

Efficient Earthquake Inversion using the Finite Element Method

van Zwieten, G.J.

DOI

[10.4233/uuid:e9420781-7c76-453b-bacb-d806c8bf5fa8](https://doi.org/10.4233/uuid:e9420781-7c76-453b-bacb-d806c8bf5fa8)

Publication date

2023

Document Version

Final published version

Citation (APA)

van Zwieten, G. J. (2023). *Efficient Earthquake Inversion using the Finite Element Method*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:e9420781-7c76-453b-bacb-d806c8bf5fa8>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

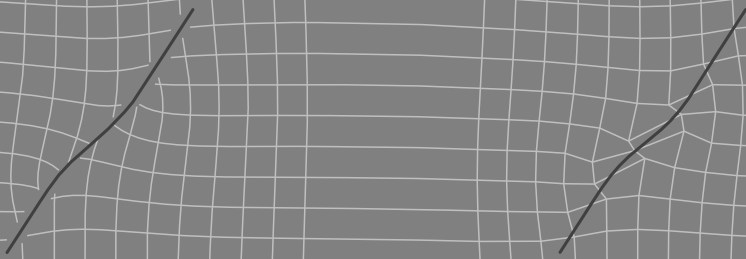
Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Efficient Earthquake Inversion using the Finite Element Method



GJ VAN ZWIETEN

Efficient Earthquake Inversion using the Finite Element Method

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
dinsdag 21 maart 2023 om 15:00 uur

door

Gerrit Johannes VAN ZWIETEN
ingenieur in de Technische Wiskunde,
Technische Universiteit Delft, Nederland
geboren te Moordrecht, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. R.F. Hanssen	Technische Universiteit Delft, promotor
Prof. dr. ir. E.H. van Brummelen	Technische Universiteit Eindhoven, promotor

Onafhankelijke leden:

Prof. dr. S. Jónsson	King Abdullah University of Science and Technology
Prof. dr. D. Volkov	Worcester Polytechnic Institute
Prof. dr. H.G. Matthies	Technische Universität Braunschweig
Prof. dr. ir. R.M.A. Govers	Universiteit Utrecht
Prof. dr. ir. L.J. Sluys	Technische Universiteit Delft
Prof. dr. ir. E.C. Slob	Technische Universiteit Delft, reservelid



Gerrit Johannes VAN ZWIETEN

Department of Geoscience and Remote Sensing, Delft University of Technology

Keywords: earthquake, surface deformation, elastic dislocation, synthetic aperture radar interferometry (InSAR), Finite Element Method (FEM), Weakly-enforced Slip Method (WSM), inverse problem.

ISBN 978-94-6366-673-2

Copyright © 2023 by Gerrit Johannes VAN ZWIETEN

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without the prior permission of the author.

Type set by the author with the L^AT_EX Documentation System.

An electronic version of this dissertation is available at
<http://repository.tudelft.nl/>.

For Amal

Preface

As I write this introduction, attempts are still ongoing to rescue survivors of the magnitude 7.8 earthquake that struck Türkiye and Syria in the night of February 6th, and death tolls are adjusted upward day after day. It is the latest tragedy in a long history of tragedies, and a grim reminder that for those living in fault zones there is still not much in the way of warning that disaster is afoot.

The work presented in this thesis will, I am sad to say, not change much about this state of affairs. It presents no means to predict earthquakes, which to date remains an unsolved problem. Instead, it offers a technique to improve on the established practice of looking back at past events, such as this Gaziantep earthquake, to understand in detail what happened at the fault level. The hope is that such understanding informs experts about what remains in store for the area, and that this, in turn, informs policy makers in their efforts to mitigate the consequences of the next event in the cycle. A cog in a machine, therefore — which is not to diminish the value of the cog.

I am grateful to have been given excellent circumstances to learn. When I embarked on this project in 2006, newly graduated from mathematics, I lacked backgrounds in both geophysics and satellite interferometry. Today, while still far from an expert in either topic, I like to think that I gained sufficient understanding to be able to see my own work in context. I have Prof. Rob Govers and Prof. Rinus Wortel to thank for teaching me the basics of tectonophysics, in courses that I enjoyed both for the subject matter and for the engaging style of teaching. I would like to thank Prof. Sigurjón Jónsson for welcoming me to KAUST, and allowing me to participate in a memorable field trip to the fault zones of Harrat Lunayyir. And I thank my supervisor Prof. Ramon Hanssen and my colleagues from the MGP group to whom I owe my knowledge of radar interferometry, learning through osmosis by watching science progress in our weekly radar meeting.

I am also lucky to have found myself in a great environment for my own research. Due to the nature of the topic, the gravity of this project has rested from the start on the domain of solid mechanics, for which reason I was made member of the Engineering Mechanics group, as well as the EM graduate school. I would like to thank my then supervisor Miguel Gutiérrez, and my many colleagues there for the interesting discussions and for making my time there a pleasurable and interesting one. In particular I would like to mention Clemens Verhoosel, Kris van der Zee, and naturally Wijnand Hoitinga, with whom I would later found HvZ Engineering (now Evalf Computing).

When my PhD contract ended with no thesis in hand, it was Prof. Harald van Brummelen who took me under his wing by offering me a research position at Eind-

hoven university. We had many insightful discussions on the topic of this thesis, and his help has been instrumental to develop the convergence proofs that form the heart of Chapter 3. I owe a great debt of gratitude for creating the circumstances in which I could not only progress on this thesis, but also lay the foundations of the Nutils framework, the open source computing library that ended up powering the simulations of Chapter 4.

Mathematicians study a problem in abstraction, and it can be very fulfilling to find elegance in this space that is stripped of the specifics of reality. Yet, as much as I enjoyed studying the problem for its own sake, I did find it motivating to think that the intended application was a lofty one. As four years turned to eight, and eight to sixteen, this has been an increasingly important factor, and I am not certain that this thesis would have seen completion without the benefit of this driving thought. However modest the scope of this work, I believe that the techniques presented herein are practical, and I sincerely hope they will find some use towards the goal of improving earthquake preparedness in high risk areas.

Gertjan van Zwieten
February 12, 2023

Contents

Preface	v
1 Introduction	1
2 Classic Solutions to Volterra’s Dislocation Problem	9
2.1 Problem definition	11
2.2 Analytical solutions	14
2.2.1 Volterra’s equation	14
2.2.2 The homogeneous halfspace	16
2.2.3 Closed form solutions	17
2.3 Propagator matrix	19
2.4 Series expansion	23
2.4.1 Material heterogeneity	23
2.4.2 Topography	24
2.5 Finite Elements	25
2.5.1 Weak formulation	25
2.5.2 Finite Element Method	27
2.6 Conclusions	29
3 The Weakly-enforced Slip Method	31
3.1 Problem formulation	33
3.1.1 The strong form	34
3.1.2 The weak form	36
3.1.3 Finite element approximation	38
3.2 The Weakly-enforced Slip Method	39
3.2.1 Collapsing the lift	40
3.2.2 Alternative derivation via Nitsche’s variational principle	41
3.3 Approximation properties of WSM	43
3.3.1 Traces and tractions	43
3.3.2 Global approximation properties of WSM	45
3.3.3 Local approximation properties	52
3.4 Numerical results	56
3.4.1 Test case I: 2D plane strain	57
3.4.2 Test case II: 3D traction-free halfspace	60
3.4.3 Test case III: 3D traction-free rupturing halfspace	62

3.5	Conclusions	64
4	Inverting elastic dislocations using the WSM	67
4.1	Forward Problem: Linear Elastic Dislocation	69
4.1.1	Analytical solutions	71
4.1.2	The Weakly-enforced Slip Method	71
4.2	Inverse Problem: Bayesian Formulation	72
4.2.1	Prior Distribution	73
4.2.2	The Likelihood	75
4.2.3	Posterior Distribution	76
4.3	Methodology	77
4.3.1	Constructing fault parameters	78
4.3.2	Constructing the slip distribution	79
4.3.3	Synthesizing observation data	81
4.3.4	Sampling the posterior distribution	83
4.3.5	Evaluating the posterior expected slip	85
4.4	Results	86
4.4.1	Linear inversion: slip distribution	86
4.4.2	Nonlinear inversion: fault parameters	91
4.5	Conclusions	96
5	Conclusions and Outlook	99
5.1	Research Questions	99
5.2	Recommendations	102
	Bibliography	105
	Summary	111
	Samenvatting	113
	Curriculum Vitae	115

1 | Introduction

Throughout human history, people have known to treat the ground under their feet with a degree of mistrust. Land slides, volcanic eruptions, earthquakes — all examples of violent, instantaneous events that, while infrequent, have the potential to cause widespread devastation and loss of life with very little in the way of warning. Frustratingly, while not all locations are equally prone to disaster, it is often the areas that present the highest risk that provide the most favourable conditions for settlement — volcanic ash makes for highly fertile soil [22]; tectonic structures may form natural containers for ground water [41]. When the best place to live is simultaneously the worst place to be, it is all societies can do to mitigate the consequence of the inevitable.

