximations at the integration points are

$$\boldsymbol{\varepsilon}_\mathrm{h}^{①} = \begin{bmatrix} (u - \alpha_\mathrm{u}\, u)/(2h) \\ (v + \alpha_\mathrm{v}\, v)/(2h) \\ (u + \alpha_\mathrm{u}\, u + v - \alpha_\mathrm{v}\, v)/(2h) \end{bmatrix} \quad , \tag{3.294}$$

$$\boldsymbol{\varepsilon}_\mathrm{h}^{②} = \begin{bmatrix} (-u + \alpha_\mathrm{u}\, u)/(2h) \\ (v + \alpha_\mathrm{v}\, v)/(2h) \\ (u + \alpha_\mathrm{u}\, u - v + \alpha_\mathrm{v}\, v)/(2h) \end{bmatrix} \quad , \tag{3.295}$$

with

$$\operatorname{tr} \boldsymbol{\varepsilon}_\mathrm{h}^{①} = (u - \alpha_\mathrm{u}\, u + v + \alpha_\mathrm{v}\, v)/(2h) \quad , \tag{3.296}$$

$$\operatorname{tr} \boldsymbol{\varepsilon}_\mathrm{h}^{②} = (-u + \alpha_\mathrm{u}\, u + v + \alpha_\mathrm{v}\, v)/(2h) \quad . \tag{3.297}$$

For simple shear ($v = 0$) it is by assuming $\alpha_\mathrm{u} < 1$ (the horizontal displacement cannot be constant across the crack)

$$\operatorname{tr} \boldsymbol{\varepsilon}_\mathrm{h}^{①} = (u - \alpha_\mathrm{u}\, u)/(2h) > 0 \qquad \text{and} \qquad \operatorname{tr} \boldsymbol{\varepsilon}_\mathrm{h}^{②} = (-u + \alpha_\mathrm{u}\, u)/(2h) < 0 \quad . \tag{3.298}$$

Due to the negative trace of the strain tensor at the second integration point the stress response of the volumetric-deviatoric decomposition (3.232) is non-zero.

For the discretization with one quadrilateral element as depicted in Fig. 3.37c (no edge corresponds to the direction of the crack plane) the strain approximations at the integration points are

$$\boldsymbol{\varepsilon}_\mathrm{h}^{①} = \begin{bmatrix} 0 \\ (v + v/\sqrt{3} + \alpha_\mathrm{v}\, v - \alpha_\mathrm{v}\, v/\sqrt{3})/(2h) \\ (u + u/\sqrt{3} + \alpha_\mathrm{u}\, u - \alpha_\mathrm{u}\, u/\sqrt{3})/(2h) \end{bmatrix} \quad , \tag{3.299}$$

$$\boldsymbol{\varepsilon}_\mathrm{h}^{②} = \begin{bmatrix} (u/\sqrt{3} - \alpha_\mathrm{u}\, u/\sqrt{3})/(2h) \\ (v + \alpha_\mathrm{v}\, v)/(2h) \\ (u + \alpha_\mathrm{u}\, u + v/\sqrt{3} - \alpha_\mathrm{v}\, v/\sqrt{3})/(2h) \end{bmatrix} \quad , \tag{3.300}$$

$$\boldsymbol{\varepsilon}_\mathrm{h}^{③} = \begin{bmatrix} 0 \\ (v - v/\sqrt{3} + \alpha_\mathrm{v}\, v + \alpha_\mathrm{v}\, v/\sqrt{3})/(2h) \\ (u - u/\sqrt{3} + \alpha_\mathrm{u}\, u + \alpha_\mathrm{u}\, u/\sqrt{3})/(2h) \end{bmatrix} \quad , \tag{3.301}$$

$$\boldsymbol{\varepsilon}_\mathrm{h}^{④} = \begin{bmatrix} (-u/\sqrt{3} + \alpha_\mathrm{u}\, u/\sqrt{3})/(2h) \\ (v + \alpha_\mathrm{v}\, v)/(2h) \\ (u + \alpha_\mathrm{u}\, u - v/\sqrt{3} + \alpha_\mathrm{v}\, v/\sqrt{3})/(2h) \end{bmatrix} \quad , \tag{3.302}$$

with

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{①} = (v + v/\sqrt{3} + \alpha_{\mathrm{v}}\, v - \alpha_{\mathrm{v}}\, v/\sqrt{3})/(2h) \quad , \tag{3.303}$$

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{②} = (u\sqrt{3} - \alpha_{\mathrm{u}}\, u/\sqrt{3} + v + \alpha_{\mathrm{v}}\, v)/(2h) \quad , \tag{3.304}$$

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{③} = (v - v/\sqrt{3} + \alpha_{\mathrm{v}}\, v + \alpha_{\mathrm{v}}\, v/\sqrt{3})/(2h) \quad , \tag{3.305}$$

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{④} = (-u/\sqrt{3} + \alpha_{\mathrm{u}}\, u/\sqrt{3} + v + \alpha_{\mathrm{v}}\, v)/(2h) \quad . \tag{3.306}$$

For simple shear ($v=0$) and by assuming $\alpha_{\mathrm{u}} < 1$ it is

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{①} = 0 \quad , \qquad \operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{②} = (u\sqrt{3} - \alpha_{\mathrm{u}}\, u/\sqrt{3})/(2h) > 0 \quad , \tag{3.307}$$

$$\operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{③} = 0 \quad , \qquad \operatorname{tr} \boldsymbol{\varepsilon}_{\mathrm{h}}^{④} = (-u/\sqrt{3} + \alpha_{\mathrm{u}}\, u/\sqrt{3})/(2h) < 0 \quad . \tag{3.308}$$

Due to the negative trace of the strain tensor at the fourth integration point the stress response of the volumetric-deviatoric decomposition (3.232) is non-zero.

As demonstrated above, the volumetric-deviatoric decomposition does not yield homogeneous tension or compression states after discretization. This applies also to all other decompositions. In general, the strain approximation in finite elements yields parasitic strains, which make it impossible to apply any kind of decomposition, if the element edge is not aligned with the crack as in Fig. 3.36. In all other cases, although the parasitic strains might be small, it should be kept in mind that they are multiplied by the full stiffness, which eventually results in significant deviations from the traction-free state. The application of higher order Lagrangian shape functions does not improve the situation in general. In addition, neither the application of an Assumed Natural Strain (ANS) approach, where stresses are constructed via sampling points with smaller errors (Dvorkin and Bathe, 1984), nor an Enhanced Assumed Strain (EAS) approach, where additional degrees of freedom[16] are introduced to enable a partial decoupling of strains (Simo and Rifai, 1990), can reliably tackle the deformation modes depicted in Fig. 3.37b and c. Only by degrading all stress components (see Sect. 3.5.4) or by remeshing the partially and fully degraded zone with elements aligned with the crack (thus the crack orientation is required, see Sect. 3.6.1), traction-free cracks can be guaranteed.

3.7.2 Surface energy "consumption" at crack tip

The smeared crack representation is introduced in (3.15) in such a way that it should consume the same amount of energy as a discrete crack surface. The internal length ℓ in (3.52) acts as a regularization parameter and for $\ell \to 0$ the phase field recovers the discrete

[16] The additional degrees of freedom are introduced on the element level and can the be condensed out. Thus, the additional deformation modes are not continuous across element boundaries, which would be beneficial in the present situation. The enhancement of finite element deformation modes, based on the Hu-Washizu functional (2.105), to model fracture is a key component of the embedded discontinuity approach, see the pioneering works of Dvorkin et al. (1990; 1991), Klisinski et al. (1991) and Simo et al. (1993).

crack Γ, i.e. the so-called Γ-convergence is obtained, see e.g. Braides (1998), and (3.14) holds. However, to retain the corresponding surface energy in the finite element discretization the phase field $s(\boldsymbol{x})$ of the smeared crack has to be adequately resolved. Therefore, a lower bound of the internal length is given by the element size h_e, e.g. by $\ell/h_e \geq 2$ (Miehe et al., 2010c). The energy consumption of a smeared crack is approximated quite well for a small, yet non-vanishing, internal length ℓ as long as the latter is several times smaller than the domain size. However, the diffuse phase field representation of the crack tip consumes additional energy which is not accounted for in the approximation (3.52).

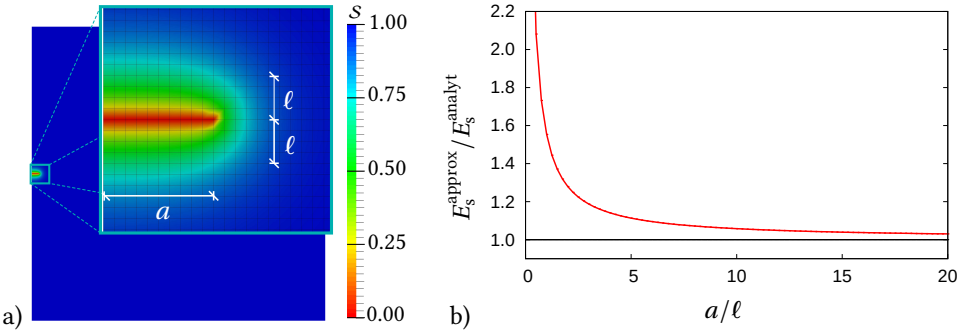


Figure 3.38: a) Crack of length a from the free surface modeled by imposing Dirichlet boundary condition $s=0$ on a single row of nodes, b) accuracy of crack surface energy approximation in absence of mechanical load depending on crack length a.

To illustrate the influence of the additional energy consumed by the crack tip for a non-vanishing internal length ℓ, the plane 2D situation in Fig. 3.38a in absence of mechanical load is considered. A single edge crack of length a from the free surface is modeled in the phase field setting by imposing a Dirichlet boundary condition $s = 0$ on a single row of finite element nodes to approximate the crack surface energy (in absence of the mechanical load) as good as possible. In Fig. 3.38b the surface energy approximated through the phase field

$$E_s^{approx} = \int_\Omega \mathcal{G}_c \, \gamma_\ell(s, \nabla s) \, dV \tag{3.309}$$

is compared with the surface energy (per unit thickness) of a discrete crack

$$E_s^{analyt} = \mathcal{G}_c \, a \quad . \tag{3.310}$$

The deviation of the approximated surface energy from the theoretical value is quite significant unless the crack is several times longer than the internal length ℓ, see also Sicsic and Marigo (2013). Obviously, this requirement cannot be fulfilled during crack formation in absence of a pre-crack. The inevitable consequence is an inaccurate approximation of the surface energy during crack nucleation. In general, this effect also appears for kinking and bifurcation, see the detailed discussion by Freddi (2019). However, latter cases usually appear in conjunction with comparatively long cracks. Thus, the amount of erroneously dissipated energy is much smaller.

3.7.3 Parameter recalibration

In the discretized (FE) setting the purely *geometric* phase field representation of a crack (i.e. absence of mechanical load) with $s=0$ in single (crack) nodes is altered by the application of mechanical loading in such a way that the crack becomes *mechanically* represented by zero stiffness of finite elements lying on the crack. As already noted by Bourdin et al. (2008) and further discussed by Strobl and Seelig (2018), this leads to a broadening of the phase field profile, i.e. the broken state $s \approx 0$ expands over at least one element, and hence to an overestimation of the crack surface energy. This effect is illustrated in Fig. 3.39 for the 1D example of a tensile bar with an initial through-crack. By contrast to the preceding section, this effect is no inherent shortcoming of the phase field method but attributed to the numerical solution procedure. More precisely, as the phase field solution localizes to $s = 0$ at nodal values, there is still non-zero stress computed at neighboring integration points. The overestimation of the surface energy can be almost eliminated, as suggested by Bourdin et al. (2008), by replacing the actual material fracture toughness $\mathcal{G}_c$ by an alleviated value

$$\tilde{\mathcal{G}}_c := \frac{\mathcal{G}_c}{1 + 0.5 h_e / \ell} \quad , \tag{3.311}$$

where h_e denotes the finite element size at the location of the crack. Of course, such a modification is non-trivial in case of an unstructured mesh with different element sizes. Since a too coarse mesh can considerably delay or even prevent crack initiation. Thus mesh refinement must be applied.

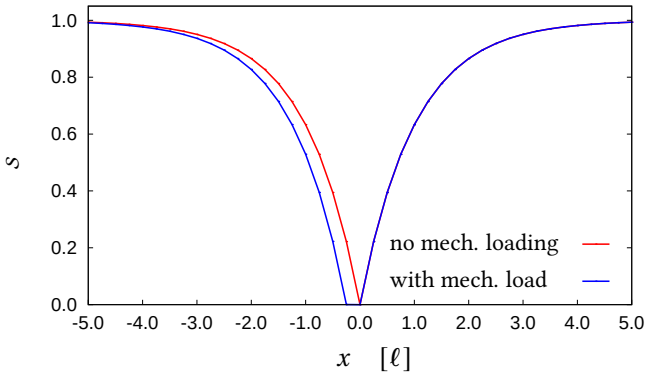


Figure 3.39: Phase field solution $s(x)$ in a 1D bar in the vicinity of a crack (at $x = 0$), comparison of the case without mechanical loading and the case with mechanical load after the broken state $s \approx 0$ expands over one element ($h_e = \ell/4$) to reduce the stiffness at the crack to zero.

The extension of the broken state over two elements, e.g. provoked by perfectly symmetric conditions (as shown by Li et al., 2016) or in the case of exploited symmetry at the crack

path, the recalibration can be done from a theoretical point of view with the double correction

$$\tilde{\tilde{\mathcal{G}}}_{\mathrm{c}} := \frac{\mathcal{G}_{\mathrm{c}}}{1 + h_{\mathrm{e}}/\ell} \quad . \tag{3.312}$$

In this specific case, the choice of the parameter $\eta_\epsilon > 0$ as introduced in Sect. 3.2 is essential. However, such situations should be avoided, also because of the fact that the width of the zone with minimum (or zero) stiffness spans over two elements, with the center nodes being (almost) freely movable if not constrained by boundary conditions.

ℓ/h_{e}	2	4	8	16	32
Adjusted fracture toughness acc. to (3.311)					
$\tilde{\mathcal{G}}_{\mathrm{c}}/\mathcal{G}_{\mathrm{c}}$	0.800	0.889	0.941	0.970	0.985
Adjusted fracture toughness acc. to (3.312)					
$\tilde{\tilde{\mathcal{G}}}_{\mathrm{c}}/\mathcal{G}_{\mathrm{c}}$	0.667	0.800	0.889	0.941	0.970
Adjusted internal length and fracture toughness acc. to (3.313)					
$\ell^*/\ell = 1 - 0.5\, h_{\mathrm{e}}/\ell$	0.750	0.875	0.934	0.969	0.984
$\mathcal{G}_{\mathrm{c}}^*/\mathcal{G}_{\mathrm{c}} = 1 - 0.5\, h_{\mathrm{e}}/\ell$	0.750	0.875	0.934	0.969	0.984
Comparison of discretizations					
h_{e}/ℓ	0.500	0.250	0.125	0.063	0.031
h_{e}/ℓ^*	0.667	0.286	0.133	0.065	0.032

Table 3.1: Numerical recalibration of parameters $\mathcal{G}_{\mathrm{c}}$ and ℓ for varying discretization h_{e}.

In the case that the tensile strength σ_{t} is used as an independent material parameter and thus is related to a non-vanishing internal length ℓ the modified toughness (3.311) affects the internal length parameter ($\ell \to \ell^*$) which, again, reenters (3.311). Hence, keeping the degradation function $g(s)$ as well as the parameters E, $\mathcal{G}_{\mathrm{c}}$, σ_{t}, h_{e} fixed, we end up with (3.184) two equations for the two modified parameters

$$\mathcal{G}_{\mathrm{c}}^* := \frac{1}{1 + h_{\mathrm{e}}/2\ell^*}\, \mathcal{G}_{\mathrm{c}} = \frac{2\ell - h_{\mathrm{e}}}{2\ell}\, \mathcal{G}_{\mathrm{c}} \quad , \qquad \ell^* := \underbrace{(f_{\mathrm{gw}})^2\, \frac{\mathcal{G}_{\mathrm{c}}\, E}{\sigma_{\mathrm{t}}^2}}_{=\,\ell} - \frac{h_{\mathrm{e}}}{2} \quad , \tag{3.313}$$

see also Strobl and Seelig (2019). From the latter expression it is obvious that the modification is only possible if the mesh is finer than one half of the original length parameter ℓ to get $\ell^* > 0$. Table 3.1 gives values for typical discretizations (quantified by the ratio ℓ/h_{e}) as used later on in Sect. 4.3. Such a readjustment of the fracture toughness and the internal length parameter guarantees that in numerical simulations the prescribed tensile strength as well as the specific crack surface energy (fracture toughness) are correctly reproduced.

This is shown in Fig. 3.40 for the 1D example of a fracturing tensile bar (cf. Fig. 3.14a) with $h_e \approx \ell/4$. While in the numerical simulation either the surface energy is overestimated when using the original parameters $\mathcal{G}_c$ and ℓ (Fig. 3.40a) or the bar breaks before reaching the material strength when using only the alleviated toughness $\tilde{\mathcal{G}}_c$ (Fig. 3.40b), the calibration of both parameters according to (3.313) results in perfect agreement with the intended behavior.

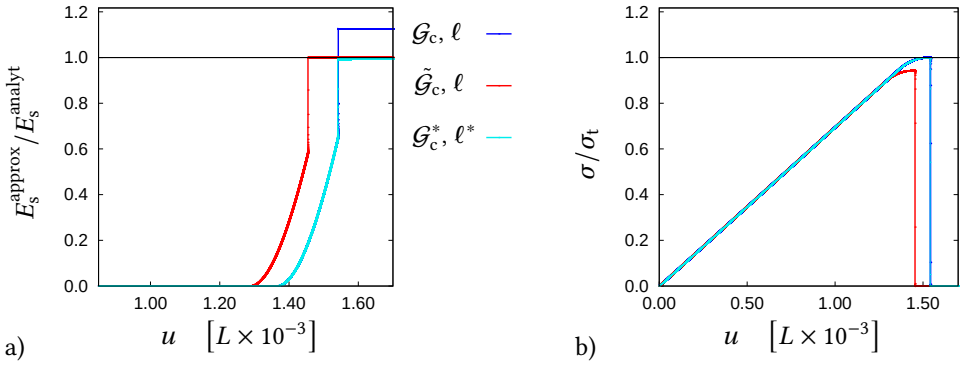


Figure 3.40: Numerical simulation of 1D tensile bar (see Fig. 3.14a) using the quartic degradation function (3.74) with $h_e \approx \ell/4$. Comparison of results obtained by using original parameters ($\mathcal{G}_c$ and ℓ), using only the alleviated toughness $\tilde{\mathcal{G}}_c$ according to (3.311) in combination with ℓ, and both parameters $\mathcal{G}_c^*$ and ℓ^* calibrated according to (3.313): a) surface energy evolution, b) load-displacement curves.

3.7.4 Limiting crack velocity and crack branching

After initiation, cracks can quickly accelerate depending on the actual conditions. However, inertia effects limit the crack tip velocity. Fast cracks can become unstable regarding their propagation direction or branch into multiple macroscopic cracks or initiate additional microscopic cracks. Although these effects of dynamic crack propagation have been investigated for decades a generally accepted theory, including, e.g., a reliable crack branching criterion, is still missing. As already discussed in Sect. 2.1.5, the wave propagation at free surfaces is governed by the Rayleigh wave speed c_R. By regarding the dynamic crack tip field, see Gross and Seelig (2011), the ratio of the normal stress normal to the direction of crack progress to the normal stress in the direction of crack progress decreases ahead of the crack tip with increasing crack speed. As the crack speed approaches c_R crack propagation straight ahead (deflection angle $\varphi = 0°$) becomes impossible. Based on these theoretical considerations, the Rayleigh wave speed is an upper limit for crack propagation. However, from experiments with many different materials it is known that the mode I crack tip velocity usually does not exceed $0.6\, c_R$, see Ravi-Chandar and Knauss (1984). Only by constraining crack propagation to a weak plane (e.g. an interface) higher crack tip velocities can be observed. The limiting velocity depends on both the material behavior and the experimental setup. At constant (high) crack tip velocity the stress

intensity factor and surface roughness increase (Broberg, 1996). Both effects, as well as the formation of microcracks, can be attributed to increased energy dissipation and balance the energy flux into the process zone (see Sect. 2.2.4) caused by, e.g., increased energy release. If the energy release rate is sufficient to create two or multiple new cracks one necessary condition for macroscopic crack branching is fulfilled (Ramulu and Kobayashi, 1985). Since the phase field approach does not require any specific assumptions regarding the direction of crack advance the method seems to be well suited to investigate crack tip speed limits and branching phenomena.

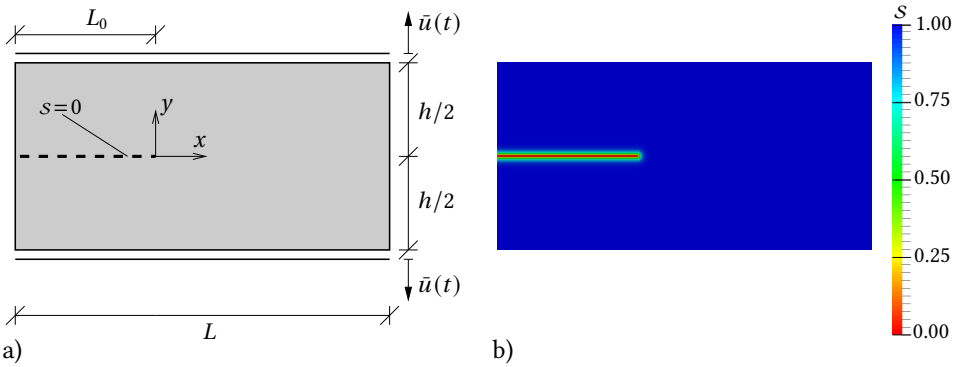


Figure 3.41: Pre-cracked plate used to investigate the limiting speed of crack propagation: a) geometry and loading conditions, b) initial crack modeled by imposing $s \approx 0$ along one continuous row of finite elements.

Inspired by the dynamic crack branching problem in a pre-cracked glass sheet discussed by Song et al. (2008), Borden et al. (2012) investigated the limiting speed of crack propagation and crack branching by using the phase field method. In the following, a similar problem is studied by using a rectangular plate (plane strain conditions) including a pre-crack as depicted in Fig. 3.41a. The pre-crack is modeled by imposing $s \approx 0$ along one continuous row of finite elements (Fig. 3.41b), see also App. C.4. In order to compare the results to the results obtained by Borden (2012), the same set of material parameters for "glass" (Song et al., 2008) is used, i.e.

- Young's modulus $E = 32\,\text{GPa}$,
- Poisson's ratio $\nu = 0.20$,
- density $\rho = 2.45\,\text{g/cm}^3$,
- crack resistance $\mathcal{G}_\text{c} = 3\,\text{J/m}^2$. (3.314)

With the relations of linear elastodynamics discussed in Sect. 2.1.5, the dilatational (2.126), shear (2.127), and Rayleigh wave speeds (2.132) are

$$c_\text{d} = 3809.5\,\text{m/s} \quad , \qquad c_\text{s} = 2332.0\,\text{m/s} \quad , \qquad c_\text{R} = 2126.8\,\text{m/s} \quad , \tag{3.315}$$

respectively. The dimensions of the plate are $L = 80\,\text{mm}$, $h = 40\,\text{mm}$, and $L_0 = 30\,\text{mm}$. The regularization length is chosen to be $\ell = 0.5\,\text{mm}$, which is two times the finite element size $h_\text{e} = 0.25\,\text{mm}$ of the regular mesh. According to (3.311) the overestimation of

surface energy is eliminated by using $\tilde{\mathcal{G}}_c = 2.4\,\mathrm{J/m^2}$ instead. Since cracks remain open once they have been formed the stiffness is fully degraded by the quartic degradation function (3.74) according to (3.249), while phase field evolution is only driven by positive normal strains (3.250). Note, that in the present problem, the plate is loaded by imposed displacement boundary conditions $\bar{u}(t) = v_0\, t$ at the top and bottom (Fig. 3.41a), similar to Hofacker and Miehe (2013), instead of applying traction boundary conditions $\bar{t}(t)$ typically used, c.f. Borden (2012), Schlüter et al. (2014) and Schlüter (2018). The resulting stress is $\sigma_{yy} = v_0 \sqrt{(\lambda + 2\mu)\,\rho}$ (e.g. Meyers, 1994). To obtain the intended tensile stress $\sigma_{yy} = 1\,\mathrm{MPa}$ (Borden et al., 2012), which is according to (3.148) clearly below the material's tensile strength $\sigma_t \approx 5.33\,\mathrm{MPa}$, the applied velocity on the boundary is $v_0 = 0.113\,\mathrm{mm/ms}$. For time integration the implicit Newmark method with the standard set of parameters ($\beta = 0.25$ and $\gamma = 0.50$) is used (see App. C.2), although a modified version with $\beta = 0.5$ might allow for larger times steps at the cost of introducing additional dissipation (Schlüter et al., 2017). To obtain accurate solutions the alternate minimization procedure (Sect. 3.2.4) with multiple iterations between both field solutions and a very small time increment (in the context of implicit methods) is chosen with $\Delta t = h_e / c_R \approx 1 \times 10^{-7}\,\mathrm{s}$. In addition, the maximum x-coordinate and the minimum and maximum y-coordinates where $s < 0.1$ are tracked in each time step to obtain the crack tip position(s). The speed of crack propagation is then calculated according to

$$v_{ct} \approx \frac{\mathrm{d}}{\mathrm{d}t} \sqrt{\mathrm{d}x_{ct}{}^2 + \mathrm{d}y_{ct}{}^2} \quad . \tag{3.316}$$

Figures 3.42 shows the crack path predicted by the phase field method and the corresponding velocity of the crack tip. At the beginning the crack accelerates and propagates without any deflection from its initial direction ($\varphi = 0°$). After the velocity reaches a plateau ($t \approx 17\,\mu\mathrm{s}$) the crack starts to widen at $t \approx 26\,\mu\mathrm{s}$. It further accelerates until a maximum velocity is obtained ($t \approx 31\,\mu\mathrm{s}$). At this point in time the crack branches symmetrically. The branched cracks propagate at an almost constant velocity significantly below the maximum crack tip velocity, but the paths of the upper and lower crack branches oscillate. One reason for oscillating crack paths might be reflected waves traveling from the upper and lower boundary with $c_d \gg v_{ct}$. The deflection of the upper crack shortly after branching is ($\varphi_1 \approx 16.3°$) which is in the range of experimental observations, see the table including theoretical predictions and experimental findings provided by Kobayashi and Ramulu (1985). Despite somewhat different boundary conditions, the crack tip velocity obtained by Borden (2012) by extracting and interpolating iso-curves in a post-processing step (see Fig. 3.43) are quite similar regarding the maximum crack speed and the average values. It should be noted, that during the whole event of crack propagation and branching, the crack velocity v_{ct} is clearly below $0.6\, c_R = 1276\,\mathrm{m/s}$, which is the limit found by experiments (see above), significantly lower than the Rayleigh wave speed.

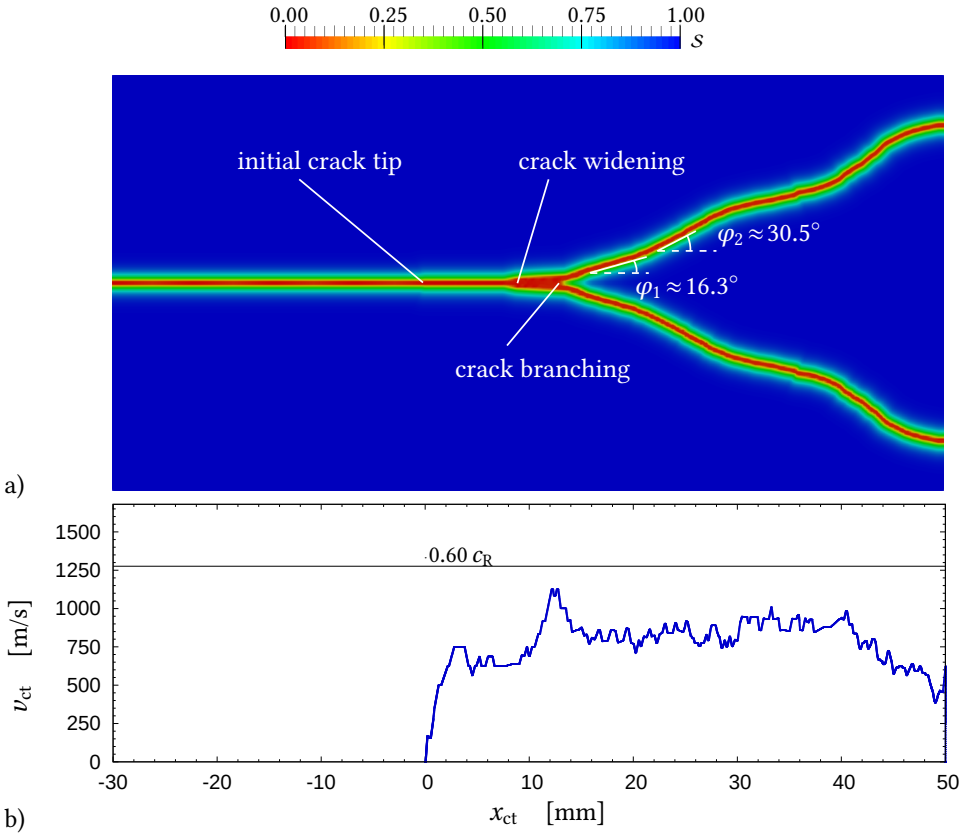


Figure 3.42: Investigation of limiting speed of crack advance and dynamic crack branching, a) crack path predicted by phase field approach, b) corresponding crack tip velocity as function of crack advance in horizontal direction.

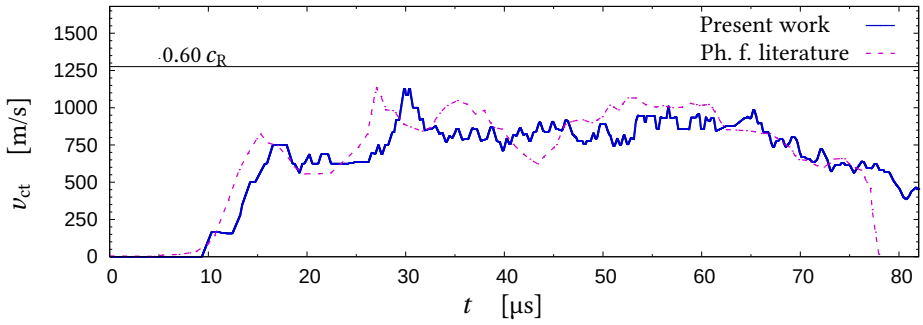


Figure 3.43: Comparison of computed crack tip velocity v_{ct} with phase field study from literature (Borden, 2012).

3.7.5 Mode I+II fracture

Further numerical investigations are carried out in order to verify the phase field response under mixed mode I-mode II loading conditions. Similar to the investigation of crack tip velocities (Sect. 3.7.4), both the critical load at which crack progress starts and the direction of crack advance must be predicted by the phase field method. Different well-established concepts are introduced in Sect. 2.2.5, all based on specific assumptions concerning the energy density or the stress field in the vicinity of a crack tip without taking the micro-mechanisms of fracture into account. Inspired by the study proposed by Pham (2015), the (in-plane) mixed mode loading is studied by using a square plate (plane strain conditions) including a pre-crack as depicted in Fig. 3.44.

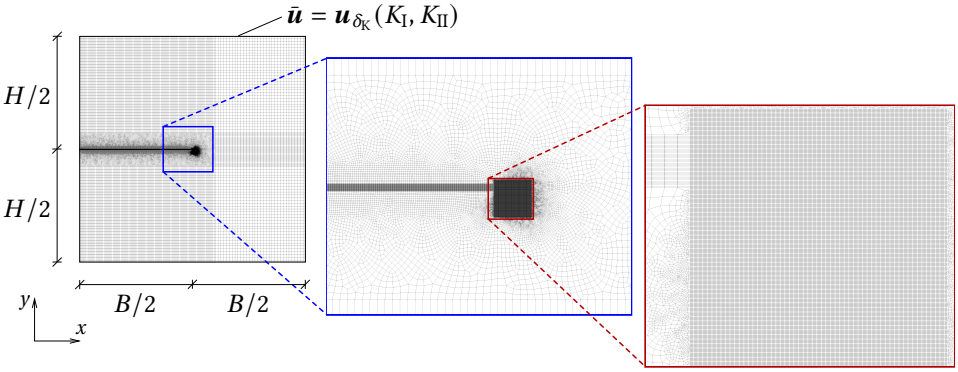


Figure 3.44: Geometry and discretization of the square plate used to investigate fracture caused by (in-plane) mixed mode loading. Different zoom levels show the refined mesh with ($h_e = 0.0025$ mm) in the zone of interest.

The pre-crack is modeled by the phase field by imposing $s \approx 0$ along one continuous row of finite elements (Fig. 3.45a), see also App. C.4, with $\ell = 0.01$ mm. In relation to $\ell = 0.01$, the problem dimensions are chosen to be $H/2 = B/2 = 1000\ell = 10$ mm. Different combinations of mode I and II displacements, see (A.93) and (A.99) provided in App. A.8, are applied at the outer boundary according to

$$\bar{\boldsymbol{u}} = \boldsymbol{u}_{\delta_K}(K_I, K_{II}) := \begin{bmatrix} u_x(K_I) \\ u_y(K_I) \end{bmatrix} + \begin{bmatrix} u_x(K_{II}) \\ u_y(K_{II}) \end{bmatrix} \quad . \tag{3.317}$$

According to Pham (2015), this is "adequate to recover the elastic solution at large distances from the local perturbations near the crack tip" caused by (constitutive and modeling choices) of the phase field method. The ratio of mode II to mode I loading is varied by

$$\delta_K := \frac{K_{II}}{K_I + K_{II}} \quad , \tag{3.318}$$

with $\delta_K = 0$ corresponding to pure mode I and $\delta_K = 1$ corresponding to pure mode II loading. Simulations are carried out quasi-statically by proportionally increasing K_I

and K_{II} (at constant δ_{K}) with very small increments until the crack propagates. The set of material parameters includes

- Young's modulus $E = 100\,\mathrm{MPa}$,
- Poisson's ratio $\nu = 0.25$,
- crack resistance $\mathcal{G}_{\mathrm{c}} = 10\,\mathrm{J/m^2}$. (3.319)

In the region of interest, i.e. in a zone around the crack tip (see Fig. 3.44), the mesh is regular and refined with elements of the size $h_{\mathrm{e}} = 0.0025\,\mathrm{mm}$ (more than 96000 quadrilateral elements in total). According to (3.311) the overestimation of surface energy is eliminated by using $\tilde{\mathcal{G}}_{\mathrm{c}} = 8.88\,\mathrm{J/m^2}$ instead. Since cracks remain open once they have been formed the stiffness is fully degraded by the quadratic (3.70), cubic (3.71) or quartic degradation functions (3.74) according to (3.249), while phase field evolution is only driven by positive normal strains (3.250). In this setting, the alternate minimization procedure (Sect. 3.2.4) with multiple iterations between both field solutions is used to obtain accurate solutions.

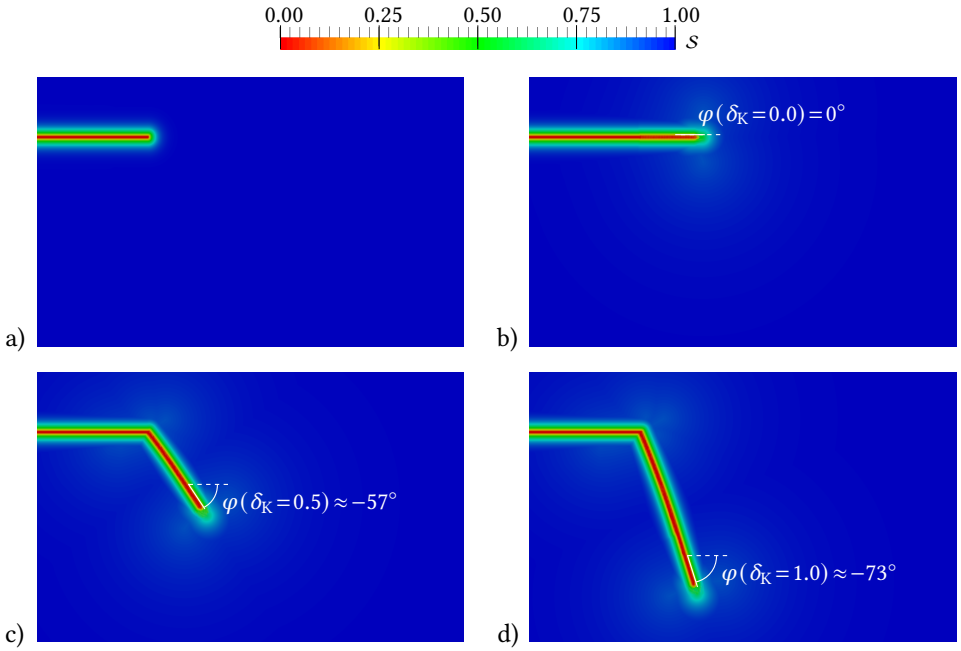


Figure 3.45: Crack progress for mixed mode I-mode II loading computed by the phase field method. a) Initial crack modeled by imposing $s \approx 0$ on one continuous row of finite elements. Crack propagation for b) $\delta_{\mathrm{K}} = 1.0$, c) $\delta_{\mathrm{K}} = 0.5$, d) $\delta_{\mathrm{K}} = 0.0$ with corresponding deflection angles.

Crack progress is investigated for $\delta_{\mathrm{K}} = \{0.0, 0.25, 0.50, 0.75, 1.0\}$. Figures 3.45b to d show the crack path predicted by the phase field method (with quadratic degradation). As expected, with increased mode II loading the crack deflection angle increases (in absolute values). For pure mode II loading the deflection angle is $\varphi \approx -73°$ (with negative

sign due to the direction of deflection). In Figure 3.46a, all deflection angles obtained in the study (somewhat varying for different degradation functions) are compared with theoretical predictions based on the concepts introduced in Sect. 2.2.5 and the results obtained by Pham (2015). Here, common abbreviations are used, i.e. MTS denotes the maximum hoop/tangential stress criterion, ERR the maximum energy release rate criterion, SED the strain-energy-density criterion (depending on ν), and PLS the principle of local symmetry. For the evaluation of the ERR criterion and the PLS criterion the stress intensity factors of a kinked crack calculated by Leblond (1999) are used, while the SED criterion is evaluated by using the expressions provided by Gross and Seelig (2011). In addition, the critical loads at which crack progress start (normalized by the fracture toughness K_{Ic}) are compared in Fig. 3.46a. In this context, crack initiation is assumed as soon as $s < 0.2$ occurs on average at an element outside the pre-crack.

Both Figure 3.46a and b clearly indicate that results from phase field simulations driven by positive normal strains (3.250) are in good agreement with theoretical predictions for (in-plane) mixed mode loading, especially with the ERR criterion. Moreover, the study shows that phase field approaches based on full stiffness degradation do not fail to model traction-free cracks, see the discussion in Sect. 3.5, and thus can reasonably predict crack growth emanating from phase field modeled initial cracks. Deviations in the present study, according to Fig. 3.46b the critical load predicted by phase field simulations is underestimated for all cases, have to be seen in the light of modeling choices and different possible criteria to detect crack initiation.

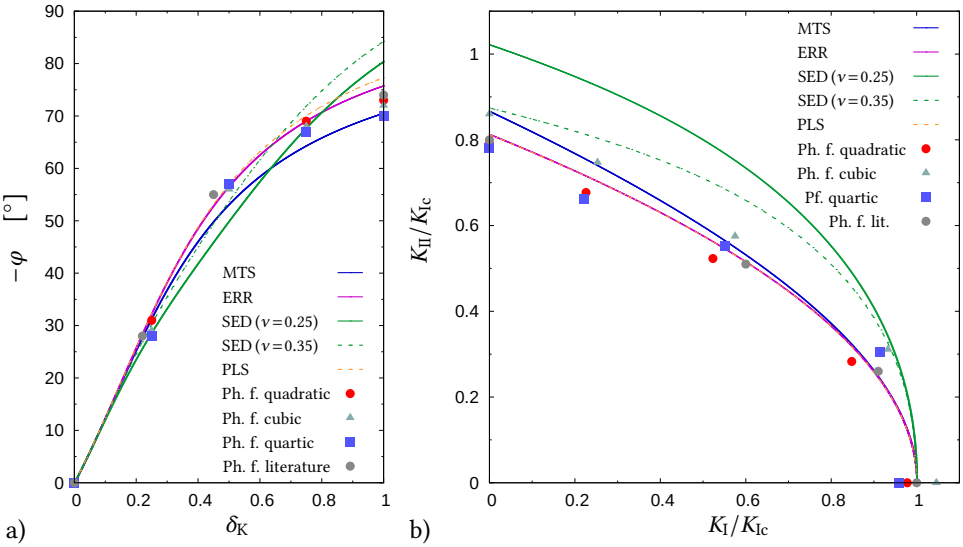


Figure 3.46: Comparison of analytical results with predictions computed by the phase field method for mixed mode I-mode II loading, a) crack deflection angle φ, b) critical combinations of mode I and mode II loading in terms of the fracture toughness K_{Ic}. For comparison, also results from the literature (Pham, 2015) are given.

4 Analysis of Hertzian indentation fracture

The formation and further evolution of cracks caused by the compression of a stiff indenter onto the surface of a brittle solid is a fascinating problem of fracture mechanics and was already investigated by H. Hertz at the end of the nineteenth century. In absence of initial defects, however, its prediction is still a challenging task since crack nucleation is caused by a rather weak stress concentration in the contact near-field which cannot be described by classical fracture mechanics concepts. To predict crack initiation, a hybrid fracture criterion in the concept of finite fracture mechanics (FFM) is applied to the indentation fracture problem. The method is used to analyze the influence of geometric parameters on the initial crack formation at some a priori unknown location outside of the contact region. In a further step, a phase field approach is utilized to holistically describe the entire indentation fracture process, including crack formation and the subsequent growth of a cone crack. Motivated by the FFM analysis, particular emphasis in the phase field formulation is placed on the significance of the material's tensile strength and fracture toughness as independent physical properties necessary for the quantitative description of crack formation. Results obtained by phase field simulations are compared with FFM results as well as experimental findings from the literature.

4.1 Indentation fracture problem setting

The boundary value problem investigated in this chapter is specified as follows. The flat and initially defect-free surface of an isotropic linear elastic body with Young's modulus E and Poisson's ratio ν is subjected to indentation loading by a rigid cylindrical punch of radius a (Fig. 4.1). Axisymmetry is assumed (thus a cylindrical coordinate system with r, z and ϕ is used) and loading by the axial indenter force F is imposed through a prescribed vertical displacement d in the contact region $r \leq a$, which is assumed frictionless. In case of a half-space, the relation between force and indenter displacement prior to fracture has already been analyzed by Boussinesq (1885) and is given by

$$F = \frac{2\,E\,a\,d}{1-\nu^2} \quad . \tag{4.1}$$

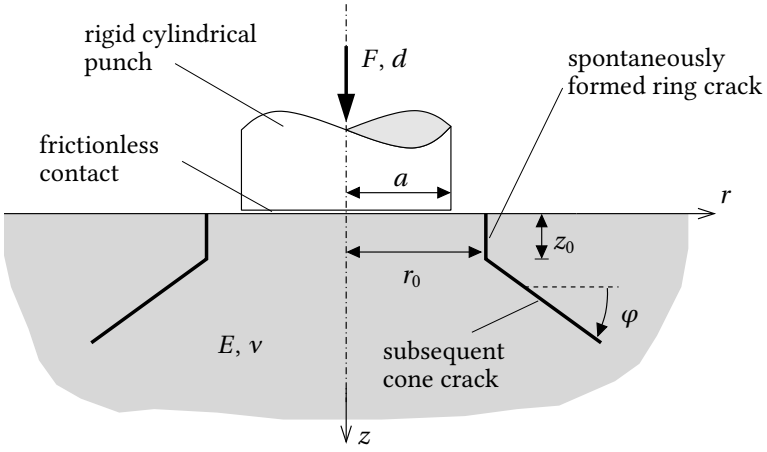


Figure 4.1: Schematic of axisymmetric boundary value problem with isotropic elastic solid subjected to indentation by rigid cylindrical punch; the two stages of fracture (ring crack formation at radius r_0 with extension z_0, cone crack propagation with inclination φ) are indicated.

4.1.1 Idealization of fracture problem

According to numerous experimental studies, e.g. Tillett (1956) and Mouginot and Maugis (1985), the fracture process is understood to proceed in two stages (Fig. 4.1):

1) A ring crack of radius $r_0 > a$ and small extension z_0 from the free surface *spontaneously* forms somewhat outside the contact region at a critical load.

2) Afterwards and with increasing load, it turns into a cone-shaped crack that grows in a stable manner at an approximately constant angle φ which depends on Poisson's ratio ν of the material.

This observation holds for cylindrical as well as spherical indenters. Considering here a cylindrical indenter (instead of a spherical one as in most studies) bears the advantage that the contact region remains constant and thus can not extend over the ring crack with increasing load so that difficulties discussed e.g. by Chen et al. (1995) and Chai (2006) are avoided. The usefulness of a flat punch was already recognized by Roesler (1956a; 1956b) and was later exploited, e.g., by Mouginot and Maugis (1985).

The present model set-up involves a number of idealizations with respect to real indentation tests which are considered to be not very severe for this fundamental study, yet deserve some comments. In contrast to the here assumed rigid punch, the influence of a finite stiffness of (spherical) indenters was analyzed by Johnson et al. (1973) who state that more compliant indenters lead to ring cracks of smaller radii and ascribe this to reduced frictional effects with smaller elastic mismatch between indenter and substrate.

The influence of frictional contact on the indentation stress field was (likewise for spherical indenters) studied in detail by Andersson (1996). In case of a cylindrical punch, the effect of rounded corners (as prevailing in reality) on the indentation stress field was discussed by Ciavarella et al. (1998). Finally, it should be mentioned that the here assumed axisymmetry may be not fully established in the initial stage of ring crack formation as reported by Chai (2006).

With regard to the analysis of indentation fracture in the following sections, it appears instructive to provide already here some measures relevant in conjunction with the occurrence of Hertzian indentation fracture in real materials. In particular, material properties representative of borosilicate glass — investigated by Mouginot and Maugis (1985) whose experimental data will be considered later on — are

- Young's modulus $E \approx 80\,\mathrm{GPa}$,
- Poisson's ratio $\nu \approx 0.22$,
- specific (per area) work of fracture $\mathcal{G}_\mathrm{c} \approx 9\,\mathrm{J/m^2}$
- tensile strength $\sigma_\mathrm{t} \approx 50 \ldots 150\,\mathrm{MPa}$. (4.2)

The fracture strength σ_t of glass is not considered in Mouginot and Maugis (1985) and is generally known to be rather unclear. Therefore, a range of values is accounted for in view of the strong scatter of this property in case of glass (e.g. Lawn, 1993).

4.1.2 Initial stress field

Analogous to previous work on the above problem, e.g. Frank and Lawn (1967), Mouginot and Maugis (1985) and Fischer-Cripps (1997), the stress field in an uncracked half-space under indentation loading plays a key role in the following analyses. Due to linearity of the boundary value problem prior to fracture the stress components can be written (in the cylindrical coordinate system given in Fig. 4.1) as

$$\sigma_{ij}(r,z) = E\,\frac{d}{a}\,\psi_{ij}\Big(\frac{r}{a},\frac{z}{a},\nu\Big) \tag{4.3}$$

where ψ_{ij} are dimensionless functions. While the complete axisymmetric stress field σ_{ij} (and corresponding dimensionless functions ψ_{ij}) can be found in App. D.1 only those components relevant for the analysis of the indentation problem are briefly recapitulated here. Figure 4.2 shows the variation of the normal stress components, normalized according to (4.3), at the surface ($z=0$) in radial ψ_{rr} and circumferential direction $\psi_{\phi\phi}$ and normal to the free surface ψ_{zz}. Inside the contact region ($r \le a$) all normal components are compressive and even singular at the indenter edge. According to Fig. 4.2, the only positive (tensile) stress is the moderate radial stress σ_{rr} outside the contact region ($r > a$), which is at the free surface $\sigma_{rr}(r, z=0) = E\,d/a\,\psi_0$ with

$$\psi_0\Big(\frac{r}{a},\nu\Big) := \psi_{rr}\Big(\frac{r}{a},\frac{z}{a}=0,\nu\Big) = \frac{1}{\pi}\left(\frac{1-2\nu}{1-\nu^2}\right)\frac{1}{(r/a)^2} \quad . \tag{4.4}$$

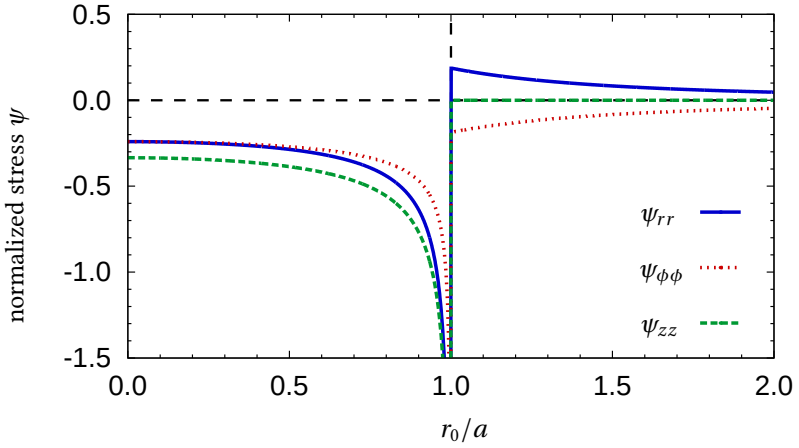


Figure 4.2: Variation of pre-fracture normal stress components at substrate surface $z = 0$ for $\nu = 0.22$. Stress components are normalized as $\psi_{ij} := \sigma_{ij}\, a/(E\, d)$.

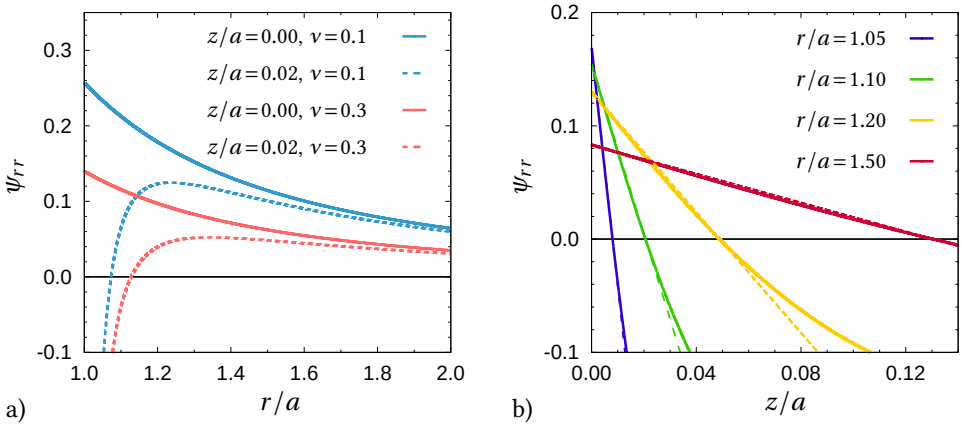


Figure 4.3: Normalized radial stress ψ_{rr} outside the contact region ($r > a$): a) radial stress depending on radial position r/a for different depths $z/a = \{0,\ 0.02\}$ and varying ν; b) radial stress depending on depth z/a for different radii $r/a = \{1.05 \ldots 1.5\}$ and $\nu = 0.22$ (dashed lines illustrate linear stress approximation used in Sect. 4.2).

Note, that the variation of σ_{ij} with the coordinates r and z as well as its dependence on the indenter displacement d in (4.3) scales with the indenter radius a which is prior to fracture the only characteristic length in the boundary value problem. According to (4.4) the radial stress at the free surface is positive for $\nu < 0.5$ and thus understood to be responsible for the formation of the ring crack observed in experiments. Some details of the axisymmetric radial stress distribution outside the contact region ($r \geq a$) where crack initiation is expected are presented in Fig. 4.3 in terms of the dimensionless function ψ_{rr}

introduced in (4.3). Figure 4.3a shows the radial variation of the radial stress at the free surface ($z=0$) and at a small distance below (inside the half-space). It can be seen that, except for $z = 0$, a local stress maximum exists and that the stress level strongly depends on Poisson's ratio with $\sigma_{rr}(r, z=0) = 0$ for $\nu \to 0.5$ according to (4.4). The rapid decay of the radial stress with distance z from the free surface, especially close to the contact radius $r = a$, is depicted in Fig. 4.3b. It should be noted that σ_{rr} ($\sim \psi_{rr}$) decreases almost linearly with increasing z in the range where σ_{rr} is positive. This fact will be exploited below in the FFM analysis (Sect. 4.2) to evaluate the stress state and to calculate the energy release rate.

4.2 Analysis of indentation fracture using the framework of FFM

In this section only the first stage of the indentation problem (Fig. 4.1) — the spontaneous ring crack formation — is analyzed by applying a hybrid fracture initiation criterion as introduced in Sect. 2.2.6. The question to be analyzed is *when*, i.e. at which critical load d, and *where*, i.e. at which radius r_0, this spontaneous crack formation of likewise unknown *length* z_0 takes place. The subsequent cone crack growth is outside the scope of the FFM study.

To be more precise, by assuming the formation of a crack with a finite length $z > 0$ we look for the minimum load d_c (in terms of indenter displacement) which satisfies both (i.e. stress and energy) parts of the hybrid fracture criterion for the same radial position r_0 and length z_0 of the ring crack. Based on experimental observations, the crack orientation thereby is assumed normal to the free surface (cylindrical crack), though this is a principal stress direction only at the free surface. Thus the stress part of the hybrid criterion (2.160) is formulated in terms of the radial stress σ_{rr}, which at the free surface is indeed the maximum principal stress (Fig. 4.2). By applying this assumption to the pointwise fulfillment of the stress criterion (2.161) it is

$$\sigma_{rr}(z) \geq \sigma_t \quad , \tag{4.5}$$

here referred to as *local stress criterion*, which means that the radial stress must reach or exceed the tensile strength σ_t at every point $z \in [0,\ z_0]$ along the hypothetical crack extension z_0 (measured from the free surface $z = 0$).

Alternatively, by applying the above assumption to (2.163) it is

$$\bar{\sigma}_{rr}(z) = \frac{1}{z_0} \int_0^{z_0} \sigma_{rr}(z) \mathrm{d}z \geq \sigma_t \quad . \tag{4.6}$$

In this case the radial stress must reach or exceed the tensile strength only in the mean, in the following this type of approach is referred to as *average stress criterion*. The energy part of the hybrid criterion (2.160) for a cylindrical crack is

$$\bar{\mathcal{G}}(z_0) = \frac{1}{\Delta A_\mathrm{c}} \int_0^{\Delta A_\mathrm{c}} \mathcal{G}\,\mathrm{d}A = \frac{1}{z_0} \int_0^{z_0} \mathcal{G}\,\mathrm{d}z \geq \mathcal{G}_\mathrm{c} \quad , \tag{4.7}$$

here called the *average energy release rate*, where $\mathcal{G} = -\mathrm{d}\Pi/\mathrm{d}A$ is the standard energy release rate (2.149) and $\mathrm{d}A = 2\pi\, r_0 \mathrm{d}z$ and $\Delta A_\mathrm{c} = 2\pi\, r_0\, z_0$ are the increment of cylindrical crack area at radius r_0 and the total crack area, respectively.

In the following, both parts of the criterion are first analyzed separately before they are combined in Sect. 4.2.3 and finally compared with experimental data in Sect. 4.2.4. It is noteworthy with regard to the present boundary value problem that the hybrid fracture criterion is also capable to predict size effects as pointed out in Sect. 2.2.6. In the context of indentation fracture, one such size effect manifests itself in the so-called Auerbach law (Sect. 4.2.4).

4.2.1 Stress criterion

The stress part of the hybrid criterion can either be fulfilled by a local criterion (4.5) or on average (4.6). From Figure 4.3b it can be seen that the radial stress σ_{rr} ($\sim \psi_{rr}$) decreases monotonically with increasing z at fixed r. It is hence sufficient for satisfying the local stress criterion (4.5) at some radius $r = r_0$ that $\sigma_{rr}(r_0, z_0) \geq \sigma_\mathrm{t}$ where z_0 denotes the length of the fictitious ring crack measured from the free surface ($z = 0$). Using the representation (4.3), the stress part of the hybrid criterion can thus be written as

$$E\,\frac{d}{a}\,\psi\Big(\frac{r_0}{a}, \frac{z_0}{a}, \nu\Big) \geq \sigma_\mathrm{t} \tag{4.8}$$

or, by extracting the indenter displacement (normalized by a),

$$\frac{d}{a} \geq \frac{\sigma_\mathrm{t}}{E}\,\frac{1}{\psi\Big(\frac{r_0}{a}, \frac{z_0}{a}, \nu\Big)} \tag{4.9}$$

with

$$\psi := \langle \psi_{rr} \rangle \tag{4.10}$$

for the local stress criterion (4.5). Correspondingly, for the average stress criterion (4.6), it is

$$\frac{d}{a} \geq \frac{\sigma_\mathrm{t}}{E}\,\frac{1}{\overline{\psi}\Big(\frac{r_0}{a}, \frac{z_0}{a}, \nu\Big)} \tag{4.11}$$

with (by using the substitution $\zeta := z/a$ and thus $\mathrm{d}z = a\mathrm{d}\zeta$)

$$\overline{\psi} := \left\langle \frac{1}{z_0/a} \int_0^{z_0/a} \psi_{rr}\left(\frac{r_0}{a}, \zeta, \nu\right) \mathrm{d}\zeta \right\rangle \quad . \tag{4.12}$$

The bracket operator $\langle x \rangle := (|x| + x)/2$ (see also App. A.9), is introduced in the above equations to preclude negative stresses (found at some distance below the surface, see Fig. 4.3b) from contributing to the stress criterion. As depicted in Fig. 4.3b the radial stress decreases almost linearly with z from the surface value given by (4.4) to $\psi_{rr}(z/a = z^*/a) = 0$ at some depth z^*/a (relative to the indenter radius) which depends on the radial position r_0/a (and on ν). Instead of evaluating the radial stress ψ_{rr} for each z/a separately the almost linear stress distribution is exploited in the following by using the linear stress approximation

$$\psi_{rr} \approx \psi_{\mathrm{lin}} := \psi_0 \left(1 - \frac{z/a}{z^*/a}\right) \quad . \tag{4.13}$$

Only z^*/a needs to be determined numerically for each radial position r_0/a. Figure 4.4 shows the corresponding results for different Poisson's ratios $\nu = \{0.1,\ 0.22,\ 0.3\}$.

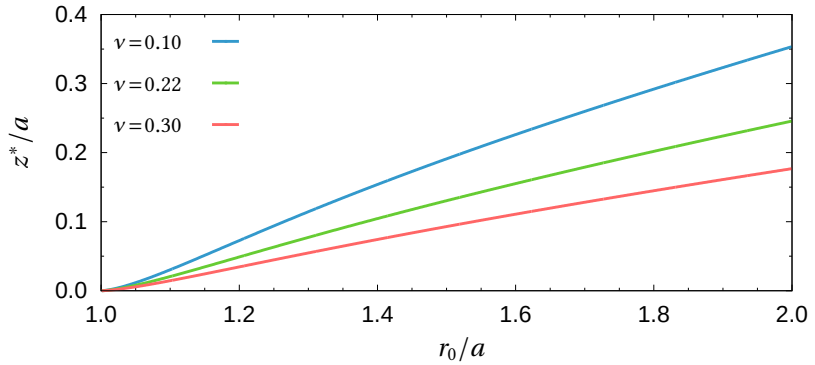


Figure 4.4: Variation of (normalized) depth z^*/a of zero radial stress $\psi_{rr}(z^*/a) = 0$ with radial position r_0/a.

By using the linear stress approximation (4.13) the dimensionless function in the local stress criterion (4.9) is

$$\psi := \frac{1}{\pi} \left(\frac{1 - 2\nu}{1 - \nu^2}\right) \frac{1}{(r_0/a)^2} \left\langle 1 - \frac{z_0/a}{z^*/a} \right\rangle \tag{4.14}$$

with $\psi > 0$ for $z < z^*$, while the dimensionless function in the average stress criterion (4.11) by evaluating (4.12) is

$$\overline{\psi} := \frac{1}{\pi} \left(\frac{1 - 2\nu}{1 - \nu^2}\right) \frac{1}{(r_0/a)^2} \left\langle 1 - \frac{1}{2} \frac{z_0/a}{z^*/a} \right\rangle \tag{4.15}$$

with $\overline{\psi} > 0$ for $z < 2z^*$. Note, that for $z > z^*$ where the radial stress is negative the stress distribution might not be linear any more (Fig. 4.3b). Corresponding deviations between the exact stress evaluation and its linear approximation used here are discussed in App. D.6. Equality in (4.9) and (4.11) determine the load (in terms of indenter displacement) $d_\sigma(r_0, z_0)$ and $d_{\overline{\sigma}}(r_0, z_0)$ necessary to form a cylindrical crack of length z_0 at radius r_0 according to the local or the average stress criterion, respectively.

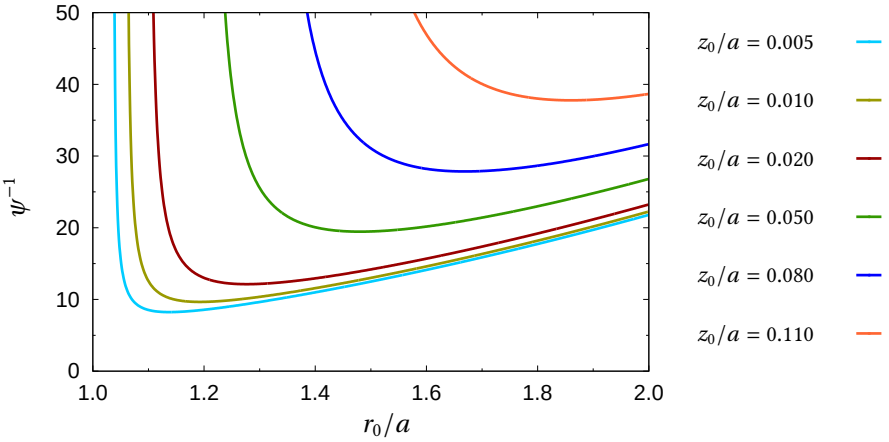


Figure 4.5: Dimensionless function ψ^{-1} given in (4.14) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$.

The dependence of d_σ on the radius r_0 and the length z_0 of the crack is shown in terms of the dimensionless quantity

$$d_\sigma^* := \frac{d_\sigma E}{a \sigma_t} = \psi^{-1} \tag{4.16}$$

in Fig. 4.5 for different crack lengths $z_0/a = \{0.005 \ldots 0.11\}$. The value of Poisson's ratio $\nu = 0.22$ is chosen here already in correspondence to real material data for borosilicate glass (4.2) considered in Sect. 4.2.4. The analogous relation based on the average stress criterion

$$d_{\overline{\sigma}}^* := \frac{d_{\overline{\sigma}} E}{a \sigma_t} = \overline{\psi}^{-1} \tag{4.17}$$

is shown in Fig. 4.6 for the same set of crack lengths z_0/a. Both versions of the stress criterion yield qualitatively similar results and indicate that the critical load to cause fracture according to the stress criterion increases with increasing length of the fictitious crack. Note, that the average stress criterion is fulfilled for smaller loads since $1/\overline{\psi} < 1/\psi$ for each r_0 and z_0. Moreover, with the linear stress approximation (4.13) utilized in both formulations it is

$$\overline{\psi}^{-1}\left(\frac{r_0}{a}, 2\frac{z_0}{a}, \nu\right) = \psi^{-1}\left(\frac{r_0}{a}, \frac{z_0}{a}, \nu\right) \quad , \tag{4.18}$$

which means that the critical load to create a crack of length $2z_0$ with the average stress criterion is equal to the critical load to create a crack of half the length z_0 with the

local stress criterion. It should also be noted that the critical load as a function of crack radius r_0/a displays local minima well outside the contact region $r = a$. In the evaluation of the hybrid fracture criterion in Sect. 4.2.3 and 4.2.4 the average stress criterion (4.6) with results depicted in Fig. 4.6 is used while the effect of utilizing the local stress criterion in the version (4.5) is discussed in App. D.4.

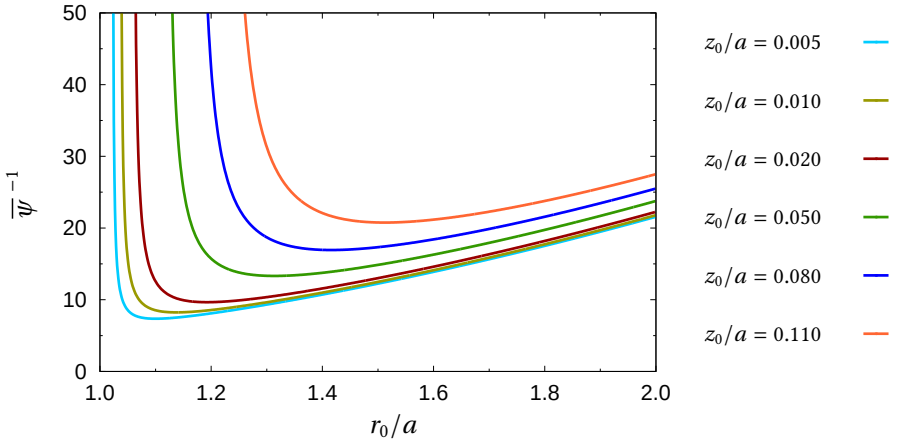


Figure 4.6: Dimensionless function $\overline{\psi}^{-1}$ given in (4.15) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$.

It should be mentioned that the idea of a critical tensile stress acting over the entire length of some (fictitious) flaw being necessary for initial ring crack formation was already suggested in Mouginot and Maugis (1985) as an explanation for the observation that the ring crack forms at some distance ($r_0 > a$) outside the contact region where the stress gradient (see, e.g., Fig. 4.3) is less severe.

4.2.2 Energy release criterion

The evaluation of the energy criterion (4.7) in the present problem of indentation fracture is carried out by two different approaches. First, the energy release is computed semi-analytically by imposing the fictitious ring crack to the pre-existing stress field in the uncracked half-space from Sect. 4.1.2. Second, in order to overcome (and study) the simplifying assumptions made in the semi-analytical approach, detailed finite element analyses of the boundary value problem of the half-space including the ring crack (Fig. 4.1) are performed and the energy release is computed numerically.

With regard to the application of the concept of finite fracture mechanics in general, it appears noteworthy that the assumption of finite crack formation (or extension) yields results which depend on the "type" of loading, i.e. prescribed force *or* displacement. While in case of ("classical") infinitesimal crack advance the type of loading enters only in the

analysis of crack propagation stability, this issue is anticipated in FFM by the assumed finite crack extension. This effect is tacitly neglected in the evaluation of the average energy release rate $\bar{\mathcal{G}}$ in the following by utilizing the pre-fracture stress field in the semi-analytical approach and by considering only the case of prescribed indenter displacement in the numerical analysis. As briefly worked out in App. D.7 the average energy release rate of FFM is generally larger when a prescribed force is considered than in case of a prescribed displacement.

Semi-analytical evaluation

Similar to earlier approaches, e.g. Frank and Lawn (1967), Mouginot and Maugis (1985) and Fischer-Cripps (1997), the calculation of the energy released by the formation of the ring crack is here based on the key assumption that the stress field prevailing in the uncracked half-space (Sect. 4.1.2) can be considered as an imposed external loading and is not altered by the presence of the crack. This allows to compute the mode I stress intensity factor approximately from the relation

$$K_{\mathrm{I}}(z) = 2\sqrt{\frac{z}{\pi}} \int_0^z \frac{\sigma_{rr}(\bar{z})}{\sqrt{z^2 - \bar{z}^2}} f_{\mathrm{sf}}(\bar{z}) \mathrm{d}\bar{z} \tag{4.19}$$

with

$$f_{\mathrm{sf}}(\bar{z}) := 1.3 - 0.3 \left(\frac{\bar{z}}{z}\right)^{5/4} \tag{4.20}$$

for a crack of length z emanating from a free surface and subjected to a spatially varying normal load $\sigma_{rr}(\bar{z})$, e.g. Tada et al. (2000, p. 197). Note, that in earlier approaches (mentioned above) the factor $f_{\mathrm{sf}}(\bar{z})$ is taken equal to 1 which corresponds to half an internal crack of length $2z$, thus neglecting the effect of the free surface which is analyzed in App. D.3.

The *average energy release rate* for a crack of length z_0 according to (4.7) is then computed by using the relation

$$\mathcal{G} = \frac{1 - \nu^2}{E} \langle K_{\mathrm{I}} \rangle^2 \quad , \tag{4.21}$$

which takes only positive stress intensity factors into account, as

$$\bar{\mathcal{G}}(z_0) = \frac{1 - \nu^2}{E} \frac{1}{z_0} \int_0^{z_0} \langle K_{\mathrm{I}}(z) \rangle^2 \mathrm{d}z \quad . \tag{4.22}$$

Since the radial stress decreases almost linearly with z (as depicted in Fig. 4.3) the stress $\sigma_{rr}(z)$ can be approximated by a linear function (4.13), as shown in for the normalized stress. This allows, together with the simplification of the factor (4.20) according to $f_{\mathrm{sf}}(\bar{z}) \approx 1 + 0.3(1 - \bar{z}/z)$, the evaluation of the integrals in (4.19) and (4.22) in closed

form as carried out in App. D.2. Analogous to the stress field (4.3), the average energy release rate can then be written as

$$\bar{\mathcal{G}} = E\,a\left(\frac{d}{a}\right)^2 \phi_{\text{ana}}\left(\frac{r_0}{a}, \frac{z_0}{a}, \nu\right) \tag{4.23}$$

where ϕ_{ana} is a dimensionless function. By using the auxiliary expressions

$$\chi_{\text{K}} := 1.3\,\frac{\pi}{2} - 0.3 \approx 1.742 \qquad \text{and} \qquad \eta_{\text{K}} := 1.3 - 0.3\,\frac{\pi}{4} \approx 1.064 \tag{4.24}$$

the distance from the free surface where K_{I} becomes negative is

$$z^{**} := z^*\,\frac{\chi_{\text{K}}}{\eta_{\text{K}}} \approx 1.637\,z^* \quad . \tag{4.25}$$

As analyzed in Appendix D.2, the dimensionless function used to express the average energy release rate (4.23) of the present boundary value problem is

$$\phi_{\text{ana}} = \begin{cases} \dfrac{\Phi_\nu}{(r_0/a)^4}\left[\dfrac{\chi_{\text{K}}^2}{2}\,\dfrac{z_0}{a} - \dfrac{2}{3}\chi_{\text{K}}\,\eta_{\text{K}}\,\dfrac{1}{z^*/a}\left(\dfrac{z_0}{a}\right)^2 + \dfrac{\eta_{\text{K}}^2}{4}\,\dfrac{1}{(z^*/a)^2}\left(\dfrac{z_0}{a}\right)^3\right] & \text{for } z \le z^{**} \\ \dfrac{\Phi_\nu}{(r_0/a)^4}\,\dfrac{1}{12}\,\dfrac{1}{z_0/a}\left(\dfrac{\chi_{\text{K}}^2}{\eta_{\text{K}}}\,\dfrac{z^*}{a}\right)^2 & \text{for } z > z^{**} \end{cases} \tag{4.26}$$

with the prefactor

$$\Phi_\nu = \frac{4}{\pi^3}\,\frac{(1-2\nu)^2}{1-\nu^2} \quad . \tag{4.27}$$

Results are shown in Fig. 4.8 along with those obtained from finite element analyses discussed in the following.

Numerical evaluation of average energy release rate

The average energy release rate due to formation of the cylindrical ring crack (Fig. 4.1) can be calculated from the change of the total potential energy Π between the uncracked and the cracked configuration at the same load. With the latter specified in terms of prescribed indenter displacement d and due to linearity of the boundary value problem one has

$$\Delta\Pi = \frac{d}{2}\,\Delta F = \frac{d}{2}\,(F_{\text{crack}} - F_0) \tag{4.28}$$

where F_{crack} is the indentation force on the cracked half-space and F_0 that on the uncracked half-space according to (4.1). In both (linear elastic) configurations the force is proportional to E, a and d so that

$$F_{\text{crack}} = E\,a\,d\,\tilde{F}_{\text{crack}}\left(\frac{r_0}{a}, \frac{z_0}{a}, \nu\right) \tag{4.29}$$

and

$$F_0 = E\, a\, d\, \tilde{F}_0(\nu) \tag{4.30}$$

where $\tilde{F}_{\text{crack}}$ and $\tilde{F}_0$ are dimensionless functions. The average energy release rate according to (4.7) can thus be written as

$$\bar{\mathcal{G}} = -\frac{\Delta\Pi}{\Delta A_{\text{c}}} = \frac{E\, a}{4\pi}\left(\frac{d}{a}\right)^2 \frac{a^2}{r_0\, z_0}\left[\tilde{F}_0(\nu) - \tilde{F}_{\text{crack}}\left(\frac{r_0}{a}, \frac{z_0}{a}, \nu\right)\right] \tag{4.31}$$

or in compact form analogous to (4.23) as

$$\bar{\mathcal{G}} = E\, a\left(\frac{d}{a}\right)^2 \phi_{\text{num}}\left(\frac{r_0}{a}, \frac{z_0}{a}, \nu\right) \tag{4.32}$$

where ϕ_{num} is a dimensionless function to be determined numerically (e.g. by finite element analyses) from the force difference ΔF on the cracked and the uncracked half-space.

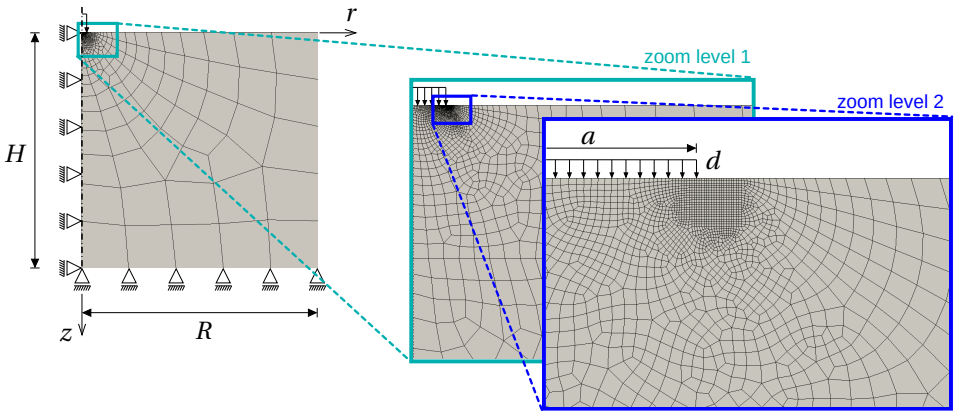


Figure 4.7: Axisymmetric finite element model used for the evaluation of average energy release rate with magnification of contact region.

For the numerical evaluation of the axisymmetric boundary value problem including the ring crack (Fig. 4.1), the finite element package ABAQUS 6.13 is used and the half-space is approximated by a finite domain of height H and radius R. Some aspects of the model consisting of about 10^5 four-node linear displacement elements with reduced integration (of type CAX4R, see Abaqus (2013) for further details) are shown in Fig. 4.7. The discretization is highly refined in the region comprising the contact edge $r = a$ and the ring crack ($r = r_0 > a$). There, a regular mesh is used (see Fig. 4.7) in order to accurately represent the crack by introducing double nodes over a length z_0 extending perpendicular to the free surface by using the "seam" function, see Abaqus (2013). Generation of the variety of finite element meshes containing a thus modeled crack with a wide range of radii r_0 and lengths z_0 as well as the computations are automatized using Python scripting (Puri, 2011). Frictionless loading by the indenter displacement d is modeled by prescribing constant

vertical displacements to nodes inside the contact region ($r \leq a$, $z = 0$). Convergence studies are performed to ensure that numerical results are not affected by the finite domain size, for further details see Dowgiałło (2016), and a dimension of $H = R = 100\,a$ is finally used in the simulations. Nevertheless, for consistency the indenter force F_0 on the uncracked domain is — instead of using the analytical relation (4.1) — also computed numerically (using the same discretization as for the cracked domain, yet without double nodes).

Numerical results for the average energy release rate along with those from the semi-analytical approach above are collectively shown in Fig. 4.8 in terms of the dimensionless functions ϕ_{num} and ϕ_{ana} introduced in (4.32) and (4.23), respectively. For clarity, results are depicted only for a small number of crack lengths $z_0/a = \{0.005 \ldots 0.115\}$. Obviously, the deviations between ϕ_{num} and ϕ_{ana} are rather small which indicates that the assumption made in the semi-analytical approach, that the presence of the ring crack does not much affect the pre-fracture indentation stress field, is not very severe. As expected, however, the deviations become larger for larger crack lengths. In this case the semi-analytical approach (employing the stress field unaffected by the crack) leads to a small overestimation of the energy release (see maxima of curves in Fig. 4.8). The local maxima in the energy release rate with respect to crack radius r_0 — present for all crack lengths z_0, yet at different r_0 — indicates a propensity to crack formation at the respective radii.

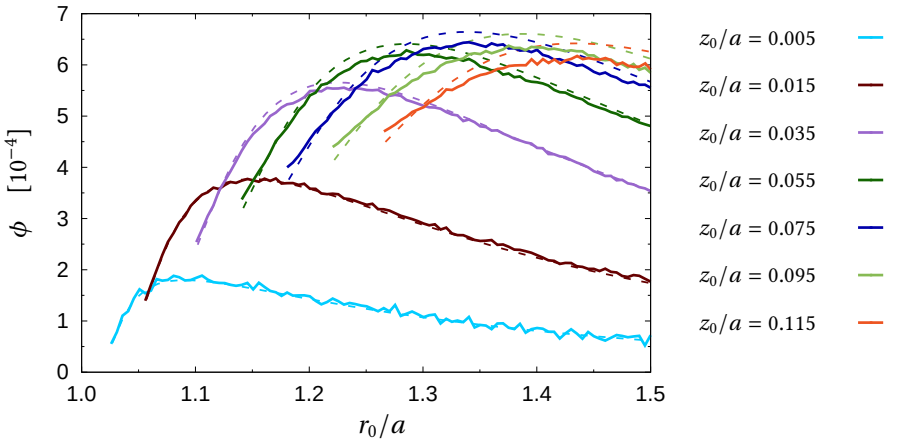


Figure 4.8: Dimensionless measure ϕ of average energy release rate for crack lengths $z_0/a = \{0.005 \ldots 0.115\}$ and $\nu = 0.22$. Solid lines indicate results computed from finite element analyses while dashed lines represent analytical results according to (4.26).

Critical load from energy criterion

With the average energy release rate given by (4.23) or (4.32) the energetic criterion $\bar{\mathcal{G}} \geq \mathcal{G}_\mathrm{c}$ in (4.7) can be written as

$$\frac{d}{a} \geq \sqrt{\frac{\mathcal{G}_\mathrm{c}}{E\,a}\,\frac{1}{\phi\left(\frac{r_0}{a}, \frac{z_0}{a}, \nu\right)}} \quad . \tag{4.33}$$

Equality in (4.33) determines the load (indenter displacement) $d_\pi(r_0, z_0)$ necessary to form a cylindrical crack of length z_0 at radius r_0 according to the energy criterion. Analogous to (4.16) this critical load can be represented in dimensionless form as

$$d_\pi^* := \frac{d_\pi\,E}{a\,\sigma_\mathrm{t}} = \sqrt{\frac{\mathcal{G}_\mathrm{c}\,E}{a\,\sigma_\mathrm{t}^2}}\,\frac{1}{\sqrt{\phi}} \quad . \tag{4.34}$$

The dependence of this critical load on the radius r_0 and the length z_0 of the crack is shown in terms of the dimensionless quantity $1/\sqrt{\phi}$ in Fig. 4.9. Note, that in view of the small difference between the semi-analytical and the numerical evaluation of the average energy release rate (Fig. 4.8) these are taken equal $\phi_\mathrm{ana} \approx \phi_\mathrm{num}$ and $\phi = \phi_\mathrm{ana}$ is used in the following. From Figure 4.9 it can be seen that the critical load to cause fracture according to energy criterion decreases with increasing length z_0 of the fictitious crack (contrary to the effect of the stress criterion in Fig. 4.6) and that local minima of this critical load exist with respect to the crack radius r_0.

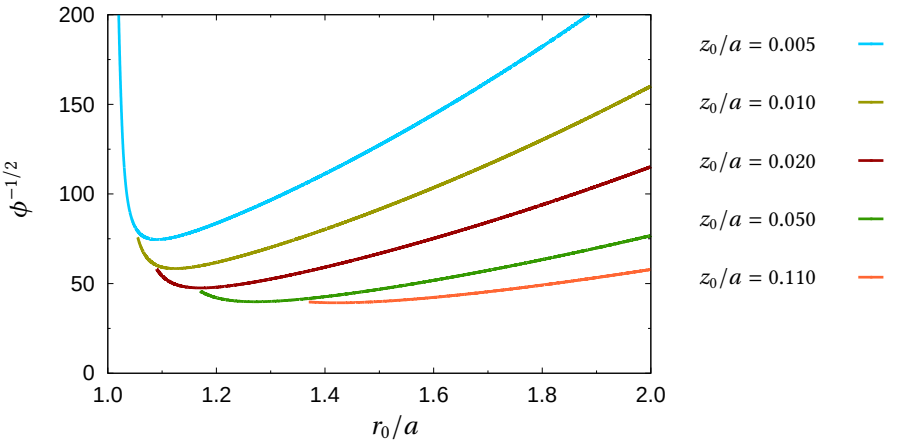


Figure 4.9: Dimensionless function $\phi^{-1/2}$ given in (4.26) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$.

For typical values of the material parameters E, $\mathcal{G}_c$ and σ_t given in (4.2) and an indenter radius of $a = 1\,\text{mm}$ the prefactor of $1/\sqrt{\phi}$ in (4.34) is approximately

$$\sqrt{\frac{\mathcal{G}_c\, E}{a\, \sigma_t^2}} \approx 0.18 \ldots 0.54 \quad . \tag{4.35}$$

Considering this in conjunction with Fig. 4.9 and comparison with Fig. 4.6 indicates that the critical loads (indenter displacements normalized in the same way) according to the stress criterion and the energy criterion are of similar magnitude; a quantitative evaluation is subject of Sect. 4.2.4.

4.2.3 Combined (hybrid) criterion

The hybrid fracture initiation criterion (2.160) formulated by means of the critical indenter displacements $d_{\overline{\sigma}}$ and d_π satisfying the stress (4.11) and the energy criterion (4.33), respectively, can be expressed as

$$d_c = \min_{\{r_0,\, z_0\}} \max \left\{ d_{\overline{\sigma}}(r_0,\, z_0),\, d_\pi(r_0,\, z_0) \right\} \tag{4.36}$$

where d_c is the indenter displacement that actually causes fracture. By taking the maximum of $d_{\overline{\sigma}}$ and d_π in (4.36) it is guaranteed that both criteria are fulfilled, while minimizing over the whole range of r_0 and z_0 determines the critical load as well as the radius and the length of the crack.

Prior to a quantitative evaluation of the criterion for real material data and a comparison with experiments in Sect. 4.2.4, it appears worthwhile to look at the mode of operation of the hybrid criterion in principle when applied to the present situation of indentation fracture. Therefore, the critical indenter displacement (4.36) is written in dimensionless form, analogous to (4.17) and (4.34),

$$d_c^* := \frac{d_c\, E}{a\, \sigma_t} = \min_{\{r_0,\, z_0\}} \max \left\{ \frac{1}{\overline{\psi}},\, \underbrace{\sqrt{\frac{\mathcal{G}_c\, E}{a\, \sigma_t^2}}\, \frac{1}{\sqrt{\phi}}}_{= \sqrt{l_c/(a\, \varepsilon_c)}} \right\} \tag{4.37}$$

where $\overline{\psi}$ and ϕ depend only on r_0/a, z_0/a and ν. The characteristic length

$$l_c := \frac{\mathcal{G}_c}{\sigma_t} \tag{4.38}$$

introduced in (4.37) can be taken as a measure of toughness while

$$\varepsilon_c := \frac{\sigma_t}{E} \tag{4.39}$$

is a measure of strength, so that the dimensionless parameter $\sqrt{l_c/(a\,\varepsilon_c)}$ can be interpreted as the *ratio of toughness to strength* for a certain indenter radius a. The normalized critical load d_c^* hence depends on material properties and indenter geometry only through the parameter $\sqrt{l_c/(a\,\varepsilon_c)}$.

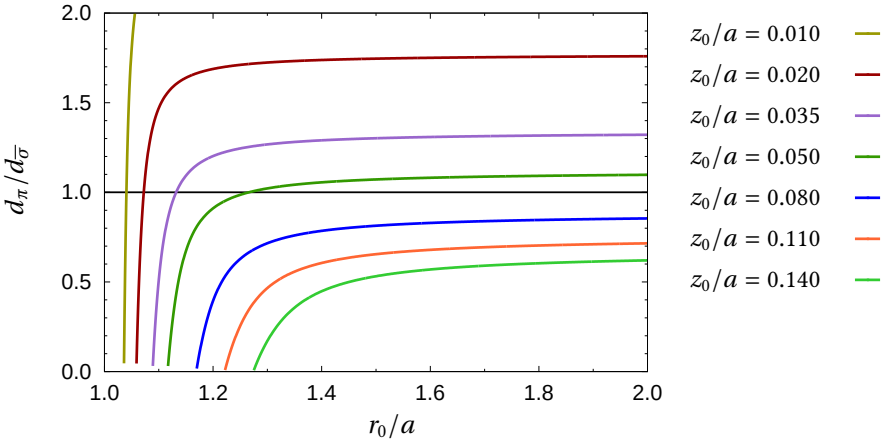


Figure 4.10: Ratio of critical loads $d_\pi/d_{\overline{\sigma}}$ from energy and stress criterion for $\sqrt{l_c/(a\,\varepsilon_c)} = 0.34$ and $\nu = 0.22$.

It also seems instructive to look at the ratio of critical loads according to the energy and stress criteria

$$\frac{d_\pi}{d_{\overline{\sigma}}} = \sqrt{\frac{l_c}{a\,\varepsilon_c}}\,\frac{\overline{\psi}}{\sqrt{\phi}}\left(\frac{r_0}{a},\,\frac{z_0}{a},\,\nu\right) \tag{4.40}$$

which determines the range in which either of the two criteria controls crack initiation. For $d_\pi/d_{\overline{\sigma}} > 1$ the energy criterion is dominant (and vice versa) so that relation (4.40) predicts for larger values of $\sqrt{l_c/(a\,\varepsilon_c)}$ (toughness/strength) a stronger influence of the energy criterion — as expected. In addition, for fixed material properties the energy criterion is favored by smaller indenter radii a. The ratio $d_\pi/d_{\overline{\sigma}}$ of critical loads vs. crack radius r_0/a shown in Fig. 4.10 for various crack lengths z_0/a indicates that at a fixed crack radius the formation of short cracks is more prone to be controlled by the energy criterion ($d_\pi/d_{\overline{\sigma}} > 1$) whereas longer cracks are likely to be controlled by the stress criterion ($d_\pi/d_{\overline{\sigma}} < 1$). The application of the hybrid fracture criterion in principle is illustrated in the following by considering for clarity a limited set of only three different crack lengths $z_0/a = \{0.02, 0.05, 0.08\}$, whereas its exact evaluation in Sect. 4.2.4 involves a continuous variation of z_0/a over a wide range. Figure 4.11 shows the normalized critical loads (indenter displacements) according to the stress criterion (dashed lines) and the energy criterion (solid lines) separately for values of $\sqrt{l_c/(a\,\varepsilon_c)} = 0.34$ (Fig. 4.11a) and $\sqrt{l_c/(a\,\varepsilon_c)} = 0.44$ (Fig. 4.11b). Points A ($z_0/a = 0.02$), B ($z_0/a = 0.05$) and C ($z_0/a = 0.08$) in Fig. 4.11a indicate the minimum load by which *both* criteria are satisfied for each of the crack lengths considered here. They also determine the corresponding radius r_0/a of the ring crack.

Among the three crack lengths considered in Fig. 4.11a the lowest critical load is found for $z_0/a = 0.05$ (point B). That means that according to the hybrid fracture criterion, crack initiation would take place at the critical load $d^* = 13.7$ by the formation of a crack of length $z_0 = 0.05a$ and radius $r_0 \approx 1.27a$, i.e. well outside the punch contact zone ($r = a$) which is in qualitative agreement with experimental observations (as mentioned in the introduction, see Fig. 4.1). Note, that in this case the critical load (point B) is found at the intersection point of the $d^*_{\overline{\sigma}}$- and d^*_{π}-curves (there having slopes of different sign), which means that for crack initiation both stress criterion and energy criterion are satisfied exactly and thus both criteria are active. The critical load in case of the shorter crack ($z_0/a = 0.02$) is found at a local minimum of the d^*_{π}-curve (point A) whereas the critical load in case of the longer crack ($z_0/a = 0.08$) is found at a local minimum of the $d^*_{\overline{\sigma}}$-curve (point C). At the former point (point A) the stress criterion is clearly over-satisfied and crack initiation hence would be controlled solely by the energy criterion. By contrast, at the latter point (point C) the energy criterion is over-satisfied so that crack initiation would be controlled solely by the stress criterion.

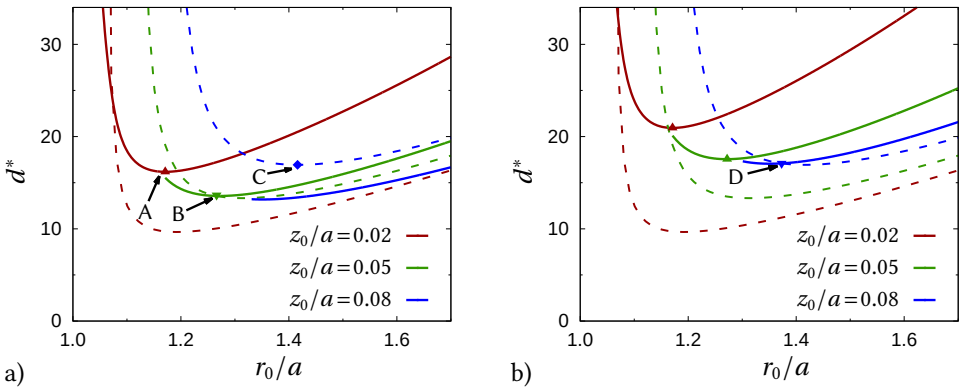


Figure 4.11: Normalized critical loads in terms of indenter displacements d^* from stress criterion (dashed lines $d^*_{\overline{\sigma}}$) and energy criterion (solid lines d^*_{π}) for varying crack lengths $z_0/a = \{0.02,\ 0.05,\ 0.08\}$ and $\nu = 0.22$; a) $\sqrt{l_c/(a\,\varepsilon_c)} = 0.34$, b) $\sqrt{l_c/(a\,\varepsilon_c)} = 0.44$.

Increasing the parameter $\sqrt{l_c/(a\,\varepsilon_c)}$ from 0.34 to 0.44 leads to the situation depicted in Fig. 4.11b where now the minimum critical load (indicated by point D) is found for the largest of the three considered crack lengths (again at the intersection point of the $d^*_{\overline{\sigma}}$- and d^*_{π}-curves). Comparison of Figures 4.11a and b shows that increasing the parameter $\sqrt{l_c/(a\,\varepsilon_c)}$ leads to ring crack formation with a larger (normalized) radius r_0/a and length z_0/a at higher critical loads.

4.2.4 Comparison with experiments

For a quantitative evaluation of the hybrid fracture initiation criterion material data of borosilicate glass given in (4.2) is used. Figure 4.12 lists the consecutive evaluation steps

based on the considerations of the previous sections and App. D. First, a finite number of possible crack lengths $\tilde{z}_0/a$ and radial crack positions $\tilde{r}_0/a$ (both quantities normalized by a) are chosen. Then, for each of these combinations and a specific Poisson's ratio ν, the normalized depth z^*/a where the radial stress is zero must be found (step 2), and functions ψ (representing the stress criterion, step 3) and ϕ (representing the energetic criterion, step 4) are calculated. After the choice of further material parameters (E, σ_t and $\mathcal{G}_c$, step 5), the minimum load where both the stress criterion and the energetic criterion are fulfilled must be found. This is done for a finite number of indenter radii a (step 6).

FFM analysis of Hertzian indentation fracture initiation:

1. Choose specific ν and select a discrete set of m initial crack lengths $\mathcal{I}_Z = \{(\tilde{z}_0/a)_1, \ldots, (\tilde{z}_0/a)_m\}$ and n radial crack positions $\mathcal{I}_R = \{(\tilde{r}_0/a)_1, \ldots, (\tilde{r}_0/a)_n\}$ with $(\tilde{z}_0/a)_i \in [0,\ z_{max}/a]$, $(\tilde{r}_0/a)_j \in (1,\ r_{max}/a]$ and z_{max} and r_{max} sufficiently large.
2. Compute depths z^*/a of zero stress from initial radial stress field (D.13) for set $\mathcal{I}_R$.
3. Calculate critical load from stress criterion (4.16) or (4.17) by evaluation of (4.10), (4.14), (4.15), (D.35), (D.38) or (D.40) for each $\tilde{z}_0/a \in \mathcal{I}_Z$ and $\tilde{r}_0/a \in \mathcal{I}_R$.
4. Evaluate dimensionless function $\phi^{-1/2}$ which represents the fulfillment of the energy criterion by using (4.26), (D.45) or corresponding values obtained by FE analyses (4.32) for each $\tilde{z}_0/a \in \mathcal{I}_Z$ and $\tilde{r}_0/a \in \mathcal{I}_R$.
5. Select a discrete set of p indenter radii $\mathcal{I}_A = \{a_1, \ldots, a_p\}$ with $a_k > 0$ and choose specific material parameters E, σ_t and $\mathcal{G}_c$.
6. Search minimum load d_c/a according to (4.37) in sets of values from **step 3** ($1/\psi$) and **step 4** ($1/\sqrt{\phi}$) over all $\mathcal{I}_Z \times \mathcal{I}_R$ for each $a_k \in \mathcal{I}_A$. Position of critical load in sets of values yield both initial crack length $z_0/a = (\tilde{z}_0/a)_i$ and radial crack position $r_0/a = (\tilde{r}_0/a)_j$.

Figure 4.12: Workflow of FFM analysis to determine indentation fracture initiation (critical load d_c, radial crack position r_0 and initial crack length z_0).

The critical load in terms of normalized indenter displacement d_c/a thus computed from the criterion (4.36) is depicted in Fig. 4.13 as a function of indenter radius a. Obviously, the fracture strength σ_t becomes irrelevant for small indenter radii, i.e. in the range where solely the energetic criterion controls fracture initiation as discussed above. This can also be anticipated from (4.37) resolved for d_c/a where σ_t cancels out when the energy criterion ($\phi^{-1/2}$-term) is dominant. Moreover, Figure 4.13 shows good agreement between FFM results for a high tensile strength of $\sigma_t = 150\,\text{MPa}$ and the experimental data from Mouginot and Maugis (1985).

In Figure 4.14 the normalized radius r_0/a and length z_0/a of the ring crack predicted from (4.36) are shown as functions of the indenter radius a. Both quantities decrease with increasing indenter radius. This means that indenters with a larger radius give rise to the formation of (relatively) shorter cracks (relatively) closer to the contact edge ($r = a$) which is in qualitative agreement with experimental observations by, e.g., Mouginot and Maugis (1985) and Chai (2006). For a quantitative comparison, experimental data from Mouginot and Maugis (1985) are included in Fig. 4.14a. Deviations between those and the predicted

values have to be seen in the light of various idealizations made in the present FFM analysis as well as the scatter in the experimental data (Mouginot and Maugis, 1985, Fig. 19).

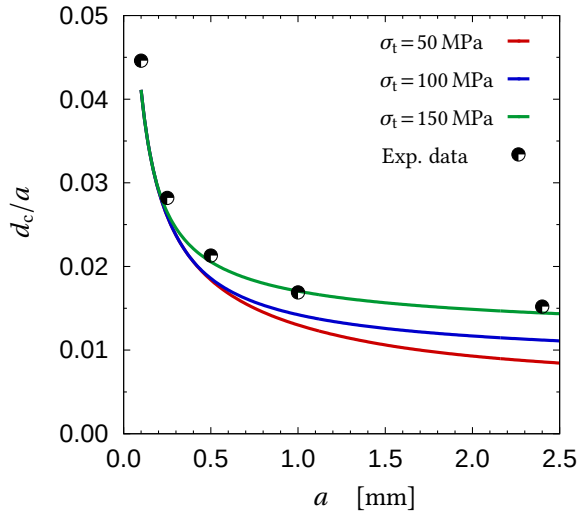


Figure 4.13: Variation of normalized critical indenter displacement d_c/a with indenter radius a, experimental data taken from Mouginot and Maugis (1985).

Rather striking, however, is the predicted "saturation" value of $r_0/a \approx 1.35$ for small indenter radii which is even more pronounced for small tensile strengths. It underestimates the experimental data in Mouginot and Maugis (1985) where r_0/a is found to increase monotonically with decreasing indenter radii. This saturation regime for small indenter radii — also seen in Fig. 4.14b in terms of the relative crack length z_0/a with the saturation value of $z_0/a \approx 8.1 \times 10^{-2}$ — is not present when the local stress criterion (4.9) is employed as depicted in Fig. D.10d of App. D.4 (at least for the same set of material parameters and indenter radii). Thus, it appears to be an artifact of the hybrid fracture criterion applied here. It corresponds to the range where crack initiation is solely controlled by the energy criterion $d_c/d_\pi \approx 1$ (with the stress criterion oversatisfied $d_c/d_{\overline{\sigma}} > 1$), while otherwise both criteria are active. A case where only the stress criterion controls crack initiation does not exist. By using the dimensionless parameter $\sqrt{l_c/(a\,\varepsilon_c)}$ introduced in Sect. 4.2.3 the range where only the energy criterion is active is $\sqrt{l_c/(a\,\varepsilon_c)} \gtrapprox 0.450$, which corresponds to the indenter radii

$$a \lessapprox \frac{1}{0.450^2}\,\frac{\mathcal{G}_c\,E}{\sigma_t^2} = \begin{cases} 1.42\,\text{mm} & \text{for} \quad \sigma_t = 50\,\text{MPa} \\ 0.36\,\text{mm} & \text{for} \quad \sigma_t = 100\,\text{MPa} \\ 0.16\,\text{mm} & \text{for} \quad \sigma_t = 150\,\text{MPa} \end{cases} \tag{4.41}$$

shown in Fig. 4.14. The critical load, when solely determined by the energy criterion, displays a local minimum with respect to the relative crack radius and length (local

maximum of $\phi(r_0/a, z_0/a, \nu)$, Fig. 4.8) at $r_0/a \approx 1.35$ and $z_0/a \approx 0.08$ (also discussed in conjunction with Fig. 4.10). These are the saturation values seen in Fig. 4.14 which depend only on Poisson's ratio ν. A change of the indenter radius a or of material parameters (except ν) in this regime only affects the critical load d_c but not the "optimal" values of r_0/a and z_0/a.

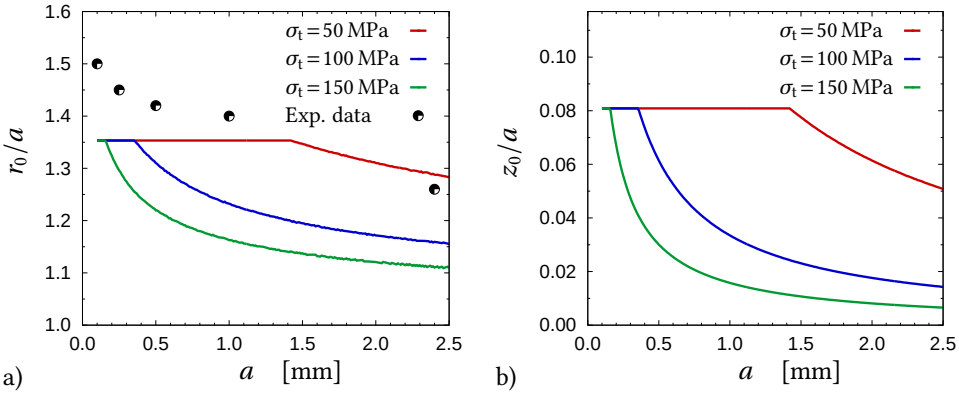


Figure 4.14: Variation of a) normalized crack radius r_0/a and b) normalized crack length z_0/a with indenter radius a for different tensile strengths $\sigma_t = \{50, 100, 150\}$ MPa, experimental data taken from Mouginot and Maugis (1985).

A measure for comparison of critical loads obtained from theoretical and experimental analyses is suggested by the experimental findings of Auerbach (see, e.g., Fischer-Cripps, 2007) who shortly after Hertz stated that the critical force to initiate indentation fracture is approximately proportional to the radius ρ of the (at that time used) spherical indenter (Auerbach, 1891), i.e.

$$F \sim \rho \quad . \tag{4.42}$$

The contact radius a of a spherical indenter depends on the radius ρ of the latter and the indentation depth δ according to

$$a = \sqrt{\rho\,\delta} \quad . \tag{4.43}$$

As calculated by Hertz (1882a) the applied indentation force F is related to the indentation depth (with the assumption of a rigid indenter) by

$$F = \frac{4}{3}\,\frac{E}{1-\nu^2}\,\sqrt{\rho}\,\sqrt{\delta^3} \quad , \tag{4.44}$$

so that, by inserting (4.43) and (4.44) into (4.42), the critical load in Auerbach's experiments scales with the contact radius as

$$F \sim a^{3/2} \quad . \tag{4.45}$$

In the (experimental) literature hence the range of indenter radii where $F_c/a^{3/2} \approx$ const. is referred to as the *Auerbach range*, e.g. Fischer-Cripps (2007). Using this measure of

critical load, results obtained from the hybrid fracture criterion (4.36) are depicted in Fig. 4.15 as a function of indenter radius a. Note, that in case of the cylindrical flat punch considered here (Fig. 4.1) the indenter radius is identical to the contact radius in the above considerations, and that the stress fields outside the contact zone due to spherical or flat indentation are very similar, e.g. Fischer-Cripps (2007). Also included in Fig. 4.15 are experimental data from Mouginot and Maugis (1985) which (see Fig. 21 in the original paper) display some amount of scatter. The stationary value of about 70 MN/m$^{3/2}$ in Fig. 4.15 predicted for small indenter radii is in reasonable quantitative agreement with the experimental findings. This range of constant $F_c/a^{3/2}$ which here extends up to 0.16 mm for the fracture strength $\sigma_t = 150$ MPa or 1.42 mm for $\sigma_t = 50$ MPa corresponds to the above mentioned *Auerbach range*. Also well predicted according to Fig. 4.15 is the increase of the apparent critical load to about 120 MN/m$^{3/2}$ for larger indenter radii up to $a = 2.5$ mm. With regard to experimentally observed critical loads causing indentation fracture it should be noted that some amount of uncertainty also arises from the load increase between the onset of fracture and its complete encirclement of the contact area as reported in Chai (2006).

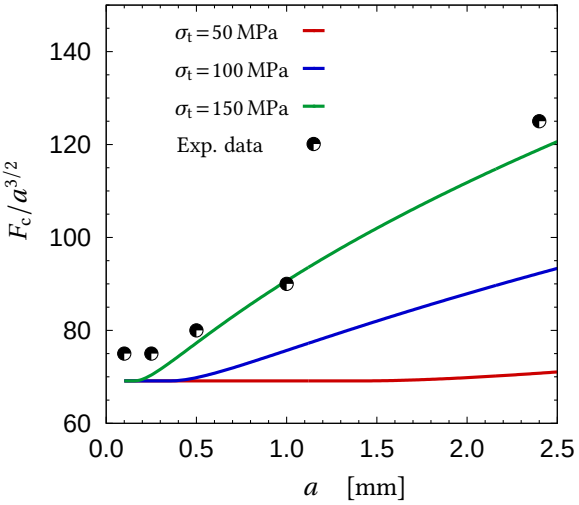


Figure 4.15: Variation of apparent critical load $F_c/a^{3/2}$ with indenter radius a, experimental data taken from Mouginot and Maugis (1985).

4.3 Simulation of indentation fracture using the phase field method

A particular challenge in the theoretical analysis of Hertzian indentation fracture is the initial ring crack formation from a defect-free surface which cannot be described by classical fracture mechanics concepts as discussed in Sect. 4.1. The shortcoming of classical fracture mechanics to describe crack formation can be overcome by utilizing a coupled two-part fracture criterion in the concept of finite fracture mechanics (FFM). Its application to the present problem in Sect. 4.2 has shown reasonable qualitative and quantitative predictive capabilities with respect to the critical load as well as the location and extension of initial ring crack formation — without any additional assumptions. However, FFM provides essentially a fracture initiation criterion but not a tool for tracking subsequent crack growth. This section provides a holistic description of the entire fracture process by employing a phase field methodology.

Crack formation takes place in a complex stress field as worked out in Sect. 4.1.2. Outside the indenter at the free surface the radial stress is the only positive stress component (Fig. 4.2) and thus is understood to be responsible for crack initiation. Below the surface ($z>0$) the stress field varies continuously, see App. D.1, yet still with severe gradients, and the maximum radial stress is found somewhat outside the contact region ($r>a$), see Fig. 4.3a, where in fact crack initiation takes place as sketched in Fig. 4.1. This means that the initial ring crack formation is caused by a rather weak tensile stress concentration within a strongly varying stress field, which renders the quantitative prediction of fracture initiation a difficult task.

The reason for the predictive capabilities of FFM can be seen in the fact that the methodology involves (tensile) strength σ_t and specific fracture energy (toughness) $\mathcal{G}_\mathrm{c}$ as two independent, physically well defined materials parameters. Phase field modeling of brittle fracture in the present section hence is strongly guided by the claim to likewise prescribe strength and toughness as independent quantities. As thoroughly discussed in Sections 3.4 and 3.7.3, this requirement essentially determines the intrinsic length ℓ of the non-local phase field representation. From the material data (4.2) a material specific characteristic length may be deduced as $l_\mathrm{mat} := E\,\mathcal{G}_\mathrm{c}/\sigma_\mathrm{t}^2 \approx 0.032 \ldots 0.29\,\mathrm{mm}$. For indenter radii in the range of $a = 0.5 \ldots 2.5\,\mathrm{mm}$ as considered by Mouginot and Maugis (1985) as well as in the present numerical simulations, an extension of the spontaneously formed ring crack (see Fig. 4.1) of

$$z_0 \approx 0.04\,\mathrm{mm} \tag{4.46}$$

is estimated by the FFM study in Sect. 4.2.4. Hence, a strong correlation between this initial ring crack length (to be resolved in the numerical simulations), the above material specific length l_mat and the characteristic length introduced by the phase field modeling is anticipated and will be subject of thorough analyses in this section.

4.3.1 Critical issues and current phase field approach

As already addressed in Chapter 3, a severe shortcoming of common versions of the phase field approach to fracture is the unphysical occurrence of fracture under compression and the inadequate treatment of traction-free open cracks. Since in the present boundary value problem (see Sect. 4.1) large regions, especially below the indenter, are loaded by compression this becomes particularly relevant here. Therefore, the introduction of an asymmetric tension-compression behavior in the phase field method plays a crucial role. However, as discussed in Sect. 3.5 common isotropic tension-compression splits do not account for the crack orientation. Thus, they generally violate crack boundary conditions and may lead to non-physical crack growth. In the following the severe effects of these shortcomings in application to indentation fracture are shown.

Utilizing the volumetric-deviatoric decomposition (3.230) phase field evolution and corresponding stiffness degradation starts right at the indenter edge (Fig. 4.16a). The zone of broken material then further grows below the indenter (Fig. 4.16b), since deviatoric strains drive the phase field evolution even under compression as already noted in Sect. 3.5.2.

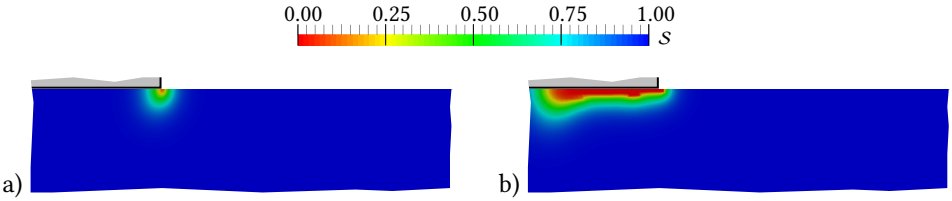


Figure 4.16: Simulation of indentation fracture using the volumetric-deviatoric split of strains (3.230): a) onset of degradation right at the indenter edge, b) unphysical phase field evolution $s \rightarrow 0$ in the compressive zone below the indenter.

Using the spectral decomposition (3.204), stiffness degradation also starts right at the indenter edge (Fig. 4.17a). Further loading yields a larger zone of broken material in this region. Since the spectral split does not result in traction-free (open) cracks for $s \rightarrow 0$, see Sect. 3.5.1 for further details, especially in regions where shear loading dominates like at the indenter edge significant tensile stresses are still present and lead to unphysical broadening of the broken zone (Fig. 4.17b). Obviously, common phase field approaches based on the above splits totally fail in reproducing the experimentally observed behavior.

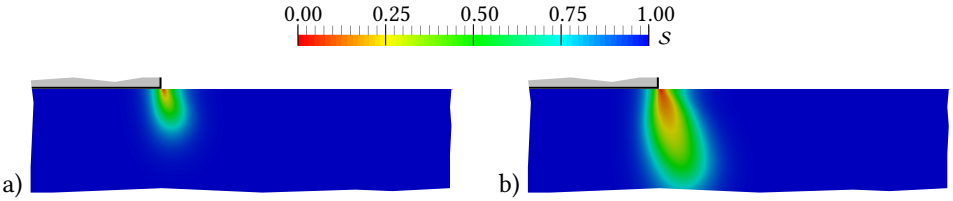


Figure 4.17: Simulation of indentation fracture using the spectral decomposition of strains (3.204): a) onset of degradation right at the indenter edge, b) further diffuse phase field evolution and unphysical broadening.

The physically correct treatment of crack boundary conditions (including closure) generally requires a tension-compression split based on the normal direction of the crack surface as discussed in Sect. 3.6. The incorporation of this split, however, is rather laborious and still leads to problems in conjunction with the FE mesh orientation, see Sect. 3.7.1 for more details. In the present situation of indentation fracture, fortunately, cracks once formed remain open (i.e. traction-free). As a pragmatic approach, therefore, it is justified here to omit any split of the stress tensor and consider the degradation of all stiffness components by a degradation function $g(s)$ according to (3.249) which automatically guarantees a traction-free crack for $s \to 0$. Nevertheless, a tension-compression split is still required in the phase field evolution equation (3.33) in order to avoid the phase field to be driven by compressive elastic energy. Since brittle materials tend to fail normal to the maximum principal tensile stress it appears reasonable to assume phase field evolution (starting here from the intact state) to be driven by positive principal stresses rather than strains. This assumption is in the present context supported by the pre-fracture distribution of the maximum principal strain ε_1 and stress σ_1 in the vicinity of the indenter depicted in Fig. 4.18. Except for the peak values of both quantities prevailing at the surface ($z=0$) at the indenter edge, the contour plots in Fig. 4.18 illustrate that large values of σ_1 are found in the near surface region outside the contact zone (Fig. 4.18b) where indentation fracture actually is expected (see Sect. 4.1). In contrast, large values of ε_1 prevail only at the indenter edge (Fig. 4.18a) where approaches assuming phase field evolution to be driven by the maximum principal strain erroneously predict fracture initiation (Fig. 4.17a). In conjunction with the quadratic single-well function (3.58), the phase field evolution equation (3.33) is thus re-written as

$$\frac{\mathrm{d}g(s)}{\mathrm{d}s} D_{\mathrm{s}}(\breve{\boldsymbol{\sigma}}) - \mathcal{G}_{\mathrm{c}} \left(\frac{1-s}{\ell} + \ell \, \Delta s \right) = 0 \tag{4.47}$$

with the crack driving energy according to (3.251).

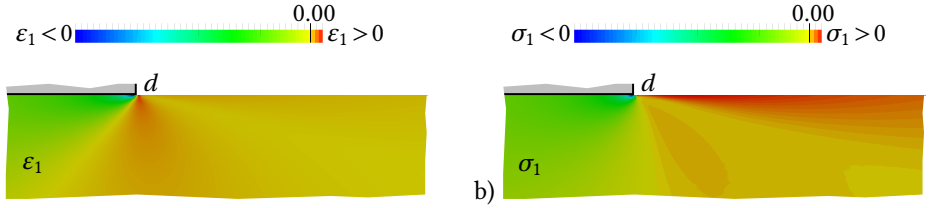


Figure 4.18: Pre-fracture contour plots of a) maximum principal strain ε_1 with concentration at the indenter edge and b) maximum principal stress σ_1 with maximum values at the free surface outside of the indenter. Red and orange colors indicate positive values.

As a pragmatic approach, therefore, it is justified here to omit any split of the stress tensor and consider the degradation of all stiffness components by a degradation function $g(s)$ according to (3.249). According to Sect. 3.3.2, the degradation function $g(s)$ establishes the coupling between mechanical fields and the phase field. With the quadratic single-well function (3.58) and σ_{t} (along with E and $\mathcal{G}_{\mathrm{c}}$) fixed as material parameters according

to (4.2), the relation (3.126) yields the internal length ℓ in dependence on the degradation function $g(s)$, quantified by f_{gw} introduced in Sect. 3.4.2. For instance, in case of the cubic degradation function (3.71) one obtains

$$\ell = \frac{2187}{6250}\,\frac{\mathcal{G}_{\mathrm{c}}\,E}{\sigma_{\mathrm{t}}^2} \approx 0.350\,\frac{\mathcal{G}_{\mathrm{c}}\,E}{\sigma_{\mathrm{t}}^2} = \begin{cases} 0.1008\,\mathrm{mm} & \text{for} \quad \sigma_{\mathrm{t}} = 50\,\mathrm{MPa} \\ 0.0112\,\mathrm{mm} & \text{for} \quad \sigma_{\mathrm{t}} = 150\,\mathrm{MPa} \end{cases} . \tag{4.48}$$

In this regard, as discussed in Sect. 3.7.2, the regularized representation of the crack erroneously captures the surface energy at the crack tip depending on the size of ℓ. Especially during crack nucleation (with the crack length $z \approx 0$) and for short cracks ($z \approx \ell$) this energy portion plays a crucial role. As a compromise, the crack surface energy should be reasonably approximated at least in the mean over the fracture initiation process. Therefore, the internal length parameter ℓ has to be chosen sufficiently small compared to the extension z_0 of the "spontaneously" formed ring crack (see Fig. 4.1). However. the internal length calculated in (4.48) is too large to be suitable as a regularization parameter in view of the expected initial crack length z_0 for the studied indenter radii (4.46). A poor approximation of the surface energy would be a serious consequence (see Fig. 3.38b).

According to Fig. 3.38b a value of at least

$$\ell < \frac{z_0}{10} \ldots \frac{z_0}{4} \tag{4.49}$$

is required to approximate the crack surface energy during crack initiation in a satisfactory manner. From the 1D example in Fig. 3.14 it is obvious that the quadratic degradation function (3.70) does not lead to a linear material behavior prior to fracture and is therefore not an option here. The relation (3.183) with the prefactor f_{gw} discussed in Sect. 3.4.2 for different formulations can be used to find the optimal degradation function. Since the cubic degradation function (3.71) yields too large values of the internal length (4.48) and, on the other side, degradation functions of much higher order are inappropriate as pointed out in Sect. 3.4.2, the quartic degradation function (3.74) with $a_{\mathrm{s}} = 10^{-4}$ is used here as a compromise. Evaluation of (3.126) yields the internal length as

$$\ell = \frac{32768}{177147}\,\frac{\mathcal{G}_{\mathrm{c}}\,E}{\sigma_{\mathrm{t}}^2} \approx 0.185\,\frac{\mathcal{G}_{\mathrm{c}}\,E}{\sigma_{\mathrm{t}}^2} \quad . \tag{4.50}$$

For E and $\mathcal{G}_{\mathrm{c}}$ fixed according to (4.2) the internal length thus varies monotonically as a function of the tensile strength with values in the range

$$\ell \approx \begin{cases} 0.0533\,\mathrm{mm} & \text{for} \quad \sigma_{\mathrm{t}} = 50\,\mathrm{MPa} \\ 0.0059\,\mathrm{mm} & \text{for} \quad \sigma_{\mathrm{t}} = 150\,\mathrm{MPa} \end{cases} . \tag{4.51}$$

It should be noted that the requirement (4.49) in conjunction with the estimate $z_0 \approx 0.04\,\mathrm{mm}$ is still violated for low tensile strengths. Therefore, in the following, the numerical investigation is focused on tensile strengths above $\sigma_{\mathrm{t}} = 115\,\mathrm{MPa}$ for which $\ell \leq 0.01\,\mathrm{mm}$.

In the present problem the irreversibility of the fracture process is accounted for using the "crack-like" constraint (3.36) on the phase field, since convergence studies by Borden (2012) and a detailed investigation by Linse et al. (2017) revealed that the application of the "damage-like" irreversibility (3.40) is in general not suitable to reproduce the crack surface energy. Details on the implementation of the irreversibility constraint are provided in App. C.3.

4.3.2 Simulation of indentation fracture process

According to the present boundary value problem (Sect. 4.1) an axisymmetric finite element formulation with bilinear shape functions of Lagrangian type is employed, with details provided in App. C.2. To solve the resulting coupled set of algebraic equations the alternate minimization procedure described in Sect. 3.2.4 is preferred, where displacements $\boldsymbol{u}$ and phase field s are computed in an alternating manner. This choice prevents unsymmetric entries in the tangent matrix caused by the physically motivated separate choice of stiffness degradation (3.249) and crack driving energy (3.252). Reliable solutions are achieved by small load steps and by performing several iterations between the solutions of both equation systems in each step. Due to the absence of a strong stress concentration and the comparatively small zone which is affected by crack formation, convergence between both field solutions is checked by comparing the phase field solution between consecutive iterations instead of using a global energetic stopping criterion as proposed, e.g., by Ambati et al. (2015a). The maximum number of iterations between switching from one field computation to the other (200 per load increment), is very seldom attained so that in general an equilibrated solution is obtained. Despite the "spontaneous" character of initial ring crack formation, inertia is neglected in the present study. This is motivated from the FFM study of indentation fracture (Sect. 4.2) which has shown that the quasi-static energy balance is satisfied in the mean during ring crack formation and subsequent crack growth is stable.

The size of the discretized domain (see Fig. 4.19) with $100a \times 100a$ (a = indenter radius) is chosen large compared to the region where crack initiation and propagation takes place, since results are partly to be compared with analytical solutions obtained for a half-space. Boundary conditions of zero normal displacement and zero shear stress are applied at the bottom and on the left hand side (symmetry axis) as depicted in Fig. 4.19. On the right hand side the boundary is traction-free. Loading (indenter displacement d) is imposed by prescribed vertical nodal displacements in the friction-less contact region ($r \leq a$). The load increment $\Delta d \lesssim 10^{-6} a$ is adaptively controlled to accurately capture fracture initiation.

The finite element mesh is refined in such a way that the phase field at the crack is accurately resolved. Small ℓ / h_e-ratios (with h_e being the finite element size) lead to a very poor approximation of the regularized crack and an overestimation of the corresponding

crack surface energy. Experience shows, that the ratio $\ell/h_e \approx 2$ usually given in the literature is not sufficient. Therefore the element size is chosen to be

$$h_e \approx \frac{\ell}{4} \tag{4.52}$$

in the region of interest. Since the position of the initial crack is a priori unknown, the refined area must be sufficiently large as shown in Fig. 4.19. The area below the indenter is also refined by elements of the size $h_e \leq a/100$. One main problem is that the element size is coupled via ℓ to material parameters (E, $\mathcal{G}_c$, σ_t) as discussed in Sect. 3.4.2 while the domain size scales with the indenter radius a. Since the indenter radii to be studied range from 0.5 mm to 2.5 mm, i.e. vary by a factor of 5, the size of the refined area varies as well, while the elements are of a constant size. This makes the numerical analysis very expensive. It should be emphasized that the mesh is not refined locally, since it should not trigger the position of the initial crack.

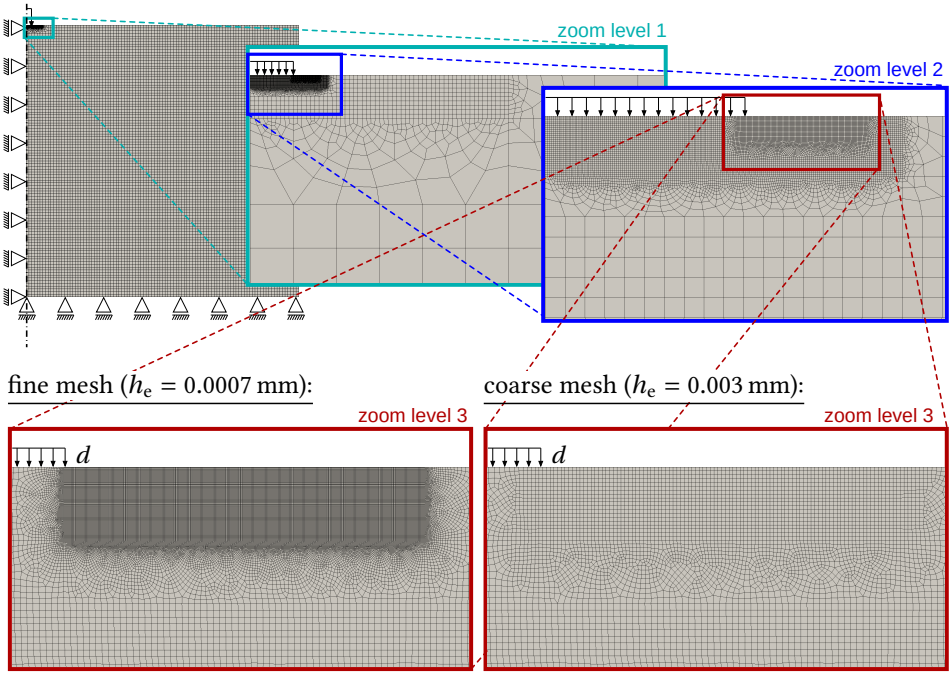


Figure 4.19: Discretization of boundary value problem at different zoom levels showing refined mesh (h_e = {0.0007, 0.003} mm) in the zone of interest.

To eliminate the overestimation of the crack surface energy caused by the discretization as discussed in Sect. 3.7.3 a recalibration of both the fracture toughness $\mathcal{G}_c$ and the internal length ℓ is carried out according to (3.313). This readjustment of parameters with values given in Tab. 4.1 guarantees that in numerical simulations the prescribed tensile strength as well as the specific crack surface energy (fracture toughness) are correctly

reproduced. Note, that this calibration does not eliminate the error caused by the phase field approximation of the crack tip (see Sect. 3.7.2).

Given parameters:	$\mathcal{G}_c = 9 \times 10^{-3}$ N/mm		
σ_t [MPa]	115	125	150
h_e [10^{-3} mm]	2.5	2.0	1.5
Corresponding internal length according to (4.50)			
ℓ [10^{-3} mm]	10.071	8.524	5.919
Adjusted internal length and fracture toughness acc. to (3.313)			
ℓ^* [10^{-3} mm]	8.821	7.524	5.169
h_e/ℓ^*	0.283	0.266	0.290
$\mathcal{G}_c^*$ [10^{-3} N/mm]	7.883	7.944	7.860

Table 4.1: Numerical recalibration of parameters $\mathcal{G}_c$ and ℓ to simulate Hertzian indentation fracture.

With the modifications discussed above and in Section 4.3.1 the phase field method is in the following applied to the simulation of the entire process of indentation fracture. This includes crack formation at some a priori unknown location outside of the contact region as well as the development of a cone crack with increasing load, as observed in experiments and discussed in Sect. 4.2.

The progression of the fracture process with increasing load (indenter displacement d) is illustrated in terms of the evolution of the phase field ($s = 1 \rightarrow 0$) in Fig. 4.20. Here, an indenter radius of $a = 0.5$ mm and a tensile strength of $\sigma_t = 115$ MPa are considered, with the remaining material parameters given in (4.2). The spatial phase field evolution initially takes place in a rather diffuse manner (Fig. 4.20a) with values far above the irreversibility threshold $s_0 = 0.05$. Yet, with increasing load the phase field localizes, recognizable by an intense phase field drop. By falling below the threshold $s < s_0$, the localized phase field features a crack which extends almost normal from the free surface well outside the contact region (Figs. 4.20b–d). Note, that during phase field localization and formation of this ring crack, a recovery with $\dot{s} > 0$ of the initially diffuse phase field seen in Fig. 4.20a takes place outside the vicinity of the crack. Quantitatively, the fracture process can also be tracked from the accompanying evolution of the fracture surface energy depicted in Fig. 4.21. For instance, the "spontaneous" character of the ring crack formation (discussed in Sect. 4.2) can be seen in Fig. 4.21 from the abrupt increase of the fracture surface energy at an indenter displacement slightly above $d \approx 0.009$ mm. After that, the crack grows into the well-known cone shape with increasing load (Figs. 4.20d–g). The fracture surface energy (Fig. 4.21) then displays a progressive growth due to the nonlinear increase of cone area with crack extension.

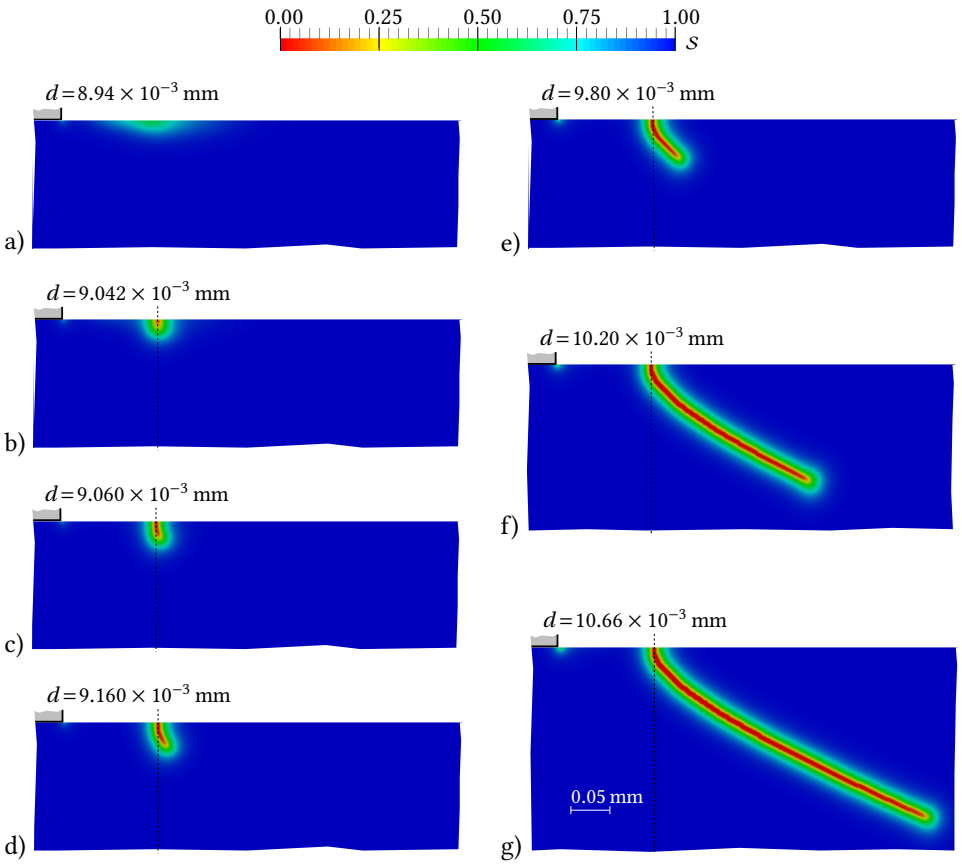


Figure 4.20: Simulation of Hertzian indentation fracture: phase field contour plots display a) slight phase field evolution at the free surface outside the contact region, b)–d) localization of the phase field within a few load steps resulting in crack initiation, e)–g) subsequent crack growth and formation of a cone. The radial position of the ring crack is $r_0 = 0.62$ mm (indicated by the vertical line). Note, that only a small part of the discretized domain (Fig. 4.19) is shown here.

For the situation considered above ($a = 0.5$ mm, $\sigma_t = 115$ MPa), a radius of crack initiation $r_0 = 0.62$ mm is obtained (vertical line in Fig. 4.20). Much less clear is the vertical extension z_0 of the spontaneously formed ring crack (cf. Fig. 4.1) which, however, can be estimated from the sudden jump ΔE_s of the crack surface energy in Fig. 4.21. By assuming a cylindrical crack surface of radius r_0 its axial extension is approximated as

$$z_0 \approx \frac{\Delta E_s}{2\pi\, r_0\, \mathcal{G}_c} \quad . \tag{4.53}$$

With $\Delta E_s \approx 8.8 \times 10^{-7}$ J from Fig. 4.21 and $\mathcal{G}_c \approx 9$ J/m^2 one obtains

$$z_0 \approx 2.51 \times 10^{-2}\,\text{mm} \quad , \tag{4.54}$$

a value not too much different from (4.46) estimated using FFM in Sect. 4.2. Due to partial recovery of the phase field outside the vicinity of the crack and a therefore unclear contribution of the corresponding surface energy to crack initiation, one might instead use $\Delta E_s^* \approx 1.14 \times 10^{-6}$ J as also indicated in Fig. 4.21 which leads to

$$z_0 \approx 3.25 \times 10^{-2}\,\text{mm} \quad . \tag{4.55}$$

These values for z_0 also correspond well to the amount of initial vertical crack growth reported by Kocer (2003b).

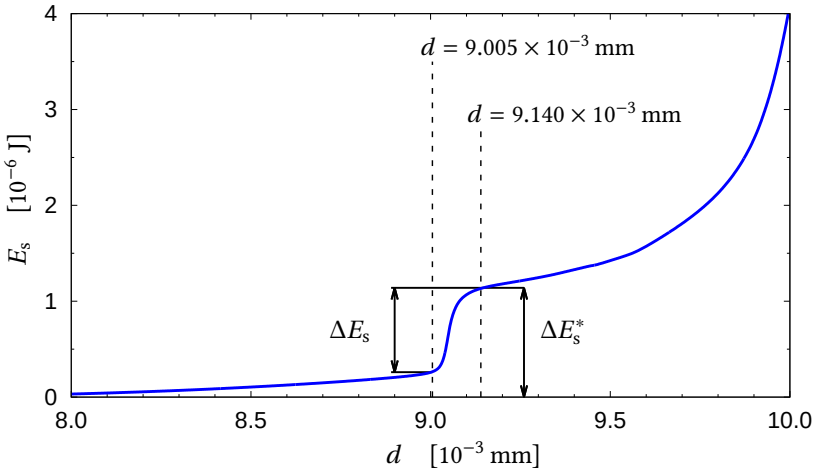


Figure 4.21: Evolution of crack surface energy E_s. Fracture initiation is indicated by the sudden jump ΔE_s which is utilized to estimate the spontaneous ring crack extension z_0.

While the above results indicate that the present phase field method allows a reasonable reproduction of the indentation fracture process, a more detailed study of physical parameter influences and a comparison with experimental findings is provided in the following two sections.

4.3.3 Initial ring crack formation

The influence of the indenter radius on fracture initiation, in particular the required critical load but also the location of the initially forming ring crack, has been subject of extensive research as discussed in the Introduction. For comparison with the experimental study by Mouginot and Maugis (1985) on borosilicate glass, in the following, simulations are focused on the range of indenter radii $a = 0.5 \ldots 2.5$ mm considered by these authors. Moreover, since the tensile strength of glass σ_t is a rather unclear material parameter, a range of reasonable values is adopted here.

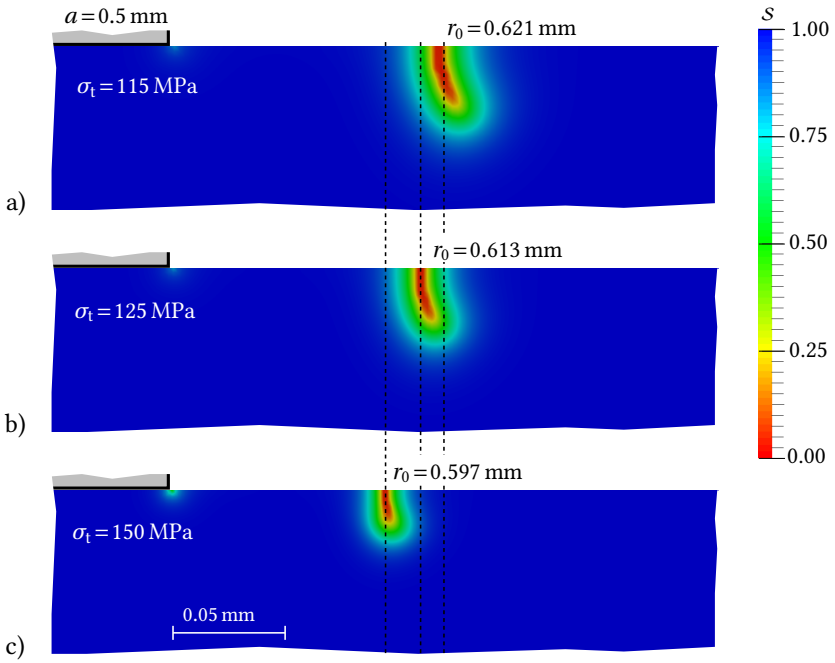


Figure 4.22: Fracture initiation for small indenter with $a = 0.5$ mm and different tensile strengths: a) $\sigma_t = 115$ MPa, b) $\sigma_t = 125$ MPa, c) $\sigma_t = 150$ MPa.

For different indenter radii and strength values, the contour plots in Figs. 4.22 and 4.23 show ring cracks shortly after initiation, i.e. somewhat after the sudden jump of the surface energy ΔE_s (cf. Fig. 4.21) discussed in Sect. 4.3.2. These figures indicate that the radial crack position r_0 decreases with increasing material strength σ_t. Furthermore, from the comparison of Fig. 4.22 with Fig. 4.23 one can deduce that (for constant material strength) a larger indenter provokes ring crack formation at a relatively smaller radial position r_0/a. For further quantitative insight, Fig. 4.24 shows the relative radial positions r_0/a of fracture initiation depending on the indenter radius a and the material's tensile strength σ_t. Present simulation results (solid lines with symbols) are displayed along with those obtained from FFM (dashed lines) presented in Sect. 4.2 and experimental data by Mouginot and Maugis (1985). The general trend of smaller relative crack radii r_0/a with increasing indenter radii a is clearly visible. The phase field prediction of smaller r_0/a with increasing material strength σ_t also quantitatively agrees with results from FFM (Sect. 4.2). However, quantitative deviations between the two theoretical/numerical studies and the experimental data may be attributed to increasing influences of the idealizations in the theoretical model discussed in Sect. 4.1.1 as well as a decreasing accuracy of the smeared crack representation in the phase field approach for smaller indenters.