The subject matter of the present thesis is earthquakes: the instantaneous release of potential energy when a section of a tectonic fault gives way to accumulated strain [42]. The characteristic seismic waves can last for seconds or minutes depending on the strength of the quake. As the released energy dissipates, the earth assumes a new configuration of static equilibrium to accommodate the internal change. This new configuration can contain elevation changes, or the lateral displacement of land masses, depending on the original stress conditions. Over a large period of time, the process is often seen to repeat with a recurrence rate that is specific to the fault segment, in what is known as the earthquake cycle [69].

Aiming to mitigate the consequences of seismicity, an obvious field of study is the search for precursor events: anomalous phenomena preceding large earthquakes that can serve as a warning of imminent danger. Commonly known in this category are anecdotal reports of anomalous animal behaviour in the moments before an earthquake; other observables that have been studied in recent history include changes in gas emissions, electro-magnetism, ground temperature and micro-seismicity [32]. Unfortunately, a universally reliable precursor has so far proved elusive, leading some to believe that the search will remain fruitless [33]. Early warning systems that are in operation today warn instead of earthquakes that have just occurred, giving residents in adjoining areas precious moments to seek shelter before the seismic wave arrives [3].

The other field of study that aims to mitigate the consequences of earthquakes is that of seismic hazard analysis, which aims to inform long term mitigation strategies through products such as seismic hazard maps. Urban planning authorities, for instance, may use these to develop building codes, assure emergency preparedness and plan for post-disaster recovery [30]. Though many approaches to seismic hazard analysis exist and continue to evolve [5], what collectively sets them apart from precursor studies is that they are rooted in past events, operating from the point of view that

the past informs the future. As such, they benefit greatly from data gathering efforts.

While the present thesis stops short of actual seismic hazard analysis, it is in this broader context that it sets out to develop novel techniques for the analysis of past seismic events. In particular, this thesis will aim to contribute to the analysis of satellite based InSAR observations.

Interferometric Synthetic Aperture Radar (InSAR)

In June 1978, NASA launched the Seasat satellite mission, aimed at showing the feasibility of remote monitoring of the oceans. In addition to sensors to measure height, wind speed, surface temperature and cloud features, Seasat was one of the first satellite missions to carry a Synthetic Aperture Radar (SAR) sensor, which brought the ability to monitor surface wave fields and sea ice unencumbered by cloud cover, rain, or darkness. Despite the relatively short duration of the mission, which was cut short at 105 days, the collected data showed sufficient promise that other SAR missions soon followed, and from 1992 onward the earth's surface, both sea and land, has been continuously monitored by an ever growing constellation of satellites [6].

Synthetic Aperture Radar operates by the same principles as classical radar, sending electromagnetic radiation in the direction of the area under observation and forming an image from the reflected signal, in which the arrival time is a measure of distance. Unlike classical radar, a SAR antenna utilises its own motion by processing a stream of data from continuously changing positions, and 'focussing' this data numerically into an image with much higher resolution than warranted by its physical aperture size. In satellite missions, the orbital motion allows for synthetic aperture sizes of many kilometers, resulting in a spatial resolution in the order of meters or less despite orbiting at a height of 500-800 km for typical missions [36].

While it was Seasat's mission to monitor the oceans, it soon became apparent that SAR's real strength is over land. Here, SAR data can be interpreted in a way that allows for very high precision deformation measurements, by using the electromagnetic phase, rather than the intensity, of the returned signal. Though the phase information of a single image holds essentially no information, the phase difference between two subsequently taken images may directly relate to the change in distance to the satellite. Extracting this information is known as interferometry [36]. Its primary data product is the interferogram, an example of which can be seen in Figure 1.1.

While measuring in fractions of the radar wavelength holds the incredible promise of millimeter precision data, there are a few important caveats to be made. First of all, while we might know the distance change to a fraction of the wavelength, we cannot know the number of full wavelengths to which the fraction is to be added. This problem of phase ambiguities means that the interferogram can only be interpreted as relative data, under the assumption that neighbouring pixels move similarly to within a single wavelength. By accumulating the smallest phase difference from neighbour to neighbour, the relative deformation component can be computed between any two points in the radar image. This process is known as phase unwrapping, resulting in an unwrapped interferogram as seen in Figure 1.1.

The data is not perfect. Beyond its inherent relativeness, the necessary assumption of smooth deformations, and the fact the deformation is measured only in the line of sight component, there are noise sources such as heterogeneous atmospheric delay [35]

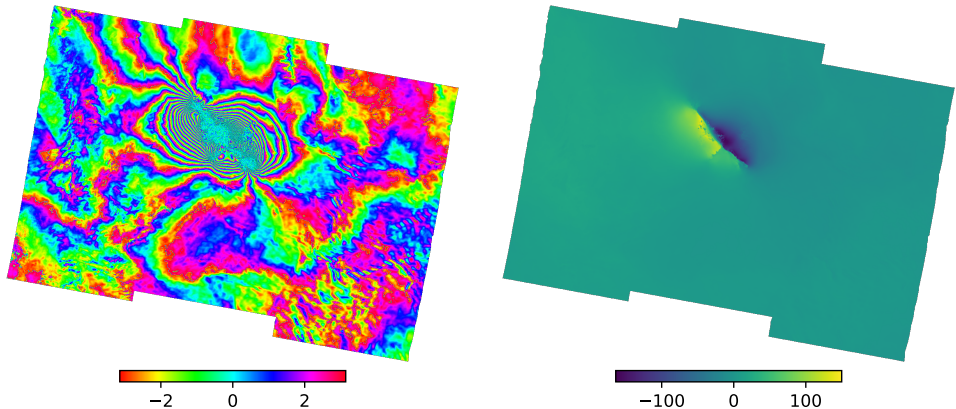


Figure 1.1 – An example of co-seismic deformation data from satellite borne Synthetic Aperture Radar, showing deformation caused by the 2019 Ridgecrest, California earthquake as observed by the Sentinel-1 satellite mission. Left: the interferogram, showing the differential phase of the returned signal, with every cycle of 2π corresponding to a line-of-sight deformation of 2.77 cm. Right: the unwrapped interferogram, showing line-of-sight deformation relative to a point in the far field. Credit: ASF DAAC HyP3 2022. Contains modified Copernicus Sentinel data 2019, processed by ESA.

and changes in the scattering behaviour that complicate interpretation, any of which necessitate their own mitigating processing steps. But none of that takes away from the power of having high density measurements with high revisit times of the entire globe without having to install active surveying equipment, offering unprecedented data to study complex deformation processes.

Deformation based Earthquake Inversions

Inversion studies aim to infer from observations the rupture mechanism behind an earthquake. Arguably the most established branch of inversion science is that of seismology, which uses seismicity as the primary observable. By measuring the arrival time of the seismic waves at different geographical locations, it is possible to triangulate the hypocenter of the quake. Including other characteristics of the seismic waves it is possible to infer additional information such as the orientation of the fault plane and the slip vector [24]. Nonetheless, the nature of the data largely limits results to aggregate quantities, rather than a detailed picture of the event.

More detailed results are promised by deformation data, which is closer to the source than most seismic stations and therefore subject to less noise and requiring fewer model assumptions. When Massonnet et al. used InSAR data from the recently launched ERS-1 satellite for their analysis of the 1992 Landers earthquake [49], the first of its kind, this promptly made the cover of the *Nature* issue in which it appeared. In his accompanying view in the same issue, William Prescott observes [58]:

“This map may look familiar to some readers. We have seen similar maps showing the theoretical displacement field of earthquakes. But no other

technique has been able to map the actual displacement field with anything like this kind of detail.”

Heralding InSAR as an important new tool for studying earth deformation, he concludes:

“Provided that the difficulties caused by ‘uninteresting’ changes [...] are not too serious, the impact of the technique could be huge.”