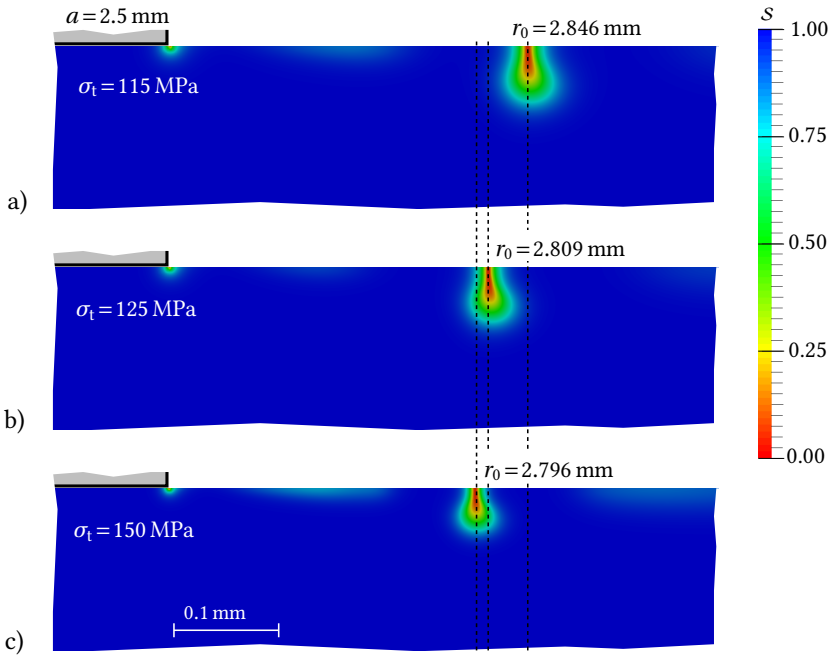


Figure 4.23: Fracture initiation for large indenter with $a = 2.5$ mm and different tensile strengths: a) $\sigma_t = 115$ MPa, b) $\sigma_t = 125$ MPa, c) $\sigma_t = 150$ MPa.

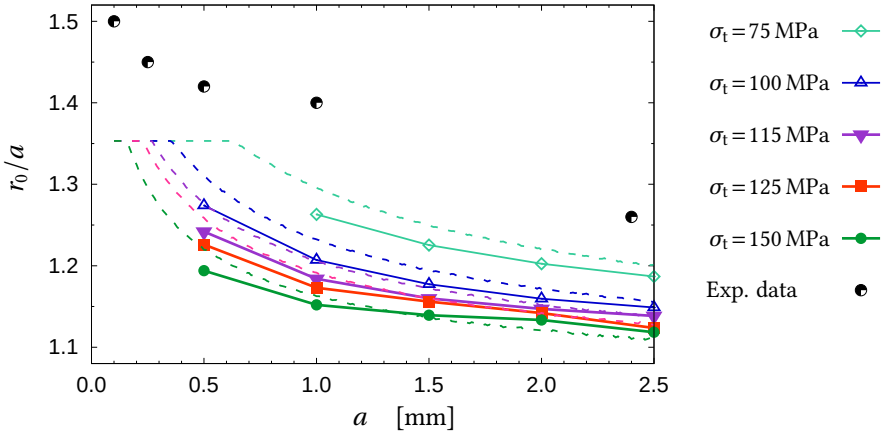


Figure 4.24: Relative radial positions r_0/a of fracture initiation depending on indenter radius a shown for different tensile strengths σ_t in comparison with FFM results (dashed lines) from Sect. 4.2 and experimental data by Mouginot and Maugis (1985).

It should be mentioned that results for $\sigma_t = 100\,\text{MPa}$ and $\sigma_t = 75\,\text{MPa}$, which strictly speaking fall outside the range of applicability of the present phase field method as discussed in Sect. 4.3.1, are shown here only for the sake of completeness and are to be taken with caution. Moreover, the phase field result for $\sigma_t = 75\,\text{MPa}$ and $a = 0.5\,\text{mm}$ has been omitted in Fig. 4.24 (as well as in the following Fig. 4.25) since in this case the ring crack position r_0/a could not be identified as discussed in Sect. 4.3.5 below.

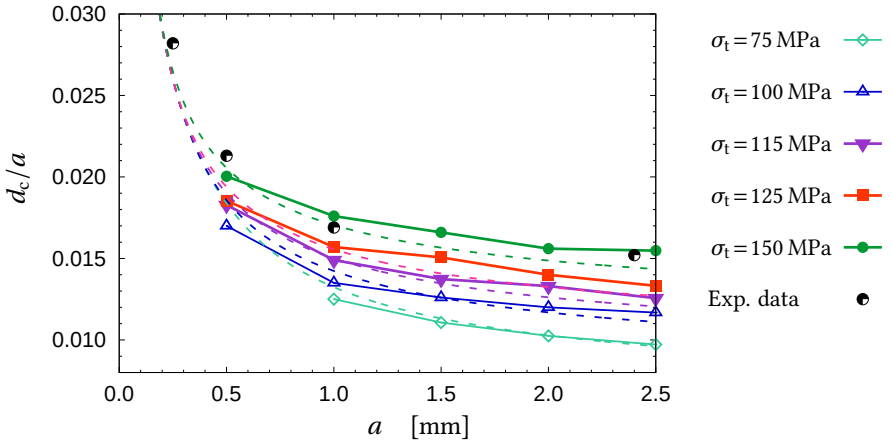


Figure 4.25: Critical indenter displacement d_c/a for fracture initiation depending on indenter radius a shown for different tensile strengths σ_t in comparison with FFM results (dashed lines) from Sect. 4.2 and experimental data by Mouginot and Maugis (1985).

In Figure 4.25 the critical load (normalized indenter displacement) d_c/a to initiate fracture is depicted as a function of the indenter radius. Presented in this way, a very good agreement between the present phase field results and those by FFM from Sect. 4.2 as well as the experimental data by Mouginot and Maugis (1985) is found. Lower values of the critical load, of course, are obtained for smaller tensile strengths σ_t. Moreover, a better agreement between numerical and experimental results in terms of a global quantity such as d_c/a than in terms of a local one such as r_0/a is not very surprising.

Just for comparison of predictions by the present phase field approach with FFM, the influence of Poisson's ratio on crack initiation is explored, again in terms of the initial crack position r_0 and critical load d_c. Therefore a fictitious material with varying ν, but constant E and $\mathcal{G}_c$ according to (4.2) and $\sigma_t = 115\,\text{MPa}$, is considered. Figure 4.26a shows that the initial crack position r_0 (normalized by a fixed $a = 0.5\,\text{mm}$) increases with increasing Poisson's ratio. This relation qualitatively corresponds to the results obtained from FFM (dashed line in Fig. 4.26a), however, the initial crack position predicted by the phase field simulations is closer to the indenter, similar to the results shown in Fig. 4.24. Furthermore, results from phase field simulations show that the critical load d_c increases with increasing Poisson's ratio, which is qualitatively and quantitatively in good agreement with the FFM results as shown in Fig. 4.26b.

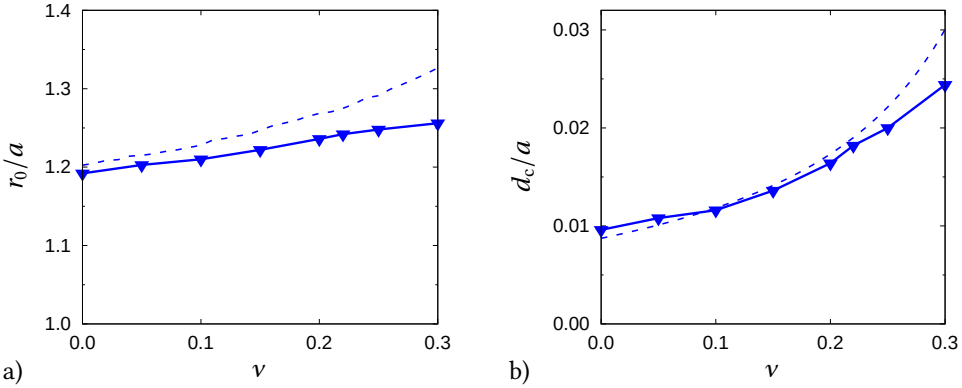


Figure 4.26: Effect of Poisson's ratio ν on crack initiation (keeping E, $\mathcal{G}_c$, $\sigma_t = 115$ MPa and $a = 0.5$ mm fixed): a) radial position r_0/a of initial crack, b) critical load d_c/a for fracture initiation. Solid lines with symbols indicate phase field results while dashed lines represent results obtained by FFM following Sect. 4.2.

4.3.4 Cone crack propagation — effect of Poisson's ratio

Once the initial ring crack has formed, subsequent fracture evolution turns into a cone-shaped crack as discussed in Sect. 4.1. The angle of this cone depends on Poisson's ratio ν as reported, e.g., by Lawn et al. (1974) and Kocer (2003b). Individual experimental data for different materials can be found in the literature, e.g. $\varphi = 24.5° \pm 1°$ for fused silica with $\nu = 0.16$ (Benbow, 1960), $\varphi = 24°$ for Pyrex glass with $\nu = 0.20$ (Conrad et al., 1979), $\varphi = 22.4° \pm 2.8°$ for soda-lime glass with $\nu = 0.21$ (Kocer and Collins, 1998), $\varphi = 22°$ for borosilicate glass with $\nu = 0.22$ (Mouginot and Maugis, 1985) and $\varphi = 21.5° \pm 1°$ for silicate glass with $\nu = 0.25$ (Roesler, 1956a). The ordering of smaller angles for larger Poisson's ratio was also obtained in earlier numerical studies with FEM (Kocer, 2003b) and X-FEM (Tumbajoy-Spinel et al., 2013), both assuming pre-existence of an initial ring crack. In the following, the continuous evolution of the cone-shaped crack is simulated with the phase field method. The focus is on an indenter radius of $a = 0.5$ mm and a tensile strength of $\sigma_t = 115$ MPa. The assumption that the cone angle depends essentially on Poisson's ratio ν allows to compare different materials without taking their different values of E, $\mathcal{G}_c$ and σ_t into account.

For Poisson's ratio in the range $\nu = 0.0 \ldots 0.3$ and E and $\mathcal{G}_c$ according to (4.2), Fig. 4.27 shows the simulated cone cracks with the inclination denoted by the angle φ. In qualitative agreement with the above mentioned experimental and numerical literature findings, larger values of ν give rise to smaller cone angles. Note, that some minor amount of phase field evolution is observed at the indenter edge with increasing Poisson's ratio (Figs. 4.27c and d) owing to the dependence of the driving energy (3.252) on ν.

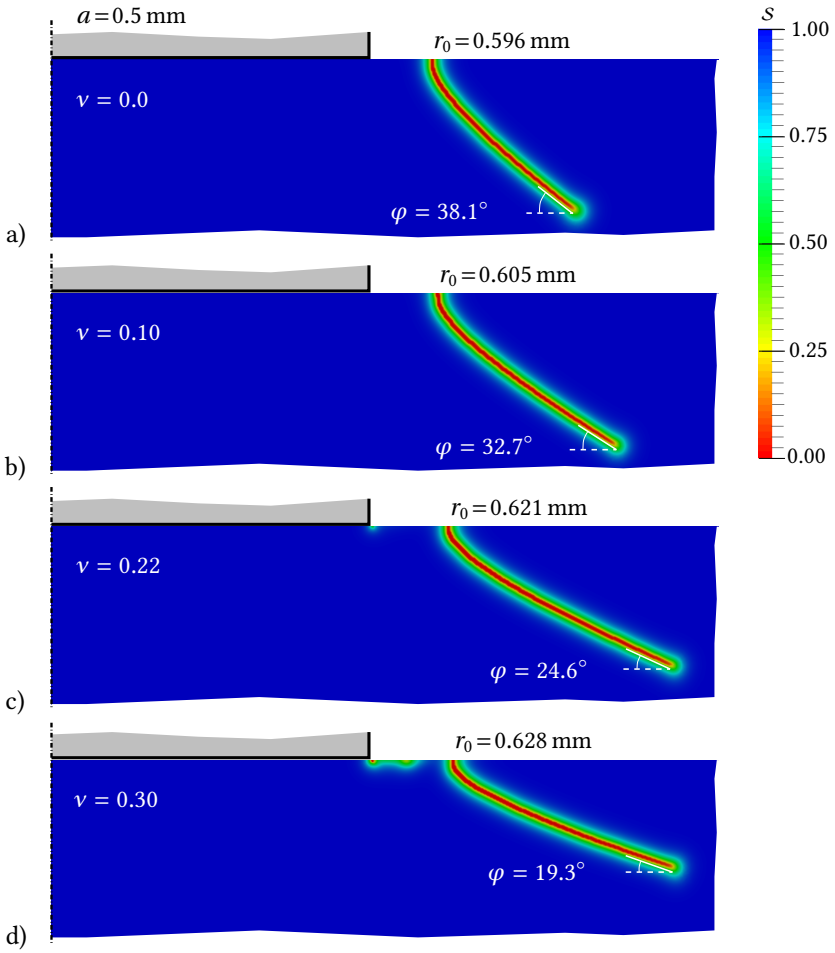


Figure 4.27: Simulated cone cracks with inclination angle φ shown for different values of Poisson's ratio: a) $\nu = 0$, b) $\nu = 0.1$, c) $\nu = 0.22$, d) $\nu = 0.3$. The area below the indenter up to the symmetry axis is shown to illustrate the initial crack position r_0 and cone crack extension in relation to the indenter radius (here $a = 0.5$ mm).

In Figure 4.28 the present simulation results are quantitatively compared with those from literature (see above) and a reasonable agreement can be observed. In particular, an almost linear variation of the cone angle φ with Poisson's ratio is found which corresponds to the FEM results by Kocer (2003b). It appears that the simulated cone angles from the present study somewhat overestimate those from the literature. However, in the latter, see also Mikosza and Lawn (1971), the cone angles have been determined at significantly larger amounts of crack extension than in the present simulations (see Fig. 4.27) where the cone cracks are still in an early stage. Extending the phase field simulations beyond the situation in Fig. 4.27 to more "mature" cone cracks would yield somewhat smaller inclination angles

and thus a better agreement than in Fig. 4.28. Early phase field investigations of the indentation fracture problem (Morand, 2015; Strobl et al., 2016) suffer from the choice of split in the momentum balance equation not resulting in traction-free (opened) cracks. Thus, cone crack angles are predicted significantly larger (Fig. 4.28).

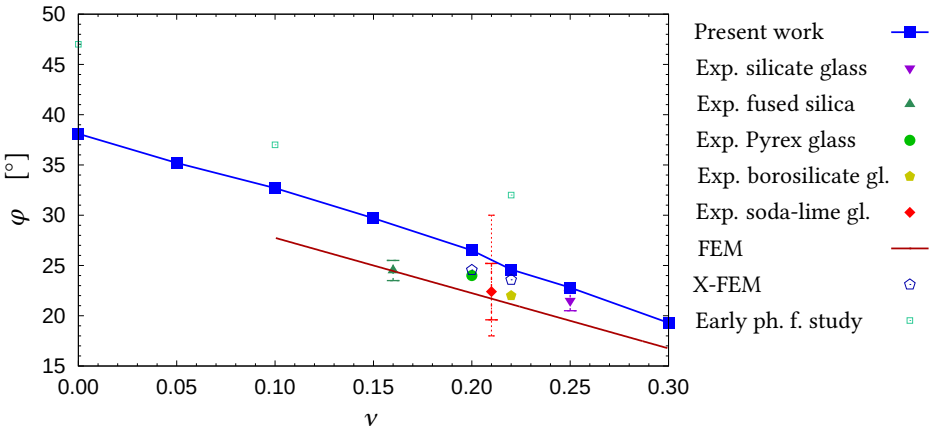


Figure 4.28: Cone crack angle φ depending on Poisson's ratio ν in comparison with experimental findings for silicate glass (Roesler, 1956a), fused silica (Benbow, 1960), Pyrex glass (Conrad et al., 1979), borosilicate glass (Mouginot and Maugis, 1985), soda-lime glass (Kocer and Collins, 1998), numerical studies using FEM (Kocer, 2003b) and X-FEM (Tumbajoy-Spinel et al., 2013) and an early phase field investigation by Morand (2015).

4.3.5 Further issues of phase field modeling of indentation fracture

Results presented in the preceding sections illustrate that the phase field methodology relying on the modifications discussed in Sect. 4.3.1 well captures a number of physical aspects of indentation fracture. However, particularly in the context of the present application, some general issues of phase field modeling deserve a closer inspection as discussed in the following subsections.

Influence of crack driving energy formulation

In Section 4.3.1 the crack driving energy $D_s(\tilde{\boldsymbol{\sigma}})$ is introduced as a constitutive choice based on positive effective stresses. Since such an expression offers many possibilities the influence of different choices for D_s on crack initiation as well as subsequent cone crack propagation is investigated here.

As an alternative to (3.252) with $D_s(\tilde{\boldsymbol{\sigma}}) \sim \psi(\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3)$, a crack driving energy is introduced where the phase field is driven by only the largest positive effective stress $\tilde{\sigma}_1$ in the r-z-plane in conjunction with a positive hydrostatic stress according to

$$D_s := \frac{1+\nu}{2E} \langle \tilde{\sigma}_1 \rangle^2 - \alpha_v \frac{\nu}{2E} \langle \mathrm{tr}(\tilde{\boldsymbol{\sigma}}) \rangle^2 \quad . \tag{4.56}$$

This formulation is in the following referred to as $D_s \sim \psi(\tilde{\sigma}_1, \alpha_v)$ and incorporates the factor $\alpha_v \in [0, 1]$ to describe the influence of positive hydrostatic effective stress on the phase field evolution, which bears some additional arbitrariness. It should be noted that, by applying (4.56), according to (3.190) generally negative values of the driving energy may occur[1] which then have to be suppressed. This, however, is not an issue in the present situation. Simulation results for the different choices of D_s are shown in the following for an indenter of radius $a = 0.5$ mm and a tensile strength of $\sigma_t = 115$ MPa with the remaining parameters according to (4.2).

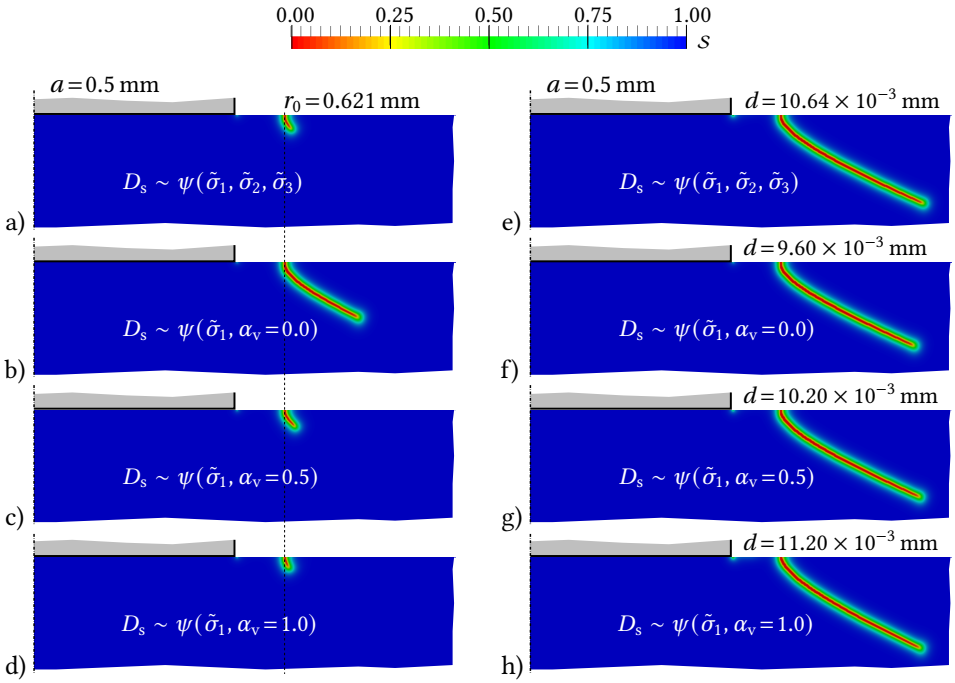


Figure 4.29: Effect of different crack driving energies D_s a)–d) somewhat after crack initiation (at equal load $d = 9.4 \times 10^{-3}$ mm), e)–h) on further crack propagation (equal amount of crack progress).

[1] See also the discussion in Sect. 3.5.4 about the construction of thermodynamically "consistent" crack driving energies.

Figures 4.29a–d show the phase field at a fixed indenter displacement of $d = 9.4 \times 10^{-3}$ mm briefly after crack initiation for four different choices of the crack driving energy D_s. For the same cases of D_s, Figs. 4.29e–h illustrate the situation at later instants with a common amount of crack advance (at different load levels). Additionally, Fig. 4.30 displays the evolution of the fracture surface energy with indenter displacement. Obviously, the location (radius $r_0 \approx 0.62$ mm) of crack initiation as well as the critical load ($d_c \approx 9.14 \times 10^{-3}$ mm) are fairly independent of the choice of D_s. This is not surprising since the portions of the pre-fracture stress field (cf. Sect. 4.1.2) entering D_s are identical for all choices of D_s, i.e. prior to crack initiation $\mathrm{tr}\,\tilde{\boldsymbol{\sigma}} = 0$ at the free surface and only the radial stress σ_{rr} is positive there. After crack initiation, the redistribution of stress leads to different contributions to the different crack driving energy formulations thus giving rise to different amounts of crack growth for the same loading stage (Figs. 4.29a–d). In particular the effect of hydrostatic stress plays a crucial role as visible for the different α_v values in Figs. 4.29a–d. Remarkably, despite the strong influence of the choice of D_s on the crack growth rate after initiation, it does hardly affect the crack path (i.e. cone angle) as seen in Figs. 4.29e–h. The fracture surface energy in Fig. 4.30 accordingly evolves with the indenter displacement for the different choices of D_s.

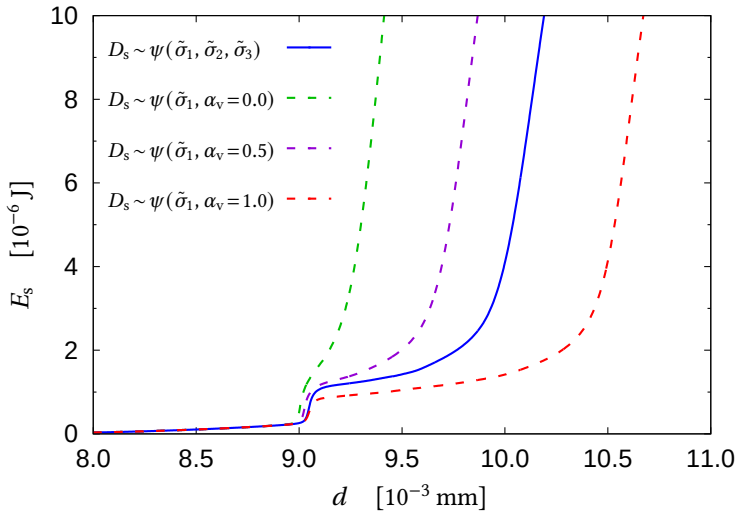


Figure 4.30: Evolution of surface energy E_s for different crack driving energy formulations D_s.

Effect of irreversibility constraint

Another critical issue in the phase field modeling of fracture is the treatment of the irreversibility of cracking. This becomes particularly subtle in the present situation of

indentation fracture where the site of crack initiation is not pre-determined and fracture takes place somewhere in the region of only weak stress concentration outside the indentation contact zone (cf. Sect. 4.1.2). The continuous transition of the material's state from "intact" to "broken" in the phase field description of fracture gives rise to a non-proportional stress redistribution and, at the same time, induces some arbitrariness to the definition of an irreversibly broken state. As mentioned in Sect. 3.2.3, two ways of accounting for the irreversibility of fracture in its phase field approximation are used in the literature: either a "crack-like" constraint (3.36) is imposed on the phase field once it becomes close enough to the limit $s = 0$, or its evolution is "damage-like" constrained to $\dot{s} \leq 0$ (3.40). Both approaches may cause problems, particularly in the present situation of an a priori undetermined crack initiation site, as will be illustrated in the following by simulating indentation fracture with an indenter radius $a = 0.5$ mm and a tensile strength of 115 MPa.

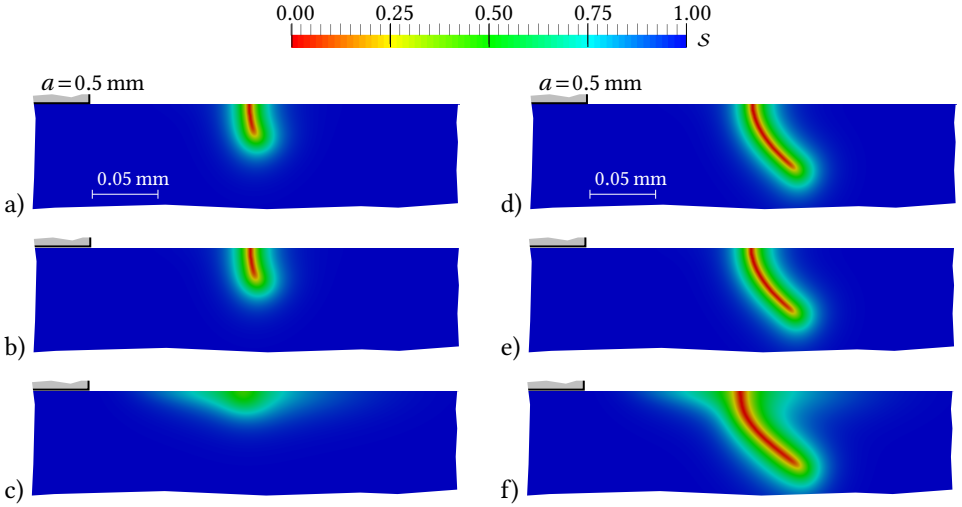


Figure 4.31: Phase field evolution for different irreversibility constraints: at crack initiation ($d = 9.3 \times 10^{-3}$ mm) a) without irreversibility constraint, b) using the "crack-like" irreversibility constraint with $s_0 = 0.05$ and c) using the "damage-like" irreversibility constraint $\dot{s} < 0$; after some crack progress ($d = 1.0 \times 10^{-2}$ mm) d) without irreversibility constraint, e) using the "crack-like" irreversibility constraint with $s_0 = 0.05$ and f) using the "damage-like" irreversibility constraint $\dot{s} < 0$.

Figure 4.31 shows the phase field solution under the effect of different irreversibility constraints at two loading stages, i.e. Figs. 4.31a–c at crack initiation ($d = 9.3 \times 10^{-3}$ mm) and Figs. 4.31d–f after further crack progress ($d = 1.0 \times 10^{-2}$ mm), respectively. More precisely, the progress of indentation fracture is compared for three cases: without any irreversibility constraint (Figs. 4.31a and e), the "crack-like" irreversibility constraint with a threshold of $s_0 = 0.05$ (Figs. 4.31b and e) and the "damage-like" irreversibility constraint $\dot{s} < 0$ (Figs. 4.31c and f). Results for the phase field approach without any irreversibility and the one with the "crack-like" irreversibility constraint are identical in

terms of the initial crack position $r_0 = 0.62$ mm and the critical load $d_c = 9.14 \times 10^{-3}$ mm for crack initiation. Neither the shape of the crack shortly after fracture initiation (Figs. 4.31a and b) nor after further crack progress (Figs. 4.31d and e) show any difference. The "crack-like" irreversibility is not activated during fracture initiation until the phase field reaches the threshold of $s_0 = 0.05$. Thus, no difference compared to the case without irreversibility occurs. During localization, in both cases the phase field can recover to larger values $\dot{s} > 0$ outside the localized zone, whereas it tends to zero at the crack. Due to the monotonic loading the localized zone remains stable without applying an irreversibility condition, so that no difference to the "crack-like" constraint occurs in the range of crack advance shown in Fig. 4.31. However, much further crack growth leads to massive stress redistribution with local unloading, so that without irreversibility constraint the phase field recovers even at a fully evolved crack $s = 0$. Beside the fact that this behavior is unphysical, it usually ends up in convergence problems.

Simulation results using the "damage-like" irreversibility constraint look quite different since the phase field evolution is constrained ($\dot{s} \leq 0$) already from the beginning on. This causes a softer material response in a large area at the free surface which delays localization significantly (Fig. 4.31c), so fracture occurs at a much higher load ($d_c = 9.35 \times 10^{-3}$ mm). In addition, the zone of fully evolved phase field at the crack is much broader (Fig. 4.31f) compared to the solution with the "crack-like" constraint in Fig. 4.31e. This typically leads to an erroneous computation of the crack surface energy as mentioned at the end of Sect. 4.3.1

The above analysis clearly shows that the phase field — in order to describe brittle fracture and prevent significant permanent stiffness reduction in large areas outside the crack — should behave irreversible only when a "sufficiently mature" crack ($s \approx 0$) has formed. For that reason the "crack-like" constraint has been utilized in the simulations in Sects. 4.3.2, 4.3.3, and 4.3.4. However, a drawback of this approach shall be discussed in the following. The focus is on the special case of a low tensile strength of $\sigma_t = 75$ MPa and a small indenter with radius $a = 0.5$ mm. Though rejected from the numerical studies in Sect. 4.3.3 because of the resulting large internal length this case is considered here for illustration purposes. The succession of snap-shots in Fig. 4.32 shows the phase field evolution with increasing indenter displacement d. At a value of $d = 8.15 \times 10^{-3}$ mm the phase field starts to localize at a radial position of approximately $r = 0.68$ mm (Fig. 4.32a). Since the phase field has not yet decreased to the threshold $s_0 = 0.05$ the "crack-like" irreversibility constraint is not activated. With further loading, instead of evolving to s_0 at the aforementioned position and forming a fully developed crack, the localized zone is relocated in outward direction (Fig. 4.32b). This behavior continues (Fig. 4.32c) until at some point of the localized zone the irreversibility constraint takes effect, in the present case this happens at $r = 0.77$ mm (Fig. 4.32d). Quantitatively, this is also shown in Fig. 4.33 by the radial variation of the phase field at the free surface for the different load levels. The main reason for this unphysical "moving crack phenomenon" can be seen in the lack of a sufficiently strong initial stress concentration (in the present situation of indentation fracture) that

would fix the crack initiation site from the beginning on. But it is also attributed to the smeared description of the crack and the fact that the phase field is able to attain any value between 0 and 1. In the particular case considered here, the localized phase field gets stuck at a small but finite value (see, e.g., Fig. 4.32a). With further loading the minimum value of the phase field decreases but also changes its position. This effect is also provoked by the non-vanishing internal length ℓ, which penalizes the formation of sharp gradients and thus hampers the phase field to decrease to zero. The occurrence of this "moving crack phenomenon" is the reason why results for the present case ($\sigma_t = 75\,\text{MPa}$, $a = 0.5\,\text{mm}$) had to be excluded from Figs. 4.24 and 4.25.

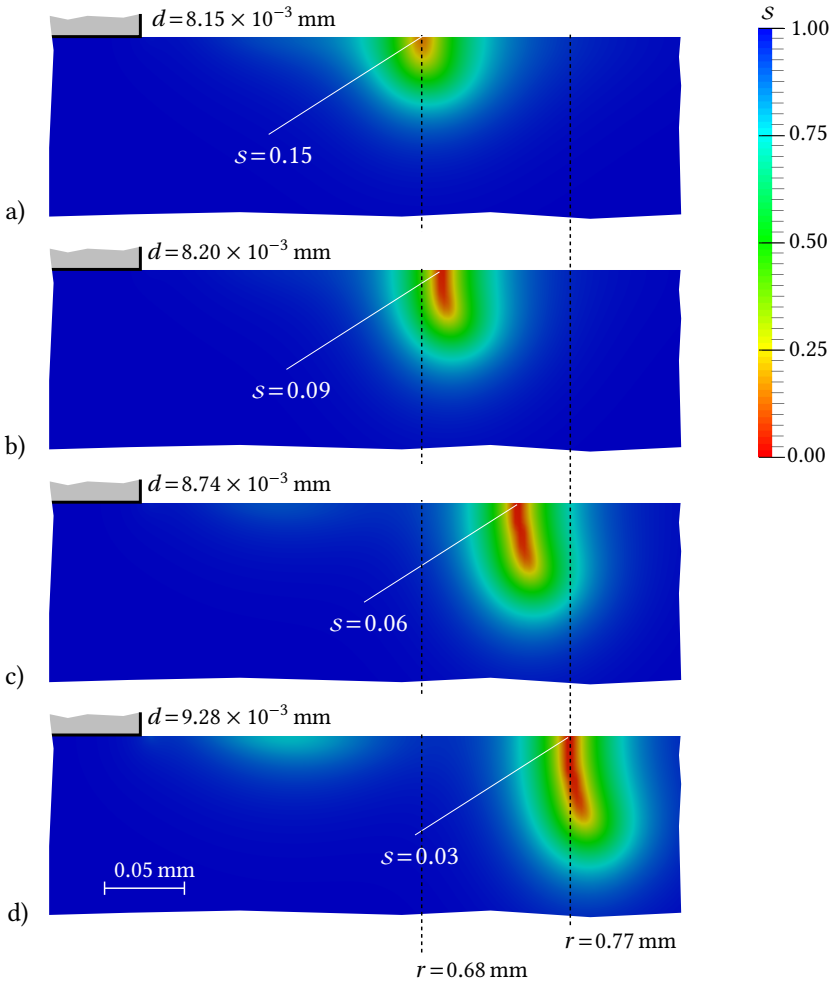


Figure 4.32: Increasing indenter displacement d provokes horizontal movement of localized phase field with $0 < s \ll 1$ until "crack-like" irreversibility constraint is activated for $s \leq s_0 = 0.05$.

Choosing a larger irreversibility threshold s_0 would, of course, fix the position of the localized phase field earlier (e.g. right at the onset of localization) and thus would prevent such a "moving crack phenomenon". However, a larger irreversibility threshold would also affect the phase field solution in the vicinity of the crack so that this is not suitable to numerically reproduce the fracture surface energy $\mathcal{G}_c$ as discussed above for the "damage-like" constraint (3.40). The irreversibility threshold of $s_0 = 0.05$ used in the present study is a compromise to avoid horizontally moving cracks for tensile strength above 115 MPa (Sect. 4.3.3) and to allow phase field recovery outside the crack to adequately reproduce $\mathcal{G}_c$.

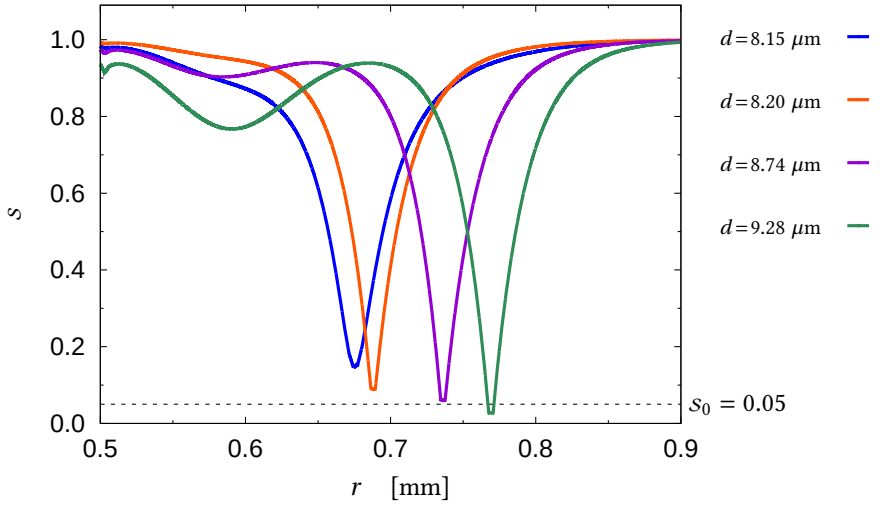


Figure 4.33: Radial variation of phase field s at free surface ($z = 0$) for different load levels d. Localized zone does not drop immediately to s_0 or below and, as a consequence, moves horizontally with increasing load.

Mesh independence

The main advantage of non-local approaches in general is to avoid dependence of the numerical solution of localization problems on the discretization. This concerns both the size and the orientation of the mesh. It should be noted, that utilizing a tension-compression split to formulate the crack driving energy, e.g. (3.252) or (4.56), might lead to a deviation of the crack path caused by the alignment of the mesh as reported by Pham et al. (2017). Preliminary investigations using triangular and unstructured quadrilateral finite elements (i.e. with a mesh different from Fig. 4.19), however, haven't shown a deviation of the initial crack position nor the cone crack angle, as long as the element size is small enough to accurately resolve the smeared crack. In the following, the influence of different finite element sizes in the range between $h_e = 0.0007$ mm and $h_e = 0.003$ mm (see Fig. 4.19) is investigated for the case of a material strength $\sigma_t = 115$ MPa and an indenter

radius $a = 0.5\,\text{mm}$ utilizing the "crack-like" irreversibility constraint as in Sects. 4.3.2–4.3.4.

Figure 4.34 shows contour plots of the phase field solution after crack initiation ($d = 9.4 \times 10^{-3}\,\text{mm}$) and after some further crack progress ($d = 10^{-2}\,\text{mm}$), respectively. Note, that due to the varying ℓ / h_e-ratio, both the internal length ℓ and the fracture toughness $\mathcal{G}_c$ have to be recalibrated according to (3.313) in conjunction with the particular element size h_e (see the discussion in Sects. 3.7.3 and 4.3.2). Corresponding values are given in Tab. 4.2.

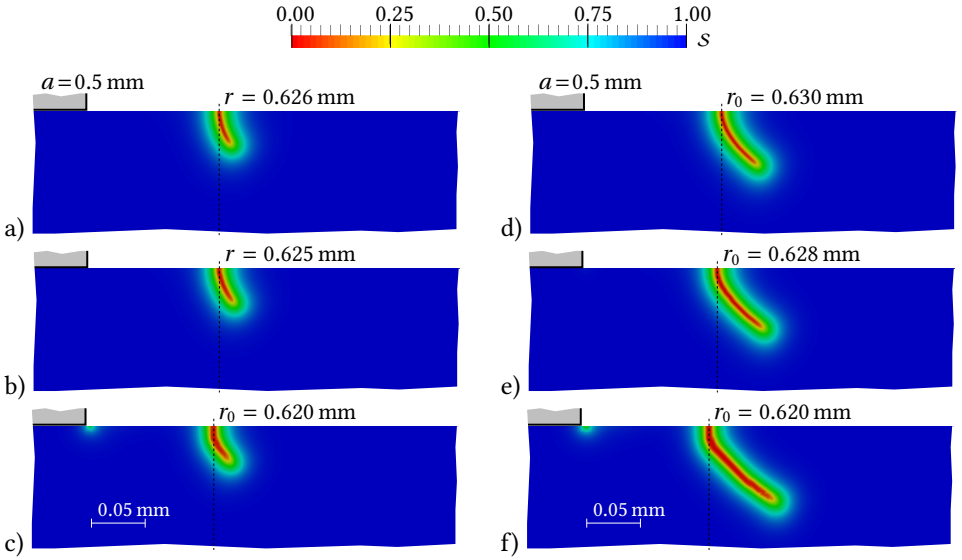


Figure 4.34: Phase field solutions for different discretizations: briefly after crack initiation ($d_c = 9.4 \times 10^{-3}\,\text{mm}$) for mesh with a) $h_e = 0.0007\,\text{mm}$, b) $h_e = 0.0015\,\text{mm}$ and c) $h_e = 0.003\,\text{mm}$; after further crack progress ($d_c = 10^{-2}\,\text{mm}$) for mesh with d) $h_e = 0.0007\,\text{mm}$, e) $h_e = 0.0015\,\text{mm}$ and f) $h_e = 0.003\,\text{mm}$.

Given parameters:	$\sigma_t = 115\,\text{MPa}$, $\mathcal{G}_c = 9 \times 10^{-3}\,\text{N/mm}$				
h_e [10^{-3} mm]	0.7	1.5	2.0	2.5	3.0
Adjusted internal length and fracture toughness according to (3.313)					
ℓ^* [10^{-3} mm]	9.72	9.32	9.07	8.82	8.57
h_e / ℓ^*	0.07	0.16	0.22	0.28	0.35
$\mathcal{G}_c^*$ [10^{-3} N/mm]	8.69	8.33	8.11	7.88	7.66

Table 4.2: Numerical recalibration of parameters $\mathcal{G}_c$ and ℓ to simulate Hertzian indentation fracture for varying element size h_e.

Numerical results for both loading stages show only minor influences of the discretization; nevertheless, these shall be briefly discussed in the following. In case of the coarse mesh (Figs. 4.34c and d) the zone of phase field localization to $s \approx 0$ (i.e. the approximated crack) is slightly broader than for the finer meshes. This, however, is quantitatively compensated by the parameter recalibration according to Tab. 4.2. For finer meshes, a more smooth transition from fracture initiation to the cone crack is observed. A small effect of the "moving crack phenomenon" (discussed in Sect. 4.3.5) is visible between the two loading stages (Fig. 4.34) in case of the fine meshes; this effect, however, is of the order of the element size. After some amount of crack growth the crack in the coarse mesh (Fig. 4.34f) is somewhat longer than those in the finer meshes at equal load. However, this has no influence on the cone angle.

In Fig. 4.35 the ratios of crack initiation radius r_0 and the critical load d_c to the results obtained for the mesh with element size $h_e = 0.0025\,\text{mm}$ (used in the previous sections) is depicted. The deviation of both parameters is in the range of $1 \ldots 3\%$.

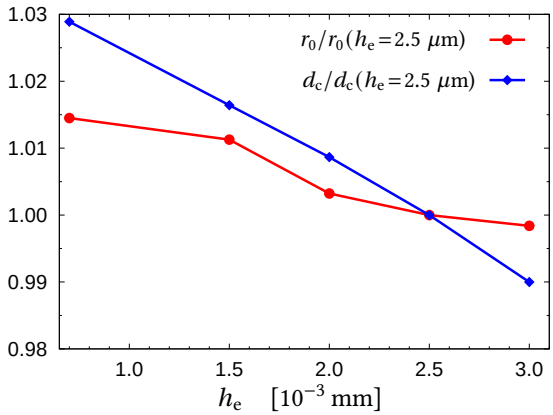


Figure 4.35: Deviation of the radial position r_0 and critical load d_c depending on the finite element size h_e (relative to mesh with element size $h_e = 0.0025\,\text{mm}$).

4.3.6 Discussion of phase field results on indentation fracture

In the present section, the entire process of indentation fracture — starting from instantaneous crack nucleation at an initially defect-free surface up to the ultimate stable growth of a cone-shaped crack — has been studied numerically. A particular challenge lay in the fact that indentation fracture initiates from a rather weak stress concentration and both the location of crack initiation and the subsequent crack path are a priori unknown. This non-standard problem of fracture mechanics has been treated in the framework of a phase field regularization. Yet, despite the wide-spread application of this computational technique, extensive modifications compared to common phase field fracture models were

necessary for a reliable analysis of indentation fracture. In particular, the approach used in the present section was guided by the following key requirements:

- In order to feature brittle fracture, the phase field affected material response should be essentially linear elastic up to failure.
- Motivated by the success of techniques such as finite fracture mechanics or cohesive zone models in predicting crack nucleation in absence of a perfect pre-crack, the material's tensile strength *and* fracture toughness (specific crack surface energy) should be incorporated as two independent, physically well-defined quantities.
- Crack boundary conditions (e.g. traction-free crack faces) should be correctly represented in the phase field setting.
- Crack nucleation and propagation should be driven in a physically reasonable manner.

These requirements are typically not all fulfilled in approaches to phase field fracture modeling found in the literature and, moreover, the commonly used frameworks do not allow to incorporate all aspects at once as thoroughly analyzed in Sect. 4.3.1. Correspondingly, it has been shown that those approaches fail in application to indentation fracture. The aim of this study, therefore, was not only the analysis of indentation fracture but also a careful assessment of phase field fracture modeling with indentation fracture serving as a paradigm.

A key aspect in the present work is the incorporation of a material strength parameter, necessary to capture crack nucleation, which together with the fracture toughness introduces a material length scale. This inevitably leads to a non-vanishing internal length ℓ of the phase field model, which in 1D can be explicitly computed from the strength and toughness in conjunction with some choice for the degradation function and the local crack surface function. With the phase field model's internal length being thus fixed, it does no longer play the role of a mere regularization parameter and the limit process $\ell \to 0$ (in the sense of so-called Γ-convergence) is ruled out.

The internal length in the present application, however, has to be sufficiently small for an energetically reasonable approximation of the surface energy of the small crack nucleating at the onset of indentation fracture. This was to some extent accomplished by carefully selecting the degradation function, which besides controlling the smooth transition from intact to broken material also affects the energy driving the phase field evolution.

This two-fold role of the degradation function stems from the joint derivation of the mechanical (momentum) balance equation and the phase field evolution equation from a variational principle. It has been shown that the particular forms of stiffness degradation and phase field (crack) driving energy so obtained in common phase field fracture models do not comply with the above-listed requirements. Rather, the present approach was focused on a phenomenological crack description instead of persisting in consistency of the total energy. Thus, to enable both traction-free crack faces and a physically reasonable crack evolution one is forced to give up the commonly preferred variational structure by

independently prescribing the stiffness degradation and the phase field driving energy portion. In particular, physically motivated crack driving energies based on positive principal effective stresses (in the sense of damage mechanics) have been introduced. The basic structure of the phase field evolution equation (i.e. occurrence of degradation function), however, has been kept for a consistent transfer of elastic energy to crack surface energy.

Another issue of the phase field description of cracks by a smooth transition from an intact to a broken material state is the realization of the irreversibility of fracture. This has been shown to be particularly critical in the present application where the crack initiation site is not fixed by some pre-existing crack or notch. The continuous evolution of the phase field localization zone accompanied by a continuous change of the stress field may either lead to significant stiffness degradation in a large region (in case of a "damage-like" irreversibility constraint) or to a relocation of the localization zone (in case of a "crack-like" irreversibility condition) prior to the full establishment of the crack. Both phenomena are unphysical and require special attention. The present approach relied on a "crack-like" irreversibility constraint to accurately approximate the surface energy. Thus, at the beginning of its evolution the phase field has been used in its role as a crack indicator, not to model damage, i.e. subsequent stiffness degradation. However, to improve the robustness of the numerical approach a combination of both types of irreversibility constraint has been developed which allows for the reliable simulation of indentation fracture initiation at least in some range of geometric and material parameters.

Moreover, to overcome in the discrete finite element setting the incongruity between the "geometric" description of cracks with nodal phase field values approaching zero and the "mechanical" manifestation of cracks by traction-free crack faces, a recalibration of the fracture toughness *and* the internal length parameter was necessary. The proposed strategy guarantees consistency of the numerically computed fracture surface energy with the physically prescribed value.

With the phase field model tailored as discussed above it was possible to simulate the entire process of indentation fracture consistently from nucleation to ultimate cone crack propagation. As carefully analyzed, the range of the reliable applicability of the present model so far is restricted to indenter radii and material tensile strength values above certain limits. In this range, however, a reasonable qualitative — and to some extent quantitative — agreement between numerical results and experimental data as well as other theoretical findings from the literature was obtained. That is, the present simulations well capture various aspects of fracture initiation such as the influence of the indenter radius on the location and extension of the initially forming ring crack as well as the critical load. A good agreement is observed between phase field and finite fracture mechanics (FFM) results which are both based on the same boundary value problem and material data. Absolute differences between numerical results and experimental data, of course, have to be seen in the light of uncertain material parameters and idealizations in the theoretical model (e.g. absence of micro-defects) which become more pronounced for smaller indenters. Also well predicted was the almost linear variation of the ultimate cone crack angle with

Poisson's ratio of the fracturing material. The rate of growth of the cone crack was shown to depend on the particular constitutive choice for the phase field driving energy which might be further calibrated by (presently not available) experimental data on the crack extension with increasing load. The critical load, the location of crack initiation and the cone angle, however, are not affected by this choice.

The present approach of modeling indentation fracture is a purely macroscopic one and the considered values of material parameters are in fact effective (*homogenized*) properties in view of distributed micro-flaws prevailing on a finer length scale. On the length scale of the actual boundary value problem (Sect. 4.1) this appears justified since the flaw size in as-received glass is typically below 1 μm, e.g. Lawn (1993). These flaws, however, manifest themselves in a rather strong scatter of the *macroscopic strength* of glass with values significantly smaller than the *intrinsic strength* of the unflawed material. Different from the present modeling approach, it is also conceivable to explicitly consider a pre-existing population of flaws and to describe the surrounding material by its *intrinsic* properties. In case of crack-like initial flaws this would avoid the need to model crack nucleation and allow the application of linear elastic fracture mechanics or a phase field model based on $\mathcal{G}_c$ alone; the strength σ_t then would not occur and the intrinsic length ℓ would retain its role as regularization parameter. Such an approach, however, besides bearing the difficulty to determine the intrinsic material parameters, would cause an enormous computational expense due to the necessarily extremely fine mesh and the need to do repeated simulations with different (yet statistically equivalent) realizations of the flaw distribution to get results independent (in the sense of ensemble averaging) from its particular choice.

5 Conclusions and further perspectives

In the present work, different constitutive ingredients of the phase field approach to fracture have been analyzed. Based on the outcome of this analysis, the method has been used to study the problem of Hertzian indentation fracture revealing critical issues and shortcomings of the method in general. For comparison, the application of hybrid fracture criteria (in the framework of finite fracture mechanics) on indentation fracture *initiation* provided new insights into the process of "spontaneous" crack formation.

At the beginning, the work outlines the successive development of the phase field approach to fracture, including the minimization of energy (showing remarkable parallels to Griffith's idea) by applying a variational principle, the regularization of cracks emanating from image segmentation and its application on elastic field equations. In this approach, fracture is indicated by a scalar field equation — the phase field — varying between zero and one. The application of Hamilton's principle of stationary action finally yields a set of coupled partial differential equations which can be solved with standard techniques in time and space to simulate fracture. Although the story has often been told, it is of practical relevance to understand the motives behind each step. In its original version (Bourdin, 1998) the phase field model already contained specific constitutive choices, which have been reworked in recent years.

Thus, the present work is concerned with a comprehensive analysis of constitutive ingredients of the phase field approach to fracture. For instance, the degradation of elastic energy with phase field evolution offers some flexibility regarding the constitutive behavior of the material prior to fracture. In this context, the local crack surface function, i.e. the distribution of the surface energy across the regularized crack, also plays a crucial role. With both constitutive choices at hand, the transition from one limiting state to the other has been analyzed, e.g. to identify quantitative effects at crack formation.

Another key aspect of the work is the different behavior of the phase field model subjected to tension or compression. This concerns the mechanical response caused by crack opening or closing in the vicinity of already existing cracks, so-called crack closure effects, and the tension-compression asymmetric behavior at crack progress. It was found, that phase field models including the widely-used approaches based on the spectral or volumetric-deviatoric decomposition are in general not able to describe crack boundary conditions appropriately. Either they fail to reproduce traction-free crack surfaces when subjected to tangential movement/shear or fail to reproduce compressive forces in cases

where no hydrostatic response is possible, e.g. close to traction-free outer boundaries. It has been shown, that no isotropic strain energy decomposition exists which fulfills all requirements at the same time. Since macroscopic cracks induce anisotropy in nature, they cannot be described by the underlying isotropic degradation, i.e. independent of the crack orientation, which is present in most phase field models. By taking the crack orientation into account, it has been shown to reproduce accurate results, leading to an anisotropic degradation of the elasticity tensor. Although the problem has been identified and the approach to modify the degradation of the stiffness tensor has been derived in the present work, its theoretical incorporation into the general framework, mainly in conjunction with the phase field evolution, is still missing. In this regard, at least three questions have occurred:

- How can the crack orientation be identified?
  Besides taking the (only piecewise continuous) phase field gradient ∇s into account, a phase field approximation with higher order continuity (Borden, 2012), a least-square projection of the phase field gradient (see Sect. 3.6.1) or its separate approximation as used for, e.g., hydraulic fracturing (Yoshioka et al., 2020), can be used to identify the crack orientation in the vicinity of an already existing crack. But, apart from the technical treatment, the question arises, what does the inhomogeneous crack orientation ahead of the crack front/tip mean?
- Is the approach influenced by the mesh?
  Due to the only piecewise continuous strain field approximation in standard finite elements, both the identification of the crack orientation and decomposition of strains/stresses suffer from mesh orientation effects (Sect. 3.7.2). This applies for the isotropic decompositions based on eigenvalues or the volumetric-deviatoric split, see Strobl and Seelig (2018), but also for the crack orientation dependent degradation. Thus, mesh orientation dependency is reintroduced.
- How does the crack evolution have to be formulated, so that fracture under mixed-mode loading conditions can occur?

From a conceptual point of view, especially the last question is still open. A concurrent incorporation of the crack orientation dependent degradation and a physically meaningful (anisotropic) phase field evolution was not possible so far in the present structure of the phase field approach. The introduction of anisotropic degradation in conjunction with reasonable crack propagation under mixed-mode loadings (brittle materials are still dominated by mode I crack growth) adds a large amount of complexity to the method, far beyond common practices. Steinke and Kaliske (2019), for example, proposed a more technically motivated approach, which does not answer all questions from above. It is also based on the crack orientation dependent degradation, but still includes an isotropic crack driving force. Therefore, the pragmatic approach of neglecting the effect of crack-closure and decomposing only the crack driving energy is strongly encouraged, however, at the price of introducing some amount of artificial energy dissipation (see Sect. 3.5.4). This approach is valid as long as cracks, once they have formed, remain open. In general, challenges regarding the tension-compression asymmetric behavior conceptually seem to

emanate from the transfer of the phase field method from image segmentation, where it governs the scalar gray level field, to mechanics, where it governs tensorial fields.

The implementation of the phase field approach into a finite element code is quite straightforward regarding most of its constituents. In particular, the additional field equations including irreversibility constraints can be incorporated at the element level. Thus, all approaches of interest have been implemented into the finite element code FEAP. The application of special solution techniques, e.g. bound-constrained solvers, would require further modifications, which is out of the scope of this work.

Another part of the work is focused on the study of the non-standard problem of Hertzian indentation fracture, including both stages of indentation fracture, the initial crack formation and subsequent cone crack propagation. First, it was illustrated that the concept of finite fracture mechanics in the format suggested by Leguillon (2002) is a reasonable tool for analyzing the indentation fracture *initiation*. Using a hybrid fracture criterion comprising a strength (stress) and a toughness (energy) condition it is hence possible to reproduce experimental findings without any assumptions beyond the material properties "stiffness" (E, ν), "strength" (σ_t) and "fracture toughness" ($\mathcal{G}_c$). In particular, the experimentally observed dependence of the critical load as well as the crack location (radius) on the indenter radius could be qualitatively and — to some extent — also quantitatively reproduced. This comprises the size effect commonly referred to as *Auerbach's law*. Quantitative deviations from experimental findings have to be seen in the light of various assumptions made in the theoretical analysis. For instance, indentation was idealized as frictionless with a rigid and perfectly sharp-edged punch; effects of rounded indenter edges and friction on the indentation stress field have been examined, e.g., in Ciavarella et al. (1998). Interestingly, it has turned out that the application of the hybrid (two-part) criterion of finite fracture mechanics to indentation fracture gives rise to two regimes: one in which both the stress and the energy criterion are active, and a second regime where only the energy criterion is active while the stress criterion is oversatisfied. The existence of two regimes with respect to the dependence of the critical load on the indenter radius has experimentally been shown already by Tillett (1956). However, its relation to the two regimes predicted by the present theoretical analysis is not yet fully understood. The comparison of the semi-analytical determination of the average energy release rate with its numerical computation from detailed finite element simulations has shown that the simplification of using the pre-fracture indentation stress field for analyzing the initial ring crack formation is not very severe. However, during subsequent cone crack propagation the alteration of the initial indentation stress field by the presence of the crack is known to be much stronger and needs to be accounted for (Lawn et al., 1974).

In a second step, the entire process of Hertzian indentation fracture from nucleation to ultimate cone crack propagation was analyzed by employing a phase field method. On the one hand, it was shown that the phase field method is — if constitutive ingredients have been selected carefully — capable of reproducing both stages of the indentation fracture problem, i.e. crack initiation at some a priori unknown location as well as the subsequent evolution of the cone-shaped crack. Thereby, the influence of geometrical

and material parameters was studied. On the other hand, this non-standard fracture problem served as a paradigm for a thorough assessment of the phase field methodology. Thus, the presented computational study allows to draw several conclusions listed below. These concern mainly the utilized phase field methodology and only to a lower extent the fracture process itself, since the latter is studied in the framework of a strongly idealized boundary value problem. In fact, tailoring a phase field model capable to capture crack nucleation as well as propagation in a quantitative manner has provided a deep insight into this computational technique and has revealed a number of additional critical issues and shortcomings:

- Common phase field models based on isotropic splits (i.e. independent of crack orientation, like the spectral or volumetric-deviatoric decomposition) of the degraded and crack driving energy portions in the framework of a variational treatment — as pursued in the majority of the literature — have proven to be not suitable for the simulation of indentation fracture since they do generally not provide traction-free faces of open cracks and/or predict an unphysical phase field evolution. One may ask how strongly this shortcoming takes effect in other applications.
- The internal length parameter of the phase field model is physically determined (and hence prohibited from being sent to zero) once the material's strength and toughness are prescribed independently in order to capture crack nucleation *and* propagation.
- The interplay of the degradation function $g(s)$ and the local crack surface function $w(s)$, which are constitutively chosen ingredients of any phase field fracture model and affect the internal length as well as the exchange of mechanical and crack surface energy, is not yet fully explored and requires further investigation. This concerns the impact of these functions on both the state of crack nucleation as well as propagation of an existing crack. Besides some quantitative effects, the application of other local crack surface functions, e.g. the linear crack surface function introduced in Sect. 3.3.1 and further discussed in Sect. 3.4.5, might also imply qualitative differences with respect to crack initiation.
- The smeared representation of cracks in the phase field setting with a continuous transition from intact to broken material raises the issue of how to establish the irreversibility of fracture. This is particularly critical in situations (as the present) where crack initiation takes place in the absence of a stress singularity and, hence, without a fixed location of phase field localization.
- Another open issue in phase field fracture modeling is how to take into consideration the energy consumed by the diffuse crack tip which is not accounted for in the crack surface density and becomes crucial in the case of short cracks and crack nucleation.
- Owing to the above-discussed limitations and particularly in view of the rather large internal length of the phase field model it was not possible to extend the present study to very small indenter radii. As a consequence, the so-called Auerbach

range (see Sects. 4.1 and 4.2.4) could not be explored in the present study. This could perhaps be accomplished by further modifications of the phase field evolution equation which, however, still has to provide consistency between elastic and crack surface energy.

- On the other hand, a physically more realistic extension of theoretical/numerical studies of indentation fracture (by phase field or any other methods) to very small indenter radii presumably should be based on a less idealized boundary value problem than the one considered here (Sect. 4.1).

It has been shown in the previous chapter, that the requirements stated at the beginning of Sect. 4.3.6 are essential for the physically sound phase field modeling of indentation fracture. Yet, it seems plausible that these requirements need to be fulfilled also in other applications of phase field fracture modeling. Moreover, only by being aware of the various open questions regarding constitutive choices and appropriate numerical treatment, the phase field approach to fracture should be applied to complex crack phenomena.

All in all, the framework of phase field modeling offers great potential to simulate fracture, including a physically sound prediction of crack formation. The suitable choice of constitutive models and corresponding parameters, however, does require much attention to reproduce the fracture problem consistently. From the theory standpoint, there are still many pressing questions, such as if Γ-convergence is required in general and, if so, it also holds for all the modifications presented in this work. One main issue still is the formulation of a tension-compression asymmetric response which is both able to describe the effects of (macroscopic) crack closure and different behavior of material subjected to tension or compression. In the context of Hertzian indentation fracture, the impact of more realistic conditions, e.g., including friction between the indenter and the solid body, would be of interest. The extension of the study to spherical indenters introduces a contact area varying with the indenter load. This, however, would require a much more accurate treatment of irreversibility in conjunction with the gradual formation of the initial ring crack. For this purpose, the application of a different local crack surface function might succeed.

A Some mathematical and mechanical preliminaries

A.1 Tensor notation, algebra and analysis

Throughout this work, a common bold face notation is used, with lowercase letters indicating a tensor of first order (vector) and capital letters indicating a second order tensor. Tensors of higher order or deviations from the common notation are introduced/indicated accordingly. Normal font letters denote scalars (tensors of zero order) and, equipped with subscripts, components of tensors. In some cases, it is convenient to use index notation, where Latin subscripts take the range $\{1, 2, 3\}$. Repeated indices imply summation (*Einstein's summation convention*) while a comma indicates partial differentiation with respect to the corresponding coordinate.

To introduce vectors and tensors of higher order, a right-handed and orthonormal coordinate system is introduced, which consists of three base vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ called a (Cartesian) basis. Their properties are $|\boldsymbol{e}_1| = |\boldsymbol{e}_2| = |\boldsymbol{e}_3| = 1$ and $\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3$, where the latter expression is equivalent to $\det([\boldsymbol{e}_1\, \boldsymbol{e}_2\, \boldsymbol{e}_3]) = 1$. Any vector $\boldsymbol{a}$ in a three dimensional Euclidean space $\mathbb{E}^3$ is represented by a unique linear combination as

$$\boldsymbol{a} = a_1\, \boldsymbol{e}_1 + a_2\, \boldsymbol{e}_2 + a_3\, \boldsymbol{e}_3 = a_i\, \boldsymbol{e}_i \quad , \tag{A.1}$$

with $a_i = \boldsymbol{a} \cdot \boldsymbol{e}_i$ being the components of vector $\boldsymbol{a}$. These components are equivalent with the projections of the vector onto the basis vectors. Accordingly, tensors of second order

$$\boldsymbol{A} = A_{11}\, \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + A_{12}\, \boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + \cdots + A_{33}\, \boldsymbol{e}_3 \otimes \boldsymbol{e}_3 = A_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \tag{A.2}$$

or any higher order are introduced in an analogous manner. Tensors of first order $\boldsymbol{a}$ and second order $\boldsymbol{A}$ can be represented (e.g. according to a basis $\boldsymbol{e}_i$) as (column) vector and matrix

$$\left[\boldsymbol{a}\right]_{\{\boldsymbol{e}_i\}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \text{and} \qquad \left[\boldsymbol{A}\right]_{\{\boldsymbol{e}_i\}} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \quad , \tag{A.3}$$

respectively. The transpose of the second order tensor is

$$\boldsymbol{A}^{\mathrm{T}} \coloneqq (\boldsymbol{A}^{\mathrm{T}})_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = A_{ji}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad . \tag{A.4}$$

It is used to additively decompose a second order tensor $\boldsymbol{A} = \boldsymbol{A}_{\mathrm{sym}} + \boldsymbol{A}_{\mathrm{skw}}$ into a symmetric tensor and an antisymmetric or skew tensor. Both the symmetric tensor and the skew tensor are unique for any second order tensor and are

$$\boldsymbol{A}_{\mathrm{sym}} = (\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}})/2 \qquad \text{and} \qquad \boldsymbol{A}_{\mathrm{skw}} = (\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}})/2 \quad , \tag{A.5}$$

respectively. The identity or unit tensor of second order is denoted (in Cartesian coordinates) by

$$\boldsymbol{I} := \delta_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad , \tag{A.6}$$

with $\delta_{ij} := 1$ for $i = j$ and $\delta_{ij} := 0$ for $i \neq j$ being the *Kronecker delta*.

To multiply two tensors of order $m > 0$ and $n > 0$ with each other, there are different types of multiplication available. The dot operator, e.g. between two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ according to

$$\boldsymbol{a} \cdot \boldsymbol{b} = \underbrace{a_i\, b_j}_{= c}\ \underbrace{\boldsymbol{e}_i \cdot \boldsymbol{e}_j}_{=: \delta_{ij}} = c \quad , \tag{A.7}$$

denotes the inner product or a (single) contraction. Contraction means, that one sums over the involved indices, the right index of the expression on the left and the left index of the expression on the right, which then disappear. Thus they are called summation or dummy indices, in contrast to the indices which remain (free indices). The operation can be applied to any combination of tensors with $m \geq 1$ and $n \geq 1$, e.g. between two second order tensors according to

$$\boldsymbol{A} \cdot \boldsymbol{B} = \underbrace{A_{ik}\, B_{lj}}_{= C_{ij}}\, (\boldsymbol{e}_i \otimes \underbrace{\boldsymbol{e}_k) \cdot (\boldsymbol{e}_l}_{=: \delta_{kl}} \otimes \boldsymbol{e}_j) = C_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \boldsymbol{C} \quad . \tag{A.8}$$

The result is a tensor of order $n + m - 2$. Powers of a second order tensor are defined by using the single contraction according to

$$\boldsymbol{A}^1 = \boldsymbol{A} \quad , \qquad \boldsymbol{A}^2 = \boldsymbol{A} \cdot \boldsymbol{A} \quad , \qquad \boldsymbol{A}^3 = \boldsymbol{A} \cdot \boldsymbol{A} \cdot \boldsymbol{A} \quad , \qquad \ldots \tag{A.9}$$

(e.g. Başar, 2000). If a second order tensor is nonsingular ($\det(\boldsymbol{A}) \neq 0$), it is invertible and a unique inverse $\boldsymbol{A}^{-1}$ exists with

$$\boldsymbol{A} \cdot \boldsymbol{A}^{-1} = \boldsymbol{A}^{-1} \cdot \boldsymbol{A} = \boldsymbol{I} \quad , \tag{A.10}$$

which also constitutes the zero power second order tensor

$$\boldsymbol{A}^0 = \boldsymbol{A}^c \cdot \boldsymbol{A}^{-c} = \boldsymbol{I} \tag{A.11}$$

with $c \in \mathbb{R}$. The determinant is defined according to the determinant of a matrix, i.e. $\det \boldsymbol{A} = \det[\boldsymbol{A}]$. In absence of whole numbers, powers $c \in \mathbb{R}$ can be evaluated by

using the spectral decomposition of a symmetric second order tensor $\boldsymbol{A}$ (with real eigenvalues λ_α and principal directions $\boldsymbol{n}_\alpha$), i.e.

$$\boldsymbol{A}^c = \sum_{\alpha=1}^{3} (\lambda_\alpha)^c \, \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha \quad , \tag{A.12}$$

see App. A.5. A second order tensor or matrix $\boldsymbol{R}$ is orthogonal if $\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{R}^{-1}$, which is equal to $\boldsymbol{R}^{\mathrm{T}} \cdot \boldsymbol{R} = \boldsymbol{I}$. Further, it is proper orthogonal if $\det(\boldsymbol{R}) = 1$, which preserves the sense of direction, and is improper if $\det(\boldsymbol{R}) = -1$.

Double contraction is denoted by two (vertically aligned) dots[1] and defined as the pairwise summation, e.g. double contraction two second order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ results in

$$\boldsymbol{A} : \boldsymbol{B} = \underbrace{A_{ij}\, B_{kl}}_{= c} \, \underbrace{(\boldsymbol{e}_i \otimes \boldsymbol{e}_j) : (\boldsymbol{e}_k \otimes \boldsymbol{e}_l)}_{= \delta_{ik}\, \delta_{jl}} = c \quad . \tag{A.13}$$

The operation can be applied to any combination of tensors with $m \geq 2$ and $n \geq 2$ and results in a tensor of order $n + m - 4$. Double contraction can be also used to extract the sum of the diagonal elements of a second order tensor according to

$$\operatorname{tr} \boldsymbol{A} := \boldsymbol{I} : \boldsymbol{A} = \boldsymbol{A} : \boldsymbol{I} = A_{ii} = A_{11} + A_{22} + A_{33} \quad , \tag{A.14}$$

which is known as the trace of the tensor. The norm of a second order tensor is defined by the square root of the double contraction by itself and is non-negative, i.e.

$$||\boldsymbol{A}|| := \sqrt{\boldsymbol{A} : \boldsymbol{A}} \geq 0 \quad . \tag{A.15}$$

If the result of a contraction is a scalar (tensor of order zero), as in the examples above, the product is denoted as scalar product.

The cross (or vector) product is defined (by considering two vectors) according to

$$\boldsymbol{a} \times \boldsymbol{b} = a_i \, b_j \underbrace{\boldsymbol{e}_i \times \boldsymbol{e}_j}_{=: \varepsilon_{ijk}\, \boldsymbol{e}_k} = \begin{bmatrix} a_2\, b_3 - a_3\, b_2 \\ a_3\, b_1 - a_1\, b_3 \\ a_1\, b_2 - a_2\, b_1 \end{bmatrix} = \boldsymbol{c} \quad , \tag{A.16}$$

with the permutation symbol (or *Levi-Civita symbol*) $\varepsilon_{ijk} := 1$ for even permutations $(1, 2, 3)$, $(2, 3, 1)$ or $(3, 1, 2)$, $\varepsilon_{ijk} := -1$ for odd permutations $(3, 2, 1)$, $(1, 3, 2)$ or $(2, 1, 3)$ and $\varepsilon_{ijk} := 0$ for all other cases (containing repeated indices). According to (A.16), it can be used to obtain the vector $\boldsymbol{c}$, which is perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$. In addition, it is introduced as right-handed cross product, i.e. the three vectors form a right-handed

[1] Note, that no unique symbolic notation exists. For instance, the double dot product $\boldsymbol{A} : \boldsymbol{B}$ is defined according to Reddy (2017) as $A_{ij}\, B_{ji}$, while in Holzapfel (2000) its definition is $A_{ij}\, B_{ij}$.

system. The cross product can be applied to any combination of tensors with $m \geq 1$ and $n \geq 1$ and results in a tensor of order $n + m - 1$. Although the permutation symbol ε_{ijk} is defined only in three dimensions, a cross product-like procedure to calculate perpendicular vectors exists for dimensions n other than three, i.e. for $n > 1$. By using the property $\boldsymbol{A} \cdot \operatorname{adj}(\boldsymbol{A}) - \det(\boldsymbol{A})\, \boldsymbol{I} = 0$, one can build up the square matrix

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_{n-1} & \tilde{\boldsymbol{a}} \end{bmatrix}^{\mathrm{T}} \quad , \tag{A.17}$$

with the n-dimensional given vectors $\boldsymbol{a}_i$ and a vector with arbitrary components $\tilde{\boldsymbol{a}}$. Finally, from $\tilde{\boldsymbol{G}} = \operatorname{adj}(\boldsymbol{G})$, the vector $\boldsymbol{c} = \tilde{\boldsymbol{G}}_{in}$ is obtained (n here indicates the last column), which is perpendicular to the other vectors.

The tensor product, outer product, dyadic product or dyad is defined (by considering two vectors) according to

$$\boldsymbol{a} \otimes \boldsymbol{b} = \underbrace{a_i\, b_j}_{= C_{ij}}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \boldsymbol{C} \quad . \tag{A.18}$$

Any second order tensor can be expressed as a linear combination of dyads with scalar coefficients (also known as dyadic), e.g. $\boldsymbol{A} = \alpha\, \boldsymbol{a} \otimes \boldsymbol{b} + \beta\, \boldsymbol{c} \otimes \boldsymbol{d} + \dots$ with the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$, ... The tensor product can be applied to any combination of tensors with $m \geq 1$ and $n \geq 1$ and results in a tensor of order $n + m$.

Possible identities of fourth order are

$$\hat{\mathbb{I}} := \delta_{ik}\, \delta_{jl} \quad , \qquad \tilde{\mathbb{I}} := \delta_{il}\, \delta_{jk} \qquad \text{and} \qquad \mathbb{I} := \delta_{ij}\, \delta_{kl} = \boldsymbol{I} \otimes \boldsymbol{I} \quad , \tag{A.19}$$

so that any second order tensor $\boldsymbol{A}$ is projected according to

$$\hat{\mathbb{I}} : \boldsymbol{A} = \boldsymbol{A} : \hat{\mathbb{I}} = \boldsymbol{A} \qquad \text{and} \qquad \tilde{\mathbb{I}} : \boldsymbol{A} = \boldsymbol{A} : \tilde{\mathbb{I}} = \boldsymbol{A}^{\mathrm{T}} \quad . \tag{A.20}$$

In addition, the symmetric identity

$$\mathbb{I}_{\mathrm{sym}} := \frac{1}{2}(\delta_{ik}\, \delta_{jl} + \delta_{il}\, \delta_{jk}) = \frac{1}{2}(\hat{\mathbb{I}} + \tilde{\mathbb{I}}) \tag{A.21}$$

and antisymmetric identity

$$\mathbb{I}_{\mathrm{asym}} := \frac{1}{2}(\delta_{ik}\, \delta_{jl} - \delta_{il}\, \delta_{jk}) = \frac{1}{2}(\hat{\mathbb{I}} - \tilde{\mathbb{I}}) \tag{A.22}$$

are introduced. Powers of a symmetric fourth order tensor $\mathbb{A}$ are defined by using the double contraction (e.g. Itskov and Başar, 2000)

$$\mathbb{A}^1 = \mathbb{A} \quad , \qquad \mathbb{A}^2 = \mathbb{A} : \mathbb{A} \quad , \qquad \mathbb{A}^3 = \mathbb{A} : \mathbb{A} : \mathbb{A} \quad , \qquad \dots \tag{A.23}$$

with the zero power symmetric fourth order tensor

$$\mathbb{A}^0 = \mathbb{I}_{\mathrm{sym}} \tag{A.24}$$

(Betten, 1982). If a symmetric fourth order tensor is nonsingular ($\det(\mathbb{A}) \neq 0$, the determinant of a matrix representation of the tensor is evaluated, see App. A.7), it is invertible and a unique inverse $\mathbb{A}^{-1}$ exists with

$$\mathbb{A} : \mathbb{A}^{-1} = \mathbb{A}^{-1} : \mathbb{A} = \mathbb{I}_{\text{sym}} \quad . \tag{A.25}$$

In absence of whole numbers, powers $c \in \mathbb{R}$ can be evaluated via a spectral decomposition, i.e. the tensor is converted by using the Kelvin notation, see App. A.7, to a 6×6 matrix and the procedures given in App. A.5 are extended accordingly. Finally, it is

$$\mathbb{A}^c = \sum_{\alpha=1}^{6} (\lambda_\alpha)^c \, \boldsymbol{N}_\alpha \otimes \boldsymbol{N}_\alpha \quad , \tag{A.26}$$

with the real eigenvalues λ_α and principal directions $\boldsymbol{N}_\alpha = \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha$.

In the following, some useful differential operators are defined, which can be applied on tensor fields of different order to obtain their spatial derivatives. In particular, a scalar field $f(\boldsymbol{x})$, a vector field $\boldsymbol{f}(\boldsymbol{x})$, and a tensor fields $\boldsymbol{F}(\boldsymbol{x})$ (of second order) are considered. Since the divergence operator reduces the order of the tensor by one, it can only be applied on the vector and tensor field, i.e.

$$\begin{aligned} \operatorname{div}\big(\boldsymbol{f}(\boldsymbol{x})\big) &:= \frac{\partial a_i}{\partial x_i} = f_{i,i} \quad , \\ \operatorname{div}\big(\boldsymbol{F}(\boldsymbol{x})\big) &:= \frac{\partial F_{ij}}{\partial x_j} \, \boldsymbol{e}_i = F_{ij,j} \, \boldsymbol{e}_i \quad . \end{aligned} \tag{A.27}$$

In the case of $\operatorname{div} \boldsymbol{f} = 0$ the vector field $\boldsymbol{f}(\boldsymbol{x})$ is divergence-free or solenoidal. The curl operator can be applied on the vector and tensor field according to

$$\begin{aligned} \operatorname{curl}\big(\boldsymbol{f}(\boldsymbol{x})\big) &:= \varepsilon_{mni} \, \frac{\partial f_m}{\partial x_n} \, \boldsymbol{e}_i = \varepsilon_{mni} \, f_{m,n} \, \boldsymbol{e}_i \quad , \\ \operatorname{curl}\big(\boldsymbol{F}(\boldsymbol{x})\big) &:= \varepsilon_{mnj} \, \frac{\partial F_{im}}{\partial x_n} \, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \varepsilon_{mnj} \, F_{im,n} \, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad , \end{aligned} \tag{A.28}$$

with ε_{ijk} being the Levi-Civita operator as introduced above. By applying the curl operator the order of the tensor is unchanged. The vector field $\boldsymbol{f}(\boldsymbol{x})$ is irrotational or curl-free if $\operatorname{curl} \boldsymbol{f} = \boldsymbol{0}$. Gradients of the scalar, vector, and tensor field are obtained by applying the gradient operator according to

$$\begin{aligned} \operatorname{grad}\big(f(\boldsymbol{x})\big) &:= \frac{\partial f}{\partial x_i} \, \boldsymbol{e}_i = f_{,i} \, \boldsymbol{e}_i \quad , \\ \operatorname{grad}\big(\boldsymbol{f}(\boldsymbol{x})\big) &:= \frac{\partial f_i}{\partial x_j} \, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = f_{i,j} \, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad , \\ \operatorname{grad}\big(\boldsymbol{F}(\boldsymbol{x})\big) &:= \frac{\partial F_{ij}}{\partial x_k} \, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k = F_{ij,k} \, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \quad . \end{aligned} \tag{A.29}$$

By applying the gradient operator the order of the tensor is increased by one. Note further the relation $\operatorname{tr}(\operatorname{grad} \boldsymbol{f}) = \operatorname{div} \boldsymbol{f}$. The differential operators given above are based on the assumption, that the total derivative of a vector function is

$$\mathrm{d}\boldsymbol{f} = \operatorname{grad}\big(\boldsymbol{f}(\boldsymbol{x})\big)\,\mathrm{d}\boldsymbol{x} = \frac{\partial f_i\,\boldsymbol{e}_i}{\partial x_k}\,\boldsymbol{e}_k \cdot \mathrm{d}x_j\,\boldsymbol{e}_j = \frac{\partial f_i}{\partial x_j}\,\mathrm{d}x_j\,\boldsymbol{e}_i \quad , \tag{A.30}$$

so the differentiation is carried out from the right hand side (e.g. Malvern, 1969). Thus, they are more explicitly called right-gradient, right-rotation and right-divergence operator to distinguish the result from the left-operators, which, in most cases, is different. In this context, to generalize the application of differential operators from the right hand side, the Nabla or del operator is introduced as a vector according to

$$\boldsymbol{\nabla}(\bullet) := \frac{\partial(\bullet)}{\partial x_i}\,\boldsymbol{e}_i = (\bullet)_{,i}\,\boldsymbol{e}_i \quad . \tag{A.31}$$

Accordingly, the differential operators operating on a tensor of sufficient order can be written as

$$\operatorname{div}(\bullet) = \bullet \cdot \boldsymbol{\nabla} \quad , \qquad \operatorname{curl}(\bullet) = \bullet \times \boldsymbol{\nabla} \quad , \qquad \operatorname{grad}(\bullet) = \bullet \otimes \boldsymbol{\nabla} \quad . \tag{A.32}$$

In the following, the Nabla symbol is only used to indicate the gradient, i.e. $\nabla \boldsymbol{f} := \operatorname{grad} \boldsymbol{f}$. Based on the additive decomposition (A.5), symmetric and antisymmetric portions of the gradient

$$\nabla_{\mathrm{s}} \boldsymbol{f} := \frac{1}{2}(f_{i,j} + f_{j,i})\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \qquad \text{and} \qquad \nabla_{\mathrm{a}} \boldsymbol{f} := \frac{1}{2}(f_{i,j} - f_{j,i})\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \tag{A.33}$$

can be defined with $\nabla \boldsymbol{f} = \nabla_{\mathrm{s}} \boldsymbol{f} + \nabla_{\mathrm{a}} \boldsymbol{f}$.

Common second derivatives are the Laplacian with

$$\begin{aligned}
\Delta f(\boldsymbol{x}) &:= \operatorname{div}(\operatorname{grad} f) = \frac{\partial^2 f}{\partial x_i\,\partial x_i} = f_{,ii} \quad , \\
\Delta \boldsymbol{f}(\boldsymbol{x}) &:= \operatorname{div}(\operatorname{grad} \boldsymbol{f}) = \frac{\partial^2 f_i}{\partial x_j\,\partial x_j}\,\boldsymbol{e}_i = f_{i,jj}\,\boldsymbol{e}_i \quad , \\
\Delta \boldsymbol{F}(\boldsymbol{x}) &:= \operatorname{div}(\operatorname{grad} \boldsymbol{F}) = \frac{\partial^2 F_{ij}}{\partial x_k\,\partial x_k}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j = F_{ij,kk}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad ,
\end{aligned} \tag{A.34}$$

and the Hessian with

$$\begin{aligned}
\nabla^2 f &:= \operatorname{grad}(\operatorname{grad} f) = \frac{\partial^2 f}{\partial x_i\,\partial x_j} = f_{,ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad , \\
\nabla^2 \boldsymbol{f}(\boldsymbol{x}) &:= \operatorname{grad}(\operatorname{grad} \boldsymbol{f}) = \frac{\partial^2 f_i}{\partial x_j\,\partial x_k}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k = f_{i,jk}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \quad , \\
\nabla^2 \boldsymbol{F}(\boldsymbol{x}) &:= \operatorname{grad}(\operatorname{grad} \boldsymbol{F}) = \frac{\partial^2 F_{ij}}{\partial x_k\,\partial x_l}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l \\
&= F_{ij,kl}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l \quad .
\end{aligned} \tag{A.35}$$

In terms of the Nabla operator (A.31), the Laplacian and Hessian are $\Delta\bullet = (\bullet \otimes \boldsymbol{\nabla}) \cdot \boldsymbol{\nabla}$ and $\nabla^2\bullet = \bullet \otimes \boldsymbol{\nabla} \otimes \boldsymbol{\nabla}$, respectively. It should be further noted that

$$\operatorname{div}(\operatorname{curl} \boldsymbol{f}) = 0 \tag{A.36}$$

and

$$\operatorname{curl}(\nabla f) = \boldsymbol{0} \tag{A.37}$$

for any vector field $\boldsymbol{f}(\boldsymbol{x})$.

The variation of a quantity is denoted by

$$\delta_\xi(\bullet) := \left.\frac{\partial(\bullet)}{\partial \xi}\right|_{\xi=0} \quad . \tag{A.38}$$

In addition, the directional derivative, also known as the Gâteaux derivative, is introduced according to

$$\delta F(\boldsymbol{u}, \delta\boldsymbol{u}) = D_{\delta\boldsymbol{u}} F(\boldsymbol{u}) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon} F(\boldsymbol{u} + \epsilon\, \delta\boldsymbol{u})\right|_{\epsilon=0} \quad . \tag{A.39}$$

Note, that both the variational operator (A.38) and the Gâteaux operator (A.39) are linear.

A.2 Theorems and useful relations

Integral theorems

Integral theorems are essential in continuum mechanics and relate surface integrals of fields to volume integrals. The divergence theorem, also known as Gauss's divergence theorem[2], in particular establishes a relation between the flux of a vector field through a closed surface and the divergence of this field. Suppose $\boldsymbol{f}(\boldsymbol{x})$ is a continuously differentiable vector field defined in a three-dimensional Euclidean space with volume Ω, which is bounded by a smooth closed surface $\partial\Omega$ with the outward normal vector $\boldsymbol{n}$. The relation (without further proof) is

$$\int_{\partial\Omega} \boldsymbol{f} \cdot \boldsymbol{n}\, \mathrm{d}A = \int_{\Omega} \operatorname{div}(\boldsymbol{f})\, \mathrm{d}V \tag{A.40}$$

[2] As traced by Stolze (1978), it was not Carl Friedrich Gauss who stated the divergence theorem. But in Gauss's work from 1813 some relations between volume and surface integrals have been proven, which were used by James Clerk Maxwell in his treatise on electricity and magnetism in 1873. The history of integral theorems might date back to 1762, when Joseph-Louis Lagrange transformed volume integrals to surface integrals by using integration by parts. In 1826 Mikhail Vasilyevich Ostrogradskii presented the divergence theorem in Cartesian coordinate form and gave a general proof (published in 1831). At the same time Siméon-Denis Poisson, George Green and Frédéric Sarrus independently applied the theorem. Much later on, in 1893, Oliver Heaviside published the vector form of the theorem.

with $\boldsymbol{n}$ being the outward normal vector on $\partial\Omega$. Gauss's divergence theorem can be extended to other useful relations for scalars and tensors of second order by replacing the vector field in (A.40) by appropriately. For $\boldsymbol{f}(\boldsymbol{x}) = \alpha\,\boldsymbol{a}$ with a constant vector $\boldsymbol{a}$ and a continuously differentiable scalar field $\alpha(\boldsymbol{x})$ it is

$$a_i \int\limits_{\partial\Omega} \alpha\, n_i \,\mathrm{d}A = \int\limits_{\partial\Omega} (\alpha\, a_i)\, n_i \,\mathrm{d}A \overset{(A.40)}{=} \int\limits_{\Omega} \frac{\partial(\alpha\, a_i)}{\partial x_i} \,\mathrm{d}V = a_i \int\limits_{\Omega} \frac{\partial \alpha}{\partial x_i} \,\mathrm{d}V \tag{A.41}$$

to show the Green-Gauss-Ostrogradski theorem

$$\int\limits_{\partial\Omega} \alpha\, \boldsymbol{n} \,\mathrm{d}A = \int\limits_{\Omega} \operatorname{grad}(\alpha) \,\mathrm{d}V \quad . \tag{A.42}$$

For $\boldsymbol{f} = \tilde{\boldsymbol{A}}^{\mathrm{T}} \cdot \boldsymbol{a}$ with a constant vector $\boldsymbol{a}$ and a continuously differentiable tensor field $\tilde{\boldsymbol{A}}(\boldsymbol{x})$ it is

$$a_j \int\limits_{\partial\Omega} \tilde{A}_{ji}\, n_i \,\mathrm{d}A = \int\limits_{\partial\Omega} (\tilde{A}_{ji}\, a_j)\, n_i \,\mathrm{d}A \overset{(A.40)}{=} \int\limits_{\Omega} \frac{\partial(\tilde{A}_{ji}\, a_j)}{\partial x_i} \,\mathrm{d}V = a_j \int\limits_{\Omega} \frac{\partial \tilde{A}_{ji}}{\partial x_i} \,\mathrm{d}V \tag{A.43}$$

to show (with $\boldsymbol{A} = \tilde{\boldsymbol{A}}^{\mathrm{T}}$)

$$\int\limits_{\partial\Omega} \boldsymbol{A} \cdot \boldsymbol{n} \,\mathrm{d}A = \int\limits_{\Omega} \operatorname{div}(\boldsymbol{A}) \,\mathrm{d}V \quad . \tag{A.44}$$

Similar results are obtained for higher-order tensors as well as for other dimensions than three.

Partial integration

Differentiation of the product of a scalar field $\alpha(\boldsymbol{x})$ and vector field $\boldsymbol{a}(\boldsymbol{x})$ gives by using the product rule

$$\operatorname{div}(\alpha\, \boldsymbol{a}) = \frac{\partial(\alpha\, a_i)}{\partial x_i} = \frac{\partial \alpha}{\partial x_i}\, a_i + \alpha\, \frac{\partial a_i}{\partial x_i} = \operatorname{grad}\alpha \cdot \boldsymbol{a} + \alpha\, \operatorname{div} \boldsymbol{a} \tag{A.45}$$

which can be rewritten as

$$\boldsymbol{a} \cdot \operatorname{grad}\alpha = -\alpha\, \operatorname{div} \boldsymbol{a} + \operatorname{div}(\alpha\, \boldsymbol{a}) \quad . \tag{A.46}$$

Accordingly, differentiation of the product of a vector field $\boldsymbol{a}(\boldsymbol{x})$ and the transpose of a tensor field $\boldsymbol{A}(\boldsymbol{x})$ gives by using the product rule

$$\operatorname{div}(\boldsymbol{A}^{\mathrm{T}} \cdot \boldsymbol{a}) = \frac{\partial(A_{ji}\, a_i)}{\partial x_j} = \frac{\partial A_{ij}}{\partial x_j}\, a_i + A_{ij}\, \frac{\partial a_i}{\partial x_j} = \operatorname{div} \boldsymbol{A} \cdot \boldsymbol{a} + \boldsymbol{A} : \operatorname{grad} \boldsymbol{a} \tag{A.47}$$

which can be rewritten as

$$\boldsymbol{A} : \operatorname{grad} \boldsymbol{a} = -\operatorname{div} \boldsymbol{A} \cdot \boldsymbol{a} + \operatorname{div}(\boldsymbol{A}^{\mathrm{T}} \cdot \boldsymbol{a}) \quad . \tag{A.48}$$

By using the Gauss's divergence theorem we obtain the expressions for partial integration (or integration by parts) of a scalar field $\alpha(\boldsymbol{x})$ and vector field $\boldsymbol{a}(\boldsymbol{x})$ in multiple dimensions

$$\begin{aligned} \int_{\Omega} \boldsymbol{a} \cdot \operatorname{grad} \alpha \, \mathrm{d}V &= -\int_{\Omega} \alpha \, \operatorname{div} \boldsymbol{a} \, \mathrm{d}V + \int_{\Omega} \operatorname{div}(\alpha \, \boldsymbol{a}) \, \mathrm{d}V \\ &\overset{(A.40)}{=} -\int_{\Omega} \alpha \, \operatorname{div} \boldsymbol{a} \, \mathrm{d}V + \int_{\partial\Omega} \alpha \, \boldsymbol{a} \cdot \boldsymbol{n} \, \mathrm{d}A \end{aligned} \tag{A.49}$$

and

$$\begin{aligned} \int_{\Omega} \boldsymbol{A} : \operatorname{grad} \boldsymbol{a} \, \mathrm{d}V &= -\int_{\Omega} \operatorname{div} \boldsymbol{A} \cdot \boldsymbol{a} \, \mathrm{d}V + \int_{\Omega} \operatorname{div}(\boldsymbol{A}^{\mathrm{T}} \cdot \boldsymbol{a}) \, \mathrm{d}V \\ &\overset{(A.40)}{=} -\int_{\Omega} \operatorname{div} \boldsymbol{A} \cdot \boldsymbol{a} \, \mathrm{d}V + \int_{\partial\Omega} (\boldsymbol{A}^{\mathrm{T}} \cdot \boldsymbol{a}) \cdot \boldsymbol{n} \, \mathrm{d}A \end{aligned} \tag{A.50}$$

respectively.

Some useful relations

By substituting the linear elastic law (2.83) and the kinematic relations (2.17) into the equation of motion (2.47), the system of Navier equations is

$$\begin{aligned} \rho \, \ddot{u}_i &\overset{(2.47)}{=} \sigma_{ij,j} + f_i \overset{(A.76)}{=} (\lambda \, u_{k,k} \, \delta_{ij} + \mu \, (u_{i,j} + u_{j,i}))_{,j} + f_i \\ &= \lambda \, u_{j,ji} + \mu \, (u_{i,jj} + u_{j,ij}) + f_i = (\lambda + \mu) \, u_{j,ji} + \mu \, u_{i,jj} + f_i \quad . \end{aligned} \tag{A.51}$$

In addition, the identity

$$\operatorname{grad}(\operatorname{div} \boldsymbol{u}) - \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{u}) = \begin{bmatrix} u_{1,11} + u_{1,22} + u_{1,33} \\ u_{2,11} + u_{2,22} + u_{2,33} \\ u_{3,11} + u_{3,22} + u_{3,33} \end{bmatrix} = \delta \boldsymbol{u} \tag{A.52}$$

the identity

$$\nabla(\operatorname{div} \boldsymbol{u}) - \Delta \boldsymbol{u} = \operatorname{curl}(\operatorname{curl} \boldsymbol{u}) \quad , \tag{A.53}$$

is useful to derive the Helmholtz wave equations.

A.3 Integration by substitution

Integration by substitution is used to evaluate integrals of the form

$$I(x) = \int_{x_a}^{x_b} \left(f\big(g(x)\big) \, \frac{\mathrm{d}g(x)}{\mathrm{d}x} \right) \mathrm{d}x \tag{A.54}$$

by applying the substitution with the intermediate variable $u := g(x)$. Here, $g(x)$ is a differentiable function depending on x with a continuous derivative and f is also continuous. Consequently, differential expression $\mathrm{d}u/\mathrm{d}x = \mathrm{d}g(x)/\mathrm{d}x$ can be rewritten as

$$\mathrm{d}u = \frac{\mathrm{d}g(x)}{\mathrm{d}x} \, \mathrm{d}x \quad . \tag{A.55}$$

The lower and upper integration boundaries x_a and x_b are substituted accordingly, i.e.

$$u_a := g(x_a) \qquad \text{and} \qquad u_b := g(x_b) \quad , \tag{A.56}$$

respectively. Finally, the definite integral $I(x)$ can be calculated by

$$I(u) = \int_{u_a}^{u_b} f(u) \, \mathrm{d}u \quad . \tag{A.57}$$

The application on a function $w\big(s(x)\big)$, where the function argument itself depends on the coordinate $s(x)$ is then, with the given relation $\mathrm{d}x = f(s) \, \mathrm{d}s$,

$$\int_a^b w\big(s(x)\big) \, \mathrm{d}x = \int_{s(a)}^{s(b)} w(s) \, f(s) \, \mathrm{d}s \quad . \tag{A.58}$$

A.4 Invariants

The three principal invariants of a second order tensor $\boldsymbol{F}$ in a three-dimensional space are

$$\begin{aligned}
\mathrm{I_F} &:= \mathrm{tr}(\boldsymbol{F}) = F_{11} + F_{22} + F_{33} \quad , \\
\mathrm{II_F} &:= \frac{1}{2}\Big[\big(\mathrm{tr}(\boldsymbol{F})\big)^2 - \mathrm{tr}(\boldsymbol{F}^2) \Big] \\
&= F_{11} F_{22} + F_{22} F_{33} + F_{11} F_{33} - F_{12} F_{21} - F_{23} F_{32} - F_{13} F_{31} \quad , \\
\mathrm{III_F} &:= \det(\boldsymbol{F}) = F_{11} F_{22} F_{33} - F_{11} F_{23} F_{32} - F_{12} F_{21} F_{33} \\
&\quad + F_{12} F_{23} F_{31} + F_{13} F_{21} F_{32} - F_{13} F_{22} F_{31} \quad .
\end{aligned} \tag{A.59}$$

In case of a symmetric tensor $\boldsymbol{A} := (\boldsymbol{F} + \boldsymbol{F}^{\mathrm{T}})/2$ the principal invariants can be simplified to

$$\begin{aligned}
\mathrm{I}_{\mathrm{A}} &:= \mathrm{tr}(\boldsymbol{A}) = A_{11} + A_{22} + A_{33} = \lambda_1 + \lambda_2 + \lambda_3 \\
\mathrm{II}_{\mathrm{A}} &:= \big(\mathrm{tr}(\boldsymbol{A})^2 - \mathrm{tr}(\boldsymbol{A}\,\boldsymbol{A})\big)/2 = A_{11}\,A_{22} + A_{22}\,A_{33} \\
&\quad + A_{11}\,A_{33} - A_{12}^2 - A_{23}^2 - A_{13}^2 = \lambda_1\,\lambda_2 + \lambda_2\,\lambda_3 + \lambda_1\,\lambda_3 \\
\mathrm{III}_{\mathrm{A}} &:= \det(\boldsymbol{A}) = A_{11}\,A_{22}\,A_{33} - A_{11}\,A_{23}^2 - A_{22}\,A_{13}^2 \\
&\quad - A_{33}\,A_{12}^2 + 2A_{12}\,A_{23}\,A_{13} = \lambda_1\,\lambda_2\,\lambda_3 \quad .
\end{aligned} \tag{A.60}$$

Alternatively, main invariants can be defined instead, which are for the second order symmetric tensor $\boldsymbol{A}$

$$\begin{aligned}
\hat{\mathrm{I}}_{\mathrm{A}} &:= \mathrm{tr}(\boldsymbol{A}) = A_{11} + A_{22} + A_{33} \quad , \\
\hat{\mathrm{II}}_{\mathrm{A}} &:= \mathrm{tr}(\boldsymbol{A}^2) = A_{11}^2 + A_{22}^2 + A_{33}^2 + 2A_{12}^2 + 2A_{23}^2 + 2A_{13}^2 \quad , \\
\hat{\mathrm{III}}_{\mathrm{A}} &:= \mathrm{tr}(\boldsymbol{A}^3) = A_{11}^3 + A_{22}^3 + A_{33}^3 + 3\big(A_{11}\,A_{12}^2 + A_{11}\,A_{13}^2 + A_{22}\,A_{12}^2 \\
&\quad + A_{22}\,A_{23}^2 + A_{33}\,A_{23}^2 + A_{33}\,A_{13}^2\big) + 6A_{12}\,A_{23}\,A_{13} \quad .
\end{aligned} \tag{A.61}$$

The relation between the principal invariants (A.60) and main invariants (A.61) is

$$\mathrm{I}_{\mathrm{A}} = \hat{\mathrm{I}}_{\mathrm{A}} \quad , \qquad \mathrm{II}_{\mathrm{A}} = \frac{\hat{\mathrm{I}}_{\mathrm{A}}^2}{2} - \frac{\hat{\mathrm{II}}_{\mathrm{A}}}{2} \quad , \qquad \mathrm{III}_{\mathrm{A}} = \frac{\hat{\mathrm{I}}_{\mathrm{A}}^3}{6} - \frac{\hat{\mathrm{I}}_{\mathrm{A}}\,\hat{\mathrm{II}}_{\mathrm{A}}}{2} + \frac{\hat{\mathrm{III}}_{\mathrm{A}}}{3} \quad , \tag{A.62}$$

the inverse relation reads

$$\hat{\mathrm{I}}_{\mathrm{A}} = \mathrm{I}_{\mathrm{A}} \quad , \qquad \hat{\mathrm{II}}_{\mathrm{A}} = \mathrm{I}_{\mathrm{A}}^2 - 2\mathrm{II}_{\mathrm{A}} \quad , \qquad \hat{\mathrm{III}}_{\mathrm{A}} = \mathrm{I}_{\mathrm{A}}^3 - 3\mathrm{I}_{\mathrm{A}}\,\mathrm{II}_{\mathrm{A}} + 3\mathrm{III}_{\mathrm{A}} \quad . \tag{A.63}$$

Variables	Number of invariants	List of invariants
$\boldsymbol{v}$	1	$\boldsymbol{v} \cdot \boldsymbol{v}$
$\boldsymbol{A}$	3	$\mathrm{tr}\,\boldsymbol{A}$, $\mathrm{tr}(\boldsymbol{A}^2)$, $\mathrm{tr}(\boldsymbol{A}^3)$
$\boldsymbol{A}$, $\boldsymbol{v}$	6	$\boldsymbol{v} \cdot \boldsymbol{A} \cdot \boldsymbol{v}$, $\boldsymbol{v} \cdot \boldsymbol{A} \cdot \boldsymbol{A} \cdot \boldsymbol{v}$ and the invariant of $\boldsymbol{v}$ and the three invariants of $\boldsymbol{A}$
$\boldsymbol{A}$, $\boldsymbol{B}$	10	$\mathrm{tr}(\boldsymbol{A} \cdot \boldsymbol{B})$, $\mathrm{tr}(\boldsymbol{A}^2 \cdot \boldsymbol{B})$, $\mathrm{tr}(\boldsymbol{A} \cdot \boldsymbol{B}^2)$, $\mathrm{tr}(\boldsymbol{A}^2 \cdot \boldsymbol{B}^2)$ and the three invariants of $\boldsymbol{A}$ and the three invariants of $\boldsymbol{B}$
$\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{v}$	16	$(\boldsymbol{A} \cdot \boldsymbol{v}) \cdot (\boldsymbol{B} \cdot \boldsymbol{v})$ and the invariant of $\boldsymbol{v}$, the three invariants of $\boldsymbol{A}$, the three invariants of $\boldsymbol{B}$, the two additional invariants of $\boldsymbol{A}$, $\boldsymbol{v}$ the two additional invariants of $\boldsymbol{B}$, $\boldsymbol{v}$ and the four additional invariants of $\boldsymbol{A}$, $\boldsymbol{B}$

Table A.1: List of invariants of symmetric second order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ and a vector $\boldsymbol{v}$ in a three-dimensional space.

The multiplication of tensors of first and second order results in a new set if invariants, which depend on the invariants of the initial tensors. For example, the invariants of a vector $\boldsymbol{v}$ and combinations with the symmetric second order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ in a three-dimensional space are comprised in Tab. A.1, for more details see, e.g., Boehler (1977).

A.5 Principal values and directions

Regarding a symmetric second order tensor $\boldsymbol{A}$ in an arbitrary (Cartesian) coordinate system, in general all elements can be non-zero. In a particular coordinate system, however, only elements on the main diagonal can be non-zero. The tensor can then be represented by its spectral decomposition

$$\boldsymbol{A} = \sum_{\alpha=1}^{3} \lambda_\alpha \, \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha \quad , \tag{A.64}$$

with the principal values $\lambda_\alpha \in \mathbb{R}$ (elements on the main diagonal) and corresponding principal axes (principal directions) $\boldsymbol{n}_\alpha$. Principal values and axes are the solution of the homogeneous linear system of equations

$$(\boldsymbol{A} - \lambda_\alpha \, \boldsymbol{I}) \, \boldsymbol{n}_\alpha = \boldsymbol{0} \quad . \tag{A.65}$$

A non-trivial solution ($\boldsymbol{n} \neq \boldsymbol{0}$) is obtained only if the determinant $\det(\boldsymbol{A} - \lambda \, \boldsymbol{I})$ vanishes. Since this solution exists only for some specific values of λ, the problem (A.65) is also called eigenvalue problem. Accordingly, the solutions λ_α and $\boldsymbol{n}_\alpha$ are called eigenvalue and eigenvector, respectively. The vanishing determinant leads to a cubic equation, the characteristic polynomial

$$p(\lambda) = \det(\boldsymbol{A} - \lambda \, \boldsymbol{I}) = \lambda^3 - \mathrm{I}_\mathrm{A} \, \lambda^2 + \mathrm{II}_\mathrm{A} \, \lambda - \mathrm{III}_\mathrm{A} = 0 \quad . \tag{A.66}$$

Since the roots of the cubic equation are not influenced by the coordinate system, also the scalar coefficients I_A, II_A and III_A must be independent of the choice of the coordinate system. Therefore, they are the principal invariants of tensor $\boldsymbol{A}$ and are given as linear, quadratic and cubic functions of the components of $\boldsymbol{A}$ according to (A.60). In accordance with (A.66), the Cayley-Hamilton theorem can be used to obtain the characteristic equation

$$\boldsymbol{A}^3 - \mathrm{I}_\mathrm{A} \, \boldsymbol{A}^2 + \mathrm{II}_\mathrm{A} \, \boldsymbol{A} - \mathrm{III}_\mathrm{A} \, \boldsymbol{I} = \boldsymbol{0} \quad . \tag{A.67}$$

It can be used to compute the inverse of $\boldsymbol{A}$ according to

$$\boldsymbol{A}^{-1} = \frac{1}{\mathrm{III}_\mathrm{A}} \left(\boldsymbol{A}^2 - \mathrm{I}_\mathrm{A} \, \boldsymbol{A} + \mathrm{II}_\mathrm{A} \, \boldsymbol{I} \right) \quad . \tag{A.68}$$

Since the solution of eigenvalue problems of symmetric second order tensors or 3×3 matrices is a common task, a closed form solution and robust implementation is preferred. The eigenvalues λ_α and dyadic products of the eigenvectors $\boldsymbol{N}_\alpha := \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha$ can be

computed explicitly from the invariants. In the following, a slightly modified version of the approaches proposed by Morman Jr. (1986) and Franca (1989) is presented, which is also used in widely used textbooks (Simo and Hughes, 1998; Wriggers, 2008). First the auxiliary parameter

$$q := \mathrm{I}_\mathrm{A}^2 - 3\mathrm{II}_\mathrm{A} \geq 0 \tag{A.69}$$

is calculated, which quantifies the difference between the three eigenvalues. In case of equal eigenvalues ($q \to 0$) the solution is $\lambda_{1\ldots3} = \mathrm{I}_\mathrm{A}/3$. Otherwise, with

$$p := -\mathrm{I}^3 + \frac{9}{2}\mathrm{I}_\mathrm{A}\,\mathrm{II}_\mathrm{A} - \frac{27}{2}\mathrm{III}_\mathrm{A} \tag{A.70}$$

the angle[3]

$$\theta = \arccos\left(p/\sqrt{q^3}\right) \tag{A.71}$$

can be computed. The eigenvalues of $\boldsymbol{A}$ are finally (already arranged in descending order)

$$\lambda_\alpha = \frac{\mathrm{I}_\mathrm{A}}{3} - \frac{2}{3}\sqrt{q}\,\cos\left(\frac{\theta}{3} + \frac{2\pi}{3}\,\alpha\right) \quad . \tag{A.72}$$

If all eigenvalues are different ($\lambda_1 > \lambda_2 > \lambda_3$) the dyadic products of the eigenvectors/principal directions are

$$\boldsymbol{N}_\alpha := \boldsymbol{n}_\alpha \otimes \boldsymbol{n}_\alpha = \frac{\lambda_\alpha}{2\lambda_\alpha^3 - \mathrm{I}_\mathrm{A}\,\lambda_\alpha^2 + \mathrm{III}_\mathrm{A}} \left(\boldsymbol{A}^2 - (\mathrm{I}_\mathrm{A} - \lambda_\alpha)\,\boldsymbol{A} + \mathrm{III}_\mathrm{A}/\lambda_\alpha\,\boldsymbol{I}\right) \quad . \tag{A.73}$$

In the case that two eigenvalues are equal, e.g. $\lambda_1 > \lambda_2 = \lambda_3$, the dyadic product $\boldsymbol{N}_1$ is calculated according to (A.73) while

$$\boldsymbol{N}_2 + \boldsymbol{N}_3 = \boldsymbol{I} - \boldsymbol{N}_1 \quad . \tag{A.74}$$

For three equal eigenvalues ($\lambda_1 = \lambda_2 = \lambda_3$) one obtains

$$\boldsymbol{N}_1 + \boldsymbol{N}_2 + \boldsymbol{N}_3 = \boldsymbol{I} \quad . \tag{A.75}$$

In general, the spectral decomposition is not limited to 3x3 matrices, which are used to represent second order tensors.

A.6 Linear elastic relations

Isotropic linear elastic material serves as a basis for further damage and fracture modeling in this work. In the following, useful relations are provided to shorten the discussion in different chapters.

[3] For numerical reasons, instead of evaluating the arccosine, usually the function atan2($\sqrt{q^3 - r^2}, r$) is used to improve the result for $q \to 0$, see Simo and Hughes (1998).

The stress-strain relation of an isotropic linear elastic material in terms of the Lamé constants is

$$\begin{aligned}\sigma_{ij} &\overset{(2.83)}{=} C_{ijkl}\,\varepsilon_{kl} \overset{(2.17)}{=} C_{ijkl}\,(u_{k,l}+u_{l,k})/2\\ &\overset{(2.85)}{=} [\lambda\,\delta_{ij}\,\delta_{kl}+\mu\,(\delta_{ik}\,\delta_{jl}+\delta_{il}\,\delta_{jk})]\,(u_{k,l}+u_{l,k})/2\\ &= \lambda\,u_{k,k}\,\delta_{ij}+\mu\,(u_{i,j}+u_{j,i}) \quad .\end{aligned} \tag{A.76}$$

All relevant relations between elastic parameters are given in Tab. A.2.

	$\lambda,\,\mu$	K, G	E, G	$E,\,\nu$
λ	λ	$K-\frac{2}{3}G$	$\frac{G\,(E-2G)}{3G-E}$	$\frac{E\,\nu}{(1+\nu)(1-2\nu)}$
μ	μ	G	G	$\frac{E}{2(1+\nu)}$
G	μ	G	G	$\frac{E}{2(1+\nu)}$
K	$\lambda+\frac{2}{3}\mu$	K	$\frac{G\,E}{3(3G-E)}$	$\frac{E}{3(1-2\nu)}$
E	$\frac{\mu\,(3\lambda+2\mu)}{\lambda+\mu}$	$\frac{9K\,G}{3K+G}$	E	E
ν	$\frac{\lambda}{2(\lambda+\mu)}$	$\frac{3K-2G}{2(3K+G)}$	$\frac{E}{2G}-1$	ν

Table A.2: Relations between common linear elastic parameters.

Plane stress and strain

In the plane stress state all stress components on a plane, e.g. in z-direction, are equal to zero, i.e.

$$\sigma_{zz}=\sigma_{xz}=\sigma_{yz}\overset{!}{=}0 \quad . \tag{A.77}$$

Solving Hook's law (2.83) by using the conditions in (A.77) the stress components in isotropic material are

$$\begin{aligned}\sigma_{xx} &= \frac{E}{1-\nu^2}\,(\varepsilon_{xx}+\nu\,\varepsilon_{yy})\\ \sigma_{yy} &= \frac{E}{1-\nu^2}\,(\varepsilon_{yy}+\nu\,\varepsilon_{xx})\\ \sigma_{xy} &= \frac{E}{1+\nu}\,\varepsilon_{xy}\end{aligned} \tag{A.78}$$

and the non-zero strain components are

$$\begin{aligned}
\varepsilon_{xx} &= \frac{1}{E}\left(\sigma_{xx} - \nu\,\sigma_{yy}\right)\\
\varepsilon_{yy} &= \frac{1}{E}\left(\sigma_{yy} - \nu\,\sigma_{xx}\right)\\
\varepsilon_{xy} &= \frac{1+\nu}{E}\,\sigma_{xy}\\
\varepsilon_{zz} &= -\frac{\nu}{1-\nu}\left(\varepsilon_{xx} + \varepsilon_{yy}\right) \quad .
\end{aligned} \tag{A.79}$$

Equivalently, elastic relations for the plane strain state, where all strain components with respect to a plane, e.g. also in z-direction, are equal to zero, are obtained, i.e.

$$\varepsilon_{zz} = \varepsilon_{xz} = \varepsilon_{yz} \overset{!}{=} 0 \quad . \tag{A.80}$$

The non-zero stress components are

$$\begin{aligned}
\sigma_{xx} &= \frac{E}{(1+\nu)\,(1-2\nu)}\left((1-\nu)\,\varepsilon_{xx} + \nu\,\varepsilon_{yy}\right) = \frac{E'}{1-\nu'^2}\left(\varepsilon_{xx} + \nu'\,\varepsilon_{yy}\right)\\
\sigma_{yy} &= \frac{E}{(1+\nu)\,(1-2\nu)}\left((1-\nu)\,\varepsilon_{yy} + \nu\,\varepsilon_{xx}\right) = \frac{E'}{1-\nu'^2}\left(\varepsilon_{yy} + \nu'\,\varepsilon_{xx}\right)\\
\sigma_{xy} &= \frac{E}{1+\nu}\,\varepsilon_{xy} = \frac{E'}{1+\nu'}\,\varepsilon_{xy}\\
\sigma_{zz} &= \nu\left(\sigma_{xx} + \sigma_{yy}\right)
\end{aligned} \tag{A.81}$$

while the non-zero strain components are

$$\begin{aligned}
\varepsilon_{xx} &= \frac{1-\nu^2}{E}\left(\sigma_{xx} - \frac{\nu}{1-\nu}\,\sigma_{yy}\right) = \frac{1}{E'}\left(\sigma_{xx} - \nu'\,\sigma_{yy}\right)\\
\varepsilon_{yy} &= \frac{1-\nu^2}{E}\left(\sigma_{yy} - \frac{\nu}{1-\nu}\,\sigma_{xx}\right) = \frac{1}{E'}\left(\sigma_{yy} - \nu'\,\sigma_{xx}\right)\\
\varepsilon_{xy} &= \frac{1+\nu}{E}\,\sigma_{xy} = \frac{1+\nu'}{E'}\,\sigma_{xy} \quad .
\end{aligned} \tag{A.82}$$

As depicted in (A.81) and (A.82), only by substituting the elastic constants with $E' := E/(1-\nu^2)$ and $\nu' := \nu/(1-\nu)$ the relations in plane strain (A.81) and (A.82) and plane stress (A.78) and (A.79) are equivalent.

A.7 Matrix-vector notation

Tensors of orders higher than two are difficult to handle on paper, so a compact notation would be beneficial. In addition, symmetric tensors contain many equal components. The numerical performance of symmetric tensors and matrices can be improved by reducing

the evaluation of equal expressions. Therefore, some quantities and relations can be reformulated in simple matrix-vector notations. There are several practical possibilities, in the following we regard the most common two approaches.

To illustrate common approaches, we regard Hook's law (2.83), which is given by

$$\sigma_{ij} = C_{ijkl}\, \varepsilon_{kl} \quad , \tag{A.83}$$

with the symmetric fourth order elasticity tensor C_{ijkl}. The inverse relation reads

$$\varepsilon_{ij} = S_{ijkl}\, \sigma_{kl} \quad , \tag{A.84}$$

with the fourth-order compliance tensor $S_{ijkl} = C^{-1}_{ijkl}$. The corresponding matrix-vector notations would look like

$$\hat{\sigma}_i = \hat{C}_{ij}\, \hat{\varepsilon}_j \qquad \text{and} \qquad \hat{\varepsilon}_i = \hat{S}_{ij}\, \hat{\sigma}_j \quad . \tag{A.85}$$

Voigt notation

As already mentioned above, there is no unique transformation law between the (physically motivated) tensor notations (A.83) or (A.84) and a technically motivated approach (A.85). A widely used concept in finite element codes is the Voigt notation[4]. By using the Voigt notation the stress and strain tensors are transformed to vectors by skipping all symmetric components and rearranging the (not uniquely defined) order, e.g., according to

$$\longrightarrow \hat{\boldsymbol{\sigma}}^{\mathrm{V}} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{bmatrix} = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \qquad \text{and} \qquad \longrightarrow \hat{\boldsymbol{\varepsilon}}^{\mathrm{V}} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{12} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad . \tag{A.86}$$

The factor of two in the shearing components of the strain vector, often referred to as engineering strain with $\gamma_{ij} := 2\varepsilon_{ij}$ (for $i \neq j$), appears to fulfill the physical relation (A.83). The corresponding elasticity matrix is

$$\longrightarrow \hat{\boldsymbol{C}}^{\mathrm{V}} = \begin{bmatrix} C_{1111} & C_{1122} & C_{1133} & C_{1112} & C_{1123} & C_{1113} \\ C_{1122} & C_{2222} & C_{2233} & C_{1222} & C_{2223} & C_{1322} \\ C_{1133} & C_{2233} & C_{3333} & C_{1233} & C_{2333} & C_{1333} \\ C_{1112} & C_{1222} & C_{1233} & C_{1212} & C_{1223} & C_{1213} \\ C_{1123} & C_{2223} & C_{2333} & C_{1223} & C_{2323} & C_{1323} \\ C_{1113} & C_{1322} & C_{1333} & C_{1213} & C_{1323} & C_{1313} \end{bmatrix} \quad . \tag{A.87}$$

[4] Woldemar Voigt, German physicist, 2 Sep. 1850 – 13 Dec. 1919

It is important to note, that the corresponding compliance tensor, however, is

$$\longrightarrow \hat{\boldsymbol{S}}^{\mathrm{V}} = \begin{bmatrix} S_{1111} & S_{1122} & S_{1133} & 2S_{1112} & 2S_{1123} & 2S_{1113} \\ S_{1122} & S_{2222} & S_{2233} & 2S_{1222} & 2S_{2223} & 2S_{1322} \\ S_{1133} & S_{2233} & S_{3333} & 2S_{1233} & 2S_{2333} & 2S_{1333} \\ 2S_{1112} & 2S_{1222} & 2S_{1233} & 4S_{1212} & 4S_{1223} & 4S_{1213} \\ 2S_{1123} & 2S_{2223} & 2S_{2333} & 4S_{1223} & 4S_{2323} & 4S_{1323} \\ 2S_{1113} & 2S_{1322} & 2S_{1333} & 4S_{1213} & 4S_{1323} & 4S_{1313} \end{bmatrix} , \tag{A.88}$$

and also contains components with pre-factors of two and four to fulfill the physical relation (A.84). Due to this modification, $\hat{\boldsymbol{C}}^{\mathrm{V}} \neq (\hat{\boldsymbol{S}}^{\mathrm{V}})^{-1}$. In addition, the principal values and directions of $\hat{\boldsymbol{C}}^{\mathrm{V}}$ and $\hat{\boldsymbol{S}}^{\mathrm{V}}$ do not correspond to those of C_{ijkl} and S_{ijkl}, respectively. Also transformations, e.g. rotations, require adapted rotation matrices.

Kelvin notation

Another matrix-vector representation is the Kelvin notation. The above shown transformation from given tensors to vectors and matrices is different in the treatment of some components. Strains and stresses are transformed according to

$$\longrightarrow \hat{\boldsymbol{\sigma}}^{\mathrm{K}} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\,\sigma_{12} \\ \sqrt{2}\,\sigma_{23} \\ \sqrt{2}\,\sigma_{13} \end{bmatrix} = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \quad \text{and} \quad \longrightarrow \hat{\boldsymbol{\varepsilon}}^{\mathrm{K}} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \sqrt{2}\,\varepsilon_{12} \\ \sqrt{2}\,\varepsilon_{23} \\ \sqrt{2}\,\varepsilon_{13} \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} . \tag{A.89}$$

To fulfill the physical relations (A.83) and (A.84) given by Hooke's law the elasticity matrix is

$$\longrightarrow \hat{\boldsymbol{C}}^{\mathrm{K}} = \begin{bmatrix} C_{1111} & C_{1122} & C_{1133} & \sqrt{2}\,C_{1112} & \sqrt{2}\,C_{1123} & \sqrt{2}\,C_{1113} \\ C_{1122} & C_{2222} & C_{2233} & \sqrt{2}\,C_{1222} & \sqrt{2}\,C_{2223} & \sqrt{2}\,C_{1322} \\ C_{1133} & C_{2233} & C_{3333} & \sqrt{2}\,C_{1233} & \sqrt{2}\,C_{2333} & \sqrt{2}\,C_{1333} \\ \sqrt{2}\,C_{1112} & \sqrt{2}\,C_{1222} & \sqrt{2}\,C_{1233} & 2\,C_{1212} & 2\,C_{1223} & 2\,C_{1213} \\ \sqrt{2}\,C_{1123} & \sqrt{2}\,C_{2223} & \sqrt{2}\,C_{2333} & 2\,C_{1223} & 2\,C_{2323} & 2\,C_{1323} \\ \sqrt{2}\,C_{1113} & \sqrt{2}\,C_{1322} & \sqrt{2}\,C_{1333} & 2\,C_{1213} & 2\,C_{1323} & 2\,C_{1313} \end{bmatrix} . \tag{A.90}$$

and the compliance matrix is

$$\longrightarrow \hat{\boldsymbol{S}}^{\mathrm{K}} = \begin{bmatrix} S_{1111} & S_{1122} & S_{1133} & \sqrt{2}\,S_{1112} & \sqrt{2}\,S_{1123} & \sqrt{2}\,S_{1113} \\ S_{1122} & S_{2222} & S_{2233} & \sqrt{2}\,S_{1222} & \sqrt{2}\,S_{2223} & \sqrt{2}\,S_{1322} \\ S_{1133} & S_{2233} & S_{3333} & \sqrt{2}\,S_{1233} & \sqrt{2}\,S_{2333} & \sqrt{2}\,S_{1333} \\ \sqrt{2}\,S_{1112} & \sqrt{2}\,S_{1222} & \sqrt{2}\,S_{1233} & 2\,S_{1212} & 2\,S_{1223} & 2\,S_{1213} \\ \sqrt{2}\,S_{1123} & \sqrt{2}\,S_{2223} & \sqrt{2}\,S_{2333} & 2\,S_{1223} & 2\,S_{2323} & 2\,S_{1323} \\ \sqrt{2}\,S_{1113} & \sqrt{2}\,S_{1322} & \sqrt{2}\,S_{1333} & 2\,S_{1213} & 2\,S_{1323} & 2\,S_{1313} \end{bmatrix} . \tag{A.91}$$

The advantage of this specific matrix-vector notation is, that the principal values and directions of $\hat{\boldsymbol{C}}^{\mathrm{K}}$ and $\hat{\boldsymbol{S}}^{\mathrm{K}}$ correspond to those of C_{ijkl} and S_{ijkl}. Since these physical properties are kept in the transformed notation, also mathematical operations like the inversion $\hat{\boldsymbol{S}}^{\mathrm{K}} = (\hat{\boldsymbol{C}}^{\mathrm{K}})^{-1}$ are valid.

A.8 Crack tip fields

Depending on the type of crack opening the displacement field as well as strain and stress fields in the vicinity of a crack tip can be given by compex functions (Westergaard, 1939) or by factorization of stress fields into radial terms and angular functions by using the concept of stress intensity factors (see Sect. 2.2.3). In the following, only non-zero components are given. The parameter κ is used to distinguish between plane strain and plane stress conditions with

$$\kappa = \begin{cases} 3 - 4\nu & \text{plane strain} \\ (3-\nu)/(1+\nu) & \text{plane stress} \end{cases} . \tag{A.92}$$

As reported by Irwin (1960), the fields can be easily extended to three dimensions since they are insensitive with respect to the third coordinate.

For pure mode I loading ("tensile opening") the displacment field is

$$\begin{Bmatrix} u_x \\ u_y \end{Bmatrix} = \frac{K_{\mathrm{I}}}{2\mu}\sqrt{\frac{r}{2\pi}}\,(\kappa - \cos\varphi) \begin{Bmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{Bmatrix} , \tag{A.93}$$

the strain field is

$$\begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{xy} \end{Bmatrix} = \frac{K_{\mathrm{I}}}{2\mu}\sqrt{\frac{1}{2\pi\, r}}\,\cos(\varphi/2) \begin{Bmatrix} (\kappa-1)/2 - \sin(\varphi/2)\,\sin(3\varphi/2) \\ (\kappa-1)/2 + \sin(\varphi/2)\,\sin(3\varphi/2) \\ \sin(\varphi/2)\,\cos(3\varphi/2) \end{Bmatrix} \tag{A.94}$$

with $\varepsilon_{zz} = 0$ under plane strain and

$$\varepsilon_{zz} = -\frac{\nu}{1-\nu}\,(\varepsilon_x + \varepsilon_y) = -\frac{K_{\mathrm{I}}}{\mu}\sqrt{\frac{1}{2\pi\, r}}\,\frac{\nu}{1+\nu}\,\cos(\varphi/2) \tag{A.95}$$

under plane stress conditions and the stress field is

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{Bmatrix} = \frac{K_{\mathrm{I}}}{\sqrt{2\pi\, r}}\,\cos(\varphi/2) \begin{Bmatrix} 1 - \sin(\varphi/2)\,\sin(3\varphi/2) \\ 1 + \sin(\varphi/2)\,\sin(3\varphi/2) \\ \sin(\varphi/2)\,\cos(3\varphi/2) \end{Bmatrix} \tag{A.96}$$

with

$$\sigma_{zz} = \nu\,(\sigma_x + \sigma_y) = \frac{K_\mathrm{I}}{\sqrt{2\pi\,r}}\,2\nu\,\cos(\varphi/2) \tag{A.97}$$

under plane strain and $\sigma_{zz} = 0$ under plane stress conditions. From (A.93) one obtains

$$u_y^{\pm} = u_y(\pm\pi) = \pm\frac{K_\mathrm{I}}{2\mu}\sqrt{\frac{r}{2\pi}}\,(\kappa+1) \quad , \tag{A.98}$$

which represents the parabola-shaped opened crack subjected to normal tension.

For pure mode II loading ("in-plane shear") the displacment field is

$$\left\{\begin{array}{c} u_x \\ u_y \end{array}\right\} = \frac{K_\mathrm{II}}{2\mu}\sqrt{\frac{r}{2\pi}}\left\{\begin{array}{c} \sin(\varphi/2)\,(\kappa+2+\cos\varphi) \\ \cos(\varphi/2)\,(\kappa-2+\cos\varphi) \end{array}\right\} \tag{A.99}$$

and the stress field is

$$\left\{\begin{array}{c} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{array}\right\} = \frac{K_\mathrm{II}}{\sqrt{2\pi\,r}}\left\{\begin{array}{c} -\sin(\varphi/2)\big(2+\cos(\varphi/2)\,\cos(3\varphi/2)\big) \\ \sin(\varphi/2)\,\cos(\varphi/2)\,\cos(3\varphi/2) \\ \cos(\varphi/2)\big(1-\sin(\varphi/2)\,\sin(3\varphi/2)\big) \end{array}\right\} \tag{A.100}$$

with $\sigma_z = \nu\,(\sigma_x + \sigma_y)$ under plane strain and $\sigma_z = 0$ under plane stress conditions.

For pure mode III loading ("out-of-plane shear") the displacment field is

$$u_z = \frac{2K_\mathrm{III}}{\mu}\sqrt{\frac{r}{2\pi}}\,\sin(\varphi/2) \quad , \tag{A.101}$$

the strain field is

$$\left\{\begin{array}{c} \varepsilon_{xz} \\ \varepsilon_{yz} \end{array}\right\} = \frac{K_\mathrm{III}}{2\mu\,\sqrt{2\pi\,r}}\left\{\begin{array}{c} -\sin(\varphi/2) \\ \cos(\varphi/2) \end{array}\right\} \tag{A.102}$$

and the stress field is

$$\left\{\begin{array}{c} \tau_{xz} \\ \tau_{yz} \end{array}\right\} = \frac{K_\mathrm{III}}{\sqrt{2\pi\,r}}\left\{\begin{array}{c} -\sin(\varphi/2) \\ \cos(\varphi/2) \end{array}\right\} \quad . \tag{A.103}$$

A.9 Basic properties of the bracket operator

In order to extract the portion of a scalar parameter $a \in \mathbb{R}$ which is larger than a specific scalar value $p \in \mathbb{R}$, e.g. a threshold value, the bracket operator $\langle\bullet\rangle$ with

$$\langle a-p\rangle := \frac{1}{2}\big(|a-p| + a - p\big) = \left\{\begin{array}{ll} 0 & \forall a \le p \\ a-p & \forall a > p \end{array}\right. \tag{A.104}$$

is utilized in the following. In the English literature the bracket operator is often denoted as "Macaulay bracket" (Macaulay, 1919) while it actually has been introduced by Clebsch (1862) to calculate beam deflections and became publicly known by the work of Föppl (1914). Alternatively, the bracket operator can be defined by

$$\langle a - p \rangle := \max \left\{ a - p,\, 0 \right\} \quad . \tag{A.105}$$

The operator is also closely related to the Heaviside (step) function

$$\mathcal{H}\{x\} := \frac{1}{2}\big(\operatorname{sgn}(x) + 1\big) = \begin{cases} 1 & \forall x > 0 \\ 1/2 & \forall x = 0 \\ 0 & \forall x < 0 \end{cases} \tag{A.106}$$

which fulfills, by definition, the identity $\mathcal{H}\{x\} + \mathcal{H}\{-x\} \equiv 1$ for $x \in \mathbb{R}$. With (A.106) the bracket operator reads

$$\langle a - p \rangle = \mathcal{H}\{a - p\}\,(a - p) \quad . \tag{A.107}$$

Accordingly, the portion (absolute value) of a which is smaller than p is extracted by

$$\langle p - a \rangle := \frac{1}{2}\big(|p - a| + p - a\big) = \mathcal{H}\{p - a\}\,(p - a) = \begin{cases} 0 & \forall a \geq p \\ p - a & \forall a < p \end{cases} \quad . \tag{A.108}$$

With $x := a - p$ the definition (A.104) yields the widely used relation

$$\langle x \rangle = \frac{1}{2}\big(|x| + x\big) = \begin{cases} 0 & \forall x \leq 0 \\ x & \forall x > 0 \end{cases} \quad , \tag{A.109}$$

with $\langle x \rangle \in \mathbb{R}_0^+$. The operator can be considered as a non-smooth ramp function as depicted in Fig. A.1a. To additively decompose x into positive and negative portions the identity

$$x \equiv \langle x \rangle - \langle -x \rangle \tag{A.110}$$

is utilized. In addition, the property

$$\langle x \rangle \, \langle -x \rangle = 0 \tag{A.111}$$

is often used. By using (A.110) and (A.111) one finds that

$$x^2 = \big(\langle x \rangle - \langle -x \rangle\big)^2 = \langle x \rangle^2 - 2 \underbrace{\langle x \rangle \, \langle -x \rangle}_{=\,0} + \langle -x \rangle^2 = \langle x \rangle^2 + \langle -x \rangle^2 \quad . \tag{A.112}$$

With the introducing of the scalar parameters $y \in \mathbb{R}$ and $z \in \mathbb{R}$ the corresponding products can be decomposed according to

$$\begin{aligned} x\,y &= \big(\langle x \rangle - \langle -x \rangle\big)\big(\langle y \rangle - \langle -y \rangle\big) \\ &= \langle x \rangle \, \langle y \rangle - \langle -x \rangle \, \langle y \rangle - \langle x \rangle \, \langle -y \rangle + \langle -x \rangle \, \langle -y \rangle \end{aligned} \tag{A.113}$$

and

$$\begin{aligned} x\,y\,z &= \big(\langle x\rangle - \langle -x\rangle\big)\big(\langle y\rangle - \langle -y\rangle\big)\big(\langle z\rangle - \langle -z\rangle\big) \\ &= \langle x\rangle\,\langle y\rangle\,\langle z\rangle - \langle -x\rangle\,\langle y\rangle\,\langle z\rangle - \langle x\rangle\,\langle -y\rangle\,\langle z\rangle + \langle -x\rangle\,\langle -y\rangle\,\langle z\rangle \\ &\quad - \langle x\rangle\,\langle y\rangle\,\langle -z\rangle + \langle -x\rangle\,\langle y\rangle\,\langle -z\rangle + \langle x\rangle\,\langle -y\rangle\,\langle -z\rangle - \langle -x\rangle\,\langle -y\rangle\,\langle -z\rangle \quad . \end{aligned} \tag{A.114}$$

The partial derivative are

$$\frac{\partial\,\langle x\rangle}{\partial\,x} = \frac{\langle x\rangle}{x} = \mathcal{H}\{x\} \tag{A.115}$$

and

$$\frac{1}{2}\,\frac{\partial\,\langle x\rangle^2}{\partial\,x} = \frac{\langle x\rangle^2}{x} = \langle x\rangle = \mathcal{H}\{x\}\,x \quad , \tag{A.116}$$

see Fig. A.1b, and the partial derivatives of a continuous function $f(x)$ are

$$\frac{\partial\,\langle f(x)\rangle}{\partial\,x} = \frac{\langle x\rangle}{x}\,\frac{\partial\,f(x)}{\partial\,x} = \mathcal{H}(f(x))\,f'(x) \tag{A.117}$$

and

$$\frac{1}{2}\,\frac{\partial\,\langle f(x)\rangle^2}{\partial\,x} = \langle f(x)\rangle\,\frac{\partial\,\langle f(x)\rangle}{\partial\,x} = \langle f(x)\rangle\,f'(x) \quad . \tag{A.118}$$

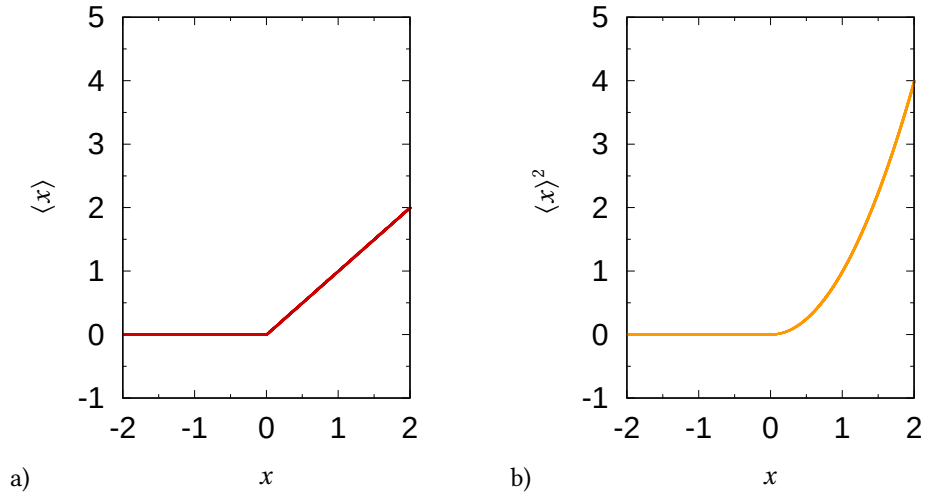


Figure A.1: Evaluation of the important ramp function a) $\langle x\rangle$ and b) $\langle x\rangle^2$.

B Additional details on phase field formulations

B.1 Useful relations

In the following some useful relations are given. The first and second spatial derivative of the phase field s can be related by using the product rule, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}s}\left[\left(\frac{\mathrm{d}s}{\mathrm{d}x}\right)^2\right] = 2\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}s}{\mathrm{d}x}\right)\cdot\frac{\mathrm{d}s}{\mathrm{d}x} = 2\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}s}{\mathrm{d}x}\right) = 2\frac{\mathrm{d}^2 s}{\mathrm{d}x^2} \quad . \tag{B.1}$$

By applying the chain rule on the inverse of the degradation function $g(s)$ one obtains

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{g(s)}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\Big(\big(g(s)\big)^{-1}\Big) = -\big(g(s)\big)^{-2}\,\frac{\mathrm{d}}{\mathrm{d}s}g(s) = -\frac{g'(s)}{\big(g(s)\big)^2} \quad . \tag{B.2}$$

B.2 Crack surface approximation — general quadratic function

The approximation of the crack surface density (3.52) depends on the surface function $w(s)$, which describes the local part (Sect. 3.3.1). In general, a polynomial surface function of n-th order is given by

$$w_{\mathrm{n}}(s) := \sum_{k=1}^{n} a_k\, s^k = a_0 + a_1\, s + a_2\, s^2 + a_3\, s^3 + \cdots + a_n\, s^n \tag{B.3}$$

with

$$w'(s) := \frac{\mathrm{d}w(s)}{\mathrm{d}s} = a_1 + 2a_2\, s + \cdots + n\, a_n\, s^{n-1} \quad . \tag{B.4}$$

With $n = 2$ and the fundamental requirements (3.54) one obtains for (B.3)

$$w(s{=}0) = a_0 = 1 \quad \Leftrightarrow \quad a_0 = 1 \tag{B.5}$$

and

$$w(s{=}1) = a_0 + a_1 + a_2 = 0 \quad \Leftrightarrow \quad a_1 = -1 - a_2 \quad . \tag{B.6}$$

By parametrizing the coefficient of the largest exponent with $\beta = a_2$ one obtains the general quadratic function

$$w_\beta(s) \coloneqq \beta\, s^2 - (\beta + 1)\, s + 1 \tag{B.7}$$

with its derivatives

$$w'_\beta(s) = 2\beta\, s - (\beta + 1) \qquad \text{and} \qquad w''_\beta(s) = 2\beta \quad . \tag{B.8}$$

Taking the requirement (3.55) into account yields

$$w'(s) = (2s - 1)\,\beta - 1 < 0 \quad \forall s \in (0,\, 1) \quad . \tag{B.9}$$

with the restrictions on β given by

$$\begin{aligned} \lim_{s \to 0^+} w'(s) = -\beta - 1 < 0 \quad &\Rightarrow \beta \geq -1 \\ \lim_{s \to 1^-} w'(s) = \beta - 1 < 0 \quad &\Rightarrow \beta \leq 1 \quad . \end{aligned} \tag{B.10}$$

B.3 Scaling factor of the crack surface density

In contrast to the Dirac distribution $\delta(x)$ which fulfills

$$\int_{-\infty}^{\infty} \delta(x)\, \mathrm{d}x = 1 \tag{B.11}$$

by definition, the phase field approximation of the crack surface must fulfill the property

$$\int_{-\infty}^{\infty} \gamma_\ell(s, \nabla s, \dots)\, \mathrm{d}x_\mathrm{n} = 1 \quad . \tag{B.12}$$

If the distribution of $s(x_\mathrm{n})$, with x_n being the coordinate normal to the crack surface, is assumed to be some kind of Dirac distribution, the scaling of the crack surface is different from the distribution (B.11) due to the additional contribution of the non-local terms. For this purpose, the scaling constant c_w is included in γ_ℓ as introduced in (3.52).