Indeed, many seismic events have since been analysed using InSAR data. And as ever more accurate sensors are launched and InSAR processing techniques continue to improve, this is only expected to increase [9].

When transient dynamics are ignored, an earthquake is characterized entirely by the relative displacement of adjoining tectonic plates. Mathematically this constitutes a dislocation: a compact, two-dimensional manifold over which the displacement field jumps discontinuously. The jump is commonly referred to as fault slip, which varies over the manifold in direction and magnitude and thus forms a vector field. Taken together, the fault geometry and slip form boundary conditions to the governing laws of physics — laws that can be modeled mathematically and simulated numerically. Inversion studies aim to use this relation to infer from observations the faulting mechanism that caused it.

For deformation based inversions, the governing law is that of static equilibrium. The main model requirements, therefore, are a description of local topography, which defines the traction free surface, and a 3D characterization of the earth’s mechanical properties. While the latter is a difficult assignment in general, the situation is helped by the common assumption that on the very short timescales that an earthquake occurs, the earth behaves to good approximation elastically (e.g. [44]), leaving only its stiffness to be determined. This material property, and its organisation in geological layers and other spatial variations, is something that field experts may be able to provide with detail and accuracy depending on available geological data.

The key element of an inversion study, then, is a means of linking any fault geometry and slip to the corresponding deformations at the surface by solving for static equilibrium, subject to the known topography and material properties. Many such solution methods exist, ranging from analytical expressions for severely simplified scenarios, to full fledged numerical techniques capable of incorporating any expert knowledge that is available.

Linear Elasticity and the Finite Element Method (FEM)

Static equilibrium refers to the configuration of a system at rest. It is a well known principle of classical mechanics that when infinitesimally small deviations are made from this equilibrium state by applying external work, this is perfectly balanced by a matching internal work as the system resists its deformation. Versions of this principle have been known since antiquity, but it was Johann Bernoulli who in 1714 was the first to introduce the notion of infinitesimal displacements. While Bernoulli named his principle that of ‘virtual velocities’, it would be with the introduction of the physics concept of ‘work’ a century later that the principle received the name ‘principle of virtual work’ by which it is known today [15].

For a deformable body, the principle of virtual work applies to deformations, rather than displacements. Internal work is due to the intra-molecular force as deformation changes the molecular arrangement of the medium, leading a rod to resist extension as it is pulled. If this process is reversible, the medium is said to be elastic. In his experiments halfway the 17th century, Robert Hooke found a spring’s extension to be approximately proportional to the applied force, culminating in the formulation of his famous empirical law. Generalized to higher dimensional continua, Hooke’s law was later found to be a useful abstraction for an entire class of materials, and came to define the field of linear elasticity theory [47].

The principle of virtual work is particularly practical as it provides a mathematical technique for solving the equilibrium problem. If both the deformation of the medium and its virtual displacements can be represented as a superposition of finitely many deformation modes, the design of which is entirely free, then the principle of virtual work directly translates into a system of equations with equally many unknowns. If, moreover, the internal virtual work is proportional to deformation, that is, if the material is linearly elastic, then this system is linear and straightforwardly solvable using standard linear algebra techniques.

For this procedure to be practical, one problem remains, which is that the system of equations produce matrices that are typically dense, consuming a lot of memory and being expensive to solve. This problem was solved in the 20th century with the development of mesh discretization techniques: by subdividing a domain in finitely many subdomains, the resulting mesh gives the required structure for both creating basis functions with very local support and efficiently integrating these functions over the domain. Expressing the deformation and virtual displacements in terms of these basis functions, the matrix becomes highly sparse, and its solution affordable. This is the basis of the Finite Element Method [75], which in the ensuing century has proved to be an incredibly successful approximation method for a wider range of differential equations.

Dislocations fit the framework very well. While the internal work functional changes due to locked in stresses, the virtual work principle still applies, as does its treatment in a discrete function space. Details of this have been worked out by Melosh and Raefsky in 1981, who proposed that the finite element mesh should conform to the dislocation so that the discontinuity cuts between elements [51]. Their *split-node* methodology still ranks high as a very powerful and versatile method for solving arbitrary dislocation problems.

Thesis Objective and Research Questions

Inverse analysis benefits from a model that is as close to what we know of reality as possible. Given that the split-node method exists, it seems the obvious choice to use as part of an earthquake inversion study. It is flexible enough to incorporate any expert knowledge that is available, has readily available implementations, and can be solved to — in principle — any level of accuracy.

The split-node method is, however, hardly ever used in this setting. The reason is that, while it is perfectly feasible to simulate a single dislocation using the split-node method, a typical inversion requires many thousands of evaluations as part of a nonlinear inversion algorithm. It is in this setting that feasibility breaks down: as

the split-node method dictates that elements align with the fault, any new fault that is considered in the inversion requires a new mesh, which in turn implies a new basis and new matrix assembly and factorization. Taken together, the costs for a single evaluation are too high to be useful in an iterative procedure based on multitudes of such evaluations.

The consequence of this situation is that studies typically resort to model simplifications, in order to gain access to cheaper solution methods. A widely used assumption, for instance, is that of a homogeneous half-space, for which many analytical solutions are available. This means that surface topography and inhomogeneous layering become model errors, with unknown but regrettable consequences for the quality of the inversion.

To improve on the status quo of forced model errors, this thesis sets out to explore suitable alternatives to the split-node method. It observes that there is room for choice within the broader framework of the Finite Element Method, and that the choices made by Melosh and Raefsky are not necessarily optimal towards the goal of efficient simulation of a multitude of different faults.

The objective of this thesis is to develop a fast but versatile numerical method for the simulation of dislocations in general elastic domains, to be used in the specific inverse setting of the inference of earthquake mechanisms from InSAR data.

In line with the above main objective of the thesis, we consider the following research questions:

- (Solution space) Compared to the split-node method, what alternative decisions could be considered in devising a method for the simulation of dislocations while staying within the broader framework of the Finite Element Method?
- (Speed) Is it possible to devise a computational method that is fast enough to be feasibly used in iterations mandated by the inversion procedure? How does computational effort compare to existing techniques?
- (Flexibility) Is it possible to make significant speed improvements while keeping the power and flexibility of the Finite Element Method intact? If not, what are the limitations of the new method?
- (Errors) If speed gains come at the expense of accuracy, essentially replacing model errors by discretization errors, how will these errors affect the accuracy of a typical inversion?

Thesis Structure

The main body of this thesis is structured in three chapters, each of which is a verbatim copy of a peer reviewed journal article, followed by conclusions.

Chapter 2 presents the article *Overview of a range of solution methods for elastic dislocation problems in geophysics* [78]. This literature study presents a formal problem definition for the dislocation problem, and proceeds to provide an overview

of existing solution methods, ranging from analytical expressions to the split-node method, with emphasis on their individual strengths and weaknesses.

Chapter 3 presents the article *Discontinuities without discontinuity: The Weakly-enforced Slip Method* [77]. This work introduces the WSM as a new method for the simulation of dislocations, with particular suitability for the modelling of multiple fault geometries due to its ability to reuse the finite element mesh and derived products. It provides two different derivations of the method, followed by convergence analyses proving optimal convergence in modified norms. Finally, numerical results in two and three dimensions provide insights in some specific properties of WSM solutions.

Chapter 4 presents the article *Inverting elastic dislocations using the Weakly-enforced Slip Method* [76]. This study subjects the WSM to a synthetic inversion to gain a better understanding of the ramifications of its error characteristics. Selecting Bayesian inference as the inversion methodology, it first develops the general formulation, and proceeds by defining the concrete workflow used in the study. While results are specific for Bayesian inference and other choices, an effort is made to generalize insights to the broader class of inverse methods where possible.

Chapter 5 concludes by revisiting the research questions and discussing the main findings of this thesis, followed by recommendations for future research.

2 | Classic Solutions to Volterra's Dislocation Problem

Tectonic faults are commonly modeled as Volterra or Somigliana dislocations in an elastic medium. Over the years, many practical solution methods have been developed for problems of this type. This work presents a concise overview in consistent mathematical notation of the most prominent of these methods, emphasizing what the various methods have in common and in what aspects they are different. No models other than that of elastic dislocations are considered. Special attention is given to underlying assumptions and range of applicability.

This chapter is a verbatim copy of "Overview of a range of solution methods for elastic dislocation problems in geophysics" by G.J. van Zwieten, R.F. Hanssen and M.A. Gutiérrez; Journal for Geophysical Research: Solid Earth, 2013 [78].

The theory of dislocations concerns the state of self-stress in a body that is discontinuously deformed. A dislocation is thought of as a two-dimensional manifold, along which the material has been subsequently cut, displaced, welded together, and released. The system assumes a new minimum energy state of self-stressed equilibrium. The magnitude and direction of the displacement jump or 'slip' is a vector field over the dislocation plane. Where a slip vector has normal components the material opens and material is added to fill the void, such as happens when a dyke intrudes a fault. Fig. 2.1 shows a (circular) dislocation with purely tangential slip, representing a downward jump of the right-hand side material with respect to the left-hand side. The displacements are shown to be discontinuous at the dislocation plane, and continuous everywhere else.

The mathematics of the Theory of dislocations date back to 1907, to the publication of Vito Volterra's "Sur l'équilibre des corps élastiques multiplement connexes" [67]. In this work Volterra lays out his dislocation theory for elastic bodies, in which multi-valued displacement functions become the natural description for material discontinuities. Volterra's theory is restricted to multiply-connected domains, such as is obtained by making a small bore-hole along the circumference of the dislocation plane of Fig. 2.1. A cut through the dislocation plane makes the domain simply connected,

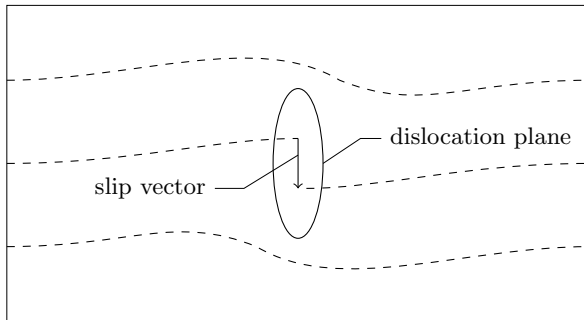


Figure 2.1 – Schematic representation of a material dislocation. The dashed lines represent straight segments in the reference configuration. In dislocated state, the entire medium is seen to be deformed, but deformation is discontinuous only at the dislocation plane. The slip vector measures the magnitude of the jump at any point on the dislocation plane.

after which a relative rigid body displacement (translation or rotation) forms the dislocation. This type of dislocation is commonly referred to as the Volterra dislocation. The theory was further developed by Volterra’s contemporary Somigliana [29], who lifted the connectivity constraint and allowed the slip vector to vary over the dislocation plane. In particular, his definition allows slip to taper off near the edges of the dislocation, which removes the stress singularities at the tip that otherwise arise from the displacement mismatch. The Somigliana dislocation is a generalization of the Volterra dislocation, and consequently allows for a closer approximation of the physical problem considered. In practice this added accuracy is not always required, and Volterra’s description is often used for sake of simplicity.

In 1958, fifty years after Volterra’s initial publication, [64] recognized Volterra’s theory of dislocations as the proper tool for a quantitative analysis of fracture zones in the earth’s crust. Steketee proposed to model the co-seismic deformation caused by an earthquake as the elastic response to a Volterra dislocation, assuming that dissipating processes such as creep are negligible in the period closely following the seismic event. He developed a Green’s function method and noted that the general Volterra problem requires six sets of elementary solutions, of which he derived one. This opened the way for a practical, quantitative analyses of co-seismic displacements and stress changes in a dislocation zone. The first to take Steketee’s results to analyse actual data was [17], who demonstrated its use to determine the depth of a fault based on surface measurements.

In 1985 Okada published a set of closed form solutions for surface displacements caused by a rectangular dislocation in a homogeneous halfspace, which was followed in 1992 by an even more elaborate set of equations capturing the entire three dimensional displacement and stress fields [55, 56]. His work was the culmination of decades of development that Steketee had set off. For a detailed overview both papers provide an excellent introduction. Okada’s equations remain a very popular tool for the analysis of co-seismic displacements, mainly because, being analytical expressions, the evaluation of (surface) displacements is fast in comparison with competing methods. This comes at the cost of severe limitations (elastic homogeneity and isotropy, absence of topography, planar fault geometries, rigid body translations) which are often crude approximations of reality at best, with obvious implications for the obtained results. Developments continued to push the balance towards a more detailed modelling.

Today a wide range of solution methods exist for the dislocation problem originally defined by Volterra and Somigliana. The computational spectrum ranges from

✓ supported ¹ ✓ through superposition ² ✓ vertical stratification ³ ✓ moderate	Heterogeneity	Anisotropy	Topography	Distrib. slip	Curved fault
2.2 Analytical solutions					
Volterra’s equation [67]	✓	✓	✓	✓	✓
Okada [55, 56]				¹ ✓	¹ ✓
Angular dislocation [20]				¹ ✓	¹ ✓
2.3 Propagator matrix					
Propagator matrix [68]	² ✓			✓	✓
2.4 Series expansion					
Moduli perturbation [21]	³ ✓	³ ✓			
Topography [70]			³ ✓		
2.5 Finite Elements					
Split node FEM [51]	✓	✓	✓	✓	✓

Table 2.1 – Overview of prominent solution methods for elastic dislocation problems, organized per corresponding section in the current study. Check marks indicate support for different aspects of the generic problem formulation.

relatively inexpensive methods that solve a limited set of problems, such as those presented by Okada, to computationally intensive methods that are capable of solving any Somigliana dislocation. To choose the proper method for a given task requires a solid understanding of the various methods and their differences. This paper aims to facilitate this by creating an overview for the most prominent of these methods. Table 2.1 shows a reference of methods to be discussed in detail in the indicated sections. We note that due to the length of many mathematical results cited in this work we can not reproduce all equations in this text. Instead we will refer to the work of the respective authors. The equations that we do choose to present should serve to bring a sufficient level of understanding.

The remainder of this text is organized as follows. Section 2.1 formally introduces the general dislocation problem that we aim to solve. The subsequent sections cover different solution methodologies. Section 2.2 derives Volterra’s equation and introduces derived results for homogeneous halfspaces. Section 2.3 introduces the propagator matrix method for vertically stratified halfspaces. Section 2.4 details two series expansion methods, one for dealing with material heterogeneities, the other for dealing with topography. Finally, Section 2.5 explains the Finite Element Method for general dislocation problems.

2.1 Problem definition

In this section we formally define the mathematical problem that is the subject of this study. We aim to find an equilibrium displacement field u on domain Ω . The domain is bounded by $\Gamma = \Gamma_1 \cup \Gamma_2$, where Γ_1 represents a traction-free (Neumann) boundary and Γ_2 a displacement-free (Dirichlet) boundary. A dislocation plane \mathcal{F} is embedded in Ω , with normal vector ν . The displacement field is discontinuous over \mathcal{F} following the slip vector b , which is free to vary over \mathcal{F} . Fig. 2.2 shows a schematic of the setup. The stress field $\sigma(u)$ is in equilibrium throughout the continuous area $\Omega \setminus \mathcal{F}$.

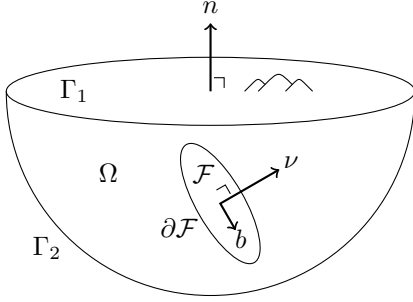


Figure 2.2 – Schematic representation of the computational setup, related to the problem definition of (2.1). The domain Ω is bounded by the traction-free boundary Γ_1 and the displacement-free boundary Γ_2 . A dislocation plane \mathcal{F} is embedded in Ω and bounded by $\partial\mathcal{F}$. The slip vector b is the local displacement jump when moving at any point through \mathcal{F} , which is often but not necessarily perpendicular to the normal ν .

Together this leads to the following system of equations for the elastic dislocation problem considered in the current study:

$$\operatorname{div} \sigma(u) + f = 0 \text{ on } \Omega \setminus \mathcal{F} \quad (2.1a)$$

$$\llbracket u \rrbracket = b \text{ on } \mathcal{F} \quad (2.1b)$$

$$\llbracket \sigma_\nu(u) \rrbracket = 0 \text{ on } \mathcal{F} \quad (2.1c)$$

$$\sigma_n(u) = 0 \text{ on } \Gamma_1 \quad (2.1d)$$

$$u = 0 \text{ on } \Gamma_2 \quad (2.1e)$$

We shall go over these equations one by one.