General relation

According to the equilibrium between the local and non-local term in the phase field evolution equation (3.39), it is

$$2\frac{\mathrm{d}^2 s}{\mathrm{d}x_\mathrm{n}^2} = \frac{1}{\ell^2}\,\frac{\mathrm{d}w(s)}{\mathrm{d}s} \quad . \tag{B.13}$$

Multiplication of both sides with $\mathrm{d}s/\mathrm{d}x_\mathrm{n}$ and integration over x_n yields

$$\int_x 2\frac{\mathrm{d}^2 s}{\mathrm{d}x_\mathrm{n}^2}\frac{\mathrm{d}s}{\mathrm{d}x_\mathrm{n}}\,\mathrm{d}x_\mathrm{n} = \int_x \frac{1}{\ell^2}\frac{\mathrm{d}w(s)}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}x_\mathrm{n}}\,\mathrm{d}x_\mathrm{n} \quad . \tag{B.14}$$

With the product rule (B.1) one obtains

$$\left(\frac{\mathrm{d}s}{\mathrm{d}x_\mathrm{n}}\right)^2 = \frac{1}{\ell^2}\,w(s) \tag{B.15}$$

or

$$\frac{\mathrm{d}s}{\mathrm{d}x_\mathrm{n}} = \pm\frac{1}{\ell}\sqrt{w(s)} \quad , \tag{B.16}$$

which relates the spatial derivative of the phase field with the (local) surface function. From (B.16) the differential relation

$$\mathrm{d}x_\mathrm{n} = \frac{\ell}{\sqrt{w(s)}}\,\mathrm{d}s \tag{B.17}$$

is obtained. The scaling parameter c_w can then be calculated by using integration by substitution (see App. A.3) from

$$\begin{aligned} c_\mathrm{w} &= \frac{1}{2}\int_{-\infty}^{\infty}\frac{w(s(x_\mathrm{n}))}{\ell}+\ell\left(\frac{\mathrm{d}s}{\mathrm{d}x_\mathrm{n}}\right)^2\mathrm{d}x_\mathrm{n} = \int_0^{\infty}\frac{w(s(x_\mathrm{n}))}{\ell}+\ell\left(\frac{\mathrm{d}s}{\mathrm{d}x_\mathrm{n}}\right)^2\mathrm{d}x_\mathrm{n} \\ &\overset{(B.15)}{=} \int_0^{\infty}\frac{w(s(x_\mathrm{n}))}{\ell}+\frac{w(s(x_\mathrm{n}))}{\ell}\,\mathrm{d}x_\mathrm{n} = 2\int_0^{\infty}\frac{w(s(x_\mathrm{n}))}{\ell}\,\mathrm{d}x_\mathrm{n} \\ &\overset{(B.17)}{=} 2\int_{s(x=0)}^{s(x=\infty)}\frac{w(s)}{\ell}\,\frac{\ell}{\sqrt{w(s)}}\,\mathrm{d}s = 2\int_0^1\sqrt{w(s)}\,\mathrm{d}s \quad . \end{aligned} \tag{B.18}$$

Application on quadratic surface function

The evaluation of (B.18) for the quadratic surface function (3.57) yields

$$c_\mathrm{w} = 2\int_0^1 \underbrace{\sqrt{\beta\,s^2-(\beta+1)\,s+1}}_{=:\,f_\beta}\,\mathrm{d}s \tag{B.19}$$

For the case $\beta = 1$ one obtains

$$\begin{aligned} c_\mathrm{w} &= 2\int_0^1 \sqrt{s^2 - 2s + 1}\,\mathrm{d}s = 2\int_0^1 \sqrt{(1-s)^2}\,\mathrm{d}s = 2\int_0^1 1 - s\,\mathrm{d}s \\ &= 2\left[s - \frac{1}{2}s^2\right]_0^1 = 2\left(\frac{1}{2} - 0\right) = 1 \quad , \end{aligned} \tag{B.20}$$

while for $\beta = 0$ it is

$$\begin{aligned} c_\mathrm{w} &= 2\int_0^1 \sqrt{-s + 1}\,\mathrm{d}s = 2\int_0^1 (1-s)^{1/2}\,\mathrm{d}s \\ &= 2\left[-\frac{2}{3}(1-s)^{3/2}\right]_0^1 = 2\left(0 + \frac{2}{3}\right) = \frac{4}{3} \quad . \end{aligned} \tag{B.21}$$

According to, e.g., Bronštejn (2015) the undetermined integral of f_β is

$$\begin{aligned} I_\beta(s) = {} & \frac{2\beta s - (\beta+1)}{4\beta}\sqrt{\beta s^2 - (\beta+1)\,s + 1} \\ & - \frac{(\beta-1)^2}{8\beta}\begin{cases} \dfrac{-1}{\sqrt{-\beta}}\,\arcsin\left(\dfrac{2\beta s - (\beta+1)}{\sqrt{(\beta-1)^2}}\right) & \forall\beta \in [-1,\,0) \\ \dfrac{1}{\sqrt{\beta}}\,\operatorname{arcosh}\left(\dfrac{|2\beta\, s - (\beta+1)|}{\sqrt{(\beta-1)^2}}\right) & \forall\beta \in (0,\,1) \quad . \end{cases} \end{aligned} \tag{B.22}$$

Finally, it is for $\beta \in [-1,\,0)$

$$\begin{aligned} c_\mathrm{w} &= 2I_\beta(1) - 2I_\beta(0) \\ &= \frac{(\beta-1)^2}{4\beta}\,\frac{1}{\sqrt{-\beta}}\,\underbrace{\arcsin\left(\frac{\beta-1}{1-\beta}\right)}_{=-\pi/2} + \frac{\beta+1}{2\beta} - \frac{(\beta-1)^2}{4\beta}\,\frac{1}{\sqrt{-\beta}}\,\arcsin\left(\frac{-(\beta+1)}{1-\beta}\right) \\ &= \frac{1}{4\beta}\left[\frac{(1-\beta)^2}{\sqrt{-\beta}}\left(\arcsin\left(\frac{1+\beta}{1-\beta}\right) - \frac{\pi}{2}\right) + 2(1+\beta)\right] \end{aligned} \tag{B.23}$$

and for $\beta \in (0,\ 1)$

$$\begin{aligned} c_\mathrm{w} &= 2I_\beta(1) - 2I_\beta(0) \\ &= \frac{(\beta-1)^2}{4\beta}\,\frac{1}{\sqrt{\beta}}\,\underbrace{\operatorname{arcosh}\left(\frac{|\beta-1|}{1-\beta}\right)}_{=0} + \frac{\beta+1}{2\beta} - \frac{(\beta-1)^2}{4\beta}\,\frac{1}{\sqrt{\beta}}\,\operatorname{arcosh}\left(\frac{|-(\beta+1)|}{1-\beta}\right) \\ &= \frac{1}{4\beta}\left[-\frac{(1-\beta)^2}{\sqrt{\beta}}\,\operatorname{arcosh}\left(\frac{1+\beta}{1-\beta}\right) + 2(1+\beta)\right] \quad . \end{aligned} \tag{B.24}$$

Note, that the latter result can also be represented in the form

$$c_\mathrm{w} = \frac{1}{4\beta}\left[\frac{(1-\beta)^2}{\sqrt{\beta}}\,\ln\left(\frac{1-\beta}{1+2\sqrt{\beta}+\beta}\right) + 2(1+\beta)\right] \tag{B.25}$$

as given by Wu (2017).

B.4 Incorporation of lower and upper bounds

In order to construct a function $f(x)$ which approaches another function $\tilde{f}(x)$ within a specific range, e.g. below a specific value $x < x_\mathrm{upp}$, and diverges above this value (for $x > x_\mathrm{upp}$) superimposition can be used to form $f(x) = \tilde{f}(x) + p_\mathrm{x}$. Based on the bracket operator as introduced in App. A.9 the superimposed function might be

$$\hat{p}_\mathrm{upp}(x) = \epsilon_\mathrm{x}\,\langle x - x_\mathrm{upp}\rangle \tag{B.26}$$

with the choice of a sufficiently large penalty parameter $\epsilon_\mathrm{x} \gg 0$. Note, that (B.26) is not smooth, and thus, not differentiable at x_upp. The same applies for the superimposition with $\hat{p}_\mathrm{low}(x) = \epsilon_\mathrm{x}\,\langle x_\mathrm{low} - x\rangle$, so that $f(x)$ diverges for $x < x_\mathrm{low}$. As an alternative, the smooth functions

$$p_\mathrm{low}(x) = a\,\exp\left(\epsilon_\mathrm{x}\,(x_\mathrm{low} - x)\right) + c \tag{B.27}$$

and

$$p_\mathrm{upp}(x) = a\,\exp\left(\epsilon_\mathrm{x}\,(x - x_\mathrm{upp})\right) + c \tag{B.28}$$

inspired by the work of Kort and Bertsekas (1972) can be used to introduce diverging terms below x_low and above x_upp, respectively. $\epsilon_\mathrm{x} \gg 0$ acts as penalty parameter, in fact it stretches the function to make it much steeper on one side and more flat on the other side of the limit value. Since the exponential function has a non-zero function value and slope at $x = 0$, the coefficients are used to adjust the correct function value and slope at x_low and x_upp, respectively.

In terms of the phase field approach, above considerations might be used with $s_\mathrm{low} = 0$ and $s_\mathrm{upp} = 1$ to keep the phase field s in the desired range $s \in [0, 1]$.

Application on surface functions

The upper bound of phase field values is ensured by the condition $w'(s{=}1) = 0$. If a surface function is used which does not possess this property, e.g. the quadratic function $w_{\beta\neq1}(s) = (1-s)$, application of (B.28) and incorporation of (3.54) and (3.56) yields the surface function

$$\begin{aligned} w_\epsilon(s) &:= w_\beta(s) + \underbrace{a\, \exp\left(\epsilon_\mathrm{w}\,(s-1)\right) + c}_{=:\, p_\mathrm{w}(s)} \\ &= \beta\, s^2 - (\beta+1)\, s + 1 + (1-\beta)\left[\, \exp\left(\epsilon_\mathrm{w}\,(s-1)\right) - 1\right]/\epsilon_\mathrm{w} \end{aligned} \tag{B.29}$$

with its first and second derivatives

$$w'_\epsilon(s) = 2\beta\, s - (\beta+1) + (1-\beta)\, \exp\left(\epsilon_\mathrm{w}\,(s-1)\right) \tag{B.30}$$

$$w''_\epsilon(s) = 2\beta + (1-\beta)\, \epsilon_\mathrm{w}\, \exp\left(\epsilon_\mathrm{w}\,(s-1)\right) \quad . \tag{B.31}$$

Table B.1 lists the scaling parameter c_w for different choices of β. In addition, it shows the influence of the penalty parameter on the scaling parameter. As discussed in Sect. 3.3.1, c_w scales the energetic contribution of the regularized crack. Thus, the deviations shown in Tab. B.1 provide information about the accuracy of crack surface energy approximation.

β	exact c_w	$\epsilon_\mathrm{w} = 10^{1.0}$	$\epsilon_\mathrm{w} = 10^{1.5}$	$\epsilon_\mathrm{w} = 10^{2.0}$	$\epsilon_\mathrm{w} = 10^{2.5}$
−1.0	$\pi/2 \approx 1.5708$	1.4612	1.5294	1.5566	1.5661
−0.5	1.4584	1.3484	1.4177	1.4446	1.4539
0.0	$4/3 \approx 1.3333$	1.2365	1.2974	1.3211	1.3293
0.5	1.1884	1.1226	1.1632	1.1797	1.1855
1.0	1.0000	1.0000			

Table B.1: Scaling parameter c_w for different choices of β and influence of the penalization to keep $s \leq 1$.

Application on degradation functions

The lower bound of phase field values is ensured by the condition $g'(s{=}0) = 0$. If a degradation function is used which does not possess this property, e.g. the linear function $\tilde{g}(s) = s$, application of (B.27) and incorporation of (3.64) and (3.68) yields the degradation function

$$g(s) = b\, \tilde{g}(\ell) + a\, \exp\left(-\epsilon_\mathrm{g}\, s\right) + c = \frac{s + \left[\exp(-\epsilon_\mathrm{g}\, s) - 1\right]/\epsilon_\mathrm{g}}{1 + \left[\exp(-\epsilon_\mathrm{g}) - 1\right]/\epsilon_\mathrm{g}} \quad . \tag{B.32}$$

It corresponds to the function introduced by Steinke and Kaliske (2019). If, in addition, also (3.69) should be fulfilled, both functions (B.27) and (B.28) are superimposed on, e.g., the linear function $\tilde{g}(s) = s$ from above. In this case, the result is obtained from

$$g(s) = b\,\tilde{g}(s) + a_1\,\exp\big(-\epsilon_\mathrm{g}\,s\big) + a_2\,\exp\big(\epsilon_\mathrm{g}\,(s-1)\big) + c \quad , \tag{B.33}$$

by incorporating (3.64), (3.68) and (3.68). For $a_\mathrm{s} = 0$ the result is

$$g(s) = \frac{s\,\big(1+\exp(-\epsilon_\mathrm{g})\big) + \big[\,\exp(-\epsilon_\mathrm{g}\,s) - \exp\big(\epsilon_\mathrm{g}\,(s-1)\big) + \exp(-\epsilon_\mathrm{g}) - 1\big]/\epsilon_\mathrm{g}}{1+\exp(-\epsilon_\mathrm{g}) + 2\big[\,\exp(-\epsilon_\mathrm{g}) - 1\big]/\epsilon_\mathrm{g}} \tag{B.34}$$

Interestingly, for the choice $\epsilon_\mathrm{g} = 1$, the degradation function

$$g(s) = \frac{s}{3}\,\big[\,\exp(-s) - \exp(s-1) + \exp(-1) - 1\big] \approx 3s^2 - 2s^3 \tag{B.35}$$

approaches the cubic degradation function (3.71), as can be shown by using the series expansion of $\exp(x) = \sum_{n=0}^{\infty} x^n/(n!)$.

B.5 Generalized polynomial degradation function

The construction of a polynomial degradation function

$$g_\mathrm{n}(s) := \sum_{k=1}^{n} a_k\,s^k = a_0 + a_1\,s + a_2\,s^2 + a_3\,s^3 + \cdots + a_n\,s^n \tag{B.36}$$

of order n offers a lot of freedom. While, at least, a polynomial of third order is required to incorporate all requirements introduced in Sect. 3.3.2, that are

$$g(s{=}0) = 0\,, \quad g(s{=}1) = 1\,, \quad g'(s{=}0) = 0\,, \quad g'(s{=}1) = a_\mathrm{s} \gtrless 0 \quad , \tag{B.37}$$

note the additional requirement $g''(s{=}1) \neq 0$ found in Sect. 3.4.2, polynomials of higher order seem to be under-determined. However, since in addition to these requirements also

$$g'(s) > 0 \quad \forall s \in (0,\,1) \tag{B.38}$$

is essential (also to keep $g(s) \in [0,\,1]$), by far not all combinations of coefficients are possible.

Degradation functions of fourth order

The incorporation of the requirements (B.37) yields a general fourth order degradation function

$$g_4(s) = a_4\,s^4 + (a_\mathrm{s} - 2a_4 - 2)\,s^3 + (a_4 - a_\mathrm{s} + 3)\,s^2 \quad . \tag{B.39}$$

Figure B.1 shows the function for different choices of the parameter a_4. For simplicity, $a_s = 0$ is assumed. Obviously, the curve for $a_4 = -5$ falls below $s=0$ in $s \in (0, \approx 0.31)$ and the curve for $a_4 = 5$ exceeds $s=1$ in $s \in (\approx 0.69, 1)$. This goes along with the function being concave $g''(s=0) < 0$ and convex $g''(s=1) > 0$, respectively. As a consequence, both functions have two inflection points in the range of interest $s \in (0, 1)$ in contrast to the other parameter choices. By allowing just a single inflection point for $s \in (0, 1)$ the valid parameter range is found to be $a_4 \in [a_s - 3, 3 - 2a_s)$. The value $a_4 = 3 - 2a_s$ is excluded to avoid $g''(s=1) = 0$ for the reason discussed in Sect. 3.4.2.

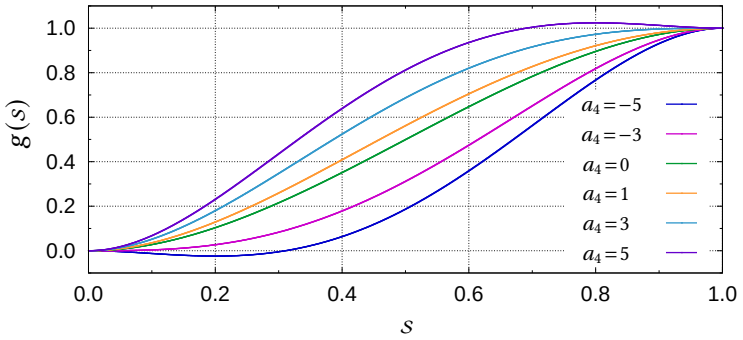


Figure B.1: Fourth order degradation functions $g_4(s)$ (with $a_s = 0$) shown for different choices of the parameter $a_4 = \{-5, -3, 0, 1, 3, 5\}$. Note, that the first and last choice are invalid since $s<0$ and $s>0$, respectively.

Degradation functions of higher order — a systematic approach

While for degradation functions of fourth and maybe fifth order the allowed coefficient ranges might be estimated, the question arises, which criteria can be used additionally to determine all parameters uniquely. Especially by regarding the inequality constraint (B.38), this does not seem to be a simple task. The most promising approach is to control the $n-2$ inflection points $\tilde{s}_i$ of the degradation function as already discussed above. More precisely, all but the single inflection point $\tilde{s}_1$, which is naturally given by the requirements (B.37) and located at $\tilde{s}_1 \in (0, 1)$, are prescribed outside of this range, i.e. at

$$\left(\tilde{s}_i \in (-\infty, 0]\right) \vee \left(\tilde{s}_i \in (1, \infty)\right) \tag{B.40}$$

with $1 < i \leq n-2$. Note, that a possible inflection point at $\tilde{s}_i = 1$ is excluded to avoid $g''(s=1) = 0$.. For this purpose, the second derivative of the degradation function is factorized to prescribe $g''(\tilde{s}_i) \stackrel{!}{=} 0$ according to

$$g''(s) = (s - \tilde{s}_1)\,(s - \tilde{s}_2)\,\ldots\,(s - \tilde{s}_{n-2})\,f_g(s, a_s) \quad , \tag{B.41}$$

while $f_g(s, a_s)$ is determined by incorporating (B.37). Note, that multiple (m-fold) inflection points can be introduced by prescribing the second to m+1-th derivative at this specific function value, i.e.

$$g''(\tilde{s}_i) = \cdots = g^{(m+1)}(\tilde{s}_i) \stackrel{!}{=} 0 \quad . \tag{B.42}$$

From a purely mathematical perspective, this approach does not provide the most general solution which satisfies the constraints (B.37–B.38). However, only oscillating solutions are excluded by this approach, which seems to be reasonable for physical reasons.

Based on this approach, specific families of degradation functions are highlighted. All are based on the assumption, that all $m := n - 2$ additional inflection points $i = 2 \dots m + 1$ are located at the same function value $\tilde{s}_m$. In this specific case, the degradation function is constructed from

$$g(s) = \underbrace{\left[\frac{(s - \tilde{s}_m)^{m+2}}{(m+1)\,(m+2)} + C_1\, s + C_2\right]}_{h_1(s)} \underbrace{(A + B\, s)}_{h_2(s)} \tag{B.43}$$

with its first derivative

$$\begin{aligned} g'(s) &= \underbrace{\left[\frac{(s - \tilde{s}_m)^{m+1}}{(m+1)} + C_1\right]}_{h_1'(s)} \underbrace{(A + B\, s)}_{h_2(s)} + \underbrace{\left[\frac{(s - \tilde{s}_m)^{m+2}}{(m+1)\,(m+2)} + C_1\, s + C_2\right]}_{h_1(s)} \underbrace{B}_{h_2'(s)} \\ &= \frac{(s - \tilde{s}_m)^{m+1}}{(m+1)} \left(A + B\, s + \frac{(s - \tilde{s}_m)}{(m+2)}\, B\right) + C_1\, (A + B\, s) + (C_1 + C_2)\, B \end{aligned} \tag{B.44}$$

and second derivative

$$\begin{aligned} g''(s) &= \underbrace{(s - \tilde{s}_m)^m}_{h_1''(s)} \underbrace{(A + B\, s)}_{h_2(s)} + 2 \underbrace{\left[\frac{(s - \tilde{s}_m)^{m+1}}{(m+1)} + C_1\right]}_{h_1'(s)} \underbrace{B}_{h_2'(s)} \\ &= (s - \tilde{s}_m)^m \left((A + B\, s) + 2\frac{(s - \tilde{s}_m)^m}{(m+1)}\, B\right) + 2 C_1\, B \quad . \end{aligned} \tag{B.45}$$

The coefficients A, B, C_1 and C_2 are then obtained from the requirements (B.37). They are

$$\begin{aligned} A &= 0 \quad , \\ B &= \frac{(a_\mathrm{s}-2)\,\left(m^2+3\,m+2\right)}{\left[-\tilde{s}_\mathrm{m}^2 - m\,\tilde{s}_\mathrm{m} + m + 1\right]\,(1-\tilde{s}_\mathrm{m})^m + (-1)^m\,\tilde{s}_\mathrm{m}^{m+2}} \quad , \\ C_1 &= \frac{F(s)\,(1-\tilde{s}_\mathrm{m})^m - (-1)^m\,\tilde{s}_\mathrm{m}^m\,\tilde{s}_\mathrm{m}^2 + (-1)^m\,a_\mathrm{s}\,\tilde{s}_\mathrm{m}^{m+2}}{(a_\mathrm{s}-2)\,\left(m^2+3\,m+2\right)} \quad , \\ C_2 &= \frac{(-1)^{m+1}\,\tilde{s}_\mathrm{m}^{m+2}}{m^2+3\,m+2} \end{aligned} \tag{B.46}$$

with

$$F(s) = (1-a_\mathrm{s})\,\tilde{s}_\mathrm{m}^2 + (2\,a_\mathrm{s}-4-m)\,\tilde{s}_\mathrm{m} - a_\mathrm{s} + m + 3 \quad . \tag{B.47}$$

For $a_\mathrm{s} = 0$, the parameters

$$\begin{aligned} B &= \frac{2m^2+6\,m+4}{\left[\tilde{s}_\mathrm{m}^2 + m\,\tilde{s}_\mathrm{m} - m - 1\right]\,(1-\tilde{s}_\mathrm{m})^m - (-1)^m\,\tilde{s}_\mathrm{m}^{m+2}} \quad , \\ C_1 &= \frac{\left[-\tilde{s}_\mathrm{m}^2 + (4+m)\,\tilde{s}_\mathrm{m} - m - 3\right]\,(1-\tilde{s}_\mathrm{m})^m + (-1)^m\,\tilde{s}_\mathrm{m}^{m+2}}{2m^2+6m+4} \end{aligned} \tag{B.48}$$

can be simplified.

The following functions seem to be promising:

- For $\tilde{s}_m = 0$ the higher-order polynomial degradation function

  $$g(s) = (a_\mathrm{s}-n+1)\,s^n + (n-a_\mathrm{s})\,s^{n-1} \tag{B.49}$$

  as introduced by Strobl and Seelig (2018) is obtained. With $a_\mathrm{s} = 0$ it can be simplified to

  $$g(s) = n\,s^{n-1} - (n-1)\,s^n = s^{n-1}\,[n-(n-1)\,s] \tag{B.50}$$

  with its first and second derivatives

  $$g'(s) = n\,(n-1)\,s^{n-2} - n\,(n-1)\,s^{n-1} = s^{n-2}\,n\,(n-1)\,(1-s) \tag{B.51}$$

  $$\begin{aligned} g''(s) &= n\,(n-1)\,(n-2)\,s^{n-3} - n\,(n-1)^2\,s^{n-2} \\ &= s^{n-3}\,n\,(n-1)\,[(n-2)-(n-1)\,s)] \quad . \end{aligned} \tag{B.52}$$

- For both limit values $\tilde{s}_m \to \pm\infty$ the polynomial degradation function approaches the cubic one, i.e.

  $$g(s) = (a_\mathrm{s}-2)\,s^3 + (3-a_\mathrm{s})\,s^2 \quad . \tag{B.53}$$

- For $m = 1$ and $\tilde{s}_\mathrm{m} = 1 + \sqrt{3}/3$ one obtains with $a_\mathrm{s} = 0$ the 4th-order degradation function
$$g(s) = s^4 - 4s^3 + 4s^2 \quad . \tag{B.54}$$
In combination with the single-well surface function (3.58) the maximum stress is obtained (in contrast to other degradation functions) at $s = 1$, see the discussion in Sect. 3.4.2.

B.6 Determination of the phase field profile

To obtain the phase field solution $s(x)$ for the quadratic single-well function in absence of mechanical loads, the inhomogeneous, linear second-order ordinary differential equation
$$s(x) - \ell^2\, s''(x) = 1 \tag{B.55}$$
must be solved. The general solution of the homogeneous equation is found by using the ansatz
$$s(x) = C\, \exp(\lambda\, x) \tag{B.56}$$
with
$$s'(x) = C\, \lambda\, \exp(\lambda\, x) \qquad \text{and} \qquad s''(x) = C\, \lambda^2\, \exp(\lambda\, x) \quad . \tag{B.57}$$
Inserting (B.56) in (B.55) yields
$$C\, \exp(\lambda\, x)\, (1 - \ell^2\, \lambda^2) = 0 \quad , \tag{B.58}$$
with its characteristic equation
$$1 - \ell^2\, \lambda^2 = 0 \qquad \Leftrightarrow \qquad \lambda = \pm\frac{1}{\ell} \quad . \tag{B.59}$$
The homogeneous solution is
$$s_\mathrm{h}(x) = C_1\, e^{x/\ell} + C_2\, e^{-x/\ell} \quad . \tag{B.60}$$
Due to the constant on the right hand side of (B.55) the particular solution is $s_\mathrm{p}(x) = 1$ and the general solution is obtained from
$$s(x) = s_\mathrm{h}(x) + s_\mathrm{p}(x) = C_1\, e^{x/\ell} + C_2\, e^{-x/\ell} + 1 \quad . \tag{B.61}$$
The specific solution for a given constraint s_0 at $x = 0$ in a finite domain with $x = \pm L$ can be obtained by applying the phase field boundary conditions $s(x{=}0) = s_0$ and $s'(x{=}\pm L) = 0$ on (B.61). By using the definitions of the hyperbolic sine, hyperbolic cosine and hyperbolic tangent
$$\begin{aligned}
\sinh(x) &:= \frac{e^x - e^{-x}}{2} \quad , \\
\cosh(x) &:= \frac{e^x + e^{-x}}{2} \quad , \\
\tanh(x) &:= \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad ,
\end{aligned} \tag{B.62}$$

one obtains the parameters (here only shown for $x \geq 0$)

$$\begin{aligned} C_1 &= \frac{(s_0-1)\,e^{-L/\ell}}{e^{L/\ell}+e^{-L/\ell}} = \frac{1}{2}(1-s_0)\frac{e^{L/\ell}-e^{-L/\ell}}{e^{L/\ell}+e^{-L/\ell}} - \frac{1}{2} + \frac{s_0}{2} \\ &= \frac{1}{2}(1-s_0)\,\tanh\left(\frac{L}{\ell}\right) - \frac{1}{2} + \frac{s_0}{2} \quad , \\ C_2 &= \frac{(s_0-1)\,e^{L/\ell}}{e^{L/\ell}+e^{-L/\ell}} = \frac{1}{2}(s_0-1)\frac{e^{L/\ell}+e^{-L/\ell}}{e^{L/\ell}+e^{-L/\ell}} - \frac{1}{2} + \frac{s_0}{2} \\ &= -\frac{1}{2}(1-s_0)\,\tanh\left(\frac{L}{\ell}\right) - \frac{1}{2} + \frac{s_0}{2} \quad . \end{aligned} \tag{B.63}$$

The specific solution finally is

$$\begin{aligned} s(x) &= 1 + \left(\frac{s_0}{2} - \frac{1}{2}\right)\left(e^{x/\ell} + e^{-x/\ell}\right) + \left(C_1 + \frac{1}{2} - \frac{s_0}{2}\right)\,e^{x/\ell} - \left(C_2 + \frac{1}{2} - \frac{s_0}{2}\right)\,(-e^{-x/\ell}) \\ &= 1 + (s_0-1)\,\cosh\left(\frac{x}{\ell}\right) + (1-s_0)\,\tanh\left(\frac{L}{\ell}\right)\,\sinh\left(\frac{x}{\ell}\right) \quad . \end{aligned} \tag{B.64}$$

In case of a crack at $x = 0$ it is $s_0 = 0$, thus the solution is

$$s(x) = 1 - \cosh\left(\frac{x}{\ell}\right) + \tanh\left(\frac{L}{\ell}\right)\,\sinh\left(\frac{x}{\ell}\right) \quad . \tag{B.65}$$

For the case of most interest $\ell \ll L$, it is

$$\lim_{\ell/L \to 0} \tanh\left(\frac{L}{\ell}\right) = 1 \tag{B.66}$$

and thus (B.65) approaches the well-known solution

$$\begin{aligned} s(x) &= 1 - \cosh\left(\frac{|x|}{\ell}\right) + \sinh\left(\frac{|x|}{\ell}\right) \\ &= 1 - \cosh\left(-\frac{|x|}{\ell}\right) - \sinh\left(-\frac{|x|}{\ell}\right) = 1 - \exp\left(-\frac{|x|}{\ell}\right) \quad . \end{aligned} \tag{B.67}$$

Note, that somewhat different results have been reported by Kuhn (2013). By applying the phase field boundary condition for the crack $s(x=0) = 0$ and constraining the phase field at the boundary $s(x=\pm L) = 1$, the latter appears to be questionable, one obtains the parameters (here only shown for $x \geq 0$)

$$\begin{aligned} C_1 &= \frac{e^{-L/\ell}}{e^{L/\ell}-e^{-L/\ell}} = \frac{1}{2}\left(\frac{e^{L/\ell}+e^{-L/\ell}}{e^{L/\ell}-e^{-L/\ell}} - 1\right) = \frac{1}{2}\left(\coth\left(\frac{L}{\ell}\right) - 1\right) \quad , \\ C_2 &= -\frac{e^{L/\ell}}{e^{L/\ell}-e^{-L/\ell}} = \frac{1}{2}\left(-\frac{e^{L/\ell}+e^{-L/\ell}}{e^{L/\ell}-e^{-L/\ell}} - 1\right) = \frac{1}{2}\left(-\coth\left(\frac{L}{\ell}\right) - 1\right) \quad . \end{aligned} \tag{B.68}$$

The solution is then

$$
\begin{aligned}
s(x) &= 1 - \frac{1}{2}\left(e^{x/\ell} + e^{-x/\ell}\right) + \left(C_1 + \frac{1}{2}\right) e^{x/\ell} - \left(C_2 + \frac{1}{2}\right)\left(-e^{-x/\ell}\right) \\
&= 1 - \cosh\left(\frac{x}{\ell}\right) + \coth\left(\frac{L}{\ell}\right) \sinh\left(\frac{x}{\ell}\right) \quad , \qquad \text{(B.69)}
\end{aligned}
$$

which corresponds to the result obtained by Kuhn (2013).

B.7 Additional analysis of inhomogeneous phase field evolution

As a continuation of the analysis provided in Sect. 3.4.3, additional results are shown here. Figure B.2a shows $V_\sigma(s)$ for different stress levels for the combination of the quadratic single-well function with the quadratic degradation function. The root representing the phase field value of the homogeneous solution and the root representing the minimum value of the inhomogeneous solution are marked by upward and downward pointing triangles, respectively. Figure B.2b shows the phase field profiles for each of the load levels, while Fig. B.3 shows the maximum increase of the phase field value for each location during the localization process.

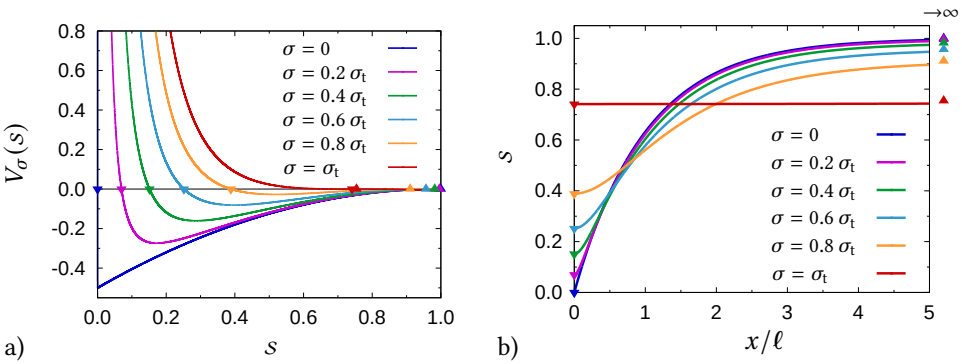


Figure B.2: Evaluation of the inhomogeneous problem for different load levels $\sigma \in [0,\ \sigma_t]$ for the single-well surface function in conjunction with the quadratic degradation function, a) evaluation of $V_\sigma(s)$ with roots indicated, b) inhomogeneous phase field solution $s(x)$.

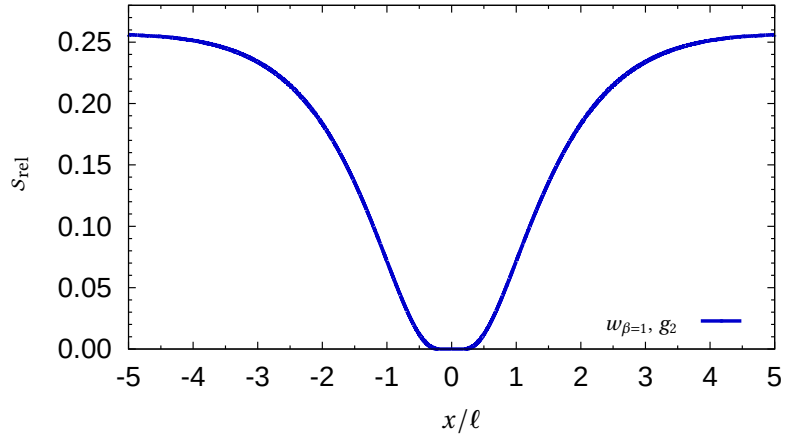


Figure B.3: Maximum release of phase field s during evolution from the unloaded intact state $\sigma = 0$, $s = 1$ via the maximum load $\sigma = \sigma_{\mathrm{t}}$ to the broken state $\sigma = 0$, $s(x_0) = 0$.

B.8 Details on the strain energy decomposition

In some cases, the strain energy decomposition (3.188) presupposing the split of the strain tensor $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^+ + \boldsymbol{\varepsilon}^-$ can be formulated according to

$$\begin{aligned}\psi_0(\boldsymbol{\varepsilon}) &= \frac{1}{2}\,(\boldsymbol{\varepsilon}^+ + \boldsymbol{\varepsilon}^-) : \mathbb{C}_0 : (\boldsymbol{\varepsilon}^+ + \boldsymbol{\varepsilon}^-) \\ &= \frac{1}{2}\Big[\boldsymbol{\varepsilon}^+ : \underbrace{\mathbb{C}_0 : \boldsymbol{\varepsilon}^+}_{=:\,\boldsymbol{\sigma}_0^+} + \boldsymbol{\varepsilon}^- : \underbrace{\mathbb{C}_0 : \boldsymbol{\varepsilon}^-}_{=:\,\boldsymbol{\sigma}_0^-} + \boldsymbol{\varepsilon}^+ : \underbrace{\mathbb{C}_0 : \boldsymbol{\varepsilon}^-}_{=:\,\boldsymbol{\sigma}_0^-} + \boldsymbol{\varepsilon}^- : \underbrace{\mathbb{C}_0 : \boldsymbol{\varepsilon}^+}_{=:\,\boldsymbol{\sigma}_0^+}\Big] \quad . \end{aligned} \tag{B.70}$$

While $\boldsymbol{\varepsilon}^- : \boldsymbol{\varepsilon}^- = 0$ is guaranteed because of (A.111), it is

$$\psi^{\pm} := \boldsymbol{\varepsilon}^+ : \boldsymbol{\sigma}_0^- = \boldsymbol{\varepsilon}^- : \boldsymbol{\sigma}_0^+ \neq 0 \tag{B.71}$$

in general. Moreover, if strain and related stress components in (B.71) have different signs, their product is negative and the related energy portion $\psi^{\pm}$ might be negative. To be able to degrade the energetic portions in (B.70) separately, all summands must be non-negative (3.189). Thus, the representation (B.70) does hold if

$$\boldsymbol{\varepsilon}^+ : \boldsymbol{\sigma}_0^- = \boldsymbol{\varepsilon}^- : \boldsymbol{\sigma}_0^+ = 0 \quad . \tag{B.72}$$

The volumetric-deviatoric decomposition discussed in Sect. 3.5.2, for example, does fulfill (B.72) since $\boldsymbol{\varepsilon}_{\mathrm{vol}} : \boldsymbol{\sigma}_{\mathrm{dev}} = 0$.

By contrast, the specific strain energy decomposition (3.229) based on a spectral strain split with (2.190) does not fulfill (B.72) for arbitrary strains. In particular, individual strain

energy portions can exceed the total strain energy. For $\nu > 0$, this is the case if not all principal strains have the same sign as demonstrated by the following example. With the principal strains $\varepsilon_1 = c_\varepsilon > 0$, $\varepsilon_2 = 0$, $\varepsilon_3 = -c_\varepsilon$, the positive strain tensor $\boldsymbol{\varepsilon}^+$ and negative strain tensor $\boldsymbol{\varepsilon}^-$ are (given according to principal directions $\boldsymbol{n}_i$)

$$[\boldsymbol{\varepsilon}^+]_{\{n_i\}} = \begin{bmatrix} c_\varepsilon & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad [\boldsymbol{\varepsilon}^-]_{\{n_i\}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -c_\varepsilon \end{bmatrix} \quad . \tag{B.73}$$

The corresponding (undegraded) positive stress tensor and negative stress tensor are

$$[\boldsymbol{\sigma}_0^+]_{\{n_i\}} := \begin{bmatrix} (\lambda + 2\mu)\, c_\varepsilon & 0 & 0 \\ 0 & \lambda\, c_\varepsilon & 0 \\ 0 & 0 & \lambda\, c_\varepsilon \end{bmatrix}, \quad [\boldsymbol{\sigma}_0^-]_{\{n_i\}} := \begin{bmatrix} -\lambda\, c_\varepsilon & 0 & 0 \\ 0 & -\lambda\, c_\varepsilon & 0 \\ 0 & 0 & -(\lambda + 2\mu)\, c_\varepsilon \end{bmatrix} . \tag{B.74}$$

Strain energy decomposition according to (B.70) results in

$$\psi_0 = \underbrace{\frac{1}{2}\boldsymbol{\varepsilon}^+ : \boldsymbol{\sigma}_0^+}_{=:\,\psi^+} + \underbrace{\frac{1}{2}\boldsymbol{\varepsilon}^- : \boldsymbol{\sigma}_0^-}_{=:\,\psi^-} + \underbrace{\frac{1}{2}\boldsymbol{\varepsilon}^+ : \boldsymbol{\sigma}_0^- + \frac{1}{2}\boldsymbol{\varepsilon}^- : \boldsymbol{\sigma}_0^+}_{=:\,\psi^\pm = -\lambda\, c_\varepsilon^2 \le 0} = 2\mu\, c_\varepsilon^2 \quad , \tag{B.75}$$

with a negative energy portion $\psi^\pm$ (not related to tension or compression) and separate energy portions ψ^+ and ψ^- being (in sum, possibly also separately) larger than the total energy, i.e. $\psi^+ = \psi^- = (\lambda/2 + \mu)\, c_\varepsilon^2$, which is unphysical. Thus, (3.229) is no valid strain energy decomposition.

B.9 Crack boundary conditions rewritten in effective stresses

Incorporation of *crack boundary conditions* as proposed by Strobl and Seelig (2015) and discussed in Sect. 3.6 can also be carried out in the effective stress space. The effective stress tensor (2.173) is decomposed according to

$$\boldsymbol{\sigma} = g(s)\, \tilde{\boldsymbol{\sigma}}^+ + \tilde{\boldsymbol{\sigma}}^- \tag{B.76}$$

with portions related to crack opening and closure, where only the first one is degraded by the degradation function $g(s)$. By using the definition of the effective stress tensor (3.13) the expressions of the *crack orientation dependent degradation* derived in (3.279–3.284) can be rewritten (in a coordinate system aligned with the crack as shown in Fig. 3.26). Crack opening is still detected by

$$\varepsilon_{\mathrm{nn}} - \tilde{\varepsilon} = \varepsilon_{\mathrm{nn}} - \frac{-\lambda}{\lambda + 2\mu}\,(\varepsilon_{\mathrm{pp}} + \varepsilon_{\mathrm{qq}}) \ge 0 \tag{B.77}$$

which can be rewritten as

$$\underbrace{(\lambda + 2\mu)\, \varepsilon_{\mathrm{nn}} + \lambda\,(\varepsilon_{\mathrm{pp}} + \varepsilon_{\mathrm{qq}})}_{=\,\tilde{\sigma}_{\mathrm{nn}}} \ge 0 \quad . \tag{B.78}$$

For an open crack ($\tilde{\sigma}_{nn} \geq 0$) normal stress components given in (3.279–3.281) are

$$\begin{aligned}
\sigma_{nn}^{open} &= g(s) \underbrace{\left[(\lambda + 2\mu)\, \varepsilon_{nn} + \lambda\, (\varepsilon_{pp} + \varepsilon_{qq})\right]}_{= \tilde{\sigma}_{nn}} \\
\sigma_{pp}^{open} &= g(s)\, \lambda\, \varepsilon_{nn} + \frac{g(s)\, \lambda^2 + 4\lambda\, \mu + 4\mu^2}{\lambda + 2\mu}\, \varepsilon_{pp} + \frac{g(s)\, \lambda^2 + 2\lambda\, \mu}{\lambda + 2\mu}\, \varepsilon_{qq} \\
&= \underbrace{\left[(\lambda + 2\mu)\, \varepsilon_{pp} + \lambda\, (\varepsilon_{nn} + \varepsilon_{qq})\right]}_{= \tilde{\sigma}_{pp}} + \big(g(s) - 1\big)\, \underbrace{\left[(\lambda + 2\mu)\, \varepsilon_{nn} + \lambda\, (\varepsilon_{pp} + \varepsilon_{qq})\right]}_{= \tilde{\sigma}_{nn}}\, \frac{\lambda}{\lambda + 2\mu} \\
\sigma_{qq}^{open} &= g(s)\, \lambda\, \varepsilon_{nn} + \frac{g(s)\, \lambda^2 + 2\lambda\, \mu}{\lambda + 2\mu}\, \varepsilon_{pp} + \frac{g(s)\, \lambda^2 + 4\lambda\, \mu + 4\mu^2}{\lambda + 2\mu}\, \varepsilon_{qq} \\
&= \underbrace{\left[(\lambda + 2\mu)\, \varepsilon_{qq} + \lambda\, (\varepsilon_{nn} + \varepsilon_{pp})\right]}_{= \tilde{\sigma}_{qq}} + \big(g(s) - 1\big)\, \underbrace{\left[(\lambda + 2\mu)\, \varepsilon_{nn} + \lambda\, (\varepsilon_{pp} + \varepsilon_{qq})\right]}_{= \tilde{\sigma}_{nn}}\, \frac{\lambda}{\lambda + 2\mu}
\end{aligned} \tag{B.79}$$

while for a closed crack ($\tilde{\sigma}_{nn} < 0$) the normal stress components (3.279–3.281) read

$$\begin{aligned}
\sigma_{nn}^{closed} &= \underbrace{(\lambda + 2\mu)\, \varepsilon_{nn} + \lambda\, (\varepsilon_{pp} + \varepsilon_{qq})}_{= \tilde{\sigma}_{nn}} \\
\sigma_{pp}^{closed} &= \underbrace{(\lambda + 2\mu)\, \varepsilon_{pp} + \lambda\, (\varepsilon_{nn} + \varepsilon_{qq})}_{= \tilde{\sigma}_{pp}} \\
\sigma_{qq}^{closed} &= \underbrace{(\lambda + 2\mu)\, \varepsilon_{qq} + \lambda\, (\varepsilon_{nn} + \varepsilon_{pp})}_{= \tilde{\sigma}_{qq}} \quad .
\end{aligned} \tag{B.80}$$

The shear stresses given in (3.282–3.284) are

$$\sigma_{np} = g(s) \underbrace{2\mu\, \varepsilon_{np}}_{= \tilde{\sigma}_{np}} \quad , \quad \sigma_{nq} = g(s) \underbrace{2\mu\, \varepsilon_{nq}}_{= \tilde{\sigma}_{nq}} \quad , \quad \sigma_{pq} = \underbrace{2\mu\, \varepsilon_{pq}}_{= \tilde{\sigma}_{pq}} \tag{B.81}$$

and thus independent of crack closure. Expressions for the open and closed crack can be combined by using the Heaviside function (A.106) to

$$\begin{aligned}
\sigma_{\mathrm{nn}} &= \mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\sigma_{\mathrm{nn}}^{\mathrm{open}} + \mathcal{H}(-\tilde{\sigma}_{\mathrm{nn}})\,\sigma_{\mathrm{nn}}^{\mathrm{closed}} = g(s)\,\underbrace{\mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\tilde{\sigma}_{\mathrm{nn}}}_{=\langle\tilde{\sigma}_{\mathrm{nn}}\rangle} + \underbrace{\mathcal{H}(-\tilde{\sigma}_{\mathrm{nn}})\,\tilde{\sigma}_{\mathrm{nn}}}_{=-\langle-\tilde{\sigma}_{\mathrm{nn}}\rangle} \\
\sigma_{\mathrm{pp}} &= \mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\sigma_{\mathrm{pp}}^{\mathrm{open}} + \mathcal{H}(-\tilde{\sigma}_{\mathrm{nn}})\,\sigma_{\mathrm{pp}}^{\mathrm{closed}} \\
&= \tilde{\sigma}_{\mathrm{pp}} + g(s)\,\underbrace{\mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\tilde{\sigma}_{\mathrm{nn}}}_{=\langle\tilde{\sigma}_{\mathrm{nn}}\rangle}\,\frac{\lambda}{\lambda+2\mu} - \underbrace{\mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\tilde{\sigma}_{\mathrm{nn}}}_{=\langle\tilde{\sigma}_{\mathrm{nn}}\rangle}\,\frac{\lambda}{\lambda+2\mu} \\
\sigma_{\mathrm{qq}} &= \mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\sigma_{\mathrm{qq}}^{\mathrm{open}} + \mathcal{H}(-\tilde{\sigma}_{\mathrm{nn}})\,\sigma_{\mathrm{qq}}^{\mathrm{closed}} \\
&= \tilde{\sigma}_{\mathrm{qq}} + g(s)\,\underbrace{\mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\tilde{\sigma}_{\mathrm{nn}}}_{=\langle\tilde{\sigma}_{\mathrm{nn}}\rangle}\,\frac{\lambda}{\lambda+2\mu} - \underbrace{\mathcal{H}(\tilde{\sigma}_{\mathrm{nn}})\,\tilde{\sigma}_{\mathrm{nn}}}_{=\langle\tilde{\sigma}_{\mathrm{nn}}\rangle}\,\frac{\lambda}{\lambda+2\mu} \quad .
\end{aligned} \tag{B.82}$$

Finally, one obtains the active effective stress tensor

$$\begin{aligned}
\tilde{\boldsymbol{\sigma}}_{\mathrm{act}} =& \langle\tilde{\sigma}_{\mathrm{nn}}\rangle\,\boldsymbol{n}_s\otimes\boldsymbol{n}_s + \tilde{\sigma}_{\mathrm{np}}\,(\boldsymbol{n}_s\otimes\boldsymbol{p} + \boldsymbol{p}\otimes\boldsymbol{n}_s) \\
&+ \tilde{\sigma}_{\mathrm{nq}}\,(\boldsymbol{n}_s\otimes\boldsymbol{q} + \boldsymbol{q}\otimes\boldsymbol{n}_s) + \langle\tilde{\sigma}_{\mathrm{nn}}\rangle\,\frac{\lambda}{\lambda+2\mu}\,(\boldsymbol{p}\otimes\boldsymbol{p} + \boldsymbol{q}\otimes\boldsymbol{q})
\end{aligned} \tag{B.83}$$

and the passive effective stress tensor

$$\begin{aligned}
\tilde{\boldsymbol{\sigma}}_{\mathrm{pas}} =& -\langle-\tilde{\sigma}_{\mathrm{nn}}\rangle\,\boldsymbol{n}_s\otimes\boldsymbol{n}_s + \tilde{\sigma}_{\mathrm{pp}}\,\boldsymbol{p}\otimes\boldsymbol{p} + \tilde{\sigma}_{\mathrm{qq}}\,\boldsymbol{q}\otimes\boldsymbol{q} \\
&+ \tilde{\sigma}_{\mathrm{pq}}\,(\boldsymbol{p}\otimes\boldsymbol{q} + \boldsymbol{q}\otimes\boldsymbol{p}) - \langle\tilde{\sigma}_{\mathrm{nn}}\rangle\,\frac{\lambda}{\lambda+2\mu}\,(\boldsymbol{p}\otimes\boldsymbol{p} + \boldsymbol{q}\otimes\boldsymbol{q})
\end{aligned} \tag{B.84}$$

with incorporated crack boundary conditions, which are identical with the expression utilized by Steinke and Kaliske (2019).

B.10 Transverse isotropy of crack orientation dependent degradation

The crack orientation dependent degradation as introduced by Strobl and Seelig (2015) displays a special form of transverse isotropy with symmetry about the crack normal $\overset{s}{\boldsymbol{n}}$. It is beneficial to rewrite all relations by using the tensor $\overset{s}{\boldsymbol{N}} = \overset{s}{\boldsymbol{n}}\otimes\overset{s}{\boldsymbol{n}}$ and invariants of the strain tensor. Since $\overset{s}{\boldsymbol{N}}$ is normalized, multiples of $\overset{s}{\boldsymbol{N}}$ are equal, i.e.

$$\overset{s}{\boldsymbol{N}} = \overset{s}{\boldsymbol{N}}\cdot\overset{s}{\boldsymbol{N}} = \overset{s}{\boldsymbol{N}}\cdot\overset{s}{\boldsymbol{N}}\cdot\overset{s}{\boldsymbol{N}} \tag{B.85}$$

and $\mathrm{tr}(\overset{s}{\boldsymbol{N}}) = 1$.

First, the general representation of transversely isotropic material is given. In this context, the work of Schröder (1995) provides an excellent overview. By assuming that the elastic energy is quadratic in strains, the elastic energy can be written by all possible combinations of invariants as

$$\psi(\boldsymbol{\varepsilon}, \overset{s}{\boldsymbol{N}}) = \alpha_{01}\,(\operatorname{tr}\boldsymbol{\varepsilon})^2 + \alpha_{02}\,\operatorname{tr}(\boldsymbol{\varepsilon}^2) + \alpha_{03}\,\operatorname{tr}\boldsymbol{\varepsilon}\,\operatorname{tr}(\boldsymbol{\varepsilon}\cdot\overset{s}{\boldsymbol{N}}) \tag{B.86}$$

$$+ \alpha_{04}\left(\operatorname{tr}(\boldsymbol{\varepsilon}\cdot\overset{s}{\boldsymbol{N}})\right)^2 + \alpha_{05}\,\operatorname{tr}(\boldsymbol{\varepsilon}^2\cdot\overset{s}{\boldsymbol{N}}) \quad , \tag{B.87}$$

equipped with five parameters $\alpha_{01}\ldots\alpha_{05}$. The related stress tensor is

$$\begin{aligned}\boldsymbol{\sigma}(\boldsymbol{\varepsilon}, \overset{s}{\boldsymbol{N}}) = {} & 2\alpha_{01}\,\operatorname{tr}(\boldsymbol{\varepsilon})\,\boldsymbol{I} + 2\alpha_{02}\,\boldsymbol{\varepsilon} + \alpha_{03}\left[(\boldsymbol{\varepsilon}:\overset{s}{\boldsymbol{N}})\,\boldsymbol{I} + \operatorname{tr}(\boldsymbol{\varepsilon})\;\overset{s}{\boldsymbol{N}}\;\right] \\ & + 2\alpha_{04}\,(\boldsymbol{\varepsilon}:\overset{s}{\boldsymbol{N}})\;\overset{s}{\boldsymbol{N}} + \alpha_{05}\left[\boldsymbol{\varepsilon}\cdot\overset{s}{\boldsymbol{N}} + \overset{s}{\boldsymbol{N}}\cdot\boldsymbol{\varepsilon}\right]\end{aligned} \tag{B.88}$$

and elasticity tensor is

$$\begin{aligned}\mathbb{C}(\overset{s}{\boldsymbol{N}}) = {} & 2\alpha_{01}\,\boldsymbol{I}\otimes\boldsymbol{I} + 2\alpha_{02}\,\mathbb{I}_{\mathrm{sym}} + \alpha_{03}\left[\;\overset{s}{\boldsymbol{N}}\otimes\boldsymbol{I} + \boldsymbol{I}\otimes\overset{s}{\boldsymbol{N}}\right] \\ & + 2\alpha_{04}\;\overset{s}{\boldsymbol{N}}\otimes\overset{s}{\boldsymbol{N}} + \alpha_{05}\,\mathbb{N}_{\mathrm{s}}\end{aligned} \tag{B.89}$$

with the 4th order symmetrized tensor

$$\left(\mathbb{N}_{\mathrm{s}}\right)_{ijkl} := \frac{1}{2}\left(\delta_{ik}\,N_{jl} + \delta_{il}\,N_{jk} + N_{ik}\,\delta_{jl} + N_{il}\,\delta_{jk}\right) \quad . \tag{B.90}$$

With the above at hand, the five parameters for the crack orientation dependent degradation as introduced in Sect. 3.6.1 can be identified. The parameters for the open crack are

$$\begin{aligned}&\alpha_{01} = \frac{g(s)\,\lambda^2 + 2\lambda\mu}{2\lambda + 4\mu}\,, \qquad \alpha_{02} = \mu\,, \qquad \alpha_{03} = \left(g(s) - 1\right)\frac{2\lambda\,\mu}{\lambda + 2\mu}\,, \\ &\alpha_{04} = \left(1 - g(s)\right)\frac{2\mu\,(\lambda + \mu)}{\lambda + 2\mu}\,, \quad \alpha_{05} = \left(g(s) - 1\right)2\mu\end{aligned} \tag{B.91}$$

and for the closed crack they are

$$\alpha_{01} = \frac{\lambda}{2}\,, \quad \alpha_{02} = \mu\,, \quad \alpha_{03} = 0\,, \quad \alpha_{04} = \left(1 - g(s)\right)2\mu\,, \quad \alpha_{05} = \left(g(s) - 1\right)2\mu \quad . \tag{B.92}$$

Since the approach describes (partial) degradation of (linear) elastic material, the parameters in both cases contain two elastic parameters (λ, μ) and the phase field s via the degradation function $g(s)$.

C Numerical implementation

The following sections provide a short overview of the finite element method (FEM) and related approximations and solution methods used in this work. More details on the method itself are provided in widely-used textbooks by, e.g., Zienkiewicz et al. (2013), Hughes (1987) and Belytschko et al. (2014). The notation here mainly follows Taylor (2013).

The main core of the FEM is the discretization in space, i.e. the subdivision of a large domain Ω into smaller parts $\Omega^e \subset \Omega$ (called finite elements), and the piecewise approximation of physical quantities by shape functions, controlled by a finite number of nodal values. In the following, the isoparametric concept is used, so the element geometry is approximated by the same shape functions. Finite elements are typically introduced as "reference element" (Wriggers, 2008) in natural coordinates (given in local Cartesian coordinates ξ_i with $i \in [1, 2, 3]$ in the domain Ω^{ref}) and then transformed to the (perhaps distorted) finite elements in the discretized geometry (given in global coordinates x_i). For example, a linear approximation in 2D (ξ, η or x, y) can be realized by a discretization with three-node triangular elements ($n_{\text{en}} = 3$), also called constant strain triangle (CST) elements (Zienkiewicz et al., 2014), or 4-node quadrilateral elements ($n_{\text{en}} = 4$). Each element is defined by the Lagrangian shape functions of the three-node triangular element

$$N_1(\xi, \eta) = 1 - \xi - \eta \quad , \qquad N_2(\xi, \eta) = \xi \quad , \qquad N_3(\xi, \eta) = \eta \tag{C.1}$$

or the bilinear shape functions of the 4-node quadrilateral element

$$\begin{aligned} N_1(\xi, \eta) &= (1 - \xi)\,(1 - \eta)/4 \quad , & N_2(\xi, \eta) &= (1 + \xi)\,(1 - \eta)/4 \quad , \\ N_3(\xi, \eta) &= (1 + \xi)\,(1 + \eta)/4 \quad , & N_4(\xi, \eta) &= (1 - \xi)\,(1 + \eta)/4 \quad , \end{aligned} \tag{C.2}$$

respectively. In 3D (ξ, η, ζ or x, y, z), the discretization can be realized by 4-node tetrahedral elements ($n_{\text{en}} = 4$) defined by

$$\begin{aligned} N_1(\xi, \eta, \zeta) &= 1 - \xi - \eta - \zeta \quad , & N_2(\xi, \eta, \zeta) &= \xi \quad , \\ N_3(\xi, \eta, \zeta) &= \eta \quad , & N_4(\xi, \eta, \zeta) &= \zeta \end{aligned} \tag{C.3}$$

or 8-node hexahedral elements ($n_{\text{en}} = 8$) defined by the trilinear shape functions

$$\begin{aligned} N_1(\xi, \eta, \zeta) &= (1 - \xi)\,(1 - \eta)\,(1 - \zeta)/8 \,, & N_2(\xi, \eta, \zeta) &= (1 + \xi)\,(1 - \eta)\,(1 - \zeta)/8 \,, \\ N_3(\xi, \eta, \zeta) &= (1 + \xi)\,(1 + \eta)\,(1 - \zeta)/8 \,, & N_4(\xi, \eta, \zeta) &= (1 - \xi)\,(1 + \eta)\,(1 - \zeta)/8 \,, \\ N_5(\xi, \eta, \zeta) &= (1 - \xi)\,(1 - \eta)\,(1 + \zeta)/8 \,, & N_6(\xi, \eta, \zeta) &= (1 + \xi)\,(1 - \eta)\,(1 + \zeta)/8 \,, \\ N_7(\xi, \eta, \zeta) &= (1 + \xi)\,(1 + \eta)\,(1 + \zeta)/8 \,, & N_8(\xi, \eta, \zeta) &= (1 - \xi)\,(1 + \eta)\,(1 + \zeta)/8 \,. \end{aligned} \tag{C.4}$$

The global coordinates in the finite element ($\boldsymbol{x} \in \Omega^{\mathrm{e}}$) are approximated by

$$\boldsymbol{x} = \sum_{I=1}^{n_{\mathrm{en}}} N_I(\xi_i)\, \boldsymbol{x}_I \quad , \tag{C.5}$$

based on the global coordinates of the finite element nodes $\boldsymbol{x}_I$. The relation between the spatial derivatives in local and global coordinates, for a single node I is given by using the chain rule

$$N_{I,\xi_j} := \frac{\partial N_I}{\partial \xi_j} = \frac{\partial N_I}{\partial x_j}\,\frac{\partial x_j}{\partial \xi_i} \tag{C.6}$$

with n_{ndm} being the number of dimensions. Its matrix-vector representation, e.g., in 3D,

$$\begin{bmatrix} \partial N_I/\partial\xi \\ \partial N_I/\partial\eta \\ \partial N_I/\partial\zeta \end{bmatrix} = \underbrace{\begin{bmatrix} \partial x/\partial\xi & \partial y/\partial\xi & \partial z/\partial\xi \\ \partial x/\partial\eta & \partial y/\partial\eta & \partial z/\partial\eta \\ \partial x/\partial\zeta & \partial y/\partial\zeta & \partial z/\partial\zeta \end{bmatrix}}_{=:\,\boldsymbol{J}} \underbrace{\begin{bmatrix} \partial N_I/\partial x \\ \partial N_I/\partial y \\ \partial N_I/\partial z \end{bmatrix}}_{=:\,N_{I,x_j}} \tag{C.7}$$

reveals the Jacobian matrix[1] $\boldsymbol{J} := (\partial \boldsymbol{x}/\partial \boldsymbol{\xi})^{\mathrm{T}}$. With the approximation (C.5), it can be computed in the finite element ($\boldsymbol{x} \in \Omega^{\mathrm{e}}$) from

$$\boldsymbol{J} = \sum_{I=1}^{n_{\mathrm{en}}} \begin{bmatrix} x_I\,\dfrac{\partial N_I(\xi_i)}{\partial \xi_i} & y_I\,\dfrac{\partial N_I(\xi_i)}{\partial \xi_i} & z_I\,\dfrac{\partial N_I(\xi_i)}{\partial \xi_i} \end{bmatrix} \quad , \tag{C.8}$$

so that the spatial derivatives in global coordinates are $N_{I,x_j} = \boldsymbol{J}^{-1}\, N_{I,\xi_j}$. Thus, $\boldsymbol{J}$ is required to be invertible, which guarantees the unique (and thus physically consistent) mapping between the global and natural coordinates. Since, the Jacobian determinant describes the volume ratio, here between global and local coordinates, see (2.8), the Jacobian matrix must be positive definite ($\det \boldsymbol{J} > 0$).

Integral expressions are computed in two steps: First the integral expression of $f(x, y, z)$ is transformed to natural coordinates

$$\int_{\Omega^{\mathrm{e}}} f(x, y, z)\, \mathrm{d}V = \int_{\Omega^{\mathrm{ref}}} \tilde{f}(\xi, \eta, \zeta)\, \det\big(\boldsymbol{J}(\xi, \eta, \zeta)\big)\, \mathrm{d}\xi\, \mathrm{d}\eta\, \mathrm{d}\zeta \quad . \tag{C.9}$$

[1] Note, that in the literature, the Jacobian matrix is often defined according to $\tilde{\boldsymbol{J}} := \partial \boldsymbol{x}/\partial \boldsymbol{\xi} = \boldsymbol{J}^{\mathrm{T}}$ (e.g. Taylor, 2013). This enables to compute the transformations by using the matrices $\boldsymbol{N}_\xi := \partial \boldsymbol{N}/\partial \xi_i$ and $\boldsymbol{N}_x := \partial \boldsymbol{N}/\partial x$ based on the vector of shape functions $\boldsymbol{N} := [N_1\ N_2\ \ldots\ N_{n_{\mathrm{en}}}]^{\mathrm{T}}$. Accordingly, it is $\tilde{\boldsymbol{J}} = \boldsymbol{X} \cdot \boldsymbol{N}_\xi$ with $\boldsymbol{X} := [\boldsymbol{x}_1\, \boldsymbol{x}_2\, \ldots\, \boldsymbol{x}_{n_{\mathrm{en}}}]$, and $\boldsymbol{N}_x = \boldsymbol{N}_\xi\, \tilde{\boldsymbol{J}}^{-1}$.

Then, integration is performed by using numerical integration, usually Gaussian quadrature with n_{int} specific integration points (ξ_q, η_q, ζ_q) and weights w_q, a comprehensive list is given in Zienkiewicz et al. (2013), i.e.

$$\int_{\Omega^{\text{ref}}} \tilde{f}(\xi,\eta,\zeta)\,\det\big(\boldsymbol{J}(\xi,\eta,\zeta)\big)\,\mathrm{d}\xi\,\mathrm{d}\eta\,\mathrm{d}\zeta \approx \sum_{q=1}^{n_{\text{int}}} \tilde{f}(\xi_q,\eta_q,\zeta_q)\;\det\big(\boldsymbol{J}(\xi_q,\eta_q,\zeta_q)\big)\,w_q\;. \tag{C.10}$$

C.1 Image segmentation

For image segmentation, the governing equations derived from the functional (3.6) to compute the gray level approximation u and edge approximation v from the original image gray level g in Ω are

$$2\beta\,(u-g) - 2(v^2+k_\epsilon)\,\Delta u = 0 \quad , \tag{C.11}$$

$$2v\,|\nabla u|^2 - \alpha\left(\frac{(1-v)}{2\epsilon} + 2\epsilon\,\Delta v\right) = 0 \quad , \tag{C.12}$$

with the natural boundary conditions $\nabla u \cdot \boldsymbol{n} = 0$ and $\nabla v \cdot \boldsymbol{n} = 0$ on $\partial\Omega$. This set of coupled partial differential equations is solved by the finite element method. By introducing test functions δu, $\nabla\delta u$, δv and $\nabla\delta v$ the weak forms of (C.11) and (C.12) are

$$-\int_\Omega \Big[\delta u\,\big(2\beta\,u\big) + \nabla\delta u\,\big(2(v^2+k_\epsilon)\,\nabla u\big)\Big]\,\mathrm{d}A + \int_\Omega \delta u\,\big(2\beta\,g(\boldsymbol{x})\big)\,\mathrm{d}A = 0 \quad , \tag{C.13}$$

$$\int_\Omega \left[\delta v\,\big(2v\,|\nabla u|^2\big) - \delta v\left(\alpha\,\frac{(1-v)}{2\epsilon}\right) - \nabla\delta v\cdot\big(2\alpha\,\epsilon\,\nabla v\big)\right]\mathrm{d}A = 0 \quad . \tag{C.14}$$

With the bilinear shape functions of Lagrangian type N_I, see (C.2), one obtains the approximation of the gray level field

$$u_{\text{h}}(\boldsymbol{x}) = \sum_{I=1}^{4} N_I(\boldsymbol{x})\,u_I, \qquad \delta u_{\text{h}}(\boldsymbol{x}) = \sum_{I=1}^{4} N_I(\boldsymbol{x})\,\delta u_I \tag{C.15}$$

and edge field

$$v_{\text{h}}(\boldsymbol{x}) = \sum_{I=1}^{4} N_I(\boldsymbol{x})\,v_I, \qquad \delta v_{\text{h}}(\boldsymbol{x}) = \sum_{I=1}^{4} N_I(\boldsymbol{x})\,\delta v_I \quad . \tag{C.16}$$

The spatial derivatives are

$$\nabla u_{\text{h}} = \begin{bmatrix} u_{,x} \\ u_{,y} \end{bmatrix} = \sum_{I=1}^{4} \boldsymbol{B}_I^{\text{u}}\,u_I\;, \quad \nabla\delta u_{\text{h}} = \sum_{I=1}^{4} \boldsymbol{B}_I^{\delta\text{u}}\,\delta u_I \quad \text{with} \quad \boldsymbol{B}_I^{\text{u}} = \boldsymbol{B}_I^{\delta\text{u}} = \begin{bmatrix} N_{I,x} \\ N_{I,y} \end{bmatrix} \tag{C.17}$$

and

$$\nabla v_\mathrm{h} = \begin{bmatrix} v_{,x} \\ v_{,y} \end{bmatrix} = \sum_{I=1}^{4} \boldsymbol{B}_I^\mathrm{v}\, v_I, \;, \quad \nabla \delta v_\mathrm{h} = \sum_{I=1}^{4} \boldsymbol{B}_I^{\delta\mathrm{v}}\, \delta v_I \quad \text{with} \quad \boldsymbol{B}_I^\mathrm{v} = \boldsymbol{B}_I^{\delta\mathrm{v}} = \begin{bmatrix} N_{I,x} \\ N_{I,y} \end{bmatrix} . \tag{C.18}$$

With the vectors of test functions

$$\delta \boldsymbol{d} = \begin{bmatrix} \begin{bmatrix}\delta \boldsymbol{d}_1\end{bmatrix}^\mathrm{T} & \begin{bmatrix}\delta \boldsymbol{d}_2\end{bmatrix}^\mathrm{T} & \dots & \begin{bmatrix}\delta \boldsymbol{d}_{n_\mathrm{nds}}\end{bmatrix}^\mathrm{T} \end{bmatrix}^\mathrm{T} \qquad \text{and} \qquad \delta \boldsymbol{d}_I = \begin{bmatrix} \begin{bmatrix}\delta u_I\end{bmatrix}^\mathrm{T} & \delta v_I \end{bmatrix}^\mathrm{T} \tag{C.19}$$

the discretization of the weak forms (C.13) and (C.14) finally reads

$$\begin{bmatrix}\delta \boldsymbol{d}\end{bmatrix}^\mathrm{T} \boldsymbol{R} = \bigcup_{e=1}^{n_\mathrm{el}} \sum_{I=1}^{n_\mathrm{en}} \begin{bmatrix}\delta \boldsymbol{d}_I\end{bmatrix}^\mathrm{T} \boldsymbol{R}_I^\mathrm{e} = 0 \quad , \tag{C.20}$$

with n_nds being the total number of nodes in the problem. The element-wise contribution of the nodal values with $I \in [1,\, 2, \dots,\, n_\mathrm{nds}]$ to the global residual vector $\boldsymbol{R}$ is

$$\begin{aligned} \boldsymbol{R}_I^\mathrm{e} = & \underbrace{- \int_{\Omega^\mathrm{e}} \begin{bmatrix} N_I \left(2\beta\, u_\mathrm{h}\right) + \begin{bmatrix}\boldsymbol{B}_I^\mathrm{u}\end{bmatrix}^\mathrm{T} \left(2(v_\mathrm{h}^2 + k_\epsilon)\, \nabla u\right) \\ -N_I \left(2 v_\mathrm{h} \left|\nabla u_\mathrm{h}\right|^2\right) + N_I \left(\alpha\, \dfrac{(1 - v_\mathrm{h})}{2\epsilon}\right) - \begin{bmatrix}\boldsymbol{B}_I^\mathrm{v}\end{bmatrix}^\mathrm{T} \left(2\alpha\, \epsilon\, \nabla v_\mathrm{h}\right) \end{bmatrix} \mathrm{d}V}_{= \boldsymbol{P}_I^\mathrm{e}} \\ & + \underbrace{\int_{\Omega^\mathrm{e}} \begin{bmatrix} N_I\, 2\beta\, g(\boldsymbol{x}) \\ 0 \end{bmatrix} \mathrm{d}A}_{= \boldsymbol{F}_I^\mathrm{e}} \quad . \end{aligned} \tag{C.21}$$

Since expression (C.20) holds for an arbitrary choice of the test functions $\delta \boldsymbol{d}$ the global system of coupled equations reads $\boldsymbol{R} = \boldsymbol{F} - \boldsymbol{P}(\boldsymbol{d}) = \boldsymbol{0}$. As proposed by Bourdin (1999), the problem can be decomposed into two subproblems which are then solved in an alternating manner, with cycles k switching from one field computation to the other. With

$$\boldsymbol{v}^{(k)} = \begin{bmatrix} v_1^{(k)} & v_2^{(k)} & \dots & v_{n_\mathrm{nds}}^{(k)} \end{bmatrix}^\mathrm{T} \quad \text{and} \quad \boldsymbol{u}^{(k)} = \begin{bmatrix} u_1^{(k)} & u_2^{(k)} & \dots & u_{n_\mathrm{nds}}^{(k)} \end{bmatrix}^\mathrm{T} \quad , \tag{C.22}$$

one field computation concerns the unknown edge field $\boldsymbol{v}^{(k)}$ by using the gray level approximation $\boldsymbol{u}^{(k-1)}$ from the previous cycle with

$$\boldsymbol{R}^\mathrm{v} = \boldsymbol{F}^\mathrm{v} - \boldsymbol{P}^\mathrm{v}(\boldsymbol{v}^{(k)}, \boldsymbol{u} \overset{!}{=} \boldsymbol{u}^{(k-1)}) = \boldsymbol{0} \tag{C.23}$$

and the other one is related to the unknown gray level approximation $\boldsymbol{u}^{(k)}$ by using the edge field values from the current cycle $\boldsymbol{v}^{(k)}$ obtained from (C.23)

$$\boldsymbol{R}^\mathrm{u} = \boldsymbol{F}^\mathrm{u} - \boldsymbol{P}^\mathrm{u}(\boldsymbol{u}^{(k)}, \boldsymbol{v} \overset{!}{=} \boldsymbol{v}^{(k)}) = \boldsymbol{0} \quad . \tag{C.24}$$

The systems of equations are solved by Newton-Raphson iterations (counted by i) according to $\boldsymbol{R}^{\mathrm{v}}_{i+1} \approx \boldsymbol{R}^{\mathrm{v}}_{i} - \boldsymbol{K}^{\mathrm{vv}}\,\Delta\boldsymbol{v}_i = \boldsymbol{0}$ and $\boldsymbol{R}^{\mathrm{u}}_{j+1} \approx \boldsymbol{R}^{\mathrm{u}}_{j} - \boldsymbol{K}^{\mathrm{uu}}\,\Delta\boldsymbol{u}_j = \boldsymbol{0}$ with the incremental updates $\boldsymbol{v}_{i+1} = \boldsymbol{v}_i + \Delta\boldsymbol{v}_i$ and $\boldsymbol{u}_{j+1} = \boldsymbol{u}_j + \Delta\boldsymbol{u}_j$ until convergence is obtained. The tangent matrices are

$$\boldsymbol{K}^{\mathrm{vv}} = -\frac{\partial \boldsymbol{R}^{\mathrm{v}}}{\partial \boldsymbol{v}} \qquad \text{and} \qquad \boldsymbol{K}^{\mathrm{uu}} = -\frac{\partial \boldsymbol{R}^{\mathrm{u}}}{\partial \boldsymbol{u}} \tag{C.25}$$

with the element contributions

$$\boldsymbol{K}^{\mathrm{vv}}_{IJ} = \int_{\Omega^{\mathrm{e}}} \left[-N_I \left(2\,|\nabla u_{\mathrm{h}}|^2\right) N_J - N_I \left(\frac{\alpha}{2\epsilon}\right) N_J - \left[\boldsymbol{B}^{\mathrm{v}}_I\right]^{\mathrm{T}} \left(2\alpha\,\epsilon\right) \boldsymbol{B}^{\mathrm{v}}_J \right] \mathrm{d}A \quad . \tag{C.26}$$

and

$$\boldsymbol{K}^{\mathrm{uu}}_{IJ} = \int_{\Omega^{\mathrm{e}}} \left[N_I \left(2\beta\right) N_J + \left[\boldsymbol{B}^{\mathrm{u}}_I\right]^{\mathrm{T}} \left(2(v^2_{\mathrm{h}} + k_\epsilon)\right) \boldsymbol{B}^{\mathrm{u}}_J \right] \mathrm{d}A \quad . \tag{C.27}$$

C.2 Phase field approach for fracture

The weak forms of the momentum balance and phase field evolution equation are given by (3.47) and (3.48), respectively. By using shape functions N_I, e.g. of Lagrangian type, one obtains the approximation of the displacement

$$\boldsymbol{u}_{\mathrm{h}}(\boldsymbol{x}) = \sum_{I=1}^{n_{\mathrm{en}}} N_I(\boldsymbol{x})\,\boldsymbol{u}_I \quad , \qquad \delta\boldsymbol{u}_{\mathrm{h}}(\boldsymbol{x}) = \sum_{I=1}^{n_{\mathrm{en}}} N_I(\boldsymbol{x})\,\delta\boldsymbol{u}_I \tag{C.28}$$

and phase field

$$s_{\mathrm{h}}(\boldsymbol{x}) = \sum_{I=1}^{n_{\mathrm{en}}} N_I(\boldsymbol{x})\,s_I \quad , \qquad \delta s_{\mathrm{h}}(\boldsymbol{x}) = \sum_{I=1}^{n_{\mathrm{en}}} N_I(\boldsymbol{x})\,\delta s_I \tag{C.29}$$

in each finite element, with n_{en} nodes per element. Following the Bubnov-Galerkin approach, the primary unknowns ($\boldsymbol{u}$, s) and corresponding virtual values ($\delta\boldsymbol{u}$, δs) are interpolated by identical shape functions between their discrete nodal values ($\boldsymbol{u}_I$, s_I and $\delta\boldsymbol{u}_I$, δs_I).

Two-dimensional (plane strain) problems

Plane strain boundary value problems of the domain Ω (with the Cartesian coordinates x, and y) are spatially discretized by a finite number of n_{el} of two-dimensional isoparametric finite elements $\Omega^{\mathrm{e}} \subset \Omega$. The displacement and phase field is approximated by (C.28) and (C.29) either by three-node triangles (C.1) or 4-node quadrilateral elements (C.2). By means of the matrices

$$\boldsymbol{B}^{\mathrm{u}}_I = \boldsymbol{B}^{\delta\mathrm{u}}_I = \begin{bmatrix} N_{I,x} & 0 & N_{I,y} \\ 0 & N_{I,y} & N_{I,x} \end{bmatrix}^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{B}^{\mathrm{s}}_I = \boldsymbol{B}^{\delta\mathrm{s}}_I = \begin{bmatrix} N_{I,x} \\ N_{I,y} \end{bmatrix} \quad , \tag{C.30}$$

corresponding spatial derivatives are approximated by

$$\boldsymbol{\varepsilon}_{\mathrm{h}} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{yy} & 2\varepsilon_{xy} \end{bmatrix}^{\mathrm{T}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\mathrm{u}}\, \boldsymbol{u}_I \quad , \qquad \delta\boldsymbol{\varepsilon}_{\mathrm{h}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\delta\mathrm{u}}\, \delta\boldsymbol{u}_I \quad , \tag{C.31}$$

and

$$\nabla s_{\mathrm{h}} = \begin{bmatrix} s_{,x} & s_{,y} \end{bmatrix}^{\mathrm{T}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\mathrm{s}}\, s_I \quad , \qquad \nabla\delta s_{\mathrm{h}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\delta\mathrm{s}}\, \delta s_I \quad . \tag{C.32}$$

Axisymmetric problems

According to the boundary value problem of indentation fracture discussed in Chapter 4 an axisymmetric formulation (with the cylindrical coordinates r, ϕ and z) is employed for the spatial discretization of the domain Ω into a finite number of n_{el} isoparametric finite elements $\Omega^{\mathrm{e}} \subset \Omega$. Accordingly, the displacement and phase field is approximated by (C.28) and (C.29) either by three-node triangles (C.1) or 4-node quadrilateral elements (C.2). By means of the matrices

$$\boldsymbol{B}_I^{\mathrm{u}} = \boldsymbol{B}_I^{\delta\mathrm{u}} = \begin{bmatrix} N_{I,r} & 0 & N_{I,z} & N_I/r \\ 0 & N_{I,z} & N_{I,r} & 0 \end{bmatrix}^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{B}_I^{\mathrm{s}} = \boldsymbol{B}_I^{\delta\mathrm{s}} = \begin{bmatrix} N_{I,r} \\ N_{I,z} \end{bmatrix} \quad , \tag{C.33}$$

corresponding spatial derivatives are approximated by

$$\boldsymbol{\varepsilon}_{\mathrm{h}} = \begin{bmatrix} \varepsilon_{rr} & \varepsilon_{zz} & 2\varepsilon_{rz} & \varepsilon_{\phi\phi} \end{bmatrix}^{\mathrm{T}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\mathrm{u}}\, \boldsymbol{u}_I \quad , \qquad \delta\boldsymbol{\varepsilon}_{\mathrm{h}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\delta\mathrm{u}}\, \delta\boldsymbol{u}_I \quad , \tag{C.34}$$

and

$$\nabla s_{\mathrm{h}} = \begin{bmatrix} s_{,r} & s_{,z} \end{bmatrix}^{\mathrm{T}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\mathrm{s}}\, s_I \quad , \qquad \nabla\delta s_{\mathrm{h}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\delta\mathrm{s}}\, \delta s_I \quad . \tag{C.35}$$

Three-dimensional problems

Three dimensional boundary value problems of the domain Ω (with the Cartesian coordinates x, y and z) are spatially discretized by a finite number of n_{el} of three-dimensional isoparametric finite elements $\Omega^{\mathrm{e}} \subset \Omega$. The displacement and phase field is approximated by (C.28) and (C.29) either by using 4-node tetrahedral elements (C.3), 6-node wedge/triangular prism elements (not considered here) or 8-node hexahedral elements (C.4). By means of the matrices

$$\boldsymbol{B}_I^{\mathrm{u}} = \boldsymbol{B}_I^{\delta\mathrm{u}} = \begin{bmatrix} N_{I,x} & 0 & 0 & N_{I,y} & 0 & N_{I,z} \\ 0 & N_{I,y} & 0 & N_{I,x} & N_{I,z} & 0 \\ 0 & 0 & N_{I,z} & 0 & N_{I,y} & N_{I,x} \end{bmatrix}^{\mathrm{T}} \text{ and } \boldsymbol{B}_I^{\mathrm{s}} = \boldsymbol{B}_I^{\delta\mathrm{s}} = \begin{bmatrix} N_{I,x} \\ N_{I,y} \\ N_{I,z} \end{bmatrix} , \tag{C.36}$$

corresponding spatial derivatives are approximated by

$$\boldsymbol{\varepsilon}_{\mathrm{h}} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{yy} & \varepsilon_{zz} & 2\varepsilon_{xy} & 2\varepsilon_{yz} & 2\varepsilon_{xz} \end{bmatrix}^{\mathrm{T}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\mathrm{u}}\, \boldsymbol{u}_I \quad , \qquad \delta\boldsymbol{\varepsilon}_{\mathrm{h}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\delta\mathrm{u}}\, \delta\boldsymbol{u}_I \quad , \tag{C.37}$$

and

$$\nabla s_{\mathrm{h}} = \begin{bmatrix} s_{,x} & s_{,y} & s_{,z} \end{bmatrix}^{\mathrm{T}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\mathrm{s}}\, s_I \quad , \qquad \nabla\delta s_{\mathrm{h}} = \sum_{I=1}^{n_{\mathrm{en}}} \boldsymbol{B}_I^{\delta\mathrm{s}}\, \delta s_I \quad . \tag{C.38}$$

Residual vector

The discretization of the weak forms (3.47) and (3.48) finally reads

$$\left[\delta\boldsymbol{d}\right]^{\mathrm{T}} \boldsymbol{R} = \bigcup_{e=1}^{n_{\mathrm{el}}} \sum_{I=1}^{n_{\mathrm{en}}} \left[\delta\boldsymbol{d}_I\right]^{\mathrm{T}} \boldsymbol{R}_I^{\mathrm{e}} = 0 \quad , \tag{C.39}$$

with

$$\delta\boldsymbol{d} = \begin{bmatrix} \left[\delta\boldsymbol{d}_1\right]^{\mathrm{T}} & \left[\delta\boldsymbol{d}_2\right]^{\mathrm{T}} & \dots & \left[\delta\boldsymbol{d}_{n_{\mathrm{nds}}}\right]^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \qquad \text{and} \qquad \delta\boldsymbol{d}_I = \begin{bmatrix} \left[\delta\boldsymbol{u}_I\right]^{\mathrm{T}} & \delta s_I \end{bmatrix}^{\mathrm{T}} \tag{C.40}$$

being vectors of test functions. The total number of nodes in the problem is n_{nds}. The element-wise contribution of the nodal values with $I \in [1, 2, \dots, n_{\mathrm{nds}}]$ to the global residual vector $\boldsymbol{R}$ is

$$\boldsymbol{R}_I^{\mathrm{e}} = \underbrace{-\int_{\Omega^{\mathrm{e}}} \begin{bmatrix} N_I \left(\rho\, \ddot{\boldsymbol{u}}\right) + \left[\boldsymbol{B}_I^{\mathrm{u}}\right]^{\mathrm{T}} \boldsymbol{\sigma}_{\mathrm{h}}(\boldsymbol{\varepsilon}_{\mathrm{h}}, s_{\mathrm{h}}) \\ N_I \left(\eta_{\mathrm{vc}}\, \dot{s}_{\mathrm{h}}\right) + N_I \left(g'(s_{\mathrm{h}})\, D_{\mathrm{s}}(\boldsymbol{\varepsilon}_{\mathrm{h}})\right) + N_I \dfrac{\mathcal{G}_{\mathrm{c}}\, w'(s_{\mathrm{h}})}{2\ell\, c_{\mathrm{w}}} + \left[\boldsymbol{B}_I^{\mathrm{s}}\right]^{\mathrm{T}} \left(\dfrac{\mathcal{G}_{\mathrm{c}}\, \ell}{c_{\mathrm{w}}}\, \nabla s_{\mathrm{h}}\right) \end{bmatrix} \mathrm{d}V}_{= \boldsymbol{P}_I^{\mathrm{e}}}$$

$$\underbrace{+ \int_{\Omega^{\mathrm{e}}} \begin{bmatrix} N_I\, \bar{\boldsymbol{f}}_{\mathrm{h}} \\ 0 \end{bmatrix} \mathrm{d}V + \int_{\partial\Omega^{\mathrm{e}}} \begin{bmatrix} N_I\, \bar{\boldsymbol{t}}_{\mathrm{h}} \\ 0 \end{bmatrix} \mathrm{d}A}_{= \boldsymbol{F}_I^{\mathrm{e}}} \quad , \tag{C.41}$$

with the internal forces $\boldsymbol{P}_I^{\mathrm{e}}$ and external forces $\boldsymbol{F}_I^{\mathrm{e}}$. Note, that in case of crack initiation problems without pre-existing crack it is suggested to replace $\mathcal{G}_{\mathrm{c}}$ and ℓ in (C.41) by their recalibrated values $\mathcal{G}_{\mathrm{c}}^*$ and ℓ^* (Sect. 3.7.3). Since expression (C.39) holds for an arbitrary choice of the test functions $\delta\boldsymbol{d}$ the global system of coupled equations reads

$$\boldsymbol{R} = \boldsymbol{F} - \boldsymbol{P}(\boldsymbol{d}, \dot{\boldsymbol{d}}, \ddot{\boldsymbol{d}}) = \boldsymbol{0} \quad . \tag{C.42}$$

Time integration

If no time-dependent terms are present in (C.41) the global system of coupled equations (C.42) becomes

$$\boldsymbol{R} = \boldsymbol{F} - \boldsymbol{P}(\boldsymbol{d}) = \boldsymbol{0} \quad . \tag{C.43}$$

This is the quasi-static case which is present, for example, in Sect. 4.3. In all other cases, the transient terms $\dot{\boldsymbol{d}}$ and $\ddot{\boldsymbol{d}}$ have to be computed by time integration schemes of first and second order, respectively. For implicit methods, the global system of coupled equations (C.42) is evaluated at $t + \Delta t$ or in time discretized form at $t_{n+1} = t_n + \Delta t$

$$\boldsymbol{R}_{n+1} = \boldsymbol{F}_{n+1} - \boldsymbol{P}(\boldsymbol{d}_{n+1}, \dot{\boldsymbol{d}}_{n+1}, \ddot{\boldsymbol{d}}_{n+1}) = \boldsymbol{0} \quad , \tag{C.44}$$

with Δt being the time step size. The implicit Newmark method, also called the Newmark β-scheme, approximates

$$\boldsymbol{d}_{n+1} = \boldsymbol{d}_n + \Delta t\, \dot{\boldsymbol{d}}_n + \frac{1}{2}\Delta t^2 \left[(1-2\beta)\, \ddot{\boldsymbol{d}}_n + 2\beta\, \ddot{\boldsymbol{d}}_{n+1} \right] \tag{C.45}$$

and

$$\dot{\boldsymbol{d}}_{n+1} = \dot{\boldsymbol{d}}_n + \Delta t \left[(1-\gamma)\, \ddot{\boldsymbol{d}}_n + \gamma\, \ddot{\boldsymbol{d}}_{n+1} \right] \quad , \tag{C.46}$$

where β and γ are parameters which control the stability and numerical dissipation (Newmark, 1959). For $0.5 \leq \gamma \leq 2\beta$ the method is unconditionally stable, while for $2\beta < \gamma$ the method is conditionally stable[2] for $\omega_{\max} \Delta t \leq (\gamma/2 - \beta)^{-1/2}$ (Hughes, 1987). With $\gamma > 0.5$ numerical dissipation is introduced. A typical choice is the trapezoidal rule (assuming constant average acceleration) with $\beta = 0.25$ and $\gamma = 0.5$, which is unconditionally stable. With (C.45) and (C.46) the global system of coupled equations (C.44) is

$$\boldsymbol{R}_{n+1} = \boldsymbol{F}_{n+1} - \tilde{\boldsymbol{P}}(\boldsymbol{d}_{n+1}, \boldsymbol{d}_n, \dot{\boldsymbol{d}}_n, \ddot{\boldsymbol{d}}_n) = \boldsymbol{0} \tag{C.47}$$

(Taylor, 2013). It is nonlinear and must be solved for the unknowns $\boldsymbol{d}_{n+1}$. Note, that if $\ddot{\boldsymbol{d}}$ is not present in (C.42), a time integration scheme of first order is sufficient, e.g. the implicit Euler scheme with the approximation

$$\dot{\boldsymbol{d}}_{n+1} = \frac{1}{\Delta t} \left(\boldsymbol{d}_n + \boldsymbol{d}_{n+1} \right) \quad . \tag{C.48}$$

Then, the global system of coupled equations (C.44) is

$$\boldsymbol{R}_{n+1} = \boldsymbol{F}_{n+1} - \hat{\boldsymbol{P}}(\boldsymbol{d}_{n+1}, \boldsymbol{d}_n) = \boldsymbol{0} \quad . \tag{C.49}$$

[2] According to Hughes (1987), the undamped critical sampling frequency $\omega_{\max}$ is taken as a conservative estimate if no modal damping coefficient is available.

Monolitic solution scheme

To solve the nonlinear equations (C.47) in every time step for the vector of unknowns (the current time increment $n+1$ is left for brevity)

$$\boldsymbol{d} = \left[\left[\boldsymbol{d}_1\right]^{\mathrm{T}} \quad \left[\boldsymbol{d}_2\right]^{\mathrm{T}} \quad \dots \quad \left[\boldsymbol{d}_{n_{\mathrm{nds}}}\right]^{\mathrm{T}} \right]^{\mathrm{T}} \qquad \text{with} \qquad \boldsymbol{d}_I = \left[\left[\boldsymbol{u}_I\right]^{\mathrm{T}} \quad s_I \right]^{\mathrm{T}} \quad , \tag{C.50}$$

the Newton-Raphson iteration scheme (iterations are counted by i)

$$\boldsymbol{R}_{i+1} = \boldsymbol{R}_i + \Delta\boldsymbol{R}_i \approx \boldsymbol{R}_i + D\boldsymbol{R}_i \, \Delta\boldsymbol{d}_{i+1} = \boldsymbol{R}_i - \boldsymbol{S}_i \, \Delta\boldsymbol{d}_{i+1} = \boldsymbol{0} \tag{C.51}$$

is applied, with the incremental update $\boldsymbol{d}_{i+1} = \boldsymbol{d}_i + \Delta\boldsymbol{d}_i$ until the residual vector (normalized by the norm of the residual vector of the first iteration) is below a tolerance, i.e. $||\boldsymbol{R}(\boldsymbol{d}_i)|| < \delta_{\mathrm{d}} \, ||\boldsymbol{R}(\boldsymbol{d}_1)||$. The tolerance δ_{d} is typically chosen to be half the machine precision (Zienkiewicz et al., 2014). The tangent matrix

$$\boldsymbol{S} = -\left[\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{d}} + \frac{\partial \boldsymbol{R}}{\partial \dot{\boldsymbol{d}}} \frac{\partial \dot{\boldsymbol{d}}}{\partial \boldsymbol{d}} + \frac{\partial \boldsymbol{R}}{\partial \ddot{\boldsymbol{d}}} \frac{\partial \ddot{\boldsymbol{d}}}{\partial \boldsymbol{d}} \right] = \boldsymbol{K} + \frac{\gamma}{\Delta t \, \beta} \boldsymbol{C} + \frac{1}{\Delta t^2 \, \beta} \boldsymbol{M} \tag{C.52}$$

includes the stiffness matrix $\boldsymbol{K}$, the damping matrix $\boldsymbol{C}$ and the mass matrix $\boldsymbol{M}$ (Taylor, 2013). The element stiffness matrix is

$$\boldsymbol{K}_{IJ} = \begin{bmatrix} \boldsymbol{K}_{IJ}^{\mathrm{uu}} & \boldsymbol{K}_{IJ}^{\mathrm{us}} \\ \boldsymbol{K}_{IJ}^{\mathrm{su}} & \boldsymbol{K}_{IJ}^{\mathrm{ss}} \end{bmatrix} \tag{C.53}$$

and contains

$$\boldsymbol{K}_{IJ}^{\mathrm{uu}} = \int_{\Omega^{\mathrm{e}}} \left[\boldsymbol{B}_I^{\mathrm{u}}\right]^{\mathrm{T}} \frac{\partial \boldsymbol{\sigma}_{\mathrm{h}}(\boldsymbol{\varepsilon}_{\mathrm{h}}, s_{\mathrm{h}})}{\partial \boldsymbol{\varepsilon}_{\mathrm{h}}} \boldsymbol{B}_J^{\mathrm{u}} \, \mathrm{d}V = \int_{\Omega^{\mathrm{e}}} \left[\boldsymbol{B}_I^{\mathrm{u}}\right]^{\mathrm{T}} \boldsymbol{C}(s_{\mathrm{h}}) \, \boldsymbol{B}_J^{\mathrm{u}} \, \mathrm{d}V \quad , \tag{C.54}$$

$$\boldsymbol{K}_{IJ}^{\mathrm{us}} = \int_{\Omega^{\mathrm{e}}} \left[\boldsymbol{B}_I^{\mathrm{u}}\right]^{\mathrm{T}} \frac{\partial \boldsymbol{\sigma}_{\mathrm{h}}}{\partial s_{\mathrm{h}}} N_J \, \mathrm{d}V \quad , \tag{C.55}$$

$$\boldsymbol{K}_{IJ}^{\mathrm{su}} = \int_{\Omega^{\mathrm{e}}} N_I \, g'(s_{\mathrm{h}}) \left[\frac{\partial D_{\mathrm{s}}(\boldsymbol{\varepsilon}_{\mathrm{h}})}{\partial \boldsymbol{\varepsilon}_{\mathrm{h}}} \right]^{\mathrm{T}} \boldsymbol{B}_J^{\mathrm{u}} \, \mathrm{d}V \quad , \tag{C.56}$$

$$\boldsymbol{K}_{IJ}^{\mathrm{ss}} = \int_{\Omega^{\mathrm{e}}} g''(s_{\mathrm{h}}) \, D_{\mathrm{s}}(\boldsymbol{\varepsilon}_{\mathrm{h}}) \, N_I \, N_J + \frac{\mathcal{G}_{\mathrm{c}}}{\ell} N_I \, N_J + \mathcal{G}_{\mathrm{c}} \, \ell \left[\boldsymbol{B}_I^{\mathrm{s}}\right]^{\mathrm{T}} \boldsymbol{B}_J^{\mathrm{s}} \, \mathrm{d}V \tag{C.57}$$

with the material tangent $\boldsymbol{C}(s_{\mathrm{h}})$. Note, that for a separate choice of the crack driving energy according to Sect. 3.2.2 it is

$$(\boldsymbol{K}_{IJ}^{\mathrm{su}})^{\mathrm{T}} \neq \boldsymbol{K}_{IJ}^{\mathrm{su}} \tag{C.58}$$

in general. Thus, a numerical solution procedure taking the unsymmetric tangent matrix into account is suggested, which however is computationally (much) more expensive. The element damping matrix is

$$\boldsymbol{C}_{IJ} = \begin{bmatrix} \boldsymbol{0}_{n\times n} & \boldsymbol{0}_{n\times 1} \\ \boldsymbol{0}_{1\times n} & C_{IJ}^{\mathrm{ss}} \end{bmatrix} \qquad \text{with} \qquad C_{IJ}^{\mathrm{ss}} = \int_{\Omega^{\mathrm{e}}} \eta_{\mathrm{vc}}\, N_I\, N_J \,\mathrm{d}V \quad , \tag{C.59}$$

and the mass matrix is

$$\boldsymbol{M}_{IJ} = \begin{bmatrix} \boldsymbol{M}_{IJ}^{\mathrm{uu}} & \boldsymbol{0}_{n\times 1} \\ \boldsymbol{0}_{1\times n} & \boldsymbol{0}_{1\times 1} \end{bmatrix} \qquad \text{with} \qquad \boldsymbol{M}_{IJ}^{\mathrm{uu}} = \int_{\Omega^{\mathrm{e}}} \rho\, N_I\, N_J\, \boldsymbol{I}_{n\times n} \,\mathrm{d}V \quad . \tag{C.60}$$

They contain the $n_{\mathrm{dim}} \times n_{\mathrm{dim}}$ matrix with ones in the diagonal $\boldsymbol{I}_{n\times n}$ and zero matrices $\boldsymbol{0}_{m\times k}$ of the indicated size (m rows and k columns). In case of crack initiation problems without pre-existing crack it is suggested to replace $\mathcal{G}_{\mathrm{c}}$ and ℓ in (C.57) by their recalibrated values $\mathcal{G}_{\mathrm{c}}^{*}$ and ℓ^{*} (Sect. 3.7.3).

Alternate minimization procedure

Instead of solving (C.42) for the vector of unknowns (C.50) in a monolithic fashion, the problem can be decomposed into two subproblems which are then solved in an alternating manner (with cycles k switching from one field computation to the other) as applied, e.g. in Sect. 4.3.2. One field computation concerns the unknown phase field

$$\boldsymbol{s}^{(k)} = \begin{bmatrix} s_1^{(k)} & s_2^{(k)} & \dots & s_{n_{\mathrm{nds}}}^{(k)} \end{bmatrix}^{\mathrm{T}} \tag{C.61}$$

by using displacements

$$\boldsymbol{u}^{(k-1)} = \begin{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^{(k-1)} \end{bmatrix}^{\mathrm{T}} & \begin{bmatrix} \boldsymbol{u}_2^{(k-1)} \end{bmatrix}^{\mathrm{T}} & \dots & \begin{bmatrix} \boldsymbol{u}_{n_{\mathrm{nds}}}^{(k-1)} \end{bmatrix}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{C.62}$$

from the previous cycle

$$\boldsymbol{R}^{\mathrm{s}} = \boldsymbol{F}^{\mathrm{s}} - \boldsymbol{P}^{\mathrm{s}}(\boldsymbol{s}^{(k)}, \boldsymbol{u} \stackrel{!}{=} \boldsymbol{u}^{(k-1)}) = \boldsymbol{0} \tag{C.63}$$

and the other one is related to the unknown displacements $\boldsymbol{u}^{(k)}$ by using the phase field values from the current cycle $\boldsymbol{s}^{(k)}$ obtained from (C.63)

$$\boldsymbol{R}^{\mathrm{u}} = \boldsymbol{F}^{\mathrm{u}} - \boldsymbol{P}^{\mathrm{u}}(\boldsymbol{u}^{(k)}, \boldsymbol{s} \stackrel{!}{=} \boldsymbol{s}^{(k)}) = \boldsymbol{0} \quad . \tag{C.64}$$

Since, in general, both subproblems are non-linear each of them has to be solved by Newton-Raphson iterations (counted by i) according to

$$\boldsymbol{R}_{i+1}^{\mathrm{s}} \approx \boldsymbol{R}_{i}^{\mathrm{s}} - \boldsymbol{S}^{\mathrm{ss}}\, \Delta \boldsymbol{s}_i = \boldsymbol{0} \tag{C.65}$$

and

$$\boldsymbol{R}^{\mathrm{u}}_{j+1} \approx \boldsymbol{R}^{\mathrm{u}}_{j} - \boldsymbol{S}^{\mathrm{uu}}\,\Delta\boldsymbol{u}_i = \boldsymbol{0} \tag{C.66}$$

with the incremental updates $\boldsymbol{s}_{i+1} = \boldsymbol{s}_i + \Delta\boldsymbol{s}_i$ and $\boldsymbol{u}_{j+1} = \boldsymbol{u}_j + \Delta\boldsymbol{u}_j$ until convergence is obtained, typically defined by $||\boldsymbol{R}^{\mathrm{s}}(\boldsymbol{s}_i)|| < \delta_{\mathrm{s}}\,||\boldsymbol{R}^{\mathrm{s}}(\boldsymbol{s}_1)||$ and $||\boldsymbol{R}^{\mathrm{u}}(\boldsymbol{u}_i)|| < \delta_{\mathrm{u}}\,||\boldsymbol{R}^{\mathrm{u}}(\boldsymbol{u}_1)||$. Note, that due to the missing second order time derivatives, the implicit Euler method accordingto (C.49) is used for time integration of (C.65). Thus, the tangent matrices are

$$\boldsymbol{S}^{\mathrm{ss}} = -\left[\frac{\partial \boldsymbol{R}^{\mathrm{s}}}{\partial \boldsymbol{s}} + \frac{\partial \boldsymbol{R}^{\mathrm{s}}}{\partial \dot{\boldsymbol{s}}}\,\frac{\partial \dot{\boldsymbol{s}}}{\partial \boldsymbol{s}} + \underbrace{\frac{\partial \boldsymbol{R}^{\mathrm{s}}}{\partial \ddot{\boldsymbol{s}}}\,\frac{\partial \ddot{\boldsymbol{s}}}{\partial \boldsymbol{s}}}_{=\,\boldsymbol{0}}\right] = \boldsymbol{K}^{\mathrm{ss}} + \frac{1}{\Delta t}\,\boldsymbol{C}^{\mathrm{ss}} \tag{C.67}$$

and

$$\boldsymbol{S}^{\mathrm{uu}} = -\left[\frac{\partial \boldsymbol{R}^{\mathrm{u}}}{\partial \boldsymbol{u}} + \underbrace{\frac{\partial \boldsymbol{R}^{\mathrm{u}}}{\partial \dot{\boldsymbol{u}}}\,\frac{\partial \dot{\boldsymbol{u}}}{\partial \boldsymbol{u}}}_{=\,\boldsymbol{0}} + \frac{\partial \boldsymbol{R}^{\mathrm{u}}}{\partial \ddot{\boldsymbol{u}}}\,\frac{\partial \ddot{\boldsymbol{u}}}{\partial \boldsymbol{u}}\right] = \boldsymbol{K}^{\mathrm{uu}} + \frac{1}{\Delta t^2\,\beta}\,\boldsymbol{M}^{\mathrm{uu}} \quad . \tag{C.68}$$

The element stiffness matrices read

$$\boldsymbol{K}^{\mathrm{ss}}_{IJ} = \int_{\Omega^{\mathrm{e}}} g''(s_{\mathrm{h}})\,D_{\mathrm{s}}(\boldsymbol{\varepsilon}_{\mathrm{h}})\,N_I\,N_J + \frac{\mathcal{G}_{\mathrm{c}}}{\ell}\,N_I\,N_J + \mathcal{G}_{\mathrm{c}}\,\ell\,\left[\boldsymbol{B}^{\mathrm{s}}_I\right]^{\mathrm{T}}\boldsymbol{B}^{\mathrm{s}}_J\,\mathrm{d}V \tag{C.69}$$

and

$$\boldsymbol{K}^{\mathrm{uu}}_{IJ} = \int_{\Omega^{\mathrm{e}}} \left[\boldsymbol{B}^{\mathrm{u}}_I\right]^{\mathrm{T}}\frac{\partial \boldsymbol{\sigma}_{\mathrm{h}}(\boldsymbol{\varepsilon}_{\mathrm{h}}, s_{\mathrm{h}})}{\partial \boldsymbol{\varepsilon}_{\mathrm{h}}}\,\boldsymbol{B}^{\mathrm{u}}_J\,\mathrm{d}V = \int_{\Omega^{\mathrm{e}}} \left[\boldsymbol{B}^{\mathrm{u}}_I\right]^{\mathrm{T}}\boldsymbol{C}(s_{\mathrm{h}})\,\boldsymbol{B}^{\mathrm{u}}_J\,\mathrm{d}V \tag{C.70}$$

with the material tangent $\boldsymbol{C}(s_{\mathrm{h}})$. The element damping and mass matrices are

$$\boldsymbol{C}^{\mathrm{ss}}_{IJ} = \int_{\Omega^{\mathrm{e}}} \eta_{\mathrm{vc}}\,N_I\,N_J\,\mathrm{d}V \qquad \text{and} \qquad \boldsymbol{M}^{\mathrm{uu}}_{IJ} = \int_{\Omega^{\mathrm{e}}} \rho\,N_I\,N_J\,\boldsymbol{I}_{n\times n}\,\mathrm{d}V \quad , \tag{C.71}$$

with $\boldsymbol{I}_{n\times n}$ being a $n_{\mathrm{dim}} \times n_{\mathrm{dim}}$ matrix with ones in the diagonal. For crack initiation problems (starting from $s=1$), solving for the phase field (C.65) becomes non-linear as soon as phase field evolution $\dot{s} < 0$ takes place. Without pre-existing crack it is suggested to replace $\mathcal{G}_{\mathrm{c}}$ and ℓ in (C.69) by their recalibrated values $\mathcal{G}^*_{\mathrm{c}}$ and ℓ^* (Sect. 3.7.3). In the special case that no tension-compression split is present in the momentum balance equation, i.e. the stress tensor is $\boldsymbol{\sigma}_{\mathrm{h}} = \boldsymbol{C}(s_{\mathrm{h}})\,\boldsymbol{\varepsilon}_{\mathrm{h}}$, solving for the displacements (C.66) is a linear problem.

Material tangent matrix

By using the Voigt notation as introduced in App. A.7, the material tangent matrices for commonly used tension-compression splits discussed in Chapter 3 are provided. They are required for implicit time integration schemes to calculate nodal displacements.

In case of full degradation the stiffness tensor (3.12) corresponds to the isotropically degraded elasticity tensor (2.85), i.e.

$$\mathbb{C}(s) = g(s)\left(\lambda\, \boldsymbol{I} \otimes \boldsymbol{I} + 2\mu\, \mathbb{I}_{\mathrm{sym}}\right) \quad . \tag{C.72}$$

For plane strain conditions the corresponding matrix notation in the discretized setting is

$$\boldsymbol{C}(s_{\mathrm{h}}) = \begin{bmatrix} g(s_{\mathrm{h}})\left(\lambda + 2\mu\right) & g(s_{\mathrm{h}})\,\lambda & 0 \\ g(s_{\mathrm{h}})\,\lambda & g(s_{\mathrm{h}})\left(\lambda + 2\mu\right) & 0 \\ 0 & 0 & g(s_{\mathrm{h}})\,\mu \end{bmatrix} \quad , \tag{C.73}$$

while for axisymmetric conditions the material tangent matrix (by using E and ν) is

$$\boldsymbol{C}(s_{\mathrm{h}}) = \frac{g(s_{\mathrm{h}})\,E}{(1+\nu)\,(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & 0 & \nu \\ \nu & 1-\nu & 0 & \nu \\ 0 & 0 & \frac{1}{2}(1-2\nu) & 0 \\ \nu & \nu & 0 & 1-\nu \end{bmatrix} \quad . \tag{C.74}$$

For the spectral strain decomposition (3.204) discussed in Sect. 3.5.1 the stiffness tensor is

$$\begin{aligned} \mathbb{C}(s) = &\sum_{a}^{3}\sum_{b}^{3} \frac{\partial\, \sigma_a}{\partial\, \varepsilon_b}\; \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_a \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_a \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_b \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_b \\ &+ \frac{1}{2}\sum_{a}^{3}\sum_{b\neq a}^{3} \frac{\sigma_a - \sigma_b}{\varepsilon_a - \varepsilon_b}\left(\overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_a \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_b \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_a \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_b + \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_a \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_b \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_b \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_a \right) \end{aligned} \tag{C.75}$$

with

$$\frac{\partial\, \sigma_a}{\partial\, \varepsilon_b} = g(s)\,\lambda\, \frac{\langle \mathrm{tr}(\boldsymbol{\varepsilon})\rangle}{\mathrm{tr}(\boldsymbol{\varepsilon})} + g(s)\, 2\mu\, \frac{\langle \varepsilon_a\rangle}{\varepsilon_a}\, \delta_{ab} - \lambda\, \frac{\langle -\mathrm{tr}(\boldsymbol{\varepsilon})\rangle}{\mathrm{tr}(\boldsymbol{\varepsilon})} - 2\mu\, \frac{\langle -\varepsilon_a\rangle}{\varepsilon_a}\, \delta_{ab} \tag{C.76}$$

and

$$\frac{\sigma_a - \sigma_b}{\varepsilon_a - \varepsilon_b} = g(s)\, 2\mu\, \frac{\langle \varepsilon_a\rangle - \langle \varepsilon_b\rangle}{\varepsilon_a - \varepsilon_b} - 2\mu\, \frac{\langle -\varepsilon_a\rangle - \langle -\varepsilon_b\rangle}{\varepsilon_a - \varepsilon_b} \quad . \tag{C.77}$$

It contains both derivatives with respect to current principal directions and terms attributed to the reorientation of principal directions. Its discretization to $\boldsymbol{C}(s_{\mathrm{h}})$ is straightforward. To avoid the costly computation of strain eigenvectors $\overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_i$, especially in 3D problems, one can directly use derivatives formulated according to $\overset{\boldsymbol{\varepsilon}}{\boldsymbol{N}}_i := \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_i \otimes \overset{\boldsymbol{\varepsilon}}{\boldsymbol{n}}_i$ without the need to explicitly compute eigenvectors as proposed by Miehe (1993). Since (C.76) is only defined for distinct eigenvalues unequal to zero a perturbation technique can be used or different cases described in App. A.5 must be considered.

For the volumetric-deviatoric decomposition (3.230) discussed in Sect. 3.5.2 the stiffness tensor is

$$\mathbb{C}(s) = g(s)\,\lambda\,\frac{\langle \mathrm{tr}(\boldsymbol{\varepsilon})\rangle}{\mathrm{tr}(\boldsymbol{\varepsilon})}\,\boldsymbol{I}\otimes\boldsymbol{I} - \lambda\,\frac{\langle -\mathrm{tr}(\boldsymbol{\varepsilon})\rangle}{\mathrm{tr}(\boldsymbol{\varepsilon})}\,\boldsymbol{I}\otimes\boldsymbol{I} + g(s)\,2\mu\,\mathbb{I}_{\mathrm{sym}} \quad . \tag{C.78}$$

For plane strain conditions the corresponding matrix notation in the discretized setting is equal to (C.73) for an opened crack and

$$\boldsymbol{C}(s_{\mathrm{h}}) = \begin{bmatrix} \left(\lambda+\frac{2}{3}\mu\right) + g(s_{\mathrm{h}})\,\frac{4}{3}\mu & \left(\lambda+\frac{2}{3}\mu\right) - g(s_{\mathrm{h}})\,\frac{2}{3}\mu & 0 \\ \left(\lambda+\frac{2}{3}\mu\right) - g(s_{\mathrm{h}})\,\frac{2}{3}\mu & \left(\lambda+\frac{2}{3}\mu\right) + g(s_{\mathrm{h}})\,\frac{4}{3}\mu & 0 \\ 0 & 0 & g(s_{\mathrm{h}})\,\mu \end{bmatrix} \tag{C.79}$$

for a closed crack.

For the crack orientation dependent degradation (Sect. 3.6.1) the elasticity tensor is given by (3.286) for an opened and by (3.287) for a closed crack. Its matrix representations in the discrete setting (the coordinate system is aligned with the crack) are

$$\boldsymbol{C}_{\mathrm{open}}(s_{\mathrm{h}}) = \begin{bmatrix} g(s_{\mathrm{h}})\,(\lambda+2\mu) & g(s_{\mathrm{h}})\,\lambda & g(s_{\mathrm{h}})\,\lambda & 0 & 0 & 0 \\ g(s_{\mathrm{h}})\,\lambda & \lambda+2\mu+\tilde{\Lambda}_{\mathrm{g}}(s_{\mathrm{h}}) & \lambda+\tilde{\Lambda}_{\mathrm{g}}(s_{\mathrm{h}}) & 0 & 0 & 0 \\ g(s_{\mathrm{h}})\,\lambda & \lambda+\tilde{\Lambda}_{\mathrm{g}}(s_{\mathrm{h}}) & \lambda+2\mu+\tilde{\Lambda}_{\mathrm{g}}(s_{\mathrm{h}}) & 0 & 0 & 0 \\ 0 & 0 & 0 & g(s_{\mathrm{h}})\,\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & g(s_{\mathrm{h}})\,\mu \end{bmatrix} \tag{C.80}$$

for the open crack and

$$\boldsymbol{C}_{\mathrm{closed}}(s_{\mathrm{h}}) = \begin{bmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & g(s_{\mathrm{h}})\,\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & g(s_{\mathrm{h}})\,\mu \end{bmatrix} \tag{C.81}$$

for the closed crack. The correction term $\tilde{\Lambda}_{\mathrm{g}}(s_{\mathrm{h}}) := (g(s_{\mathrm{h}})-1)\,\Lambda$ in the lateral directions contains the elastic portion $\Lambda := \lambda^2/(\lambda+2\mu)$ to obtain plane stress conditions in the plane of the opened crack.

C.3 Computational realization of irreversibility constraints

In order to account for the irreversible behavior or fracture additional constraints have to be incorporated into the system of equations (C.51) or (C.63) related to the phase field.

A common strategy when using the alternate minimization procedure is to incorporate a local "damage-like" irreversibility (3.40) by replacing the current crack driving energy in the residual vector (C.41) and the stiffness matrix (C.69) by the maximum crack driving energy obtained in history

$$D_{\mathrm{s}}(\boldsymbol{\varepsilon}_{\mathrm{h}}) \leftarrow \max\left[D_{\mathrm{s}}\big(\boldsymbol{\varepsilon}_{\mathrm{h}}(\boldsymbol{x},t)\big),\ \mathcal{H}(\boldsymbol{x},t)\right] \tag{C.82}$$

as proposed by Miehe et al. (2010a). This requires the storage of one history variable

$$\mathcal{H}(\boldsymbol{x},t) = \max_{\tau\in[0,\,t)}\left[D_{\mathrm{s}}\big(\boldsymbol{\varepsilon}_{\mathrm{h}}(\boldsymbol{x},\tau)\big)\right] \tag{C.83}$$

per Gauss point.

By contrast, the incorporation of a "crack-like" behavior (3.36), which is the preferred option to compute fracture, can be realized by applying Dirichlet boundary conditions

$$\mathrm{s}_I \stackrel{!}{=} 0 \tag{C.84}$$

on the phase field where a nodal phase field value s_I of the previous state is below an extremely small threshold value, e.g. $s_{\mathrm{tol}} = 10^{-8}$ as in Schlüter (2018). This can be done by manipulating both the stiffness matrix $\boldsymbol{K}$ (C.53) or $\boldsymbol{K}^{\mathrm{ss}}$ (C.69) and the residual vector $\boldsymbol{R}^{\mathrm{s}}$ (C.65) as described by Kuhn (2013).

The application of the constraint (C.84) requires the phase field to localize promptly to the broken state, at least below the threshold s_{tol}, to activate the irreversibility constraint. If, however, the smallest nodal phase field value at a crack is slightly above this threshold phase field recovery will still be possible. In view of the absence of a strong stress concentration in indentation fracture problem (Chapter 4) and due to the low propensity of the higher-order degradation functions to approach the broken stage this is particular critical. Its impact is shown in Sect. 4.3.5, where an unphysical "moving crack phenomenon" is observed since the threshold for the irreversibility constraint is not attained. Faced to this situation, directly constraining the phase field according to (C.84) by using a higher threshold value does not solve this problem, since the neighboring nodes are also affected by the constraint. To be able to introduce an irreversibility threshold s_0 of finite value with

$$s_{\mathrm{tol}} < s_0 \ll 1 \tag{C.85}$$

and to achieve a robust algorithmic treatment of the irreversibility constraint (C.84) the phase field is prevented to increase locally as soon as the threshold is reached, i.e.

$$\dot{s} \leq 0 \qquad \text{only if } s \leq s_0 \quad . \tag{C.86}$$

This is done (only for the alternate minimization procedure) by partially replacing the current crack driving energy in the residual vector (C.41) and the tangent matrix (C.69) by

$$D_{\mathrm{s}}(\boldsymbol{\varepsilon}_{\mathrm{h}}) = \begin{cases} \max\left[D_{\mathrm{s}}\big(\boldsymbol{\varepsilon}_{\mathrm{h}}(\boldsymbol{x},t)\big),\ \mathcal{H}(\boldsymbol{x},t)\right] & \text{if } s \leq s_0 \\ D_{\mathrm{s}}\big(\boldsymbol{\varepsilon}_{\mathrm{h}}(\boldsymbol{x},t)\big) & \text{otherwise} \quad , \end{cases} \tag{C.87}$$

with the history field variable $\mathcal{H}(\boldsymbol{x}, t)$ (see above). In addition, Dirichlet boundary conditions according to (C.84) are applied to keep s non-negative. This approach enables both to fix cracks once they have formed as well as phase field recovery with $\dot{s} > 0$ outside the immediate vicinity of the crack so that the fracture surface energy is correctly reproduced in the simulation (cf. Sect. 3.7.3). For some other reasons, Steinke and Kaliske (2019) proposed a similar algorithmic treatment.

C.4 Realization of pre-existing cracks

Closely related to the irreversibility of the phase field is the question of how to set crack boundary conditions in case of pre-existing cracks. There are at least three distinct options to model initial cracks:

- Initial cracks can be modeled as discrete cracks, i.e. by introducing double nodes in the finite element mesh along the crack surface. Consequently, finite element edges must be aligned with the crack surface. With applied loading, further crack progress can take place by phase field evolution from the stress concentration at the crack front/tip. Due to the regularization, the incomplete approximation of the $1/\sqrt{r}$-singular stress concentration in standard finite elements has (almost) no effect on the phase field solution (e.g. Hofacker, 2014), as long as the mesh is fine enough, e.g., $h_\mathrm{e} < \ell/4$. However, Klinsmann et al. (2015) showed that a mesh induced initial crack delays the progress significantly, since initial localization of the phase field consumes additional energy, see also Sect. 3.7.2. Nonetheless, discrete pre-cracks are often used in the literature, especially to simulate standard "benchmarks" like the single-edged notched test, see e.g. Miehe et al. (2010a), Borden (2012) and Hesch et al. (2017). One reason might be the insufficient and erroneous approximation of crack boundary conditions (Sect. 3.5), see also the comparison of initial crack modeling by May et al. (2015).
- Nodal phase field values can be prescribed by applying Dirichlet boundary conditions $s = 0$ at the crack, more specifically, along the crack set $\mathscr{S}_\mathrm{s}$. For this purpose, finite element edges must also be aligned with the crack surface. If only one row of nodes is constrained to zero, there is still a significant amount of remaining stiffness across the crack. Thus, it is preferred to constrain a double row of nodes along the crack, with sufficiently small finite elements in-between. This leads to vanishing stiffness, but results in significantly larger errors in terms of the crack surface energy, which hast to be corrected, e.g. according to the recalibration discussed in Sect. 3.7.3.
- If the phase field formulation includes a strain history field $\mathcal{H}(\boldsymbol{x})$ as introduced in Sect. 3.2.3, it is possible to modify (i.e. constrain) the crack driving energy by prescribing initial history values to form a pre-crack. In this case, the modeling of pre-cracks is independent of finite element edges. Based on the approach proposed

by Borden et al. (2012) in the context of isogeometric analysis, Strobl and Seelig (2016) proposed a slightly modified version with

$$\mathcal{H}(\boldsymbol{x}) = \frac{\mathcal{G}_\mathrm{c}\,(1-s_0)}{2\ell\,s_0}\,\frac{\langle h_\mathrm{e} - d_\mathrm{s}(\boldsymbol{x})\rangle}{h_\mathrm{e}} \quad , \tag{C.88}$$

which is independent of the mesh size. It contains the bracket operator (3.196) and the closest distance $d_\mathrm{s}(\boldsymbol{x})$ to the surface of the pre-crack. In fact, the width where the crack driving energy is imposed is limited to two times the element size h_e to reduce for $\ell > 2h_\mathrm{e}$ the width of the fully degraded zone (compared to the original proposal). Note, that the intended initial phase field value at the crack with $s_0 \to 0$ has to be chosen, e.g. $s_0 = 10^{-3}$ (Borden et al., 2012).

D Further FFM analyses of indentation fracture

D.1 Pre-fracture stress distribution

For finite fracture mechanics (FFM) analyses, see Sect. 2.2.6, the knowledge of the stress field distribution is of particular interest. The stress field caused by indentation of a rigid cylindrical punch with radius a on the surface of an elastic half-space was partially given by Boussinesq (1885), long before the full problem was solved by Sneddon (1946a). Outside the punch it is similar to the stress field generated by compression of a spherical indenter, which was explored by Hertz (1882a); Hertz (1882b). In order to analyze indentation fracture cylindrical indenters are preferred since the indenter radius is equal to the contact radius and remains, by assuming a rigid indenter, constant. Typographical mistakes in Sneddon (1946a) were corrected later on by, e.g., Barquins and Maugis (1982). In Kachanov et al. (2003) the stress field is given in the complex form

$$\begin{aligned} \sigma_a &= \sigma_{rr} + \sigma_{\phi\phi} \\ \sigma_b &= \sigma_{rr} - \sigma_{\phi\phi} + 2i\,\sigma_{r\phi} \\ \tau_z &= \sigma_{rz} + i\,\sigma_{\phi z} \end{aligned} \tag{D.1}$$

with the elementary functions (here already given for the axisymmetric case in a cylindrical coordinate system with $r \geq 0$, $z \geq 0$ and $a > 0$)

$$\begin{aligned} \sigma_a &= -\frac{F}{2\pi\, a}\left((1+2\nu)\,\frac{\sqrt{a^2 - l_1^2}}{l_2^2 - l_1^2} + \frac{z^2\, l_1^4 - z^2\, a^2\, (2a^2 - r^2 + 2z^2)}{\sqrt{a^2 - l_1^2}\,\left(l_2^2 - l_1^2\right)^3} \right) \quad , \\ \sigma_b &= \frac{F}{2\pi\, a}\left(-(1-2\nu)\,\frac{\sqrt{a^2 - l_1^2}}{l_2^2 - l_1^2} + (1-2\nu)\,\frac{2a - 2\sqrt{a^2 - l_1^2}}{r^2} \right. \\ &\qquad \left. - \frac{a\, z\, \sqrt{l_2^2 - a^2}\,\left[2l_1^4 + r^2\,(l_1^2 + 3l_2^2 - 6a^2)\right]}{l_2^2\,\left(l_2^2 - l_1^2\right)^3} \right) \quad , \\ \sigma_{zz} &= \frac{F}{2\pi\, a}\left(\frac{\sqrt{a^2 - l_1^2}}{l_2^2 - l_1^2} + \frac{z^2\, l_1^4 + z^2\, a^2\, (r^2 - 2a^2 - 2z^2)}{\sqrt{a^2 - l_1^2}\,\left(l_2^2 - l_1^2\right)^3} \right) \end{aligned} \tag{D.2}$$

and

$$\tau_z = -\frac{F}{2\pi\, a}\,\frac{z\, r\,\sqrt{a^2 - l_1^2}\,(3l_2^2 + l_1^2 - 4a^2)}{l_2^2 - l_1^2} \quad , \tag{D.3}$$

where l_1 and l_2 are length parameters defined by

$$\begin{aligned} l_1 &= \frac{1}{2}\left(\sqrt{(r+a)^2 + z^2} - \sqrt{(r-a)^2 + z^2}\right) \quad , \\ l_2 &= \frac{1}{2}\left(\sqrt{(r+a)^2 + z^2} + \sqrt{(r-a)^2 + z^2}\right) \quad , \end{aligned} \tag{D.4}$$

with $l_1 \geq 0$ and $l_2 \geq a$, see also the work of Fabrikant (1988). While σ_{zz} is already given in (D.2) other stress components in a cylindrical coordinate system can be obtained according to

$$\begin{aligned} \sigma_{rr} &= \frac{1}{2}\,\mathrm{Re}\{\sigma_a + \sigma_b\} &&= \frac{1}{2}(\sigma_a + \sigma_b) \quad , \\ \sigma_{\phi\phi} &= \frac{1}{2}\,\mathrm{Re}\{\sigma_a - \sigma_b\} &&= \frac{1}{2}(\sigma_a - \sigma_b) \quad , \\ \sigma_{rz} &= \mathrm{Re}\{\tau_z\} &&= \tau_z \quad , \\ \sigma_{\phi z} &= \mathrm{Im}\{\tau_z\} &&= 0 \quad , \\ \sigma_{r\phi} &= \mathrm{Im}\{\sigma_b\} &&= 0 \quad , \end{aligned} \tag{D.5}$$

where the last two components correspond with axisymmetric conditions.

Alternatively, by using the mean contact pressure

$$p_\mathrm{m} = \frac{F}{\pi\, a^2} = E\,\frac{d}{a}\,\frac{2}{\pi\,(1-\nu^2)} \tag{D.6}$$

the surface stress ($z = 0$) below the indenter $r \leq a$ can be written with

$$\bar{\varrho} = \sqrt{1 - \left(\frac{r}{a}\right)^2} \tag{D.7}$$

according to Maugis (2000) and Fischer-Cripps (2007)

$$\begin{aligned} \sigma_{rr} &= p_\mathrm{m}\,\frac{1-2\nu}{2}\,\frac{1}{(r/a)^2}(1-\bar{\varrho}) - \frac{1}{2}\,p_\mathrm{m}\,\frac{1}{\bar{\varrho}} \quad , \\ \sigma_{\theta\theta} &= p_\mathrm{m}\,\frac{1-2\nu}{2}\,\frac{1}{(r/a)^2}(\bar{\varrho}-1) - p_\mathrm{m}\,\nu\,\frac{1}{\bar{\varrho}} \quad , \\ \sigma_{zz} &= -\frac{1}{2}\,p_\mathrm{m}\,\frac{1}{\bar{\varrho}} \quad , \\ \sigma_{rz} &= 0 \quad . \end{aligned} \tag{D.8}$$

Using the solutions of integral expressions, see Sneddon (1946a),

$$
\begin{aligned}
J_{10} &= \frac{1}{\sqrt{R}}\,\sin(\phi/2) \quad , \\
J_{01} &= \frac{1}{r/a}\left(1-\sqrt{R}\,\sin(\phi/2)\right) \quad , \\
J_{21} &= \frac{r}{a}\,\frac{1}{\sqrt{R^3}}\,\sin(3\phi/2) \quad , \\
J_{20} &= \sqrt{1+\left(\frac{z}{a}\right)^2}\,\frac{1}{\sqrt{R^3}}\,\sin(3\phi/2-\theta) \quad , \\
J_{11} &= \sqrt{1+\left(\frac{z}{a}\right)^2}\,\frac{1}{r/a}\,\frac{1}{\sqrt{R}}\,\sin(\theta-\phi/2)
\end{aligned}
\tag{D.9}
$$

with the parameters

$$
\begin{aligned}
R &= \sqrt{\left[\left(\frac{r}{a}\right)^2+\left(\frac{z}{a}\right)^2-1\right]^2+4\left(\frac{z}{a}\right)^2} \quad , \\
\tan\theta &= \frac{1}{z/a} \quad , \\
\tan\phi &= \frac{2\,z/a}{(r/a)^2+(z/a)^2-1} \quad ,
\end{aligned}
\tag{D.10}
$$

the non-zero stress components outside the indenter $r > a$ can be written in compact form as

$$
\begin{aligned}
\sigma_{rr} &= -\frac{1}{2}\,p_{\mathrm{m}}\left[J_{10}-\frac{z}{a}\,J_{20}-(1-2\nu)\,\frac{1}{r/a}\,J_{01}+\frac{z/a}{r/a}\,J_{11}\right] \quad , \\
\sigma_{\theta\theta} &= -\frac{1}{2}\,p_{\mathrm{m}}\left[2\nu\,J_{10}+(1-2\nu)\frac{1}{r/a}\,J_{01}-\frac{z/a}{r/a}\,J_{11}\right] \quad , \\
\sigma_{zz} &= -\frac{1}{2}\,p_{\mathrm{m}}\left[J_{10}+\frac{z}{a}\,J_{20}\right] \quad , \\
\sigma_{rz} &= -\frac{1}{2}\,p_{\mathrm{m}}\,\frac{z}{a}\,J_{21} \quad .
\end{aligned}
\tag{D.11}
$$

A simple geometric interpretation of the parameters in (D.10) is given, e.g., by Maugis (2000). At the free surface ($z = 0$) these stress components are

$$\sigma_{rr}(z=0) = p_\mathrm{m}\,\frac{1-2\nu}{2}\,\frac{1}{(r/a)^2} = E\,\frac{d}{a}\,\frac{1}{\pi}\left(\frac{1-2\nu}{1-\nu^2}\right)\frac{1}{(r/a)^2} \quad , \tag{D.12a}$$

$$\sigma_{\theta\theta}(z=0) = -p_\mathrm{m}\,\frac{1-2\nu}{2}\,\frac{1}{(r/a)^2} = -E\,\frac{d}{a}\,\frac{1}{\pi}\left(\frac{1-2\nu}{1-\nu^2}\right)\frac{1}{(r/a)^2} \quad , \tag{D.12b}$$

$$\sigma_{zz}(z=0) = 0 \quad , \tag{D.12c}$$

$$\sigma_{rz}(z=0) = 0 \quad . \tag{D.12d}$$

Normalized according to (4.3), the radial stress component ψ_{rr} is (only depending on r/a, z/a and ν)

$$\psi_{rr} = \frac{\sigma_{rr}}{E}\,\frac{a}{d} = -\frac{1}{\pi\,(1-\nu^2)}\left[J_{10} - \frac{z}{a}\,J_{20} - (1-2\nu)\,\frac{1}{r/a}\,J_{01} + \frac{z/a}{r/a}\,J_{11}\right] \quad , \tag{D.13}$$

which is at the free surface

$$\psi_{rr}(z/a=0) = \frac{1}{\pi}\left(\frac{1-2\nu}{1-\nu^2}\right)\frac{1}{(r/a)^2} \quad . \tag{D.14}$$

Figures D.1–D.6 show contour lines of the stress components given in (D.5) and normalized according to (4.3). The contour line of zero radial stress $\psi_{rr} = 0$ is marked in Figs. D.1 and D.2.

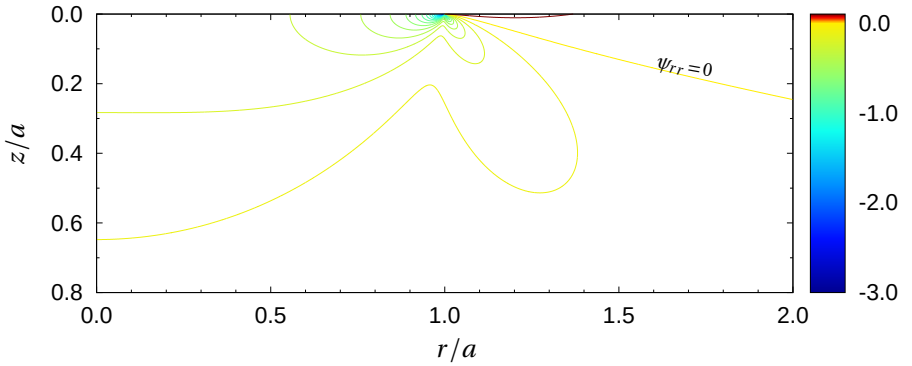


Figure D.1: Contour line of the radial stress ψ_{rr} under a cylindrical punch for $\nu = 0.22$. Red and orange colors indicate positive (tensile) values.

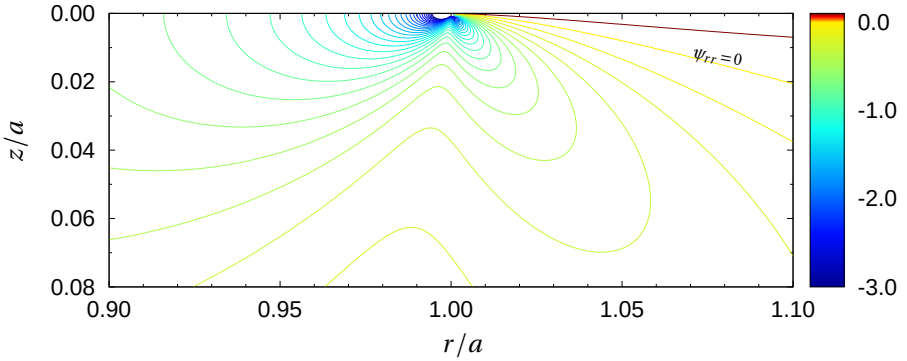


Figure D.2: Contour line of the radial stress ψ_{rr} for $\nu = 0.22$ at the indenter edge ($r/a = 1$). Red and orange colors indicate positive (tensile) values.

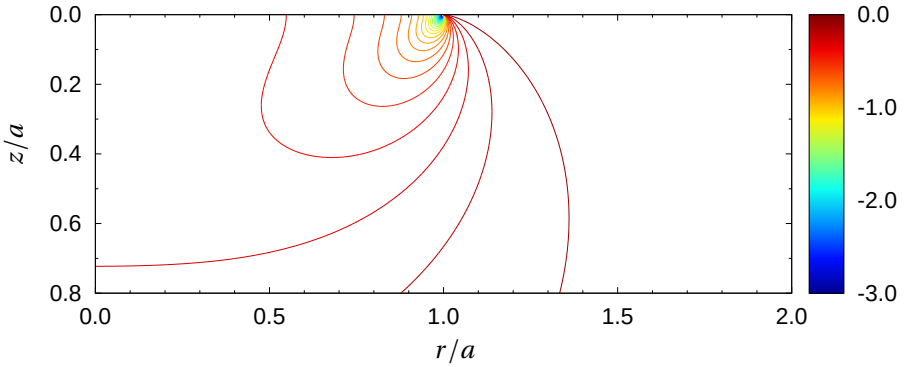


Figure D.3: Contour line of the normal stress ψ_{zz} perpendicular to the free surface under a cylindrical punch for $\nu = 0.22$.

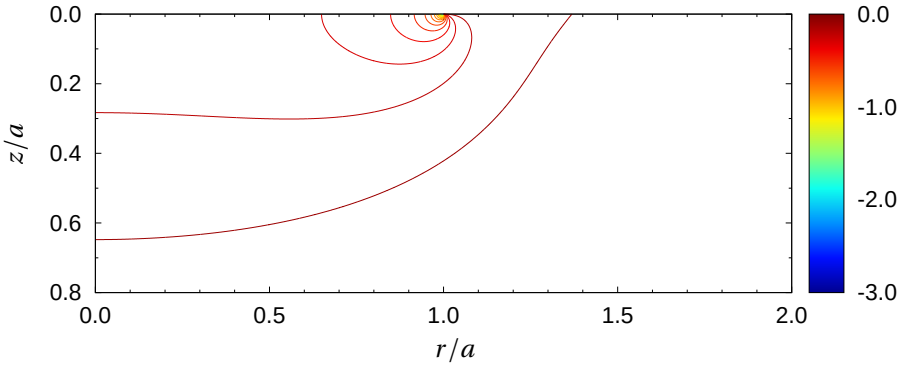


Figure D.4: Contour line of the normal stress $\psi_{\phi\phi}$ in circumferential direction under a cylindrical punch for $\nu = 0.22$.