Equation (2.1a) is the standard condition of static equilibrium, which can be found in any text book on the subject of elasticity such as [47]. The relation between stress σ and displacement u depends on the material properties of the medium. Linear elasticity theory assumes a linear relation of the form

$$\sigma_{ij} : u \rightarrow \sum_{k=1}^3 \sum_{l=1}^3 K_{ijkl} \frac{\partial u_k}{\partial x_l}, \quad (2.2)$$

where K is the fourth order *stiffness tensor* that carries the medium's elastic parameters, subject to symmetries $K_{ijkl} = K_{ijlk} = K_{klij}$ [47]. The stress σ is not defined at the discontinuity \mathcal{F} where, consequently, the equilibrium equation does not apply. The discontinuity is therefore excluded from the domain, and equilibrium of stress over \mathcal{F} is enforced separately by (2.1c).

Equations (2.1b) and (2.1c) are expressed in terms of the *jump operator*, which is defined as

$$\llbracket u \rrbracket := u^+ - u^-, \quad (2.3)$$

where u^+ denotes the displacement on the side of \mathcal{F} pointed into by ν , and u^- denotes the displacement on the opposing side of \mathcal{F} . The jump operator will be used extensively in the present work, together with the *mean operator* which is defined as

$$\{u\} := \frac{1}{2}(u^+ + u^-). \quad (2.4)$$

The two operators are connected by the following algebraic identity,

$$\llbracket u \cdot v \rrbracket = \llbracket u \rrbracket \cdot \{v\} + \{u\} \cdot \llbracket v \rrbracket. \quad (2.5)$$

Equation (2.1b) imposes a displacement jump at \mathcal{F} of magnitude and direction equal to the slip vector b . The slip vector can be tangential to \mathcal{F} , normal, or oblique,

depending on the process under study. Physically a normal component corresponds to an opening of the dislocation, and tangential components correspond to shear. Equation (2.1c) imposes equilibrium of stress by setting the normal stress difference over \mathcal{F} to zero, again making use of the jump operator. This sets the tractions on the two facing sides of \mathcal{F} equal, leaving no residual tractions.

Lastly, (2.1d) and (2.1e) set the conditions at the domain boundary. These are not specific to dislocation theory and can be found in any textbook on elasticity theory such as [47]. At Γ_1 the tractions are zero; at Γ_2 the displacements are zero. This allows for several different modelling options. The surface Γ_1 may be flat, curved, or irregularly shaped. The domain Ω may be finite or infinite. For instance, modelling the earth as a true globe, Γ_1 may represent the earth's surface in detail while Γ_2 is empty. On the other extreme, Γ_1 can be a flat surface and Γ_2 tending to infinity, thereby modelling the earth as an infinite halfspace. This is an often used simplification of reality and will be discussed in detail in Sections 2.2 and 2.3.

We note that (2.1) by itself is a direct representation of Newton's laws of physics, stating that forces are in equilibrium for a body at rest, and as such it is our best available, uncompromised description of reality. Simplifications arise from the choice of domain, boundary conditions, and material stiffness. Our baseline simplification, however, enters with the material description. It is introduced by the constitutive relation, (2.2), that incorporates the assumptions of infinitesimal strain and a linear stress-strain relation. This is the theory of linear elasticity [47], which in application to tectonophysics is generally thought a good approximation for short time scales and small deformations — but an approximation nonetheless. Another area where physical realism is lost in the vicinity of the dislocation, where a micro-cracked zone is instead modeled as a clean cut. The present model should therefore be applied only in the context of global deformation fields. Sophisticated studies of atomic level defects in the fault zone do exist, covered for example by [7], but they are beyond the scope of this study. Recall that our aim is not to validate the presented model, but to present an overview of available techniques for working with it. We do point out that some of the detailed modelling can be brought in touch with the current model via homogenization, as the model allows any anisotropic, heterogeneous material description that remains linear. Another possibility is to model a fault zone as a pileup of dislocations, for example to model creep [53]. These superposition techniques also lie in the realm of the current model.

The simplifying assumption of linearity is also of direct consequence to how we deal with body forces such as gravity, because it allows us to decouple the influence of external body forces and that of the dislocation. To see this we consider two independent solutions of (2.1). The first, denoted u_f , is the solution for f equal to gravity and b equal to zero: this is the equilibrium state before dislocation occurs. The second, denoted u_b , the solution of the same problem for f equal to zero and b the slip vector: this is the deformation change due to the act of dislocation. Then, by linearity, $u_f + u_b$ solves the combined problem. When we choose the equilibrium state u_f to be our reference state, we can disregard body forces and focus on the homogeneous problem. This is common practice, as the primary interest is usually relative displacement and stress *difference* resulting from the act of dislocation. In this study, however, we choose to maintain a body force where possible for illustrative purposes.

2.2 Analytical solutions

Our aim is to compute the displacement field u that solves the dislocation problem defined in (2.1). In this section we derive the general solution known as Volterra's equation. We will then continue to introduce the homogeneous halfspace, the Burgers equation, and finish with relevant derived methods.

2.2.1 Volterra's equation

In this section we shall derive Volterra's equation, following the construction presented by [38]. We start by defining the bilinear form

$$a(u, v) := \int_{\Omega \setminus \mathcal{F}} \sigma(u) : \nabla v \, dV, \quad (2.6)$$

where u, v are displacement fields on Ω . We wish to reformulate this through partial integration. Note that the dislocation plane \mathcal{F} is excluded from the domain of integration, which forms an internal boundary that traverses the two sides of \mathcal{F} . Partial integration yields

$$a(u, v) = \int_{\mathcal{F}^+} v \cdot \sigma_n(u) \, dS + \int_{\mathcal{F}^-} v \cdot \sigma_n(u) \, dS + \int_{\Gamma} v \cdot \sigma_n(u) \, dS - \int_{\Omega} v \cdot \operatorname{div} \sigma(u) \, dV, \quad (2.7)$$

where \mathcal{F}^+ and \mathcal{F}^- denote the opposing sides of the dislocation with consequently opposing normals. Substituting $n = \nu$ at \mathcal{F}^+ and $n = -\nu$ at \mathcal{F}^- , this result is reformulated in terms of a jump $\llbracket v \cdot \sigma_\nu \rrbracket$. It will prove convenient to split this term using (2.5), resulting in the following reformulation of (2.7):

$$a(u, v) = \int_{\mathcal{F}} \left[\llbracket v \rrbracket \cdot \{\sigma_\nu(u)\} + \{v\} \cdot \llbracket \sigma_\nu(u) \rrbracket \right] dS + \int_{\Gamma} v \cdot \sigma_n(u) \, dS - \int_{\Omega} v \cdot \operatorname{div} \sigma(u) \, dV. \quad (2.8)$$

We now introduce the fundamental point force solution, denoted $g_{x;d}$, that represents the deformation of the non-dislocated medium ($b = 0$) induced by a point force acting at location $x \in \Omega$ with (vector) magnitude d . It is the first of a number of point solutions we will use, of which Fig. 2.3 presents an overview. With this fundamental solution, and with u as the solution of the dislocation problem from (2.1), we derive the two identities

$$a(u, g_{x;d}) = \int_{\Omega} g_{x;d} \cdot f \, dV, \quad (2.9a)$$

$$a(g_{x;d}, u) = \int_{\mathcal{F}} b \cdot \sigma_\nu(g_{x;d}) \, dS - d \cdot u(x), \quad (2.9b)$$

where the last term is due to the fact that $g_{x;d}$ is a point force, hence $\operatorname{div} \sigma(g_{x;d})$ a Dirac distribution, centered at x .

Lastly, we introduce *Betti's reciprocity principle* [47], which (effectively) states that the bilinear form defined in (2.6) is symmetric. I.e., for two displacement fields u and v , the following equality applies:

$$a(u, v) = a(v, u). \quad (2.10)$$

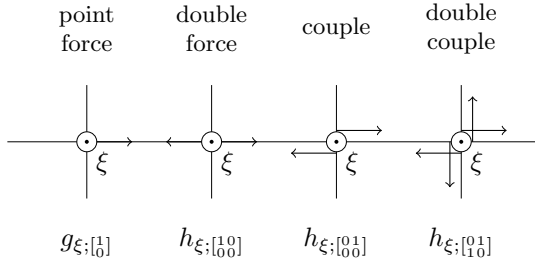


Figure 2.3 – Examples of Green's functions used in the derivations of Section 2.2.1, showing the point force g , and the dislocation kernels h defined in (2.14). Representations are 2D but serve to illustrate the nature of their 3D equivalents. All solutions g and h are in equilibrium with the presented loads at point $\xi \in \Omega$.

The property is directly inherited from the symmetries of the constitutive tensor K , which in turn stem from the existence of a strain energy functional.

With the reciprocity principle and the two identities of (2.9) we directly arrive at Volterra's equation,

$$d \cdot u(x) = \int_{\mathcal{F}} b \cdot \sigma_{\nu}(g_{x;d}) \, dS - \int_{\Omega} f \cdot g_{x;d} \, dV, \quad (2.11)$$

where d is an arbitrary direction vector. By choosing $d = e_k$, where e_k is the unit vector in direction k , one can selectively find the three Cartesian components of u . Volterra's equation describes the displacement in any point of the domain Ω in terms of a surface integral over the dislocation plane \mathcal{F} . It is usually presented without the separate volume integral, assuming a reference state that is in equilibrium with external body forces.

Although Volterra's equation is a valid result mathematically, it gives little physical understanding as it involves placing an imaginary force at the location of measurement. To aid physical interpretation we again apply the reciprocity principle (2.10), this time to two point force solutions. From $a(g_{x;d}, g_{\xi;\delta}) = a(g_{\xi;\delta}, g_{x;d})$ follows the identity

$$\delta \cdot g_{x;d}(\xi) = d \cdot g_{\xi;\delta}(x), \quad (2.12)$$

which we use to rewrite both the volume integral and the surface integral of (2.11). Substituting constitutive relation (2.2) we arrive at the following reformulation of Volterra's equation as a convolution of fundamental solutions:

$$u(x) = \int_{\mathcal{F}} h_{\xi;M(\xi)} \, d\xi - \int_{\Omega} g_{\xi;f(\xi)} \, d\xi. \quad (2.13)$$

Here, h is a new fundamental solution that is defined as the moment of two point force solutions placed infinitesimally close together,

$$h_{\xi;M} = \sum_{ij} M_{ij} \lim_{\delta \downarrow 0} \frac{g_{\xi+\delta e_i;e_j} - g_{\xi-\delta e_i;e_j}}{2\delta}, \quad (2.14)$$

and M is the *moment tensor density*, an established quantity in seismology [2] defined as

$$M_{ij} := \sum_{kl} K_{ijkl} b_k \nu_l. \quad (2.15)$$

Fig. 2.3 shows some (2D analogues of) realizations of the dislocation kernels for different values of M . Because M is a 3×3 symmetric tensor, the dislocation kernel is build up of six elementary solutions: three double forces, and three double couples. The displacement field induced by the dislocation is the cumulative effect of these six kernels integrated over the dislocation plane.

2.2.2 The homogeneous halfspace

The above analysis is generic and presents the general solution to the problem defined in (2.1). The solution, however, relies on fundamental point forces and dislocation kernels, for which solutions are not known in general. We therefore need to narrow down the problem to a more limited class for which fundamental solutions are available.

The *homogeneous halfspace* is the subclass of linear elasticity problems having a flat surface Γ_1 and a far field Γ_2 at infinity. Besides homogeneity, isotropy is commonly implied, meaning that the (constant) constitutive tensor is fully determined by two material parameters, typically

$$K_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (2.16)$$

where λ and μ are known as Lamé's first and second parameter [47]. The homogeneous halfspace is a widely used model that is generally considered a good first order approximation of reality. We emphasize again that it is not the scope of the current study to assess the quality of the model; our focus is limited to solution methods. The homogeneous halfspace is convenient primarily because analytical expressions exist for point force responses. Relevant work in this direction was carried out by [52], who presented fundamental point force solutions for a semi-infinite solid. With the dislocation kernels h from (2.14), the moment tensor density M from (2.15), and the constitutive tensor from (2.16), the only element left is to evaluate the integrals of (2.13).

Rather than using dislocation kernels for the half space directly, it is often more efficient to use the (much simpler) dislocation kernels for the infinite space, and account for the traction-free surface later. We construct the dislocation kernels from the Kelvin-Somigliana point force responses [47] through differentiation. Then, for a constant slip vector b , (2.13) evaluates to what is known as the Burgers equation [38],

$$u(x) = \frac{b}{4\pi} \int_{\mathcal{F}} \frac{r_x \cdot n}{|r_x|^3} dS + \frac{1}{4\pi} \oint_{\partial\mathcal{F}} \left[-\frac{b \times dL}{|r_x|} + \left(\frac{1/2}{1-\nu} \right) \nabla \left(\frac{(b \times r_x) \cdot dL}{|r_x|} \right) \right]. \quad (2.17)$$

In this expression, x is a point in the domain Ω , $r_x(\mathcal{F})$ is the vector pointing from x to the dislocation, and $\partial\mathcal{F}$ is the boundary of the dislocation plane that forms a one-dimensional closed loop. The dislocation plane \mathcal{F} and slip vector b are as in the problem definition (2.1), and Poisson's ratio ν relates to Lamé's parameters as $\nu = \lambda/2(\lambda + \mu)$. The surface integral represents the solid angle of \mathcal{F} subtended from x . It is this term that makes (2.17) multi-valued at \mathcal{F} , as the solid angle jumps from one half steradian (2π) to minus one half steradian (-2π) upon crossing the dislocation. It is known that the solid angle depends only on the border $\partial\mathcal{F}$ of the dislocation. Consequently, even though the displacement jump follows the geometry

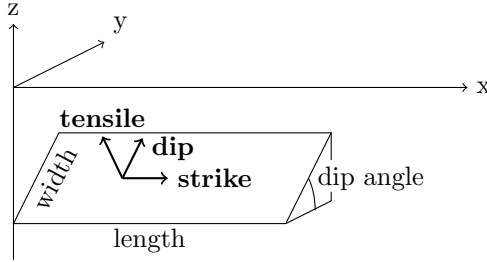


Figure 2.4 – Parameterized geometry of a rectangular dislocation plane, used in the expressions for the displacement and stress field around a rectangular source.

of \mathcal{F} , the strain and stress field are determined by its border alone, which for this reason is commonly referred to as the *dislocation line*. The stress field is computed from a single circle integral in what is known as the Peach-Koehler equation. For this and other details of dislocation theory we refer to [38].

With the Burgers equation we have obtained a solution for the infinite medium. Next is to modify the whole space solution such that it meets the boundary condition of (2.1d). A very common approach is the method of images, which involves superposing a mirrored solution such that the mirror plane corresponds with the surface. This will cancel out horizontal tractions in the mirror plane, but double the vertical tractions. What remains is the Boussinesq problem, a halfspace with a distribution of vertical tractions, which can be solved using harmonic series. [64] used this method to create one of six dislocation kernels. [60] and [18] built on this approach to account for planes of discontinuous elastic properties, however they did so only for plane strain models. For three dimensional heterogeneous mediums, the propagator matrix method of Section 2.3 is considered the more powerful method.

2.2.3 Closed form solutions

The developed framework puts us in position to derive actual closed form analytical expressions for specific dislocations. [17] was the first to take Steketee’s results and integrate them over a vertically oriented rectangular plane to represent a pure strike-slip earthquake. The developments that followed culminated in the work of [55, 56], who presented the complete set of solutions for any point source, as well as for any finite rectangular source at arbitrary depth and dip angle. The equations are lengthy and not repeated here, but we provide a copy of the parameterized geometry used by Okada as Fig. 2.4. Counting length, width, dip angle and depth of the dislocation plane, the three components of the slip vector, and Poisson’s ratio, the equations constitute a total of eight parameters. An additional three are typically required to position and orient the dislocation at distance from the origin. As it is often required to have more fine grained control it is common practice to superpose several Okada solutions, which is possible due to linearity of the problem. Distributed slip is approximated by varying the slip vector, resulting in a discontinuous step distribution. Non-planar geometries are approximated by positioning Okada sources along a curved plane. In that case, due to the rectangular shape of the building block, geometrical continuity is lost if the dislocation is doubly curved.