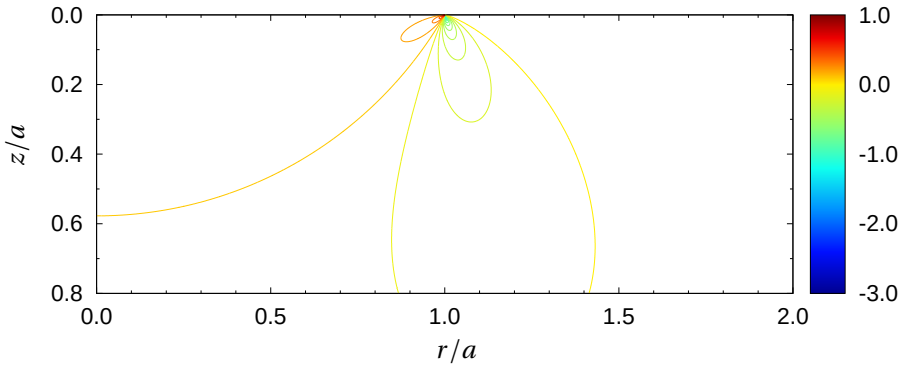


Figure D.5: Contour line of the shear stress ψ_{rz} under a cylindrical punch for $\nu = 0.22$.

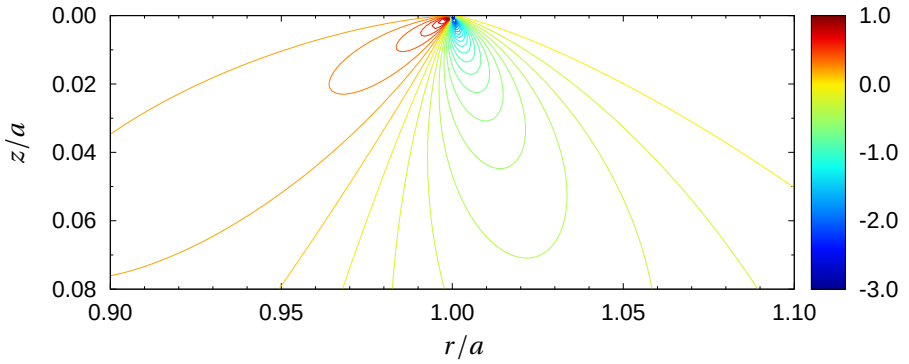


Figure D.6: Contour line of the shear stress ψ_{rz} for $\nu = 0.22$ at the indenter edge ($r/a = 1$).

D.2 Analytical evaluation of average energy release rate

The average energy release rate $\bar{\mathcal{G}}$ required for the evaluation of the hybrid fracture initiation criterion (4.7) of FFM to analyze Hertzian indentation fracture (as done in Sect. 4.2.2) can be computed analytically, after some simplifications, as follows. The mode I stress intensity factor for the edge crack[1] of length z subjected to a single normal load P (Fig. D.7) is

$$K_{\mathrm{I}}(z) = \frac{2P}{\sqrt{\pi z}} \frac{1}{\sqrt{1-(\bar{z}/z)^2}} f_{\mathrm{sf}}(\bar{z}) \tag{D.15}$$

[1] Due to $z \ll r_0$, the curvature $(1/r_0)$ of the ring crack with radius r_0 is negligible. Thus, plain strain conditions can be assumed and corresponding mode I stress intensity factor (independent of r_0) is used.

as given by, e.g., Tada et al. (2000, p. 197) with the factor

$$f_{\mathrm{sf}}(\bar{z}) = 1.3 - 0.3\left(\frac{\bar{z}}{z}\right)^{5/4} \tag{D.16}$$

to incorporate the effect of the free surface.

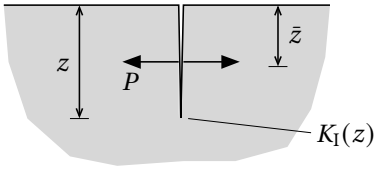


Figure D.7: Edge crack of length z in semi-infinite plane subjected to single normal load P located at position $\bar{z}$.

In order to obtain the stress intensity factor for the pre-fracture radial stress field shown in Fig. D.1 the single load in (D.15) is replaced by the normal stress $\sigma_{rr}(\bar{z})$ which is integrated over the crack length z. Thus, the corresponding mode I stress intensity factor is

$$\begin{aligned} K_{\mathrm{I}}(z) &= \frac{2}{\sqrt{\pi z}} \int_0^z \frac{\sigma_{rr}(\bar{z})}{\sqrt{1-(\bar{z}/z)^2}}\, f_{\mathrm{sf}}(\bar{z})\mathrm{d}\bar{z} \\ &= 2\sqrt{\frac{z}{\pi}} \int_0^z \frac{\sigma_{rr}(\bar{z})}{\sqrt{z^2-\bar{z}^2}}\, f_{\mathrm{sf}}(\bar{z})\mathrm{d}\bar{z} \quad . \end{aligned} \tag{D.17}$$

Replacing $f_{\mathrm{sf}}(\bar{z})$ in (D.16) by the linear approximation

$$f_{\mathrm{sf}}(\bar{z}) \approx 1 + 0.3\left(1 - \frac{\bar{z}}{z}\right) \tag{D.18}$$

the stress intensity factor is

$$K_{\mathrm{I}}(z) = 2\sqrt{\frac{z}{\pi}}\left[\int_0^z \frac{\sigma_{rr}(\bar{z})}{\sqrt{z^2-\bar{z}^2}}\mathrm{d}\bar{z} + \frac{0.3}{z}\int_0^z \frac{\sigma_{rr}(\bar{z})\,(z-\bar{z})}{\sqrt{z^2-\bar{z}^2}}\mathrm{d}\bar{z}\right] \quad . \tag{D.19}$$

By also exploiting the almost linear variation of the radial stress with depth z

$$\sigma_{rr}(\bar{z}) = \sigma_0\left(1 - \frac{\bar{z}}{z^*}\right) \tag{D.20}$$

with $\sigma_0 := \sigma_{rr}(z=0)$ from (D.12a) the mode I stress intensity factor in the present problem is

$$\begin{aligned}
K_\mathrm{I}(z) =& 2\sigma_0\sqrt{\frac{z}{\pi}}\left[\int_0^z \frac{1}{\sqrt{z^2-\bar{z}^2}}\mathrm{d}\bar{z} - \frac{1}{z^*}\int_0^z \frac{\bar{z}}{\sqrt{z^2-\bar{z}^2}}\mathrm{d}\bar{z} + \frac{0.3}{z\,z^*}\int_0^z \frac{(z^*-\bar{z})\,(z-\bar{z})}{\sqrt{z^2-\bar{z}^2}}\mathrm{d}\bar{z}\right] \\
=& 2\sigma_0\sqrt{\frac{z}{\pi}}\left[1.3\underbrace{\int_0^z \frac{\mathrm{d}\bar{z}}{\sqrt{z^2-\bar{z}^2}}}_{=:\,I_1(z)} - \left(\frac{1.3}{z^*}+\frac{0.3}{z}\right)\underbrace{\int_0^z \frac{\bar{z}\mathrm{d}\bar{z}}{\sqrt{z^2-\bar{z}^2}}}_{=:\,I_2(z)} + \frac{0.3}{z\,z^*}\underbrace{\int_0^z \frac{\bar{z}^2\mathrm{d}\bar{z}}{\sqrt{z^2-\bar{z}^2}}}_{=:\,I_3(z)}\right] \\
=& 2\sigma_0\sqrt{\frac{z}{\pi}}\left[\underbrace{1.3\,\frac{\pi}{2}-0.3}_{=:\,\chi_\mathrm{K}} - \underbrace{\left(1.3-0.3\,\frac{\pi}{4}\right)}_{=:\,\eta_\mathrm{K}}\frac{z}{z^*}\right] \quad . \qquad \text{(D.21)}
\end{aligned}$$

The solutions of integrals used in (D.21) are

$$\begin{aligned}
I_1(z) =& \arcsin\left(\frac{\bar{z}}{z}\right)\Big|_0^z = \frac{\pi}{2} \\
I_2(z) =& -\sqrt{z^2-\bar{z}^2}\Big|_0^z = z \\
I_3(z) =& \left[-\frac{\bar{z}}{2}\sqrt{z^2-\bar{z}^2} + \frac{z^2}{2}\arcsin\left(\frac{\bar{z}}{z}\right)\right]\Big|_0^z = z^2\,\frac{\pi}{4} \quad . \qquad \text{(D.22)}
\end{aligned}$$

With

$$K_\mathrm{I}^2(z) = 4\sigma_0^2\,\frac{z}{\pi}\left(\chi_\mathrm{K} - \eta_\mathrm{K}\,\frac{z}{z^*}\right)^2 = \frac{4}{\pi}\,\sigma_0^2\left(\chi_\mathrm{K}^2\,z - 2\chi_\mathrm{K}\,\eta_\mathrm{K}\,\frac{z^2}{z^*} + \eta_\mathrm{K}^2\,\frac{z^3}{z^{*2}}\right) \qquad \text{(D.23)}$$

the average energy release rate is

$$\begin{aligned}
\bar{\mathcal{G}}(z_0) =& \frac{1-\nu^2}{E}\,\frac{1}{z_0}\int_0^{z_0} K_\mathrm{I}^2(z)\mathrm{d}z \\
=& \frac{1-\nu^2}{E}\,\frac{1}{z_0}\,\frac{4}{\pi}\,\sigma_0^2\left(\frac{\chi_\mathrm{K}^2}{2}\,z^2 - \frac{2}{3}\chi_\mathrm{K}\,\eta_\mathrm{K}\,\frac{z^3}{z^*} + \frac{\eta_\mathrm{K}^2}{4}\,\frac{z^4}{z^{*2}}\right)\Bigg|_0^{z_0} \\
=& 4\frac{E}{\pi^3}\,\frac{d^2}{a^2}\,\frac{(1-2\nu)^2}{1-\nu^2}\left(\frac{a}{r_0}\right)^4\left(\frac{\chi_\mathrm{K}^2}{2}\,z_0 - \frac{2}{3}\chi_\mathrm{K}\,\eta_\mathrm{K}\,\frac{z_0^2}{z^*} + \frac{\eta_\mathrm{K}^2}{4}\,\frac{z_0^3}{z^{*2}}\right) \quad , \qquad \text{(D.24)}
\end{aligned}$$

which can also be written as

$$\bar{\mathcal{G}} = E\,a\left(\frac{d}{a}\right)^2\,\phi\Big(\frac{r_0}{a},\frac{z_0}{a},\nu\Big) \qquad \text{(D.25)}$$

by means of the dimensionless function

$$\phi\Big(\frac{r_0}{a},\frac{z_0}{a},\nu\Big)=\frac{4}{\pi^3}\,\frac{(1-2\nu)^2}{1-\nu^2}\,\frac{1}{(r_0/a)^4}\left[\frac{\chi_{\mathrm{K}}^2}{2}\,\frac{z_0}{a}-\frac{2}{3}\chi_{\mathrm{K}}\,\eta_{\mathrm{K}}\,\frac{1}{z^*/a}\,\Big(\frac{z_0}{a}\Big)^2+\frac{\eta_{\mathrm{K}}^2}{4}\,\frac{1}{(z^*/a)^2}\,\Big(\frac{z_0}{a}\Big)^3\right] \quad . \tag{D.26}$$

Due to the square of K_{I} in (D.24) a negative stress intensity factor would also contribute to the average energy release rate. Thus only positive stress intensity factors are evaluated. Regarding (D.21) the stress intensity factor is positive only for

$$\frac{z}{z^*}<\frac{\chi_{\mathrm{K}}}{\eta_{\mathrm{K}}}\approx 1.637 \qquad \text{or} \qquad z<z^*\,\chi_{\mathrm{K}}/\eta_{\mathrm{K}}=:z^{**} \quad . \tag{D.27}$$

For $z>z^{**}$ the average energy release rate is only evaluated in the range where the stress intensity factor K_{I} is positive, i.e. from $z=0$ to $z=z^{**}$, so

$$\begin{aligned}\bar{\mathcal{G}}(z_0)&=\frac{1-\nu^2}{E}\,\frac{1}{z_0}\int_0^{z^{**}}K_{\mathrm{I}}^2(z)\mathrm{d}z\\&=\frac{1-\nu^2}{E}\,\frac{1}{z_0}\,\frac{4}{\pi}\,\sigma_0^2\left(\frac{\chi_{\mathrm{K}}^2}{2}\,z^2-\frac{2}{3}\chi_{\mathrm{K}}\,\eta_{\mathrm{K}}\,\frac{z^3}{z^*}+\frac{\eta_{\mathrm{K}}^2}{4}\,\frac{z^4}{z^{*2}}\right)\Bigg|_0^{z^{**}}\\&=4\frac{E}{\pi^3}\,\frac{d^2}{a^2}\,\frac{(1-2\nu)^2}{1-\nu^2}\,\Big(\frac{a}{r_0}\Big)^4\left(\frac{1}{2}\,\frac{\chi_{\mathrm{K}}^4}{\eta_{\mathrm{K}}^2}\,z^{*2}-\frac{2}{3}\,\frac{\chi_{\mathrm{K}}^4}{\eta_{\mathrm{K}}^2}\,z^{*2}+\frac{1}{4}\,\frac{\chi_{\mathrm{K}}^4}{\eta_{\mathrm{K}}^2}\,z^{*2}\right) \quad .\end{aligned} \tag{D.28}$$

In this case the dimensionless function ϕ in (D.25) is

$$\phi\Big(\frac{r_0}{a},\frac{z_0}{a},\nu\Big)=\frac{4}{\pi^3}\,\frac{(1-2\nu)^2}{1-\nu^2}\,\frac{1}{(r_0/a)^4}\,\frac{1}{z_0/a}\,\frac{1}{12}\left(\frac{\chi_{\mathrm{K}}^2}{\eta_{\mathrm{K}}}\,\frac{z^*}{a}\right)^2 \quad . \tag{D.29}$$

D.3 Effect of free surface on average energy release rate

As discussed in Section 4.2.2, previous analyses of the indentation problem, e.g. Frank and Lawn (1967) and Warren (1978), evaluate the stress intensity factor in the present problem (4.19) by using

$$f_{\mathrm{sf}}=1 \quad , \tag{D.30}$$

thus neglecting the effect of the free surface. Its influence on crack initiation is briefly discussed here. According to Appendix D.2, the stress intensity factor using (D.30) is

$$K_{\mathrm{I}}(z)=2\sigma_0\sqrt{\frac{z}{\pi}}\int_0^z\frac{1-\bar{z}/z^*}{\sqrt{z^2-\bar{z}^2}}\mathrm{d}\bar{z}=2\sigma_0\sqrt{\frac{z}{\pi}}\Bigg[\underbrace{\frac{\pi}{2}}_{=:\,\tilde{\chi}_{\mathrm{K}}}-\underbrace{1}_{=:\,\tilde{\eta}_{\mathrm{K}}}\,\frac{z}{z^*}\Bigg] \quad , \tag{D.31}$$

which is positive for

$$\frac{z}{z^*} < \frac{\tilde{\chi}_\mathrm{K}}{\tilde{\eta}_\mathrm{K}} \approx 1.571 \qquad \text{or} \qquad z < z^*\,\tilde{\chi}_\mathrm{K}/\tilde{\eta}_\mathrm{K} =: \tilde{z}^{**} \quad . \tag{D.32}$$

Finally, the average energy release rate is represented analogous to (4.23) by the dimensionless function

$$\tilde{\phi} = \begin{cases} \dfrac{4}{\pi^3}\,\dfrac{(1-2\nu)^2}{1-\nu^2}\,\dfrac{1}{(r_0/a)^4}\left[\dfrac{\pi^2}{8}\,\dfrac{z_0}{a} - \dfrac{\pi}{3}\,\dfrac{(z_0/a)^2}{z^*/a} + \dfrac{1}{4}\,\dfrac{(z_0/a)^3}{(z^*/a)^2}\right] & \text{for } z \le \tilde{z}^{**} \\ \dfrac{\pi}{48}\,\dfrac{(1-2\nu)^2}{1-\nu^2}\,\dfrac{1}{(r_0/a)^4}\,\dfrac{(z^*/a)^2}{z_0/a} & \text{for } z > \tilde{z}^{**} \quad . \end{cases} \tag{D.33}$$

Figure D.8 shows the critical load according to the energy part of the hybrid criterion in terms of the dimensionless quantity $1/\sqrt{\tilde{\phi}}$. It is significantly larger in the whole range r_0/a in the case that the free surface is neglected for the stress intensity factor evaluation according to (4.19). Especially for short cracks the deviation is large.

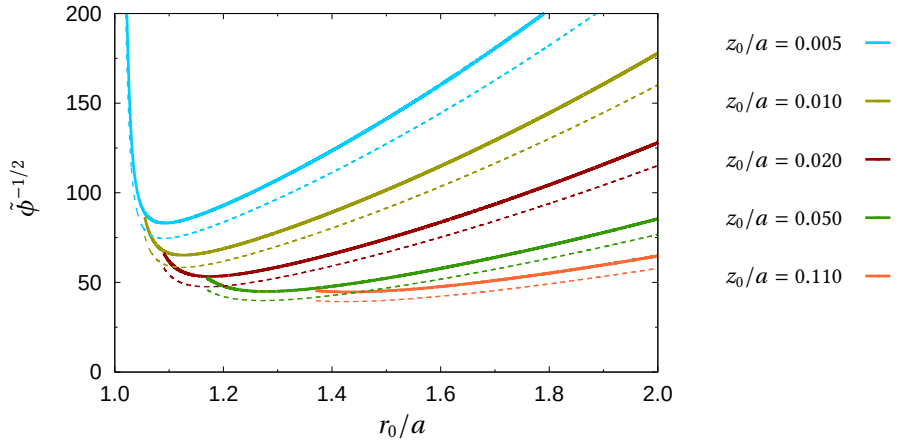


Figure D.8: Dimensionless function $\tilde{\phi}^{-1/2}$ given in (D.33) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$. Solid lines indicate results for energy release rate with $f_\mathrm{sf} = 1$ in comparison with the energy release rate accounting for the free surface $f_\mathrm{sf}(z)$ used in Sect. 4.2 (dashed lines).

Accordingly, as depicted in Figs. D.9a and b the critical load for crack initiation is larger if the effect of the free surface is neglected. Moreover, the stationary value for small indenter radii in Fig. D.9b is somewhat larger. Figures D.9c and d show that initial cracks occur at a larger radial position r_0/a and are somewhat longer where the stationary value is not reached. Interestingly, neglecting the effect of the free surface influences the range where the saturation value occurs as well as the saturation value of the initial crack length z_0/a (which is slightly shorter with $z_0/a \approx 7.8 \times 10^{-2}$) as shown by Figs. D.9c and d, but does not influence the saturation value of the radial position of the crack ($r_0/a \approx 1.35$).

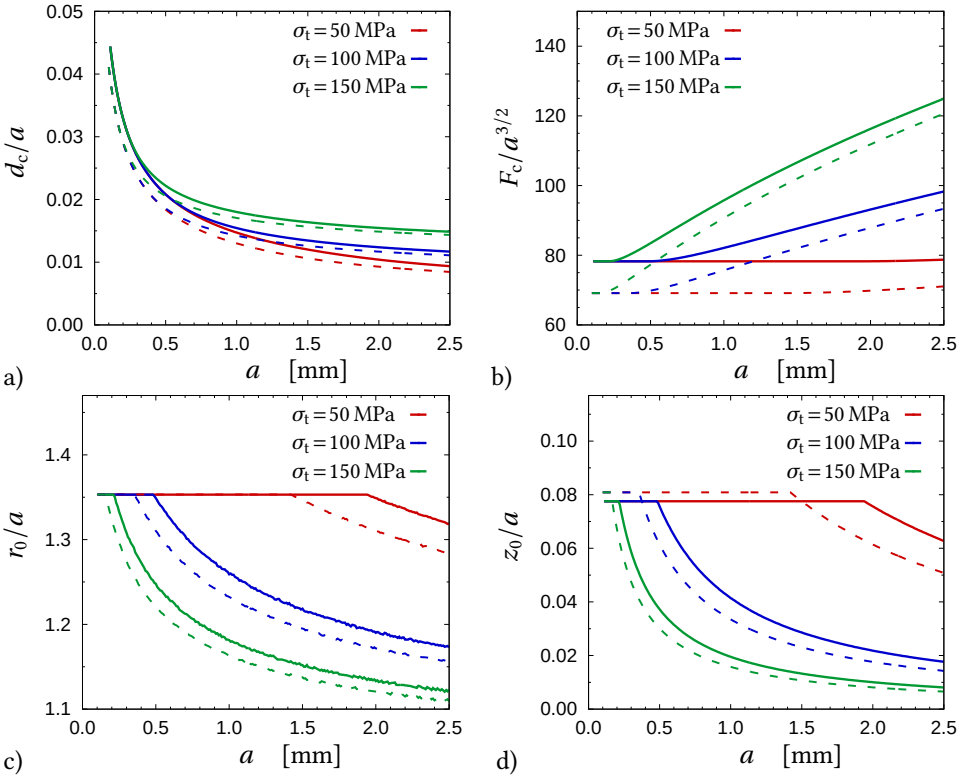


Figure D.9: Effect of neglecting the free surface for the average energy release rate evaluation by using (D.30). Variation of a) normalized critical indenter displacement d_c/a, b) critical load $F_c/a^{3/2}$, c) normalized crack radius r_0/a and d) normalized crack length z_0/a with indenter radius a for different tensile strengths $\sigma_t = \{50,\ 100,\ 150\}$ MPa. Solid lines indicate results for $f_{sf} = 1$ in comparison with the evaluation in Sect. 4.2 (dashed lines).

D.4 Effect of local stress criterion

For completeness of the evaluation, results obtained for the critical indenter displacement d_c/a, the apparent critical load $F_c/a^{3/2}$, the crack radius r_0/a and the initial crack length z_0/a (analogous to Figs. 4.13, 4.14 and 4.15) are presented when instead of the average stress criterion (4.6) the local stress criterion (4.5) is used. For instance, Figure D.10a shows that a higher critical loading (in terms of a larger indenter displacement d_c/a) is predicted by utilizing the local stress criterion instead of the average stress criterion (indicated by dashed lines). In accordance with this, Figure D.10b indicates that the use of the local stress criterion leads to higher apparent critical loads with a value of about 70 MN/m$^{3/2}$ for small indenter radii.

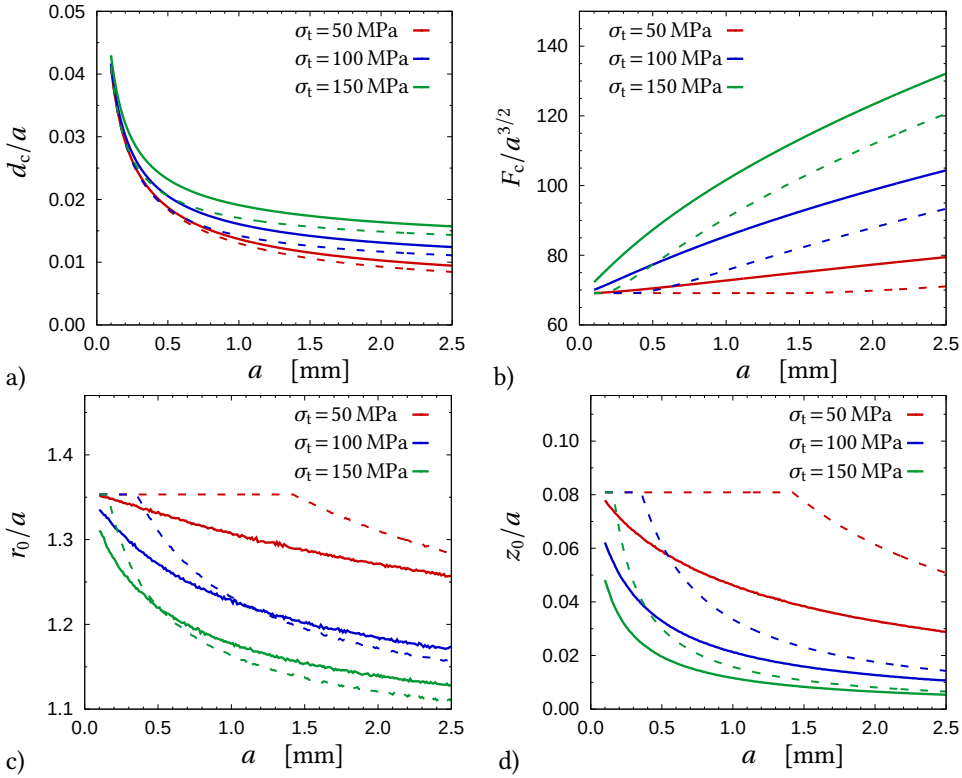


Figure D.10: Effect of local stress criterion (pointwise evaluation) on indentation fracture initiation. Variation of a) normalized critical indenter displacement d_c/a, b) critical load $F_c/a^{3/2}$, c) normalized crack radius r_0/a and d) normalized crack length z_0/a with indenter radius a for different tensile strengths $\sigma_t = \{50,\ 100,\ 150\}$ MPa. Solid lines indicate results for local stress criterion in comparison with average stress criterion used in Sect. 4.2 (dashed lines).

However, in contrast to the average stress criterion, the local stress criterion obviously does not lead to a stationary value (which might indicate the Auerbach range as discussed in Sect. 4.2.4) here. The critical load to fulfill the local stress criterion, especially for long cracks, is significantly higher (see Fig. 4.5) compared to the average stress criterion (Fig. 4.6) and thus it is more dominant for the crack initiation process. Therefore, crack initiation is controlled by both the stress and the energy criterion for all indenter radii $a \in [0.1,\ 2.5]$ mm. The absence of a saturation regime for small indenter radii is even more pronounced in Figs. D.10c and d. Figure D.10c shows the variation of the normalized radial crack position r_0/a and Fig. D.10d shows the corresponding variation of the normalized crack length z_0/a with the indenter radius a. Both figures indicate large deviations between both stress criteria, especially for small tensile strengths. Typically, by using the local stress criterion shorter cracks (Fig. D.10d) which initiate closer to the indenter (Fig. D.10c) are predicted.

Analogous to the analysis of Sect. 4.2.4, by using the dimensionless parameter $\sqrt{l_c/(a\,\varepsilon_c)}$ introduced in Sect. 4.2.3 the range where only the energy criterion controls crack initiation ($d_c/d_\pi \approx 1$) with the stress criterion being oversatisfied ($d_c/d_\sigma > 1$) is $\sqrt{l_c/(a\,\varepsilon_c)} \gtrapprox 2.097$, which is much larger than the corresponding parameter obtained for the average stress evaluation ($\sqrt{l_c/(a\,\varepsilon_c)} \gtrapprox 0.450$). For the material data given by (4.2) and a tensile strength of $\sigma_t = 50\,\mathrm{MPa}$ this corresponds to the indenter radius $a \lessapprox 6.5 \times 10^{-2}\,\mathrm{mm}$ which is much smaller than the indenter size considered here. So, from a theoretical point of view, the saturation regime also exists when using the local stress criterion as is shown next. Since for small indenter radii the influence of e.g. surface flaws is more severe, a hypothetical tensile strength of $\sigma_t = 25\,\mathrm{MPa}$ is regarded instead. The saturation regime occurs for indenter radii

$$a \lessapprox \frac{1}{2.097^2}\,\frac{\mathcal{G}_c\,E}{\sigma_t^2} = 0.26\,\mathrm{mm} \quad . \tag{D.34}$$

Figure D.11 shows the appearance of saturation values for crack initiation at the radial position $r_0/a \approx 1.35$ with the crack length $z_0/a \approx 8.1 \times 10^{-2}$. These saturation values are identical to the values obtained for the average stress criterion in Sect. 4.2.4, since they correspond to the local maximum of $\phi(r_0/a, z_0/a, \nu)$ in Fig. 4.8 and thus are independent of the stress part evaluation of the hybrid fracture initiation criterion (2.160).

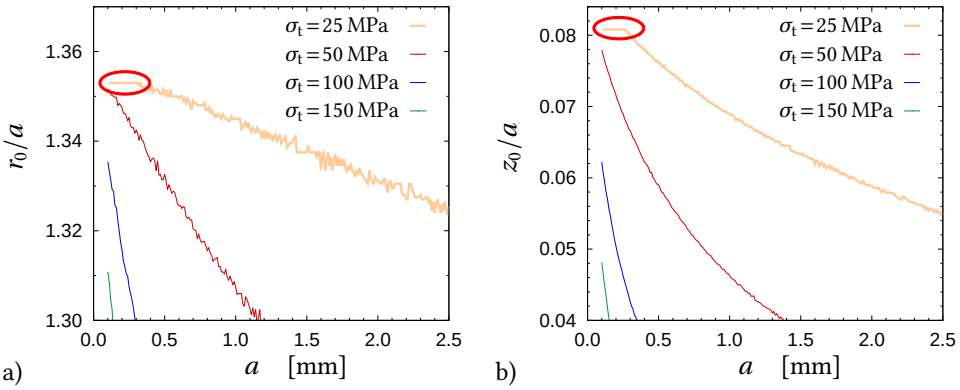


Figure D.11: Variation of a) normalized crack radius r_0/a and b) normalized crack length z_0/a with indenter radius a for different tensile strengths $\sigma_t = \{25, 50, 100, 150\}$ MPa. Saturation regime which is present for $\sqrt{l_c/(a\,\varepsilon_c)} \gtrapprox 2.097$ is indicated by red color.

D.5 Effect of cutting negative stress in average stress criterion

By considering the average stress criterion (4.6) the question arises how negative stresses should contribute to the average. While the evaluation in Sect. 4.2.1 does not distinguish between positive and negative stresses — only negative average values are excluded in (4.12) — there is also the possibility to cut negative stress values before taking the

average. By also exploiting the linear stress distribution in the present problem (4.13) one obtains the dimensionless function (by using the substitution $\zeta = z/a$)

$$\overline{\psi}_{\mathrm{pos}} := \frac{1}{z_0/a} \int_0^{\widehat{\zeta}} \psi_{rr}\left(\frac{r_0}{a}, \zeta, \nu\right) \mathrm{d}\zeta \approx \begin{cases} \psi_0 \left(1 - \frac{1}{2}\frac{z_0/a}{z^*/a}\right) & z_0/a \le z^*/a \\ \frac{1}{2}\psi_0 \frac{z^*/a}{z_0/a} & z_0/a > z^*/a \end{cases} \tag{D.35}$$

with the upper integration bound $\widehat{\zeta}$ depending on the crack length z_0/a

$$\widehat{\zeta} := \begin{cases} z^*/a & \text{for} \quad z_0/a > z^*/a \\ z_0/a & \text{for} \quad z_0/a \le z^*/a \end{cases} . \tag{D.36}$$

Obviously, excluding negative stress contribution from the integral expression as done in (D.35) only takes effect for (still hypothetical) cracks which are long enough to reach the zone where the radial stress becomes negative, see Fig. 4.4. This is the case, e.g., for long cracks (large z_0/a) at small distances from the indenter (small r_0/a). Shorter cracks are able to appear closer to the indenter without being affected by negative stresses, as depicted in Fig. D.12. Since the critical positions and crack lengths where fracture initiation occurs for present material parameters and indenter radii, see Sect. 4.2.4, are within this range the initiation process is not influenced by applying (D.35). Thus, results obtained by cutting negative stress contributions before taking the average are identical with the results shown in Figs. 4.13–4.15.

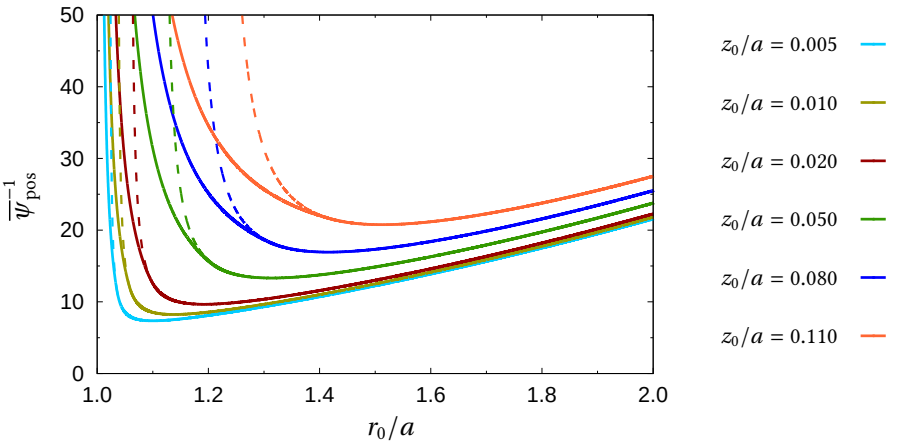


Figure D.12: Dimensionless function $\overline{\psi}^{-1}_{\mathrm{pos}}$ given in (D.35) (solid lines) in comparison with corresponding function from average stress criterion $\overline{\psi}^{-1}$ used in Sect. 4.2.1 (dashed lines, cf. Fig. 4.6) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.00 \ldots 0.11\}$.

D.6 Verification of linear stress approximation

The analysis of indentation fracture initiation in Sect. 4.2 is carried out by exploiting the almost linearly decreasing radial stress (D.20) with increasing depth z (cf. Fig. 4.3b), which finally leads to simple analytical expressions representing the local stress criterion (4.14), the average stress criterion (4.15) and the energy criterion (4.26). In the following the accuracy of these expressions is compared with exact values or values obtained by numerical integration.

Pointwise evaluation of local stress criterion

Instead of utilizing the linear approximation (4.13) the local stress criterion (4.10) is evaluated pointwise. In this case the critical load to initiate fracture is larger as depicted in Fig. D.13. Especially for long cracks (at a large distance from the indenter) the deviation between linear approximation and pointwise evaluation of the local stress criterion is large. Since the radial stress decreases monotonically with increasing z as already pointed out in Sect. 4.2.1 the local criterion takes only the stress at the end of the hypothetically formed crack at $z=z_0$ into account. Figure D.14a shows the relative error of the linear stress approximation (related to the radial stress at the free surface)

$$\delta_\psi := \frac{\psi_{\text{lin}} - \psi_{rr}}{\psi_0} \times 100\% = \left(1 - \frac{z_0/a}{z^*/a} - \frac{\psi_{rr}}{\psi_0}\right) \times 100\% \tag{D.37}$$

with depth z_0/a. Only at the surface ($z_0/a = 0$) and at the depth with zero stress ($z_0/a = z^*/a$) the linear stress approximation (4.13) is equal to the (exact) radial stress. In-between, for $0 < z_0 < z^*$ (where the radial stress is positive), the linear stress approximation overestimates the real stress, thus the critical load for fracture initiation is predicted too small (dashed lines in Fig. D.13).

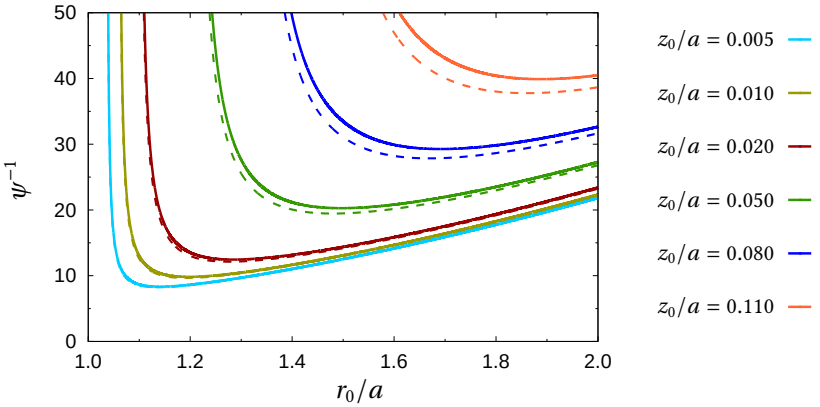


Figure D.13: Dimensionless function ψ^{-1} representing the critical load for the local stress criterion. Comparison of (exact) pointwise evaluation (solid lines) with evaluation by using linear stress approximation (dashed lines, cf. Fig. 4.5) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$.

As depicted in Fig. D.14a, the maximum deviation between the exact stress and its linear approximation is located at $z_0/z^* \approx 0.6$ and the maximum relative error increases with the radial position up to $r_0/a \approx 1.5$. In conjunction with the ratio z_0/z^* shown in Fig. D.14b it is reasonable that deviations for long cracks are larger as depicted in Fig. D.13. For completeness, Table D.1 lists the radial position r_0/a where the stress for a fixed crack length z_0/a becomes negative.

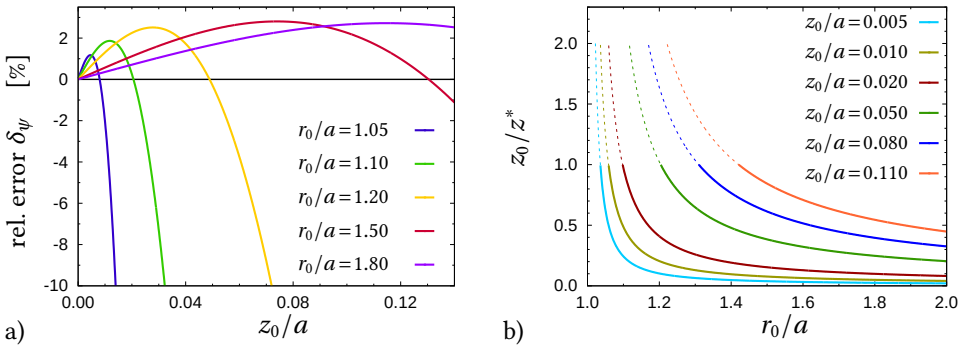


Figure D.14: a) Variation of relative error of linear stress approximation with crack length z_0/a for a set of different radial positions $r_0/a = \{1.05 \ldots 1.8\}$. b) Ratio z_0/z^* varying with radial crack position r_0/a for a set of crack lengths $z_0/a = \{0.005 \ldots 0.11\}$. Note that $z_0/z^* < 1$ for the local stress criterion (solid lines) and $z_0/z^* < 2$ for the average stress criterion (dashed lines).

z_0/a	r_0/a where $z_0 = z^*$	r_0/a where $z_0 = 2z^*$	r_0/a where $z_0 = z^{**}$
0.005	1.036	1.022	1.025
0.010	1.059	1.036	1.041
0.020	1.098	1.059	1.068
0.050	1.204	1.116	1.136
0.080	1.310	1.169	1.200
0.110	1.421	1.221	1.264

Table D.1: Radial positions where radial stress ψ_{rr}, average radial stress $\overline{\psi}_{rr}$ and stress intensity factor K_I become negative for a set of crack lengths $z_0/a = \{0.005 \ldots 0.11\}$.

Pointwise evaluation of average stress criterion

By contrast, to verify the linear approximation for the average stress criterion the dimensionless function $\overline{\psi}$ (4.12) has to be evaluated numerically. Either the trapezoidal rule

$$\overline{\psi}_{\text{trpz}} := \left\langle \sum_{i=1}^{n-1} \frac{x_{i+1} - x_i}{2} \left[\psi_{rr}\left(\frac{r_0}{a}, \frac{z_0}{a}\, x_i, \nu\right) + \psi_{rr}\left(\frac{r_0}{a}, \frac{z_0}{a}\, x_{i+1}, \nu\right) \right] \right\rangle \tag{D.38}$$

with $n - 1 \geq 1$ partitions of the integration interval

$$x_1 = 0 \;<\; x_2 \;<\; \ldots \;<\; x_{n-1} \;<\; x_n = 1 \tag{D.39}$$

or, alternatively, Gaussian quadrature

$$\overline{\psi}_{\text{Gauss}} := \left\langle \sum_{i=1}^{n} \psi_{rr}\left(\frac{r_0}{a}, \frac{z_0}{a}\, x_i, \nu\right) w_i \right\rangle \tag{D.40}$$

evaluated at n integration points adapted to the interval $0 < x_i < 1$ and weights

$$\sum_{i=1}^{n} w_i = 1 \tag{D.41}$$

is used.

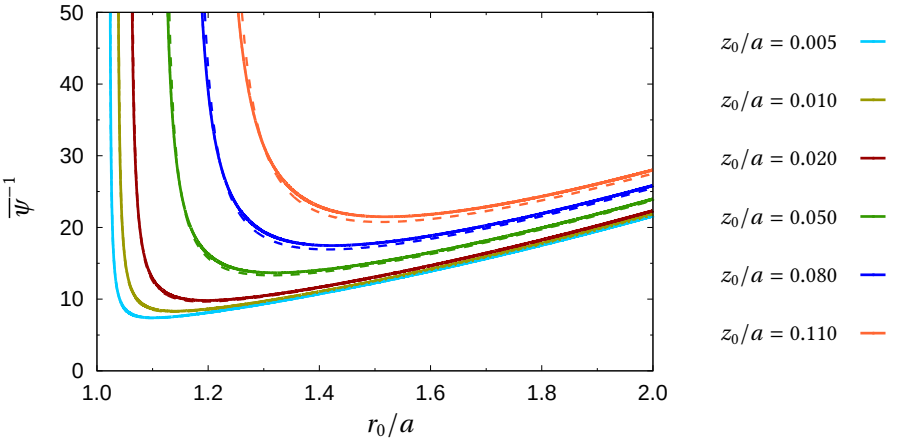


Figure D.15: Dimensionless function $\overline{\psi}^{-1}$ representing the critical load for the local stress criterion. Comparison of results obtained by Gaussian quadrature with $n = 20$ (solid lines) with evaluation by using linear stress approximation (dashed lines, cf. Fig. 4.6) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$.

Figure D.15 shows only small deviations between results obtained by Gaussian quadrature (20 integration points are used here) and the linear stress approximation. It should be

noted that deviations between the pointwise evaluation and linear approximation are much smaller compared to the local evaluation of the stress criterion discussed above (Fig. D.13) since the average of the deviation is smaller than the local value and stresses are partially overestimated (for $0 < z < z^*$) and underestimated (for $z > z^*$), see Fig. D.14.

Pointwise evaluation of average energy release rate

In a similar way, the mode I stress intensity factor

$$K_{\mathrm{I}}(z) = E\,\frac{d}{a}\,\sqrt{a}\,\boldsymbol{\mathcal{K}}_{\mathrm{I}}\Big(\frac{r_0}{a}, \frac{z}{a}, \nu\Big) \tag{D.42}$$

with the dimensionless function (using $\bar{\zeta} := \bar{z}/a$ and thus $\mathrm{d}\bar{z} = a\mathrm{d}\bar{\zeta}$)

$$\boldsymbol{\mathcal{K}}_{\mathrm{I}}(z) = \frac{2}{\sqrt{\pi}}\sqrt{z/a}\int_0^{z/a} \frac{f_{\mathrm{sf}}(\bar{\zeta})}{\sqrt{(z/a)^2 - \bar{\zeta}^2}}\,\psi_{rr}\Big(\frac{r_0}{a}, \bar{\zeta}, \nu\Big)\,\mathrm{d}\bar{\zeta} \tag{D.43}$$

and the average energy release rate (4.23) are evaluated by numerical integration. However, since the (dimensionless) stress intensity factor (D.43) is undefined in the case that load is applied right at at the crack tip $\bar{\zeta} \to z/a$ (see Fig. D.7) the integral is evaluated at n integration points by Gaussian quadrature (which avoids evaluation of the singular endpoint) according to

$$\boldsymbol{\mathcal{K}}_{\mathrm{I\,Gauss}}(z) := \frac{2}{\sqrt{\pi}}\sqrt{\frac{z}{a}}\sum_{i=1}^{n} \frac{1.3 - 0.3\,x_i^{5/4}}{\sqrt{1 - x_i}}\,\psi_{rr}\Big(\frac{r_0}{a}, \frac{z}{a}\,x_i, \nu\Big) \tag{D.44}$$

with integration points and weights adapted to the interval $0 < x_i < 1$ as done above. The numerical integration of the average energy release rate is

$$\phi_{\mathrm{Gauss}} := (1 - \nu^2)\sum_{j=1}^{m} \Big\langle \boldsymbol{\mathcal{K}}_{\mathrm{I}}\Big(\frac{r_0}{a}, \frac{z_0}{a}\,x_j, \nu\Big)\Big\rangle^2\, w_j \quad , \tag{D.45}$$

which requires the numerical evaluation of the stress intensity factor according to (D.44) at m integration points, so $m + 1$ numerical integrations are necessary for each r_0/a- and z_0/a-value to evaluate the average energy release rate.

In order to verify the accuracy of Gaussian quadrature here a convergence study is done by comparing the numerical integration (D.45) of the simplified stress intensity factor (based on the linear stress approximation)

$$\boldsymbol{\mathcal{K}}_{\mathrm{I\,simpl}}(z) := \frac{2}{\sqrt{\pi}}\sqrt{\frac{z}{a}}\sum_{i=1}^{n} \frac{1.3 - 0.3\,x_i}{\sqrt{1 - x_i}}\,\psi_0\left(1 - \frac{z/a}{z^*/a}\,x_i\right) \tag{D.46}$$

with its analytical solution given in App. D.2. Figure D.16a shows the convergence of numerical results in terms of the critical loading $\phi^{-1/2}$ for a short crack ($z_0/a = 0.005$) and a long crack ($z_0/a = 0.11$).

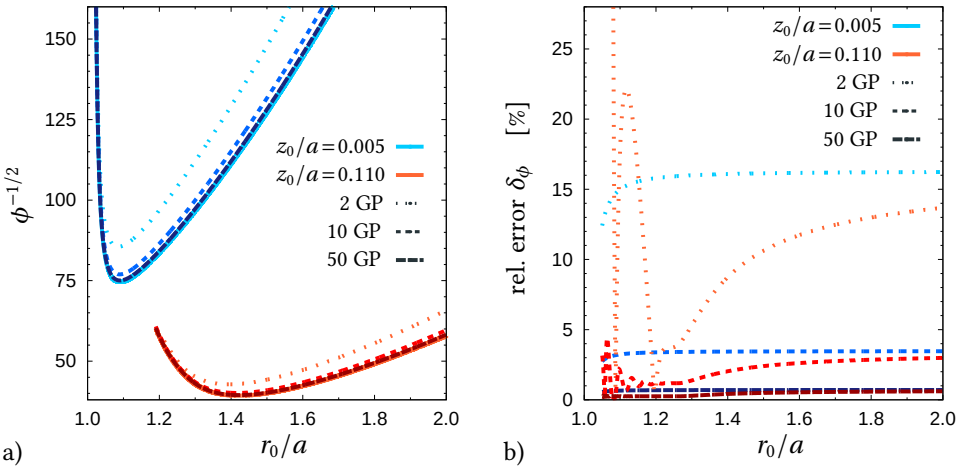


Figure D.16: Convergence study to determine accuracy of numerical integration of average energy release rate for short crack (blue/cyan) and long crack (red/orange): a) critical load in terms of $\phi^{-1/2}$ for linear approximation (solid lines) in comparison with Gaussian quadrature (dashed lines), b) relative error of Gaussian quadrature according to (D.47).

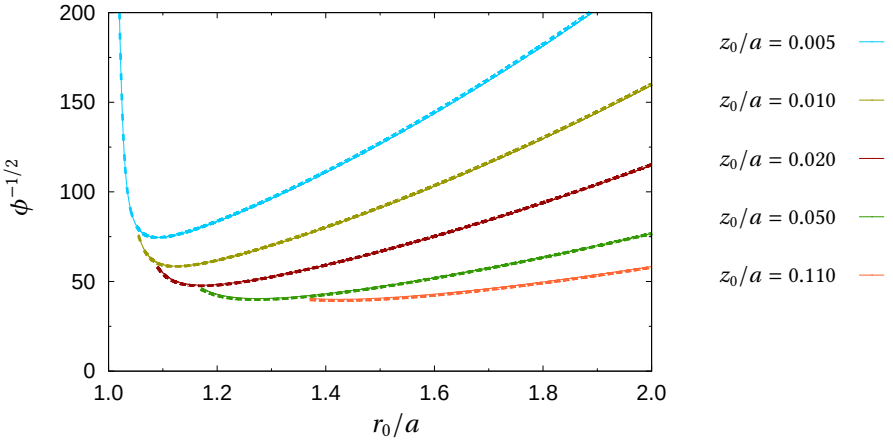


Figure D.17: Dimensionless function $\phi^{-1/2}_{\text{Gauss}}$ given in (D.45) for $\nu = 0.22$ and a set of varying crack lengths $z_0/a = \{0.005 \ldots 0.11\}$ (solid lines) in comparison with the analytical (but simplified) energy release rate used in Sect. 4.2 (dashed lines).

In both cases 50 integration points seem to be sufficient, which is also shown in Fig. D.16b in terms of the relative error

$$\delta_\phi := \frac{|\phi_{\text{Gauss}}^{-1/2} - \phi_{\text{ana}}^{-1/2}|}{\phi_{\text{ana}}^{-1/2}} \times 100\% \quad , \tag{D.47}$$

which is less than 1 %.

Finally, pointwise evaluation of the average energy release rate by using Gaussian quadrature according to (D.45) leads to the results depicted in Fig. D.17. They are almost identical with the results obtained by the simple analytical expression (4.26), which is used for the FFM study in Sect. 4.2. By taking Figures D.15 and D.17 into account one can conclude that the results discussed in Sect. 4.2.4 are not significantly influenced by the linear stress approximation.

D.7 Effect of loading type on average energy release rate

Be C_0 the compliance of some uncracked linear elastic body and C_1 with $C_1 > C_0$ that of the same body containing a small, yet finite, crack as considered in the hybrid fracture criterion. When the body is loaded by a prescribed displacement d, the change of the total potential energy Π upon formation of the crack is

$$\Delta\Pi_{\text{d}} = \frac{d}{2}\,(F_1 - F_0) \tag{D.48}$$

where F_0 and F_1 with $F_1 < F_0$ are the corresponding forces on the uncracked and on the cracked body, respectively. When, instead, the body is loaded by a prescribed force F (dead loading) one has

$$\Delta\Pi_{\text{F}} = \underbrace{\frac{F}{2}\,(d_1 - d_0)}_{\text{strain energy}} - \underbrace{F\,(d_1 - d_0)}_{\text{potential of } F} = -\frac{F}{2}\,(d_1 - d_0) \tag{D.49}$$

where d_0 and d_1 with $d_1 > d_0$ denote the displacements at the point where F acts on the uncracked and the cracked body, respectively. In terms of the compliances C_0 and C_1 one obtains

$$d_0 = C_0\,F_0 \qquad \text{and} \qquad d_1 = C_1\,F_1 \quad . \tag{D.50}$$

With $d_0 = d_1 = d$ in case of a prescribed displacement the change of the total potential energy (D.48) can thus be written as

$$\Delta\Pi_{\text{d}} = \frac{d}{2}\left(\frac{d}{C_1} - \frac{d}{C_0}\right) = \frac{d^2}{2}\left(\frac{1}{C_1} - \frac{1}{C_0}\right) \tag{D.51}$$

whereas in case of a prescribed force with $F_0 = F_1 = F$ one obtains from (D.49)

$$\Delta\Pi_{\mathrm{F}} = -\frac{F}{2}\,(C_1\,F - C_0\,F) = -\frac{F^2}{2}\,(C_1 - C_0) \quad . \tag{D.52}$$

In the uncracked configuration, however, one has

$$d_0 = d = C_0\,F_0 = C_0\,F \quad . \tag{D.53}$$

Inserting this into (D.51) yields

$$\Delta\Pi_{\mathrm{d}} = \frac{F^2}{2}\,C_0^2\left(\frac{1}{C_1} - \frac{1}{C_0}\right) = -\frac{F^2}{2}\,(C_1 - C_0)\,\underbrace{\frac{C_0}{C_1}}_{<\,1} \tag{D.54}$$

which, by comparison with (D.52), shows that the amount of energy released upon finite crack formation is larger in case of a prescribed force than in case of a prescribed displacement, i.e.

$$-\,\Delta\Pi_{\mathrm{F}} > -\Delta\Pi_{\mathrm{d}} \quad . \tag{D.55}$$

The same finding is obtained by Rosendahl et al. (2017).

List of Figures

List of Tables

Bibliography

Abaqus (2013). *ABAQUS 6.13 online documentation* (see page 170)

Achenbach, J. D. (1973). *Wave propagation in elastic solids*. Amsterdam: North-Holland (see page 33)

Alessi, R., Ambati, M., Gerasimov, T., Vidoli, S., and De Lorenzis, L. (2018a). Comparison of phase-field models of fracture coupled with plasticity. In: *Advances in Computational Plasticity: A Book in Honour of D. Roger J. Owen*. Springer International Publishing, 1–21. DOI: 10.1007/978-3-319-60885-3_1 (see page 4)

Alessi, R., Vidoli, S., and De Lorenzis, L. (2018b). A phenomenological approach to fatigue with a variational phase-field model: The one-dimensional case. *Engineering Fracture Mechanics* 190, 53–73. DOI: 10.1016/j.engfracmech.2017.11.036 (see page 4)

Altenbach, H. and Eremeyev, V. A. (2013). *Generalized continua – from the theory to engineering applications*. Vienna: Springer. DOI: 10.1007/978-3-7091-1371-4 (see page 20)

Ambati, M., Gerasimov, T., and De Lorenzis, L. (2015a). A review on phase-field models of brittle fracture and a new fast hybrid formulation. *Computational Mechanics* 55 (2), 383–405. DOI: 10.1007/s00466-014-1109-y (see pages 75, 78, 112, 118, 122, 128, 130, 184)

Ambati, M., Gerasimov, T., and De Lorenzis, L. (2015b). Phase-field modeling of ductile fracture. *Computational Mechanics* 55 (5), 1017–1040. DOI: 10.1007/s00466-015-1151-4 (see pages 4, 122)

Ambrosio, L. (1989). Variational problems in SBV and image segmentation. *Acta Applicandae Mathematicae* 17, 1–40. DOI: 10.1007/BF00052492 (see pages 59, 63)

Ambrosio, L. and Braides, A. (1995). Energies in SBV and variational models in fracture mechanics. In: *Homogenization and Applications to Material Sciences*, 1–22 (see page 53)

Ambrosio, L. and Tortorelli, V. M. (1990). Approximation of functional depending on jumps by elliptic functional via Γ-convergence. *Communications on Pure and Applied Mathematics* 43 (8), 999–1036. DOI: 10.1002/cpa.3160430805 (see pages 62, 63)

Ambrosio, L. and Tortorelli, V. M. (1992). On the approximation of free discontinuity problems. *Bollettino della Unione Matematica Italiana B* 6, 105–123 (see pages 62, 63)

Amor, H., Marigo, J.-J., and Maurini, C. (2009). Regularized formulation of the variational brittle fracture with unilateral contact: Numerical experiments. *Journal of the Mechanics and Physics of Solids* 57 (8), 1209–1229. DOI: 10.1016/j.jmps.2009.04.011 (see pages 3, 111, 122, 123)

Andersson, M. (1996). Stress distribution and crack initiation for an elastic contact including friction. *International Journal of Solids and Structures* 33 (25), 3673–3696. DOI: 10.1016/0020-7683(95)00209-X (see page 161)

Aranson, I. S., Kalatsky, V. A., and Vinokur, V. M. (2000). Continuum field description of crack propagation. *Physical Review Letters* 85, 118–121. DOI: 10.1103/PhysRevLett.85.118 (see pages 3, 4, 66, 90)

Areias, P. M. A. and Belytschko, T. (2005). Analysis of three-dimensional crack initiation and propagation using the extended finite element method. *International Journal for Numerical Methods in Engineering* 63 (5), 760–788. DOI: 10.1002/nme.1305 (see page 3)

Argyris, J. H. (1955). Energy theorems and structural analysis: A generalized discourse with applications on energy principles of structural analysis including the effects of temperature and non-linear stress-strain relations. Part I: General theory. *Aircraft Engineering and Aerospace Technology* 27 (2), 42–58 (see page 2)

ASTM (1997). *Standard test method for plane-strain fracture toughness of metallic materials*. DOI: 10.1520/E0399-90R97 (see page 39)

Auerbach, F. (1891). Absolute Härtemessung. *Annalen der Physik* 279 (5), 61–100. DOI: 10.1002/andp.18912790505 (see page 178)

Barenblatt, G. I. (1959). The formation of equilibrium cracks during brittle fracture. General ideas and hypotheses. Axially-symmetric cracks. *Journal of Applied Mathematics and Mechanics* 23 (3), 622–636. DOI: 10.1016/0021-8928(59)90157-1 (see pages 46, 47)

Barquins, M. and Maugis, D. (1982). Adhesive contact of axisymmetric punches on an elastic half-space: the modified Hertz-Hubers stress tensor for contacting spheres. *Journal de Mécanique Théorique et Appliquée* 1 (2), 331–357 (see page 269)

Başar, Y. (2000). *Nonlinear continuum mechanics of solids: Fundamental mathematical and physical concepts*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-04299-1 (see page 214)

Bažant, Z. P. (1994). Nonlocal damage theory based on micromechanics of crack interactions. *Journal of Engineering Mechanics* 120 (3), 593–617. DOI: 10.1061/(ASCE)0733-9399(1994)120:3(593) (see page 3)

Bažant, Z. P. and Cedolin, L. (2010). *Stability of structures*. World Scientific. DOI: 10.1142/7828 (see page 12)

Bažant, Z. P. and Oh, B. (1983). Crack band theory for fracture of concrete. *Matériaux et Constructions* 16, 155–177. DOI: 10.1007/BF02486267 (see page 3)

Bedford, A. (1985). *Hamilton's principle in continuum mechanics*. Pitman Advanced Publishing Program (see pages 30, 31)

Bellettini, G. and Coscia, A. (1994). Discrete approximation of a free discontinuity problem. *Numerical Functional Analysis and Optimization* 15 (3-4), 201–224. DOI: 10.1080/01630569408816562 (see pages 63, 64)

Belytschko, T. and Black, T. (1999). Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering* 45 (5), 601–620. DOI: 10.1002/(SICI)1097-0207(19990620)45:5<601::AID-NME598>3.0.CO;2-S (see page 3)

Belytschko, T., Chen, H., Xu, J., and Zi, G. (2003). Dynamic crack propagation based on loss of hyperbolicity and a new discontinuous enrichment. *International Journal for Numerical Methods in Engineering* 58 (12), 1873–1905. DOI: 10.1002/nme.941 (see page 3)

Belytschko, T., Liu, W. K., Moran, B., and Elkhodary, K. I. (2014). *Nonlinear finite elements for continua and structures*. Second edition. Chichester: Wiley (see page 253)

Ben-Menahem, A. (2009). *Historical encyclopedia of natural and mathematical sciences*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-540-68832-7 (see page 112)

Benallal, A. and Marigo, J.-J. (2007). Bifurcation and stability issues in gradient theories with softening. *Modelling and Simulation in Materials Science and Engineering* 15 (1), 283–295. DOI: 10.1088/0965-0393/15/1/S22 (see pages 4, 102)

Benbow, J. J. (1960). Cone cracks in fused silica. *Proceedings of the Physical Society* 75 (5), 697–699. DOI: 10.1088/0370-1328/75/5/308 (see pages 192, 194)

Bergmann, L. (1948). *Ultrasonics and their scientific and technical applications*. New York: Wiley (see page 33)

Berthaud, Y., La Borderie, C., and Ramtani, S. (1990). Damage modeling and crack closure effect. *Damage Mechanics in Engineering Materials, ASME Winter Annual Meeting* 109, 263–276 (see page 52)

Betten, J. (1982). Integrity basis for a second-order and a fourth-order tensor. *International Journal of Mathematics and Mathematical Sciences* 5, 354081. DOI: 10.1155/S0161171282000088 (see page 217)

Bilgen, C., Homberger, S., and Weinberg, K. (2019). Phase-field fracture simulations of the Brazilian splitting test. *International Journal of Fracture* 220 (1), 85–98. DOI: 10.1007/s10704-019-00401-w (see page 4)

Bilgen, C., Kopaničáková, A., Krause, R., and Weinberg, K. (2018). A phase-field approach to conchoidal fracture. *Meccanica* 53 (6), 1203–1219. DOI: 10.1007/s11012-017-0740-z (see page 122)

Bilgen, C. and Weinberg, K. (2019). On the crack-driving force of phase-field models in linearized and finite elasticity. *Computer Methods in Applied Mechanics and Engineering* 353, 348–372. DOI: 10.1016/j.cma.2019.05.009 (see pages 122, 129)

Bleyer, J., Roux-Langlois, C., and Molinari, J.-F. (2017). Dynamic crack propagation with a variational phase-field model: limiting speed, crack branching and velocity-toughening mechanisms. *International Journal of Fracture* 204 (1), 79–100. DOI: 10.1007/s10704-016-0163-1 (see pages 3, 79, 108)

Boehler, J. P. (1977). On irreducible representations for isotropic scalar functions. *Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 57 (6), 323–327 (see page 224)

Boltzmann, L. (1893). On the methods of theoretical physics. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* Fifth Series, 37–44 (see page i)

Borden, M. J. (2012). Isogeometric analysis of phase-field models for dynamic brittle and ductile fracture. PhD thesis. University of Texas at Austin. URL: https://hdl.handle.net/2152/ETD-UT-2012-08-6113 (see pages 4, 75, 77, 80, 89, 138, 153–155, 184, 208, 267)

Borden, M. J., Hughes, T. J. R., Landis, C. M., and Verhoosel, C. V. (2014). A higher-order phase-field model for brittle fracture: Formulation and analysis within the isogeometric analysis framework. *Computer Methods in Applied Mechanics and Engineering* 273, 100–118. DOI: 10.1016/j.cma.2014.01.016 (see pages 4, 77, 138)

Borden, M. J., Verhoosel, C. V., Scott, M. A., Hughes, T. J. R., and Landis, C. M. (2012). A phase-field description of dynamic brittle fracture. *Computer Methods in Applied Mechanics and Engineering* 217-220, 77–95. DOI: 10.1016/j.cma.2012.01.008 (see pages 3, 4, 68, 74, 75, 81, 82, 98, 102, 116, 153, 154, 268)

Born, M. (1980). *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light*. Sixth edition. Oxford: Pergamon Press (see page 30)

Bourdin, B. (1998). Une méthode variationnelle en mécanique de la rupture. PhD thesis. Université Paris-Nord. URL: https://www.theses.fr/1998PA132060 (see pages 3, 78, 207)

Bourdin, B. (1999). Image segmentation with a finite element method. *ESAIM: Mathematical Modelling and Numerical Analysis* 33 (2), 229–244. DOI: 10.1051/m2an:1999114 (see pages 62–64, 256)

Bourdin, B. (2007a). Numerical implementation of the variational formulation for quasi-static brittle fracture. *Interfaces and Free Boundaries* 9, 411–430. DOI: 10.4171/IFB/171 (see pages 3, 75)

Bourdin, B. (2007b). The variational formulation of brittle fracture: numerical implementation and extensions. In: *IUTAM Symposium on Discretization Methods for Evolving Discontinuities*. Springer Netherlands, 381–393. DOI: 10.1007/978-1-4020-6530-9_22 (see page 3)

Bourdin, B. and Chambolle, A. (2000). Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. *Numerische Mathematik* 85, 609–646. DOI: 10.1007/PL00005394 (see page 62)

Bourdin, B., Francfort, G. A., and Marigo, J.-J. (2000). Numerical experiments in revisited brittle fracture. *Journal of the Mechanics and Physics of Solids* 48 (4), 797–826. DOI: 10.1016/S0022-5096(99)00028-9 (see pages 3, 62, 64–66, 72, 75, 81, 110, 131)

Bourdin, B., Francfort, G. A., and Marigo, J.-J. (2008). The variational approach to fracture. *Journal of Elasticity* 91 (1), 5–148. DOI: 10.1007/s10659-007-9107-3 (see pages 65, 150)

Boussinesq, J. (1885). *Application des potentiels à l'étude de l'équilibre et du mouvement des solides élastiques*. Paris: Gauthier-Villars (see pages 159, 269)

Braides, A. (1998). *Approximation of free-discontinuity problems*. Berlin: Springer. DOI: 10.1007/BFb0097344 (see pages 61, 63, 149)

Braides, A. (2002). *Gamma-convergence for beginners*. Oxford: Oxford University Press (see page 63)

Brezis, H., Coron, J.-M., and Lieb, E. H. (1986). Harmonic maps with defects. *Communications in Mathematical Physics* 107, 649–705 (see page 61)

Broberg, K. B. (1996). How fast can a crack go? *Materials Science* 32 (1), 80–86. DOI: 10.1007/BF02538928 (see page 153)

Bronštejn, I. N. (2015). *Handbook of mathematics*. Sixth edition. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-46221-8 (see page 238)

Bush, M. B. (1999). Simulation of contact-induced fracture. *Engineering Analysis with Boundary Elements* 23 (1), 59–66. DOI: 10.1016/S0955-7997(98)00061-7 (see page 5)

Cajuhi, T. (2019). Fracture in porous media: phase-field modeling, simulation and experimental validation. PhD thesis. Technical University Carolo-Wilhelmina at Braunschweig. DOI: 10.24355/dbbs.084-201903071210-0 (see pages 4, 112)

Cajuhi, T., Lura, P., and De Lorenzis, L. (2017). Preliminary calibration of a phase-field model for cracks due to shrinkage in cement-based materials. *Proceedings of the 7th GACM Colloquium on Computational Mechanics* (see page 4)

Cajuhi, T., Sanavia, L., and De Lorenzis, L. (2018). Phase-field modeling of fracture in variably saturated porous media. *Computational Mechanics* 61 (3), 299–318. DOI: 10.1007/s00466-017-1459-3 (see page 4)

Chaboche, J. L., Lesne, P. M., and Maire, J. F. (1995). Continuum damage mechanics, anisotropy and damage deactivation for brittle materials like concrete and ceramic composites. *International Journal of Damage Mechanics* 4 (1), 5–22. DOI: 10.1177/105678959500400102 (see page 52)

Chai, H. (2006). Crack propagation in glass coatings under expanding spherical contact. *Journal of the Mechanics and Physics of Solids* 54 (3), 447–466. DOI: 10.1016/j.jmps.2005.10.004 (see pages 160, 161, 176, 179)

Chambolle, A. (2003). A density result in two-dimensional linearized elasticity, and applications. *Archive for Rational Mechanics and Analysis* 167, 211–233. DOI: 10.1007/s00205-002-0240-7 (see page 57)

Chambolle, A. and Dal Maso, G. (1999). Discrete approximation of the Mumford-Shah functional in dimension two. *ESAIM: Mathematical Modelling and Numerical Analysis – Modélisation Mathématique et Analyse Numérique* 33 (4), 651–672. URL: https://www.numdam.org/item/M2AN_1999__33_4_651_0 (see page 62)

Chan, S. K., Tuba, I. S., and Wilson, W. K. (1970). On the finite element method in linear fracture mechanics. *Engineering Fracture Mechanics* 2 (1), 1–17. DOI: 10.1016/0013-7944(70)90026-3 (see page 1)

Charlotte, M., Francfort, G. A., Marigo, J.-J., and Truskinovsky, L. (2000). Revisiting brittle fracture as an energy minimization problem: comparison of Griffith and Barenblatt surface energy models. In: *Symposium on Continuous Damage and Fracture*, 7–18 (see pages 55–57)

Chen, S. Y., Farris, T. N., and Chandrasekari, S. (1995). Contact mechanics of Hertzian cone cracking. *International Journal of Solids and Structures* 32 (3), 329–340. DOI: 10.1016/0020-7683(94)00127-I (see page 160)

Cherepanov, G. P. (1967). The propagation of cracks in a continuous medium. *Journal of Applied Mathematics and Mechanics* 31, 503–512 (see page 42)

Cheverton, R. D., Bolt, S. E., Holz, P. P., and Iskander, S. K. (1981). Behavior of surface flaws in reactor pressure vessels under thermal-shock loading conditions. *Experimental Mechanics* 21, 155–162. DOI: 10.1007/BF02326396 (see page 1)