Related work was performed by [73], who evaluated the Burgers equation for an *angular dislocation*; results were later corrected by [38]. The angular dislocation is

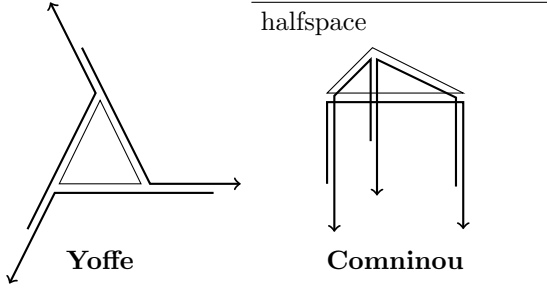


Figure 2.5 – Formation of a triangular dislocation by superposition of angular dislocations. Left: a dislocation in a whole space constructed from Yoffe’s angular dislocations. Right: a dislocation in a halfspace constructed from Comninou’s angular dislocations, where each loop is constructed of two wedges.

a wedge shaped domain bounded by two semi-infinite lines meeting in a point, of which the two sides are displaced over constant distance. Fig. 2.5 illustrates how this elementary element can be used to form any polygonal dislocation shape by adding the angular dislocations in such a way that boundary integrals over the infinite segments cancel, leaving a single integral circling the polygon. Note that the surface integral in (2.17) adds up to the solid angle of the inverted triangle, effectively forming a dislocation at the plane surrounding the triangle, rather than inside it. Except for a sign change this has no consequence for strain nor stress. From (2.17) it is clear that to obtain the correct displacement field, the solid angle over the entire containing plane should be added which shifts the two halves back by a distance b .

[20] continued on Yoffe’s work by deriving angular dislocation solutions for the homogeneous halfspace. The solutions form wedges with one leg perpendicular to the surface and one at arbitrary angle, meeting at arbitrary depth. The one perpendicular leg is no restriction as any polygonal shape can be formed by superposition, as shown in Fig. 2.5. Angular dislocations are combined to form loops with two legs pointing down, which are then combined to form any polygonal loop. Like before, the strain and stress fields are valid without need for modification. The displacement field is discontinuous at the vertical faces surrounding the polygon. This can be corrected by shifting the volume under the polygon by a distance b , transferring the discontinuity from the surrounding faces to the polygon itself. The triangular element constructed in Fig. 2.5 is of practical interest as it allows assembly of a continuous, triangulated structure; a clear advantage over Okada’s rectangular dislocation plane solution. For both Okada’s and Comninou’s elements, however, the slip vector is constant, which means that neither can be assembled to form a truly continuous slip distribution.

We finish this section by comparing Okada’s and Comninou’s solutions by looking at singularities in the equations. Apart from the dislocation plane itself, both solutions have singularities in $\Omega \setminus \mathcal{F}$ due to the way the analytical expressions are constructed. For the latter these lie on semi-infinite lines pointing down from its corners and extending one-sidedly from its edges, resulting from the infinite dislocation lines bounding the underlying wedge dislocations. Okada lacks the vertical lines, but has the four lines extending one-sidedly from its edges, two of them possibly reflecting in the surface depending on the sign of δ in Fig. 2.4, as a result of mirroring the dislocation plane to remove surface tractions. For Okada’s equations this can always be resolved by changing the parametrization such that the singular lines point in opposite direction. The same applies partly to Comninou’s equations, but the downward pointing singular lines can not be resolved. As none of this affects surface displacements, however, this is not usually an issue in practice.

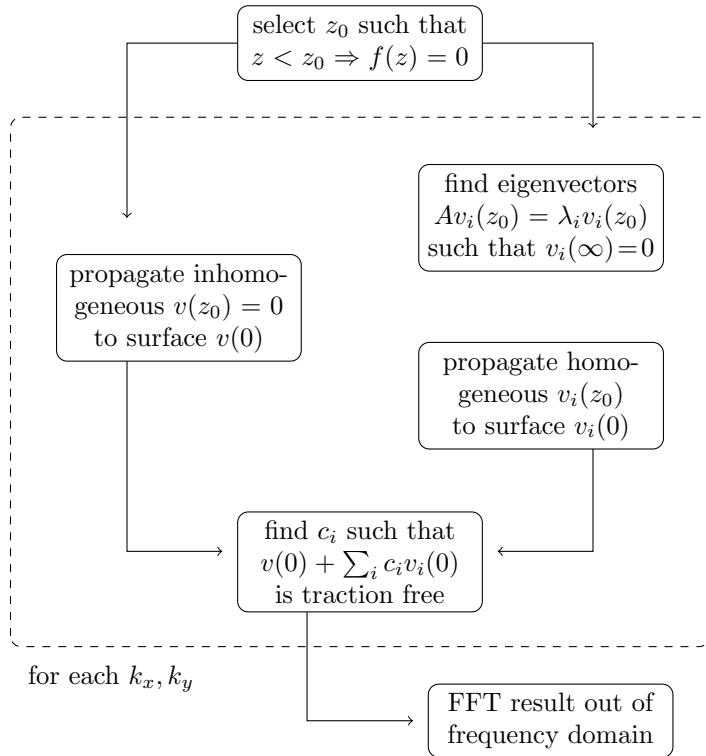


Figure 2.6 – Flow chart representation of the Propagator matrix solution method.

2.3 Propagator matrix

The propagator matrix method applies to a halfspace, but is not restricted to a homogeneous space. Rather than allowing full heterogeneity, however, the elastic properties are limited to vary with depth. This makes the propagator method suitable for modeling the earth's vertical stratification. It had been an established method in seismology since long [2], but it was not until 1985 that the method was introduced to dislocation computations by [68]. An excellent reference to the method is found in [62]. Here we shall cover the main aspects to demonstrate the principles of the method, summarized in Fig. 2.6 for reference.

We start by rewriting the general solution of the problem defined in (2.1) in terms of a single volume integral. For that we introduce the distribution $\delta_{\mathcal{F}}$ as having the property

$$\forall h : \int_{\mathcal{F}} h \, dS = \int_{\Omega} h \, \delta_{\mathcal{F}} \, dV. \quad (2.18)$$

This is similar to the standard pointwise delta distribution, but connected to a two dimensional manifold. Using this delta distribution to transform the surface integral to a volume integral, (2.13) transforms, after partial integration, to a single volume

integral

$$u = \int_{\Omega} g_{\xi; \hat{f}(\xi)} d\xi, \quad (2.19)$$

where

$$\hat{f} = f - M \nabla \delta_{\mathcal{F}}. \quad (2.20)$$

We recognize that we have obtained a body force distribution \hat{f} that is equivalent to directly imposing a displacement jump in (2.1). Note that the body forces become infinitely large at \mathcal{F} , which gives the material a finite deformation without making an actual cut. This is obviously not physical, but as a mathematical construct the obtained body force will turn out very useful.

We continue by rewriting the equilibrium (2.1a) as a first order system of unknowns u_1, u_2, u_3 and $\sigma_{13}, \sigma_{23}, \sigma_{33}$, using the previously derived body force, and using constitutive relation (2.2) with the isotropic stiffness tensor defined in (2.16). Isolating derivatives to z this leads to the following system of equations:

$$\begin{aligned} \frac{\partial u_1}{\partial z} &= \frac{1}{\mu} \sigma_{13} - \frac{\partial u_3}{\partial x} \\ \frac{\partial u_2}{\partial z} &= \frac{1}{\mu} \sigma_{23} - \frac{\partial u_3}{\partial y} \\ \frac{\partial u_3}{\partial z} &= \frac{1}{\gamma} \sigma_{33} - \frac{\lambda}{\gamma} \frac{\partial u_1}{\partial x} - \frac{\lambda}{\gamma} \frac{\partial u_2}{\partial y} \\ \frac{\partial \sigma_{13}}{\partial z} &= -\hat{f}_1 - \frac{\alpha}{\gamma} \frac{\partial^2 u_1}{\partial x^2} - \mu \frac{\partial^2 u_1}{\partial y^2} - \frac{\beta}{\gamma} \frac{\partial^2 u_2}{\partial x \partial y} - \frac{\lambda}{\gamma} \frac{\partial \sigma_{33}}{\partial x} \\ \frac{\partial \sigma_{23}}{\partial z} &= -\hat{f}_2 - \mu \frac{\partial^2 u_2}{\partial x^2} - \frac{\alpha}{\gamma} \frac{\partial^2 u_2}{\partial y^2} - \frac{\beta}{\gamma} \frac{\partial^2 u_1}{\partial x \partial y} - \frac{\lambda}{\gamma} \frac{\partial \sigma_{33}}{\partial y} \\ \frac{\partial \sigma_{33}}{\partial z} &= -\hat{f}_3 - \frac{\partial \sigma_{13}}{\partial x} - \frac{\partial \sigma_{23}}{\partial y}, \end{aligned} \quad (2.21)$$