Chukwudozie, C., Bourdin, B., and Yoshioka, K. (2019). A variational phase-field model for hydraulic fracturing in porous media. *Computer Methods in Applied Mechanics and Engineering* 347, 957–982. DOI: 10.1016/j.cma.2018.12.037 (see page 139)

Ciavarella, M., Hills, D. A., and Monno, G. (1998). The influence of rounded edges on indentation by a flat punch. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 212 (4), 319–327. DOI: 10.1243/0954406981521259 (see pages 161, 209)

Clebsch, A. (1862). *Theorie der Elasticität fester Körper*. Leipzig: Teubner (see pages 112, 232)

Cohen, L. L. (1955). *Civil aircraft accident: Report of the court of inquiry into the accidents of Comet G-ALYP on 10th January 1954 and Comet G-ALYY on 8th April 1954*. London: Her Majesty's Stationery Office, Ministry of Transport and Civil Aviation (see page 1)

Coleman, B. and Hodgdon, M. (1985). On shear bands in ductile materials. *Archive for Rational Mechanics and Analysis* 90, 219–247. DOI: 10.1007/BF00251732 (see page 3)

Comi, C. and Perego, U. (2000). A bi-dissipative damage model for concrete with applications to dam engineering. *European Congress on Computational Methods in Applied Sciences and Engineering*, 1–16 (see page 53)

Conrad, H., Keshavan, M. K., and Sargent, G. A. (1979). Hertzian fracture of Pyrex glass under quasi-static loading conditions. *Journal of Materials Science* 14 (6), 1473–1494. DOI: 10.1007/BF00549324 (see pages 192, 194)

Cornetti, P., Pugno, N., Carpinteri, A., and Taylor, D. (2006). Finite fracture mechanics: A coupled stress and energy failure criterion. *Engineering Fracture Mechanics* 73, 2021–2033. DOI: 10.1016/j.engfracmech.2006.03.010 (see page 45)

Cottrell, A. H. (1960). Theoretical aspects of radiation damage and brittle fracture in steel pressure vessels. *Steels for Reactor Pressure Circuits, A Symposium of the Iron and Steel Institute, Special Report* 69, 281. URL: https://www.osti.gov/biblio/4782070 (see page 46)

Cottrell, J. A. (2009). *Isogeometric analysis: Toward integration of CAD and FEA*. Chichester: Wiley (see page 138)

Daams, J. L. C., Rodgers, J. R., and Villars, P. (2006). Typical interatomic distances: Metals and alloys. In: *International Tables for Crystallography*. American Cancer Society, 774–777. DOI: 10.1107/97809553602060000619 (see page 9)

Dal Maso, G. (1993). *An introduction to Γ-convergence*. Boston: Birkhäuser. DOI: 10.1007/978-1-4612-0327-8 (see page 63)

Dal Maso, G., Francfort, G. A., and Toader, R. (2005). Quasistatic crack growth in nonlinear elasticity. *Archive for Rational Mechanics and Analysis* 176, 165–225. DOI: 10.1007/s00205-004-0351-4 (see page 55)

Dal Maso, G., Morel, J. M., and Solimini, S. (1992). A variational method in image segmentation: Existence and approximation results. *Acta Mathematica* 168 (1), 89–151. DOI: 10.1007/BF02392977 (see page 63)

Dal Maso, G. and Toader, R. (2002). A model for the quasi-static growth of brittle fractures: Existence and approximation results. *Archive for Rational Mechanics and Analysis* 162, 101–135. DOI: 10.1007/s002050100187 (see page 57)

Dally, T. and Weinberg, K. (2017). The phase-field approach as a tool for experimental validations in fracture mechanics. *Continuum Mechanics and Thermodynamics* 29 (4), 947–956. DOI: 10.1007/s00161-015-0443-4 (see page 122)

Daux, C., Moës, N., Dolbow, J. E., Sukumar, N., and Belytschko, T. (2000). Arbitrary branched and intersecting cracks with the extended finite element method. *International Journal for Numerical Methods in Engineering* 48 (12), 1741–1760. DOI: 10.1002/1097-0207(20000830)48:12<1741::AID-NME956>3.0.CO;2-L (see page 3)

De Giorgi, E. (1991). Free discontinuity problems in calculus of variations. In: *Frontiers in Pure and Applied Mathemathics, a collection of papers dedicated to Jacques-Louis Lions on the occasion of his sixtieth birthday*, 55–62 (see page 61)

De Giorgi, E. and Ambrosio, L. (1988). Un nuovo tipo di funzionale del calcolo delle variazioni. *Atti della Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matem-*

atiche e Naturali, Rendiconti Lincei, Matematica e Applicazioni 82 (2), 199–210. URL: https://eudml.org/doc/287202 (see page 63)

De Giorgi, E., Carriero, M., and Leaci, A. (1989). Existence theorem for a minimum problem with free discontinuity set. *Archive for Rational Mechanics and Analysis* 108, 195–218. DOI: 10.1007/BF01052971 (see page 63)

De L'Hôpital, G. F. A. (1696). *Analyse des infiniment petits, pour l'intelligence des lignes courbes*. DOI: 10.3931/e-rara-11447 (see page 92)

De Lacerda, L. and Wrobel, L. (2002). An efficient numerical model for contact-induced crack propagation analysis. *International Journal of Solids and Structures* 39, 5719–5736. DOI: 10.1016/S0020-7683(02)00465-1 (see page 5)

Del Piero, G. (1989). Constitutive equation and compatibility of the external loads for linear elastic masonry-like materials. *Meccanica* 24, 150–162. DOI: 10.1007/BF01559418 (see page 122)

Dolbow, J. E. (1999). An extended finite element method with discontinuous enrichment for applied mechanics. PhD thesis. Northwestern University (see page 2)

Dolbow, J. E., Moës, N., and Belytschko, T. (2001). An extended finite element method for modeling crack growth with frictional contact. *Computer Methods in Applied Mechanics and Engineering* 190 (51), 6825–6846. DOI: 10.1016/S0045-7825(01)00260-2 (see page 3)

Dowgiałło, P. (2016). Analyse der Rissinitiierung unter Eindruckbelastung mittels eines hybriden Kriteriums der Finiten Bruchmechanik. MA thesis. Karlsruhe Institute of Technology (see pages 46, 171)

Doyle, T. C. and Ericksen, J. L. (1956). Nonlinear elasticity. *Advances in Applied Mechanics* 4, 53–115. DOI: 10.1016/S0065-2156(08)70371-5 (see page 12)

Dugdale, D. S. (1960). Yielding of steel sheets containing slits. *Journal of the Mechanics and Physics of Solids* 8 (2), 100–104. DOI: 10.1016/0022-5096(60)90013-2 (see page 46)

Dvorkin, E. N. and Assanelli, A. P. (1991). 2D finite elements with displacement interpolated embedded localization lines: The analysis of fracture in frictional materials. *Computer Methods in Applied Mechanics and Engineering* 90 (1), 829–844. DOI: 10.1016/0045-7825(91)90186-A (see page 148)

Dvorkin, E. N. and Bathe, K.-J. (1984). A continuum mechanics based four-node shell element for general nonlinear analysis. *Engineering Computations* 1, 77–88. DOI: 10.1108/eb023562 (see page 148)

Dvorkin, E. N., Cuitiño, A. M., and Gioia, G. (1990). Finite elements with displacement interpolated embedded localization lines insensitive to mesh size and distortions. *International Journal for Numerical Methods in Engineering* 30 (3), 541–564. DOI: 10.1002/nme.1620300311 (see page 148)

Ehlers, W. and Luo, C. (2017). A phase-field approach embedded in the theory of porous media for the description of dynamic hydraulic fracturing. *Computer Methods in Applied Mechanics and Engineering* 315, 348–368. DOI: 10.1016/j.cma.2016.10.045 (see page 4)

Erdogan, F. and Sih, G. C. (1963). On the crack extension in plates under plane loading and transverse shear. *Journal of Basic Engineering* 85 (4), 519–525 (see pages 2, 44, 45)

Ericksen, J. L. (1976). Equilibrium theory of liquid crystals. *Advances in Liquid Crystals* 2, 233–298. DOI: 10.1016/B978-0-12-025002-8.50012-9 (see pages 61, 63)

Eshelby, J. D. (1951). The force on an elastic singularity. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical and Physical Sciences* 244 (877), 87–112. DOI: 10.1098/rsta.1951.0016 (see page 42)

Evans, L. C. and Gariepy, R. F. (1992). *Measure theory and fine properties of functions*. Boca Raton: CRC Press (see pages 54, 59)

Fabrikant, V. I. (1988). Elastic field around a circular punch. *Journal of Applied Mechanics* 55 (3), 604–610. DOI: 10.1115/1.3125836 (see page 270)

Fett, T., Rizzi, G., and Diegele, E. (2004). Weight functions for cone cracks. *Engineering Fracture Mechanics* 71 (16–17), 2551–2560. DOI: 10.1016/j.engfracmech.2004.01.001 (see page 5)

Finnie, I., Khatibloo, M., and Dolev, D. (1981). On the physical basis of Auerbach's law. *Journal of Engineering Materials and Technology* 103, 183–184. DOI: 10.1115/1.3224992 (see page 6)

Fischer-Cripps, A. C. (1997). Predicting Hertzian fracture. *Journal of Materials Science* 32 (5), 1277–1285. DOI: 10.1023/A:1018500522014 (see pages 6, 161, 168)

Fischer-Cripps, A. C. (2007). *Introduction to contact mechanics*. Second edition. New York: Springer. DOI: 10.1007/978-0-387-68188-7 (see pages 5, 6, 178, 179, 270)

Föppl, A. O. (1914). *Vorlesungen über Technische Mechanik*. Leipzig: Teubner (see page 232)

Franca, L. P. (1989). An algorithm to compute the square root of a 3×3 positive definite matrix. *Computers and Mathematics with Applications* 18 (5), 459–466. DOI: 10.1016/0898-1221(89)90240-X (see page 225)

Francfort, G. A. (2006). Quasistatic brittle fracture seen as an energy minimizing movement. *GAMM-Mitteilungen* 29 (2), 72–191. DOI: 10.1002/gamm.201490029 (see page 57)

Francfort, G. A. and Larsen, C. J. (2003). Existence and convergence for quasi-static evolution in brittle fracture. *Communications on Pure and Applied Mathematics* 56 (10), 1465–1500. DOI: 10.1002/cpa.3039 (see page 57)

Francfort, G. A. and Marigo, J.-J. (1998). Revisiting brittle fracture as an energy minimization problem. *Journal of the Mechanics and Physics of Solids* 46 (8), 1319–1342. DOI: 10.1016/S0022-5096(98)00034-9 (see pages 3, 53–55, 57–59)

Francfort, G. A. and Marigo, J.-J. (2005). Griffith theory of brittle fracture revisited: Merits and drawbacks. *Latin American Journal of Solids and Structures* 2, 57–64 (see pages 55, 56)

Frank, F. C. and Lawn, B. R. (1967). On the theory of Hertzian fracture. *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences* 299 (1458), 291–306. DOI: 10.1098/rspa.1967.0137 (see pages 5, 161, 168, 277)

Fraternali, F. (2007). Free discontinuity finite element models in two-dimensions for in-plane crack problems. *Theoretical and Applied Fracture Mechanics* 47 (3), 274–282. DOI: 10.1016/j.tafmec.2007.01.006 (see page 62)

Freddi, F. (2019). Fracture energy in phase field models. *Mechanics Research Communications* 96, 29–36. DOI: 10.1016/j.mechrescom.2019.01.009 (see page 149)

Freddi, F. and Royer-Carfagni, G. (2010). Regularized variational theories of fracture: A unified approach. *Journal of the Mechanics and Physics of Solids* 58 (8), 1154–1174. DOI: 10.1016/j.jmps.2010.02.010 (see pages 3, 111, 122)

Gerasimov, T. and De Lorenzis, L. (2016). A line search assisted monolithic approach for phase-field computing of brittle fracture. *Computer Methods in Applied Mechanics and Engineering* 312, 276–303. DOI: 10.1016/j.cma.2015.12.017 (see page 75)

Gerasimov, T. and De Lorenzis, L. (2019). On penalization in variational phase-field models of brittle fracture. *Computer Methods in Applied Mechanics and Engineering* 354, 990–1026. DOI: 10.1016/j.cma.2019.05.038 (see pages 4, 79, 108)

Gol'dstein, R. V. and Salganik, R. L. (1974). Brittle fracture of solids with arbitrary cracks. *International Journal of Fracture* 10 (4), 507–523. DOI: 10.1007/BF00155254 (see pages 2, 45)

Govindjee, S., Strain, J., Mitchell, T. J., and Taylor, R. L. (2012). Convergence of an efficient local least-squares fitting method for bases with compact support. *Computer Methods in Applied Mechanics and Engineering* 213-216, 84–92. DOI: 10.1016/j.cma.2011.11.017 (see page 139)

Griffith, A. A. (1921). The phenomena of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences* 221 (582-593), 163–198. DOI: 10.1098/rsta.1921.0006 (see pages 2, 40)

Gross, D. and Seelig, T. (2011). *Fracture mechanics: With an introduction to micromechanics*. Second edition. Berlin: Springer. DOI: 10.1007/978-3-642-19240-1 (see pages 34, 35, 43, 46, 152, 158)

Gugenberger, C. (2009). Free-boundary problem of crack dynamics: phase field modeling. PhD thesis. RWTH Aachen University. URL: https://publications.rwth-aachen.de/record/51360 (see page 66)

Gurson, A. L. (1977). Continuum theory of ductile rupture by void nucleation and growth. Part I: Yield criteria and flow rules for porous ductile media. *Journal of Engineering Materials and Technology* 99 (1), 2–15. DOI: 10.1115/1.3443401 (see page 49)

Gurtin, M. E. (1996). Generalized Ginzburg-Landau and Cahn-Hilliard equations based on a microforce balance. *Physica D: Nonlinear Phenomena* 92 (3), 178–192. DOI: 10.1016/0167-2789(95)00173-5 (see page 73)

Hai-Chang, H. (1954). On some variational principles in the theory of elasticity and the theory of plasticity. *Acta Physica Sinica* 10, 259–290. DOI: 10.7498/aps.10.259 (see page 29)

Hakim, V. and Karma, A. (2009). Laws of crack motion and phase-field models of fracture. *Journal of the Mechanics and Physics of Solids* 57 (2), 342–368. DOI: 10.1016/j.jmps.2008.10.012 (see pages 66, 102, 103)

Hall, F. R. and Hayhurst, D. R. (1991). Modelling of grain size effects in creep crack growth using a non-local continuum damage approach. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences* 433 (1888), 405–421. DOI: 10.1098/rspa.1991.0055 (see page 3)

Hamilton, W. R. (1834). On a general method in dynamics. *Philosophical Transactions of the Royal Society of London*, 247–308 (see page 30)

Hamilton, W. R. (1835). Second essay on a general method in dynamics. *Philosophical Transactions of the Royal Society of London* 125, 95–144 (see page 30)

Hashin, Z. (1996). Finite thermoelastic fracture criterion with application to laminate cracking analysis. *Journal of the Mechanics and Physics of Solids* 44 (7), 1129–1145. DOI: 10.1016/0022-5096(95)00080-1 (see pages 2, 45)

Hausdorff, F. (1918). Dimension und äußeres Maß. *Mathematische Annalen* 79, 157–179. DOI: 10.1007/BF01457179 (see page 54)

Hebel, J. (2010). Modellierung spröder Rissbildung an Spannungskonzentrationen mit der Bruchmechanik finiter Risse. PhD thesis. Technical University of Darmstadt. URL: https://tuprints.ulb.tu-darmstadt.de/id/eprint/2304 (see pages 4, 46)

Henry, H. and Levine, H. (2004). Dynamic instabilities of fracture under biaxial strain using a phase field model. *Physical Review Letters* 93, 105504. DOI: 10.1103/PhysRevLett.93.105504 (see pages 3, 4, 66, 90)

Hertz, H. R. (1882a). Über die Berührung fester elastischer Körper. *Journal für die reine und angewandte Mathematik* 92, 156–171 (see pages 178, 269)

Hertz, H. R. (1890). Über die Beziehungen zwischen Licht und Elektricität. *Ein Votrag gehalten bei der 62. Versammlung deutscher Naturforscher und Ärzte in Heidelberg* (see page i)

Hertz, H. R. (1882b). Über die Berührung fester elastischer Körper und über die Härte. *Verhandlungen des Vereins zur Beförderung des Gewerbfleißes*, 449–463 (see page 269)

Hesch, C., Gil, A. J., Ortigosa, R., Dittmann, M., Bilgen, C., Betsch, P., Franke, M., Janz, A., and Weinberg, K. (2017). A framework for polyconvex large strain phase-field methods to fracture. *Computer Methods in Applied Mechanics and Engineering* 317, 649–683. DOI: 10.1016/j.cma.2016.12.035 (see pages 112, 267)

Hesch, C. and Weinberg, K. (2014). Thermodynamically consistent algorithms for a finite-deformation phase-field approach to fracture. *International Journal for Numerical Methods in Engineering* 99 (12), 906–924. DOI: 10.1002/nme.4709 (see pages 75, 112, 116)

Hill, R. (1968). On constitutive inequalities for simple materials. Part I. *Journal of the Mechanics and Physics of Solids* 16 (4), 229–242. DOI: 10.1016/0022-5096(68)90031-8 (see page 12)

Hofacker, M. and Miehe, C. (2013). A phase field model of dynamic fracture: Robust field updates for the analysis of complex crack patterns. *International Journal for Numerical Methods in Engineering* 93 (3), 276–301. DOI: 10.1002/nme.4387 (see pages 4, 154)

Hofacker, M. (2014). A thermodynamically consistent phase field approach to fracture. PhD thesis. University of Stuttgart (see pages 116, 267)

Hofacker, M. and Miehe, C. (2012). Continuum phase field modeling of dynamic fracture: variational principles and staggered FE implementation. *International Journal of Fracture* 178 (1-2), 113–129. DOI: 10.1007/s10704-012-9753-8 (see pages 3, 68, 74, 75, 116)

Hofacker, M., Welschinger, F., and Miehe, C. (2009). A variational-based formulation of regularized brittle fracture. *Proceedings in Applied Mathematics and Mechanics* 9 (1), 207–208. DOI: 10.1002/pamm.200910078 (see page 116)

Holzapfel, G. A. (2000). *Nonlinear solid mechanics: A continuum approach for engineering*. Chichester: Wiley (see pages 9, 16, 26, 50, 51, 215)

Huang, R., Prévost, J., Huang, Z., and Suo, Z. (2003). Channel-cracking of thin films with the extended finite element method. *Engineering Fracture Mechanics* 70 (18), 2513–2526. DOI: 10.1016/S0013-7944(03)00083-3 (see page 3)

Hughes, T. J. R. (1987). *The finite element method: Linear static and dynamic finite element analysis*. Englewood Cliffs: Prentice-Hall (see pages 28, 253, 260)

Hussain, M. A., Pu, S. L., and Underwood, J. H. (1974). Strain energy release rate for a crack under combined mode I and mode II. *ASTM International*, 2–28 (see page 45)

Inglis, C. E. (1913). Stresses in a plate due to the presence of cracks and sharp corners. *Transactions of the Institution of Naval Architects* 55, 219–241 (see page 38)

Irwin, G. R. (1957). Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics* 24, 361–364 (see pages 2, 37, 38)

Irwin, G. R. (1960). Fracture mechanics. In: *Proceedings of the 1st Symposium on Naval Structural Mechanics*, 557–591 (see pages 47, 230)

Itskov, M. and Başar, Y. (2000). On the theory of fourth-order tensors and their application in large strain elasticity. *Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 80, 523–524. DOI: 10.1002/zamm.200008014132 (see page 216)

Ives, K. D., Shoemaker, A. K., and McCartney, R. F. (1974). Pipe deformation during a running shear fracture in line pipe. *Journal of Engineering Materials and Technology* 96 (4), 309–317. DOI: 10.1115/1.3443246 (see page 1)

Johnson, G. R. and Cook, W. H. (1985). Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures. *Engineering Fracture Mechanics* 21 (1), 31–48. DOI: 10.1016/0013-7944(85)90052-9 (see page 35)

Johnson, K. L., O'Connor, J. J., Woodward, A. C., and Tabor, D. (1973). The effect of the indenter elasticity on the Hertzian fracture of brittle materials. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences* 334 (1596), 95–117. DOI: 10.1098/rspa.1973.0082 (see page 160)

Ju, J. W. (1989). On energy-based coupled elastoplastic damage theories: Constitutive modeling and computational aspects. *International Journal of Solids and Structures* 25 (7), 803–833. DOI: 10.1016/0020-7683(89)90015-2 (see page 52)

Kachanov, L. M. (1958). Time of the rupture process under creep conditions. *Izvestiia Akademii Nauk SSSR, Otdelenie Teckhnicheskikh Nauk* 8, 26–31. DOI: 10.1023/A:1018671022008 (see pages 3, 48)

Kachanov, L. M. (1986). *Introduction to continuum damage mechanics*. Dordrecht: Nijhoff. DOI: 10.1007/978-94-017-1957-5 (see pages 34, 48)

Kachanov, M., Shafiro, B., and Tsukrov, I. (2003). *Handbook of elasticity solutions*. Dordrecht: Kluwer. DOI: 10.1007/978-94-017-0169-3 (see page 269)

Karma, A., Kessler, D. A., and Levine, H. (2001). Phase-field model of mode III dynamic fracture. *Physical Review Letters* 87, 045501. DOI: 10.1103/PhysRevLett.87.045501 (see pages 3, 4, 66, 78, 82, 85, 90, 103)

Karma, A. and Lobkovsky, A. E. (2004). Unsteady crack motion and branching in a phase-field model of brittle fracture. *Physical Review Letters* 92, 245510. DOI: 10.1103/PhysRevLett.92.245510 (see pages 3, 66)

Kiendl, J., Ambati, M., De Lorenzis, L., Gomez, H., and Reali, A. (2016). Phase-field description of brittle fracture in plates and shells. *Computer Methods in Applied Mechanics and Engineering* 312, 374–394. DOI: 10.1016/j.cma.2016.09.011 (see pages 4, 116, 118)

Kienle, D. and Keip, M.-A. (2019). Modeling of hydraulically induced fractures in elastic-plastic solids. *Proceedings in Applied Mathematics and Mechanics* 19 (1), e201900377. DOI: 10.1002/pamm.201900377 (see pages 4, 138)

Kirkaldy, D. (1863). *Experiments on wrought iron and steel*. New York: Scribner (see page 1)

Klinsmann, M., Rosato, D., Kamlah, M., and McMeeking, R. M. (2015). An assessment of the phase field formulation for crack growth. *Computer Methods in Applied Mechanics and Engineering* 294, 313–330. DOI: 10.1016/j.cma.2015.06.009 (see pages 7, 114, 116, 118, 267)

Klisinski, M., Runesson, K., and Sture, S. (1991). Finite element with inner softening band. *Journal of Engineering Mechanics* 117 (3), 575–587. DOI: 10.1061/(ASCE)0733-9399(1991)117:3(575) (see page 148)

Knauss, W. G. (1970). An observation of crack propagation in anti-plane shear. *International Journal of Fracture Mechanics* 6, 183–187. DOI: 10.1007/BF00189825 (see page 4)

Kobayashi, A. S. and Ramulu, M. (1985). A dynamic fracture analysis of crack curving and branching. *Journal de Physique* 46 (C5), 197–206. DOI: 10.1051/jphyscol:1985525 (see page 154)

Kobayashi, H. and Onoue, H. (1943). *Brittle fracture of Liberty ships – failure knowledge database*. London: Oregon State (see page 1)

Kocer, C. (2003a). An automated incremental finite element study of Hertzian cone crack growth. *Finite Elements in Analysis and Design* 39 (7), 639–660. DOI: 10.1016/S0168-874X(02)00132-4 (see page 5)

Kocer, C. (2003b). Using the Hertzian fracture system to measure crack growth data: A review. *International Journal of Fracture* 121 (3), 111–132. DOI: 10.1023/B:FRAC.0000005343.74494.7e (see pages 5, 188, 192–194)

Kocer, C. and Collins, R. E. (1998). Angle of Hertzian cone cracks. *Journal of the American Ceramic Society* 81 (7), 1736–1742. DOI: 10.1111/j.1151-2916.1998.tb02542.x (see pages 5, 6, 192, 194)

Kolosov, G. V. (1909). An application of the theory of functions of a complex variable to a planar problem in the mathematical theory of elasticity (see page 38)

Kort, B. W. and Bertsekas, D. P. (1972). A new penalty function method for constrained minimization. In: *Proceedings of the 1972 IEEE Conference on Decision and Control and 11th Symposium on Adaptive Processes*, 162–166. DOI: 10.1109/CDC.1972.268971 (see page 239)

Kuhn, C. (2013). Numerical and analytical investigation of a phase field model for fracture. PhD thesis. Technical University of Kaiserslautern. URL: https://nbn-resolving.de/urn:nbn:de:hbz:386-kluedo-35257 (see pages 4, 58, 66, 72, 73, 81, 89, 106–108, 246, 247, 266)

Kuhn, C. and Müller, R. (2008). A phase field model for fracture. *Proceedings in Applied Mathematics and Mechanics* 8 (1), 10223–10224. DOI: 10.1002/pamm.200810223 (see pages 66, 73)

Kuhn, C. and Müller, R. (2010a). A continuum phase field model for fracture. *Engineering Fracture Mechanics* 77 (18), 3625–3634. DOI: 10.1016/j.engfracmech.2010.08.009 (see pages 67, 75, 81)

Kuhn, C. and Müller, R. (2010b). Exponential finite elements for a phase field fracture model. *Proceedings in Applied Mathematics and Mechanics* 10 (1), 121–122. DOI: 10.1002/pamm.201010053 (see page 107)

Kuhn, C. and Müller, R. (2011). Exponential finite element shape functions for a phase field model of brittle fracture. *Computational Plasticity – Fundamentals and Applications: Proceedings of the 11th International Conference, Barcelona*, 478–489 (see page 107)

Kuhn, C. and Müller, R. (2013). Crack nucleation in phase field fracture models. In: *Proceedings of the 13th International Conference on Fracture, Beijing*, 910–919 (see pages 4, 102)

Kuhn, C. and Müller, R. (2016). A discussion of fracture mechanisms in heterogeneous materials by means of configurational forces in a phase field fracture model. *Computer Methods in Applied Mechanics and Engineering* 312, 95–116. DOI: 10.1016/j.cma.2016.04.027 (see page 4)

Kuhn, C., Schlüter, A., and Müller, R. (2015). On degradation functions in phase field fracture models. *Computational Materials Science* 108, Part B, 374–384. DOI: 10.1016/j.commatsci.2015.05.034 (see pages 4, 81, 82, 89, 98, 102)

Ladeveze, P. and Lemaitre, J. (1984). Damage effective stress in quasi unilateral conditions. In: *Proceedings of the 16th International Congresses of Theoretical and Applied Mechanics, Lyngby* (see page 52)

Lancioni, G. and Royer-Carfagni, G. (2009). The variational approach to fracture mechanics. A practical application to the French Panthéon in Paris. *Journal of Elasticity* 95 (1-2), 1–30. DOI: 10.1007/s10659-009-9189-1 (see pages 3, 125)

Lasry, D. and Belytschko, T. (1988). Localization limiters in transient problems. *International Journal of Solids and Structures* 24 (6), 581–597. DOI: 10.1016/0020-7683(88)90059-5 (see page 3)

Lawn, B. R. (1993). *Fracture of brittle solids*. Second edition. Cambridge: Cambridge University Press (see pages 5, 6, 161, 205)

Lawn, B. R. (1998). Indentation of ceramics with spheres: A century after Hertz. *Journal of the American Ceramic Society* 81 (8), 1977–1994. DOI: 10.1111/j.1151-2916.1998.tb02580.x (see page 5)

Lawn, B. R., Wilshaw, T. R., and Hartley, N. E. W. (1974). A computer simulation study of Hertzian cone crack growth. *International Journal of Fracture* 10 (1), 1–16. DOI: 10.1007/BF00955075 (see pages 5, 6, 192, 209)

Leblond, J.-B. (1999). Crack paths in three-dimensional elastic solids. I: Two-term expansion of the stress intensity factors – application to crack path stability in hydraulic fracturing. *International Journal of Solids and Structures* 36 (1), 79–103. DOI: 10.1016/S0020-7683(97)00276-X (see page 158)

Lee, S., Wheeler, M. F., and Wick, T. (2017). Iterative coupling of flow, geomechanics and adaptive phase-field fracture including level-set crack width approaches. *Journal of Computational and Applied Mathematics* 314, 40–60. DOI: 10.1016/j.cam.2016.10.022 (see page 139)

Leguillon, D. (1989). Calcul du taux de restitution de l'énergie au voisinage d'une singularité. *Comptes Rendus de l'Académie des Sciences, Series IIB: Mechanics* 309 (10), 945–950 (see page 57)

Leguillon, D. (2001). A criterion for crack nucleation at a notch in homogeneous materials. *Comptes Rendus de l'Académie des Sciences, Series IIB: Mechanics* 329 (2), 97–102. DOI: 10.1016/S1620-7742(01)01302-2 (see pages 2, 4, 45)

Leguillon, D. (2002). Strength or toughness? A criterion for crack onset at a notch. *European Journal of Mechanics – A/Solids* 21 (1), 61–72. DOI: 10.1016/S0997-7538(01)01184-6 (see pages 2, 6, 45, 46, 58, 209)

Leguillon, D., Quesada, D., Putot, C., and Martin, E. (2007). Prediction of crack initiation at blunt notches and cavities – size effects. *Engineering Fracture Mechanics* 74 (15), 2420–2436. DOI: 10.1016/j.engfracmech.2006.11.008 (see page 46)

Leguillon, D. and Sanchez-Palencia, E. (1987). Computation of singular solutions in elliptic problems and elasticity. *Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 68 (11), 590–590. DOI: 10.1002/zamm.19880681137 (see page 4)

Lemaitre, J. (1971). Evaluation of dissipation and damage in metals submitted to dynamic loading. *Internationale Conference on the Mechanical Behavior of Materials*, 1–16 (see page 50)

Lemaitre, J. (1985). A continuous damage mechanics model for ductile fracture. *Journal of Engineering Materials and Technology* 107 (1), 83–89. DOI: 10.1115/1.3225775 (see page 50)

Lemaitre, J. (1996). *A course on damage mechanics*. Second edition. Berlin: Springer. DOI: 10.1007/978-3-642-18255-6 (see page 52)

Lemaitre, J. and Desmorat, R. (2005). *Engineering damage mechanics: ductile, creep, fatigue and brittle failures*. Berlin: Springer (see pages 34, 53)

Levitas, V. I., Jafarzadeh, H., Farrahi, G. H., and Javanbakht, M. (2018). Thermodynamically consistent and scale-dependent phase field approach for crack propagation allowing for surface stresses. *International Journal of Plasticity* 111, 1–35. DOI: 10.1016/j.ijplas.2018.07.005 (see page 78)

Levy, N., Marcal, P. V., Ostergren, W. J., and Rice, J. R. (1971). Small scale yielding near a crack in plane strain – A finite element analysis. *International Journal of Fracture Mechanics* 7, 143–156. DOI: 10.1007/BF00183802 (see page 1)

Li, J. and Zhang, X. (2006). A criterion study for non-singular stress concentrations in brittle or quasi-brittle materials. *Engineering Fracture Mechanics* 73 (4), 505–523. DOI: 10.1016/j.engfracmech.2005.09.001 (see pages 2, 46)

Li, T., Marigo, J.-J., Guilbaud, D., and Potapov, S. (2016). Gradient damage modeling of brittle fracture in an explicit dynamics context. *International Journal for Numerical Methods in Engineering* 108 (11), 1381–1405. DOI: 10.1002/nme.5262 (see pages 112, 123, 150)

Linse, T., Hennig, P., Kästner, M., and Borst, R. de (2017). A convergence study of phase-field models for brittle fracture. *Engineering Fracture Mechanics* 184, 307–318. DOI: 10.1016/j.engfracmech.2017.09.013 (see page 184)

Lo, Y.-S., Borden, M. J., Ravi-Chandar, K., and Landis, C. M. (2019). A phase-field model for fatigue crack growth. *Journal of the Mechanics and Physics of Solids* 132, 103684. DOI: 10.1016/j.jmps.2019.103684 (see page 4)

Lorentz, E. and Godard, V. (2011). Gradient damage models: Toward full-scale computations. *Computer Methods in Applied Mechanics and Engineering* 200 (21), 1927–1944. DOI: 10.1016/j.cma.2010.06.025 (see pages 4, 84)

Luo, C. (2019). A phase-field model embedded in the theory of porous media with application to hydraulic fracturing. PhD thesis. University of Stuttgart. DOI: 10.18419/opus-10355 (see page 4)

Macaulay, W. H. (1919). Note on the deflection of beams. *The Messenger of Mathematics* 48, 129–130 (see page 232)

Malvern, L. E. (1969). *Introduction to the mechanics of a continuous medium*. Englewood Cliffs: Prentice-Hall (see page 218)

Mang, K., Wick, T., and Wollner, W. (2019). A phase-field model for fractures in nearly incompressible solids. *Computational Mechanics*. DOI: 10.1007/s00466-019-01752-w (see page 4)

Maugis, D. (2000). Rupture and adherence of elastic solids. In: *Contact, Adhesion and Rupture of Elastic Solids*. Vol. 130. Springer, 133–202. DOI: 10.1007/978-3-662-04125-3_3 (see pages 270, 272)

Mauthe, S. and Miehe, C. (2015). Phase-field modeling of hydraulic fracture. *Proceedings in Applied Mathematics and Mechanics* 15, 141–142. DOI: 10.1002/pamm.201510061 (see page 138)

Maxwell, J. C. (1865). A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London* 155, 459–512. URL: https://www.jstor.org/stable/108892 (see page i)

May, S., Vignollet, J., and Borst, R. de (2015). A numerical assessment of phase-field models for brittle and cohesive fracture: Γ-convergence and stress oscillations. *European Journal of Mechanics – A/Solids* 52, 72–84. DOI: 10.1016/j.euromechsol.2015.02.002 (see pages 4, 116, 118, 267)

Mazars, J. (1982). Mechanical damage and fracture of concrete structures. In: *Advances in Fracture Research: Proceedings of the 5th International Conference on Fracture, Cannes*. Pergamon Press, 1499–1506 (see page 52)

Mazars, J. (1986). A description of micro- and macroscale damage of concrete structures. *Engineering Fracture Mechanics* 25 (5), 729–737. DOI: 10.1016/0013-7944(86)90036-6 (see page 52)

Mazars, J., Berthaud, Y., and Ramtani, S. (1990). The unilateral behaviour of damaged concrete. *Engineering Fracture Mechanics* 35 (4), 629–635. DOI: 10.1016/0013-7944(90)90145-7 (see page 52)

Mazars, J. and Lemaitre, J. (1985). Application of continuous damage mechanics to strain and fracture behavior of concrete. In: *Application of Fracture Mechanics to Cementitious Composites*. Vol. 94. Springer Netherlands, 507–520. DOI: 10.1007/978-94-009-5121-1_17 (see page 3)

Melenk, J. M. and Babuška, I. (1996). The partition of unity finite element method: Basic theory and applications. *Computer Methods in Applied Mechanics and Engineering* 139 (1), 289–314. DOI: 10.1016/S0045-7825(96)01087-0 (see page 2)

Meyers, M. A. (1994). *Dynamic behavior of materials*. Hoboken: John Wiley and Sons. DOI: 10.1002/9780470172278 (see page 154)

Miehe, C. (1993). Computation of isotropic tensor functions. *Communications in Numerical Methods in Engineering* 9 (11), 889–896. DOI: 10.1002/cnm.1640091105 (see page 264)

Miehe, C., Hofacker, M., and Welschinger, F. (2010a). A phase field model for rate-independent crack propagation: Robust algorithmic implementation based on operator splits. *Computer Methods in Applied Mechanics and Engineering* 199, 2765–2778. DOI: 10.1016/j.cma.2010.04.011 (see pages 66, 67, 73, 75, 111, 112, 266, 267)

Miehe, C. and Schänzel, L.-M. (2014). Phase field modeling of fracture in rubbery polymers. Part I: Finite elasticity coupled with brittle failure. *Journal of the Mechanics and Physics of Solids* 65, 93–113. DOI: 10.1016/j.jmps.2013.06.007 (see page 4)

Miehe, C., Schänzel, L.-M., and Ulmer, H. (2015). 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, 449–485. DOI: 10.1016/j.cma.2014.11.016 (see pages 70, 84, 85, 129)

Miehe, C., Welschinger, F., and Hofacker, M. (2010b). A phase field model of electromechanical fracture. *Journal of the Mechanics and Physics of Solids* 58 (10), 1716–1740. DOI: 10.1016/j.jmps.2010.06.013 (see pages 4, 75)

Miehe, C., Welschinger, F., and Hofacker, M. (2010c). Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field FE implementations. *International Journal for Numerical Methods in Engineering* 83 (10), 1273–1311. DOI: 10.1002/nme.2861 (see pages 3, 66, 67, 73, 74, 81, 116, 117, 129, 149)

Mikelić, A., Wheeler, M. F., and Wick, T. (2015). Phase-field modeling of a fluid-driven fracture in a poroelastic medium. *Computational Geosciences* 19 (6), 1171–1195. DOI: 10.1007/s10596-015-9532-5 (see page 4)

Mikosza, A. G. and Lawn, B. R. (1971). Section-and-etch study of Hertzian fracture mechanics. *Journal of Applied Physics* 42 (13), 5540–5545. DOI: 10.1063/1.1659977 (see page 193)

Moës, N., Dolbow, J. E., and Belytschko, T. (1999). A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering* 46 (1), 131–150. DOI: 10.1002/(SICI)1097-0207(19990910)46:1<131::AID-NME726>3.0.CO;2-J (see page 2)

Moës, N., Gravouil, A., and Belytschko, T. (2002). Non-planar 3D crack growth by the extended finite element and level sets. Part I: Mechanical model. *International Journal for Numerical Methods in Engineering* 53 (11), 2549–2568. DOI: 10.1002/nme.429 (see page 3)

Molnár, G. and Gravouil, A. (2017). 2D and 3D Abaqus implementation of a robust staggered phase-field solution for modeling brittle fracture. *Finite Elements in Analysis and Design* 130, 27–38. DOI: 10.1016/j.finel.2017.03.002 (see page 75)

Morand, L. (2015). Entwicklung eines rotationssymmetrischen Finiten Elements mit Phasenfeldansatz für spröde Bruchvorgänge. MA thesis. Karlsruhe Institute of Technology (see pages 86, 194)

Morman Jr., K. N. (1986). The generalized strain measure with application to nonhomogeneous deformations in rubber-like solids. *Journal of Applied Mechanics* 53 (3), 726–728. DOI: 10.1115/1.3171840 (see page 225)

Mouginot, R. and Maugis, D. (1985). Fracture indentation beneath flat and spherical punches. *Journal of Materials Science* 20 (12), 4354–4376. DOI: 10.1007/BF00559324 (see pages 5, 6, 160, 161, 167, 168, 176–180, 188–192, 194)

Mumford, D. and Shah, J. (1985). Boundary detection by minimizing functionals. In: *IEEE Conference on Computer Vision and Pattern Recognition*, 22–26 (see pages 61, 62)

Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics* 42, 577–685. DOI: 10.1002/cpa.3160420503 (see page 62)

Murakami, S. (1988). Mechanical modeling of material damage. *Journal of Applied Mechanics* 55 (2), 280–286. DOI: 10.1115/1.3173673 (see page 52)

Murakami, S. (2012). *Continuum damage mechanics: A continuum mechanics approach to the analysis of damage and fracture*. Dordrecht: Springer. DOI: 10.1007/978-94-007-2666-6 (see pages 34, 52)

Needleman, A. (1988). Material rate dependence and mesh sensitivity in localization problems. *Computer Methods in Applied Mechanics and Engineering* 67 (1), 69–85. DOI: 10.1016/0045-7825(88)90069-2 (see page 3)

Newmark, N. M. (1959). A method of computation for structural dynamics. *Journal of the Engineering Mechanics Division* 85 (3), 67–94. DOI: 10.1061/JMCEA3.0000098 (see page 260)

Nguyen, T. T., Yvonnet, J., Zhu, Q.-Z., Bornert, M., and Chateau, C. (2015). A phase field method to simulate crack nucleation and propagation in strongly heterogeneous materials from direct imaging of their microstructure. *Engineering Fracture Mechanics* 139, 18–39. DOI: 10.1016/j.engfracmech.2015.03.045 (see page 4)

Nguyen, T. T., Yvonnet, J., Zhu, Q.-Z., Bornert, M., and Chateau, C. (2016). A phase-field method for computational modeling of interfacial damage interacting with crack propagation in realistic microstructures obtained by microtomography. *Computer Methods in Applied Mechanics and Engineering* 312, 567–595. DOI: 10.1016/j.cma.2015.10.007 (see page 75)

Ortiz, M. (1985). A constitutive theory for the inelastic behavior of concrete. *Mechanics of Materials* 4 (1), 67–93. DOI: 10.1016/0167-6636(85)90007-9 (see pages 3, 52)

Ostojic, P. and McPherson, R. (1987). A review of indentation fracture theory: Its development, principles and limitations. *International Journal of Fracture* 33 (4), 297–312. DOI: 10.1007/BF00044418 (see page 5)

Paris, P. C. (2014). A brief history of the crack tip stress intensity factor and its application. *Meccanica* 49, 759–764. DOI: 10.1007/s11012-014-9896-y (see pages 39, 44)

Peerlings, R. H. J., Borst, R. de, Brekelmans, W. A. M., and Vree, J. H. P. de (1995). Computational modelling of gradient-enhanced damage for fracture and fatigue problems. *Computational Plasticity – Fundamentals and Applications: Proceedings of the 4th International Conference, Barcelona*, 975–986 (see page 3)

Pham, K. H. (2015). Characterization and modeling of mixed mode I+III fracture in brittle materials. PhD thesis. University of Texas at Austin. URL: https://hdl.handle.net/2152/33391 (see pages 5, 156, 158)

Pham, K. H. and Ravi-Chandar, K. (2014). Further examination of the criterion for crack initiation under mixed-mode I+III loading. *International Journal of Fracture* 189 (2), 121–138. DOI: 10.1007/s10704-014-9966-0 (see page 4)

Pham, K. H. and Ravi-Chandar, K. (2016). On the growth of cracks under mixed-mode I+III loading. *International Journal of Fracture* 199 (1), 105–134. DOI: 10.1007/s10704-016-0098-6 (see pages 4, 5)

Pham, K. H. and Ravi-Chandar, K. (2017). The formation and growth of echelon cracks in brittle materials. *International Journal of Fracture* 206 (2), 229–244. DOI: 10.1007/s10704-017-0212-4 (see page 5)

Pham, K. H., Ravi-Chandar, K., and Landis, C. M. (2017). Experimental validation of a phase-field model for fracture. *International Journal of Fracture* 205 (1), 83–101. DOI: 10.1007/s10704-017-0185-3 (see pages 118, 200)

Pham, K., Amor, H., Marigo, J.-J., and Maurini, C. (2011). Gradient damage models and their use to approximate brittle fracture. *International Journal of Damage Mechanics* 20 (4), 618–652. DOI: 10.1177/1056789510386852 (see pages 4, 79, 88–90, 92)

Pham, K. and Marigo, J.-J. (2008). Stability of non-localized responses for damaging materials. *Vietnam Journal of Mechanics* 30 (4), 307–318. DOI: 10.15625/0866-7136/30/4/5636 (see pages 4, 102)

Pham, K. and Marigo, J.-J. (2009). Construction and analysis of localized responses for gradient damage models in a 1D setting. *Vietnam Journal of Mechanics* 31 (3-4), 233–246. DOI: 10.15625/0866-7136/31/3-4/5651 (see page 90)

Puri, G. M. (2011). *Python scripts for Abaqus: Learn by example*. Charleston: abaqus-python.com (see page 170)

Rabotnov, Y. N. (1959). On the mechanism of gradual failure, questions of strength of materials and construction. *Academy of Sciences*, 5–7 (see pages 3, 48)

Rabotnov, Y. N. (1969). Creep rupture. In: *Applied Mechanics*, 342–349 (see pages 3, 50)

Ramtani, S., Berthaud, Y., and Mazars, J. (1992). Orthotropic behavior of concrete with directional aspects: Modelling and experiments. *Nuclear Engineering and Design* 133 (1), 97–111. DOI: 10.1016/0029-5493(92)90094-C (see page 3)

Ramulu, M. and Kobayashi, A. S. (1985). Mechanics of crack curving and branching – a dynamic fracture analysis. *International Journal of Fracture* 27 (3), 187–201. DOI: 10.1007/BF00017967 (see page 153)

Ravi-Chandar, K. (2004). *Dynamic fracture*. Elsevier Science. DOI: 10.1016/B978-0-08-044352-2.X5000-5 (see pages 1, 32, 33)

Ravi-Chandar, K. and Knauss, W. G. (1984). An experimental investigation into dynamic fracture: III. On steady-state crack propagation and crack branching. *International Journal of Fracture* 26, 141–154. DOI: 10.1007/BF01157550 (see page 152)

Reddy, J. N. (2017). *Energy principles and variational methods in applied mechanics*. Third edition. Newark: John Wiley and Sons (see page 215)

Rice, J. R. (1968a). A path independent integral and the approximate analysis of strain concentration by notches and cracks. *Journal of Applied Mechanics* 35, 379–386. DOI: 10.1115/1.3601206 (see pages 42, 46)

Rice, J. R. (1968b). Mathematical analysis in the mechanics of fracture. *Fracture: An Advanced Treatise* 2, 191–311 (see page 43)

Richardson, T. J. and Mitter, S. K. (1997). A variational formulation-based edge focussing algorithm. *Sādhanā* 22, 553–574. DOI: 10.1007/BF02745579 (see page 63)

Roesler, F. C. (1956a). Brittle fractures near equilibrium. *Proceedings of the Physical Society, Section B* 69 (10), 981–992. DOI: 10.1088/0370-1301/69/10/303 (see pages 160, 192, 194)

Roesler, F. C. (1956b). Indentation hardness of glass as an energy scaling law. *Proceedings of the Physical Society, Section B* 69 (1), 55–60. DOI: 10.1088/0370-1301/69/1/307 (see pages 6, 160)

Rogers, C. A. (1998). *Hausdorff measures*. Cambridge: Cambridge University Press (see page 54)

Rosendahl, P. L., Weißgraeber, P., Stein, N., and Becker, W. (2017). Asymmetric crack onset at open-holes under tensile and in-plane bending loading. *International Journal of Solids and Structures* 113-114, 10–23. DOI: 10.1016/j.ijsolstr.2016.09.011 (see page 289)

Rosenfeld, A. (1982). *Digital picture processing*. Second edition. New York: Academic Press (see page 62)

Rossmanith, H. P. (1995). An introduction to K. Wieghardt's historical paper "On splitting and cracking of elastic bodies". *Fatigue and Fracture of Engineering Materials and Structures* 18 (12), 1367–1369. DOI: 10.1111/j.1460-2695.1995.tb00863.x (see page 38)

Sargado, J. M., Keilegavlen, E., Berre, I., and Nordbotten, J. M. (2018). High-accuracy phase-field models for brittle fracture based on a new family of degradation functions. *Journal of the Mechanics and Physics of Solids* 111, 458–489. DOI: 10.1016/j.jmps.2017.10.015 (see pages 4, 7, 83, 84)

Schänzel, L.-M. (2015). Phase field modeling of fracture in rubbery and glassy polymers at finite thermo-viscoelastic deformations. PhD thesis. University of Stuttgart. DOI: 10.18419/opus-613 (see page 4)

Schlüter, A. (2018). Phase field modeling of dynamic brittle fracture. PhD thesis. Technical University of Kaiserslautern. URL: https://nbn-resolving.de/urn:nbn:de:hbz:386-kluedo-53118 (see pages 72, 82, 89, 106, 154, 266)

Schlüter, A., Kuhn, C., Müller, R., and Gross, D. (2016). An investigation of intersonic fracture using a phase field model. *Archive of Applied Mechanics* 86 (1), 321–333. DOI: 10.1007/s00419-015-1114-4 (see page 3)

Schlüter, A., Kuhn, C., Müller, R., Tomut, M., Trautmann, C., Weick, H., and Plate, C. (2015). Phase field modelling of dynamic thermal fracture in the context of irradiation damage. *Continuum Mechanics and Thermodynamics*, 1–12. DOI: 10.1007/s00161-015-0456-z (see page 4)

Schlüter, A., Kuhn, C., Noll, T., Diewald, F., and Müller, R. (2017). Numerical solution strategies for a dynamic phase field fracture model. *Applied Mechanics and Materials* 869, 29–49. DOI: 10.4028/www.scientific.net/AMM.869.29 (see page 154)

Schlüter, A., Willenbücher, A., Kuhn, C., and Müller, R. (2014). Phase field approximation of dynamic brittle fracture. *Computational Mechanics* 54 (5), 1141–1161. DOI: 10.1007/s00466-014-1045-x (see pages 3, 68, 75, 122, 125, 154)

Schmitt, R., Kuhn, C., and Müller, R. (2015). *Martensitic transformations and damage: A combined phase field approach*. Presented at the 86th Annual Meeting of the International Association of Applied Mathematics and Mechanics, Lecce, 26th March (see page 128)

Schneider, D., Schoof, E., Huang, Y., Selzer, M., and Nestler, B. (2016). Phase-field modeling of crack propagation in multiphase systems. *Computer Methods in Applied Mechanics and Engineering* 312, 186–195. DOI: 10.1016/j.cma.2016.04.009 (see page 108)

Schröder, J. (1995). Theoretische und algorithmische Konzepte zur phänomenologischen Beschreibung anisotropen Materialverhaltens. PhD thesis. University of Hannover (see page 252)

Selvadurai, A. P. S. (2000). Fracture evolution during indentation of a brittle elastic solid. *Mechanics of Cohesive-frictional Materials* 5 (4), 325–339. DOI: 10.1002/(SICI)1099-1484(200005)5:4<325::AID-CFM98>3.0.CO;2-K (see page 5)

Seth, B. R. (1964). Generalized strain measure with applications to physical problems. *Proceedings of the IUTAM Symposium, Second-order effects in elasticity, plasticity and fluid dynamics*, 162–172 (see page 12)

Shao, J. F. and Rudnicki, J. W. (2000). A microcrack-based continuous damage model for brittle geomaterials. *Mechanics of Materials* 32 (10), 607–619. DOI: 10.1016/S0167-6636(00)00024-7 (see page 52)

Shelby, J. E. (2005). *Introduction to glass science and technology*. Second edition. Cambridge: Royal Society of Chemistry. DOI: 10.1039/9781847551160-FP005 (see page 9)

Shour, R. (2021). *Heinrich Hertz on the wisdom of equations*. DOI: 10.13140/RG.2.2.18108.97921/5 (see page i)

Sicsic, P. and Marigo, J.-J. (2013). From gradient damage laws to Griffith's theory of crack propagation. *Journal of Elasticity* 113 (1), 55–74. DOI: 10.1007/s10659-012-9410-5 (see page 149)

Sidoroff, F. (1981). Description of anisotropic damage application to elasticity. In: *Physical Non-Linearities in Structural Analysis*, 237–244 (see page 3)

Sih, G. C. (1973). Energy-density concept in fracture mechanics. *Engineering Fracture Mechanics* 5 (4), 1037–1040. DOI: 10.1016/0013-7944(73)90072-6 (see page 44)

Simo, J. C. and Hughes, T. J. R. (1998). *Computational inelasticity*. New York: Springer. DOI: 10.1007/b98904 (see page 225)

Simo, J. C., Oliver, J., and Armero, F. (1993). An analysis of strong discontinuities induced by strain softening in rate-independent inelastic solids. *Computational Mechanics* 12, 277–296. DOI: 10.1007/BF00372173 (see page 148)

Simo, J. C. and Rifai, M. S. (1990). A class of mixed assumed strain methods and the method of incompatible modes. *International Journal for Numerical Methods in Engineering* 29 (8), 1595–1638. DOI: 10.1002/nme.1620290802 (see page 148)

Singh, N., Verhoosel, C. V., and Brummelen, E. H. van (2018). Finite element simulation of pressure-loaded phase-field fractures. *Meccanica* 53 (6), 1513–1545. DOI: 10.1007/s11012-017-0802-2 (see pages 4, 138)

Skrzypek, J. J. (2015). *Mechanics of anisotropic materials*. Cham: Springer. DOI: 10.1007/978-3-319-17160-9 (see page 52)

Sluys, L. J. (1992). Wave propagation, localisation and dispersion in softening solids. PhD thesis. Delft University of Technology. URL: https://resolver.tudelft.nl/uuid:0f9b3de9-e0ec-4d9b-a42f-fcf704d5d40e (see page 3)

Sneddon, I. N. (1946a). Boussinesq's problem for a flat-ended cylinder. *Mathematical Proceedings of the Cambridge Philosophical Society* 42 (1), 29–39. DOI: 10.1017/S0305004100022702 (see pages 269, 271)

Sneddon, I. N. (1946b). The distribution of stress in the neighbourhood of a crack in an elastic solid. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences* 187 (1009), 229–260. DOI: 10.1098/rspa.1946.0077 (see page 38)

Sonar, T. (2016). *3000 Jahre Analysis: Geschichte – Kulturen – Menschen*. DOI: 10.1007/978-3-662-48918-5 (see page 92)

Song, J.-H., Wang, H., and Belytschko, T. (2008). A comparative study on finite element methods for dynamic fracture. *Computational Mechanics* 42 (2), 239–250. DOI: 10.1007/s00466-007-0210-x (see page 153)

Spatschek, R., Brener, E. A., and Karma, A. (2011). Phase field modeling of crack propagation. *Philosophical Magazine* 91 (1), 75–95. DOI: 10.1080/14786431003773015 (see page 66)

Steinke, C. and Kaliske, M. (2019). A phase-field crack model based on directional stress decomposition. *Computational Mechanics* 63 (5), 1019–1046. DOI: 10.1007/s00466-018-1635-0 (see pages 4, 84, 112, 141, 142, 208, 241, 251, 267)

Steinke, C., Zreid, I., and Kaliske, M. (2017). On the relation between phase-field crack approximation and gradient damage modelling. *Computational Mechanics* 59 (5), 717–735. DOI: 10.1007/s00466-016-1369-9 (see pages 75, 81)

Stolze, C. H. (1978). A history of the divergence theorem. *Historia Mathematica* 5 (4), 437–442. DOI: 10.1016/0315-0860(78)90212-4 (see page 219)

Storm, J., Supriatna, D., and Kaliske, M. (2019). On the analysis of crack-closure behaviour using the phase-field method together with the novel concept of Representative Crack Elements. *Proceedings in Applied Mathematics and Mechanics* 19 (1), e201900314. DOI: 10.1002/pamm.201900314 (see page 142)

Storm, J., Supriatna, D., and Kaliske, M. (2020). The concept of representative crack elements for phase-field fracture: Anisotropic elasticity and thermo-elasticity. *International Journal for Numerical Methods in Engineering* 121 (5), 779–805. DOI: 10.1002/nme.6244 (see page 142)

Streich, F. A. (2015). Analyse von Phasenfeldmodellen zur Rissausbreitung. MA thesis. Karlsruhe Institute of Technology (see page 75)

Strobl, M., Dowgiałło, P., and Seelig, T. (2017). Analysis of Hertzian indentation fracture in the framework of finite fracture mechanics. *International Journal of Fracture* 206 (1), 67–79. DOI: 10.1007/s10704-017-0201-7 (see page 7)

Strobl, M., Morand, L., and Seelig, T. (2016). Simulation of Hertzian cone cracks using a phase field description for fracture. *Proceedings in Applied Mathematics and Mechanics* 16 (1), 177–178. DOI: 10.1002/pamm.201610077 (see page 194)

Strobl, M. and Seelig, T. (2015). A novel treatment of crack boundary conditions in phase field models of fracture. *Proceedings in Applied Mathematics and Mechanics* 15 (1), 155–156. DOI: 10.1002/pamm.201510068 (see pages 7, 114, 117, 124, 135, 137, 140, 249, 251)

Strobl, M. and Seelig, T. (2016). On constitutive assumptions in phase field approaches to brittle fracture. *Procedia Structural Integrity* 2, 3705–3712. DOI: 10.1016/j.prostr.2016.06.460 (see pages 7, 86, 112, 114, 115, 118, 268)

Strobl, M. and Seelig, T. (2018). Restrictions in phase field modeling of brittle fracture. *Proceedings in Applied Mathematics and Mechanics* 18 (1), e201800157. DOI: 10.1002/pamm.201800157 (see pages 82, 108, 143, 150, 208, 244)

Strobl, M. and Seelig, T. (2019). Analysis of Hertzian indentation fracture using a phase field approach. *Proceedings in Applied Mathematics and Mechanics* 19 (1), e201900257. DOI: 10.1002/pamm.201900257 (see page 151)

Strobl, M. and Seelig, T. (2020). Phase field modeling of Hertzian indentation fracture. *Journal of the Mechanics and Physics of Solids* 143, 104026. DOI: 10.1016/j.jmps.2020.104026 (see page 7)

Sukumar, N., Moës, N., Moran, B., and Belytschko, T. (2000). Extended finite element method for three-dimensional crack modelling. *International Journal for Numerical Methods in Engineering* 48 (11), 1549–1570. DOI: 10.1002/1097-0207(20000820)48:11<1549::AID-NME955>3.0.CO;2-A (see page 3)

Tada, H., Paris, P. C., and Irwin, G. R. (2000). *The stress analysis of cracks handbook*. Third edition. New York: American Society of Mechanical Engineers Press. DOI: 10.1115/1.801535 (see pages 39, 168, 275)

Tanné, E., Li, T., Bourdin, B., Marigo, J.-J., and Maurini, C. (2018). Crack nucleation in variational phase-field models of brittle fracture. *Journal of the Mechanics and Physics of Solids* 110, 80–99. DOI: 10.1016/j.jmps.2017.09.006 (see page 4)

Taylor, D. (2007). *The theory of critical distances: A new perspective in fracture mechanics*. Amsterdam: Elsevier. DOI: 10.1016/B978-0-08-044478-9.X5000-5 (see page 46)

Taylor, D., Cornetti, P., and Pugno, N. (2005). The fracture mechanics of finite crack extension. *Engineering Fracture Mechanics* 72 (7), 1021–1038. DOI: 10.1016/j.engfracmech.2004.07.001 (see pages 2, 45)

Taylor, R. L. (2013). *FEAP – A finite element analysis program: Version 8.4 theory manual* (see pages 253, 254, 260, 261)

Taylor, R. L. (2014). *FEAP – Finite Element Analysis Program* (see page 142)

Tillett, J. P. A. (1956). Fracture of glass by spherical indenters. *Proceedings of the Physical Society, Section B* 69 (1), 47–54. DOI: 10.1088/0370-1301/69/1/306 (see pages 6, 160, 209)

Torquato, S. (2002). *Random heterogeneous materials: Microstructure and macroscopic properties*. New York: Springer. DOI: 10.1007/978-1-4757-6355-3 (see page 140)

Tracey, D. M. (1971). Finite elements for determination of crack tip elastic stress intensity factors. *Engineering Fracture Mechanics* 3 (3), 255–265. DOI: 10.1016/0013-7944(71)90036-1 (see page 1)

Truesdell, C. A. (1980). Act IV. Internal energy: The first paper of Clausius. Entropy: The first paper of Rankine. In: *The Tragicomical History of Thermodynamics, 1822–1854*. Springer, 187–218. DOI: 10.1007/978-1-4613-9444-0_8 (see page 20)

Tumbajoy-Spinel, D. Y., Feulvarch, É., Bergheau, J.-M., and Kermouche, G. (2013). 2D axisymmetric X-FEM modeling of the Hertzian cone crack system. *Comptes Rendus Mécanique* 341 (9), 715–725. DOI: 10.1016/j.crme.2013.09.004 (see pages 5, 192, 194)

Turner, M. J., Clough, R. W., Martin, H. C., and Topp, L. J. (1956). Stiffness and deflection analysis of complex structures. *Journal of the Aeronautical Sciences* 23 (9), 805–823. DOI: 10.2514/8.3664 (see page 2)

Verhoosel, C. V. and Borst, R. de (2013). A phase-field model for cohesive fracture. *International Journal for Numerical Methods in Engineering* 96 (1), 43–62. DOI: 10.1002/nme.4553 (see pages 4, 78, 138)

Verhoosel, C. V., Zwieten, G. J. van, Rietbergen, B. van, and Borst, R. de (2015). Image-based goal-oriented adaptive isogeometric analysis with application to the micro-mechanical modeling of trabecular bone. *Computer Methods in Applied Mechanics and Engineering* 284, 138–164. DOI: 10.1016/j.cma.2014.07.009 (see page 138)

Vignollet, J., May, S., Borst, R. de, and Verhoosel, C. V. (2014). Phase-field models for brittle and cohesive fracture. *Meccanica* 49 (11), 2587–2601. DOI: 10.1007/s11012-013-9862-0 (see page 4)

Wang, X.-Y., Li, L. K.-Y., Mai, Y.-W., and Shen, Y.-G. (2008). Theoretical analysis of Hertzian contact fracture: Ring crack. *Engineering Fracture Mechanics* 75 (14), 4247–4256. DOI: 10.1016/j.engfracmech.2008.03.003 (see page 5)

Warren, R. (1978). Measurement of the fracture properties of brittle solids by Hertzian indentation. *Acta Metallurgica* 26 (11), 1759–1769. DOI: 10.1016/0001-6160(78)90087-1 (see page 277)

Washizu, K. (1955). *On the variational principles of elasticity and plasticity*. Massachusetts Institute of Technology, Aeroelastic and Structures Research Laboratory (see page 29)

Washizu, K. (1975). *Variational methods in elasticity and plasticity*. Second edition. Pergamon Press (see page 29)

Weinberg, K. and Hesch, C. (2015). A high-order finite deformation phase-field approach to fracture. *Continuum Mechanics and Thermodynamics*, 1–11. DOI: 10.1007/s00161-015-0440-7 (see pages 4, 77)

Weissenburger, J. T. (1964). Integration of discontinuous expressions arising in beam theory. *AIAA Journal* 2 (1), 106–108. DOI: 10.2514/3.2224 (see page 112)

Weißgraeber, P., Leguillon, D., and Becker, W. (2016). A review of finite fracture mechanics: crack initiation at singular and non-singular stress raisers. *Archive of Applied Mechanics* 86 (1), 375–401. DOI: 10.1007/s00419-015-1091-7 (see pages 2, 46)

Wells, A. A. (1961). Unstable crack propagation in metals-cleavage and fast fracture. *Proceedings of the Crack Propagation Symposium, Cranfield*, 210–230 (see page 46)

Westergaard, H. M. (1939). Bearing pressures and cracks. *Journal of Applied Mechanics* 6, 49–53 (see pages 38, 230)

Wheeler, M. F., Wick, T., and Wollner, W. (2014). An augmented-Lagrangian method for the phase-field approach for pressurized fractures. *Computer Methods in Applied Mechanics and Engineering* 271, 69–85. DOI: 10.1016/j.cma.2013.12.005 (see pages 4, 75)

Wiederhorn, S. M. and Townsend, P. R. (1970). Crack healing in glass. *Journal of the American Ceramic Society* 53 (9), 486–489. DOI: 10.1111/j.1151-2916.1970.tb15996.x (see page 71)

Wieghardt, K. (1907). Über das Spalten und Zerreißen elastischer Körper. *Zeitschrift für Mathematik und Physik* 55 (1), 2 (see page 38)

Williams, M. L. (1952). Stress singularities resulting from various boundary conditions in angular corners of plates in extension. *Journal of Applied Mechanics* 19 (4), 526–528 (see page 38)

Williams, M. L. and Meyerson, M. R. (1949). *Investigation of fracture in the T1 tanker Capitan*. Washington: Ship structure committee (see page 1)

Willner, K. (2003). *Kontinuums- und Kontaktmechanik – Synthetische und analytische Darstellung*. Springer. DOI: 10.1007/978-3-642-55814-6 (see page 31)

Wilson, Z. A., Borden, M. J., and Landis, C. M. (2013). A phase-field model for fracture in piezoelectric ceramics. *International Journal of Fracture* 183 (2), 135–153. DOI: 10.1007/s10704-013-9881-9 (see pages 4, 83)

Wilson, Z. A. and Landis, C. M. (2016). Phase-field modeling of hydraulic fracture. *Journal of the Mechanics and Physics of Solids* 96, 264–290. DOI: 10.1016/j.jmps.2016.07.019 (see page 4)

Wriggers, P. (2008). *Nonlinear finite element methods*. Berlin: Springer. DOI: 10.1007/978-3-540-71001-1 (see pages 225, 253)

Wu, C. H. (1978). Maximum-energy-release-rate criterion applied to a tension-compression specimen with crack. *Journal of Elasticity* 8 (3), 235–257. DOI: 10.1007/BF00130464 (see page 45)

Wu, J.-Y. (2017). A unified phase-field theory for the mechanics of damage and quasi-brittle failure. *Journal of the Mechanics and Physics of Solids* 103 (Supplement C), 72–99. DOI: 10.1016/j.jmps.2017.03.015 (see pages 4, 79, 87, 88, 90, 239)

Xiang, H. P., Qian, H. J., Lu, Z. Y., Rong, M. Z., and Zhang, M. Q. (2015). Crack healing and reclaiming of vulcanized rubber by triggering the rearrangement of inherent sulfur crosslinked networks. *Green Chemistry* 17, 4315–4325. DOI: 10.1039/C5GC00754B (see page 71)

Yin, B. and Kaliske, M. (2019). Fracture simulation of viscoelastic polymers by the phase-field method. *Computational Mechanics*. DOI: 10.1007/s00466-019-01769-1 (see pages 4, 84)

Yin, B., Steinke, C., and Kaliske, M. (2020). Formulation and implementation of strain rate-dependent fracture toughness in context of the phase-field method. *International Journal for Numerical Methods in Engineering* 121 (2), 233–255. DOI: 10.1002/nme.6207 (see page 84)

Yingzhi, L. and Hills, D. A. (1991). The Hertzian cone crack. *Journal of Applied Mechanics* 58 (1), 120–127. DOI: 10.1115/1.2897138 (see page 5)

Yoshioka, K. and Naumov, D. (2019). *On the crack opening computation in phase-field models for fracture*. Presented at the 6th GAMM Workshop on Phase-Field Modeling, Karlsruhe, 8th February (see page 139)

Yoshioka, K., Naumov, D., and Kolditz, O. (2020). On crack opening computation in variational phase-field models for fracture. *Computer Methods in Applied Mechanics and Engineering* 369, 113210. DOI: 10.1016/j.cma.2020.113210 (see pages 139, 208)

Yosibash, Z., Priel, E., and Leguillon, D. (2006). A failure criterion for brittle elastic materials under mixed-mode loading. *International Journal of Fracture* 141 (1), 291–312. DOI: 10.1007/s10704-006-0083-6 (see page 46)

Zi, G. and Belytschko, T. (2003). New crack-tip elements for XFEM and applications to cohesive cracks. *International Journal for Numerical Methods in Engineering* 57 (15), 2221–2240. DOI: 10.1002/nme.849 (see page 3)

Zienkiewicz, O. C., Taylor, R. L., and Fox, D. (2014). *The finite element method for solid and structural mechanics*. Seventh edition. Oxford: Butterworth-Heinemann. DOI: 10.1016/C2009-0-26332-X (see pages 253, 261)

Zienkiewicz, O. C., Taylor, R. L., and Zhu, J. (2013). *The finite element method: Its basis and fundamentals*. Seventh edition. Oxford: Butterworth-Heinemann. DOI: 10.1016/C2009-0-24909-9 (see pages 75, 139, 253, 255)