introducing $\alpha = 4\mu(\lambda + \mu)$, $\beta = \mu(2\mu + 3\lambda)$ and $\gamma = \lambda + 2\mu$ for brevity of expression. Taking the two-dimensional Fourier transform, the above system of equations transforms to an expression of the form

$$\frac{\partial v}{\partial z} = Av - f, \quad (2.22)$$

where v is the vector of six unknowns,

$$v = (\bar{u}_1, \bar{u}_2, -i\bar{u}_3, \bar{\sigma}_{13}, \bar{\sigma}_{23}, -i\bar{\sigma}_{33})^T, \quad (2.23)$$

that describes the displacement and vertical traction at any point k_x, k_y, z , where k_x and k_y are the wave numbers of the Fourier transformed displacement \bar{u} and traction $\bar{\sigma}$. The real-valued matrix A connects these point values to the derivatives in that point with respect to depth,

$$A = \begin{pmatrix} \cdot & \cdot & k_x \frac{1}{\mu} & \cdot & \cdot \\ \cdot & \cdot & k_y \cdot & \frac{1}{\mu} & \cdot \\ \frac{\lambda}{\gamma} k_x & \frac{\lambda}{\gamma} k_y & \cdot & \cdot & -\frac{1}{\gamma} \\ \frac{\alpha}{\gamma} k_x^2 + \mu k_y^2 & \frac{\beta}{\gamma} k_x k_y & \cdot & \cdot & \frac{\lambda}{\gamma} k_x \\ \frac{\beta}{\gamma} k_x k_y & \mu k_x^2 + \frac{\alpha}{\gamma} k_y^2 & \cdot & \cdot & \frac{\lambda}{\gamma} k_y \\ \cdot & \cdot & \cdot & -k_x - k_y & \cdot \end{pmatrix}, \quad (2.24)$$

and f is the Fourier-transformed body force vector,

$$f = (0, 0, 0, \bar{\bar{f}}_1, \bar{\bar{f}}_2, \bar{\bar{f}}_3)^T, \quad (2.25)$$

We connect solutions at different depths by direct application of the second fundamental theorem of calculus,

$$v(z) = v(z_0) + \int_{z_0}^z [A(\zeta)v(\zeta) - f(\zeta)] d\zeta. \quad (2.26)$$

We would like to use this result to propagate a solution known at depth z_0 to any arbitrary depth z , but in its current form the integrand of (2.26) requires knowledge of the intermediate solutions $v(\zeta)$. This problem is solved by repeated substitution of the left hand side into the right hand side. This produces a series of nested integrals that is conveniently written as

$$v(z) = P(z, z_0)v(z_0) - \int_{z_0}^z P(z, \zeta)f(\zeta) d\zeta, \quad (2.27)$$

where P is defined as the infinite series

$$P(z, z_0) = I + \int_{z_0}^z A(\zeta_1) d\zeta_1 + \int_{z_0}^z A(\zeta_1) \left[\int_{z_0}^{\zeta_1} A(\zeta_2) d\zeta_2 \right] d\zeta_1 + \cdots. \quad (2.28)$$

It is easily verified that (2.27) solves (2.22). Furthermore (2.27) depends only on $v(z_0)$, the solution at a single two-dimensional layer at depth z_0 , from which it can reconstruct the entire solution $v(z)$ at any depth. For this reason the matrix P is commonly referred to as the propagator matrix.

We will use (2.27) to construct the entire displacement field corresponding to the body forces f . For simplicity we assume that the body force \hat{f} is supported only at depth range containing \mathcal{F} , meaning that our reference state is in equilibrium with external body forces. As a consequence \hat{f} equals zero at some depth z_0 below the dislocation, where we set the solution vector $v(z_0)$ to zero. It is directly clear from (2.27) that the solution will remain zero when propagated downwards. Upwards it will start to accumulate value when it reaches the dislocation, where \hat{f} is nonzero, and consequently will reach the surface with nonzero displacement and traction. This is of course in violation of (2.1d), the boundary condition that prescribes that the traction at the surface Γ_1 be zero.

To meet the boundary condition we superpose a solution for the homogeneous problem ($\hat{f} = 0$) to cancel the tractions at the surface. To construct this solution one might be tempted to take a solution vector at zero depth containing the spurious tractions, and propagate it downwards using (2.27) with body force set to zero. This fails for arbitrary surface displacements because displacements and stresses will diverge with depth, violating the boundary condition at the far field Γ_2 , (2.1e). The correct surface displacements are found by forming a basis of eigenvectors of matrix A for the lowest (infinite) layer, choosing the three vectors that correspond to vanishing displacements at infinity — the remaining three eigenvectors diverge. That number matches the three traction components that are to be made zero at the surface, which closes the system.

To propagate a solution from one depth to a next using the propagator matrix defined in (2.28), we need to evaluate an infinite series of integrals. This is considerably simplified when the half space is built of homogeneous layers, such that the elastic

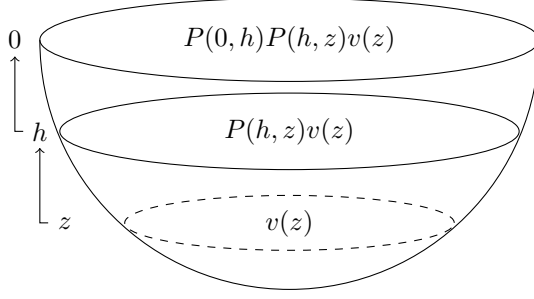


Figure 2.7 – Example of a solution propagated from depth z to the surface through two homogeneous layers.

parameters are step functions. Note that we can propagate a solution in arbitrary many substeps as

$$P(z, z_0) = P(z, z_n)P(z_n, z_{n-1}) \cdots P(z_1, z_0). \quad (2.29)$$

If the substeps z_i correspond with the layer interfaces, then matrix A in (2.27) is independent of depth for the intervals of integration and can be taken out of the integrals. The propagator matrices then evaluate to

$$P(z, z_0) = \sum_{n=0}^{\infty} \frac{A^n (z - z_0)^n}{n!} = \exp\left(A(z - z_0)\right), \quad (2.30)$$

with z limited to the layer that contains z_0 , and A specific to that layer. Using (2.29) we can propagate any solution of the homogeneous equation ($f = 0$) by a finite number of matrix multiplications. Fig. 2.7 demonstrates this with a two-layer example.

Propagation of the inhomogeneous solution through (2.27) requires integration of Pf . The delta operator in the body force \hat{f} allows us to transform the integral to a point evaluation through partial integration. For details of this operation we refer to [62]. Alternatively it is possible to take ready solutions such as the Burgers equation, and propagate them to the surface using (2.29). In case the dislocation is contained in the lowest layer the solutions vanish at infinity, and surface tractions can be cancelled using the standard eigenvector basis. If the dislocation is at a higher layer, then propagating it downward to the lowest layer cause the solution to diverge. Using the full six eigenvectors, however, it is possible to cancel displacements at the lowest layer, after which the three vanishing eigenvectors cancel the tractions at the surface.

We finally note that it is possible to decompose the 6×6 propagator matrix into two decoupled systems, a 2×2 system that is identified with an anti-plane problem, and a 4×4 system that is identified with a plane strain problem. This turns out to be a very practical implementation, but it is beyond the scope of this work to discuss the details of this decomposition. We again refer to [62] for details of this decomposition.

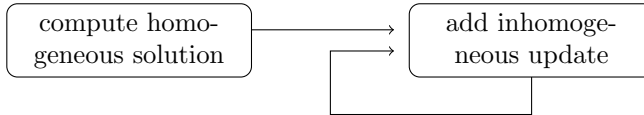


Figure 2.8 – Flow chart representation of the series expansion solution methods.

2.4 Series expansion

One can find an approximate solution of a problem by forming a Taylor series around another, related problem. This is useful in situations where the problem considered is very similar to one for which a solution method is readily available. We illustrate this technique for two situations, the first a problem with spatially varying elastic properties, the second a problem with a mild topography. The methodology presented in this section can be used to stretch the applicability of other solution methods such as those covered in Sections 2.2 and 2.3. Both methods essentially follow the work flow of Fig. 2.8.

2.4.1 Material heterogeneity

We first consider the moduli perturbation method of [21] for heterogeneous problems. This method forms a Taylor expansion around a closely related (typically homogeneous) reference problem for which solution methods are available, in order to approach the solution of the actual problem. We decompose the constitutive tensor as

$$K = \bar{K} + \hat{K}, \quad (2.31)$$

where \bar{K} is the stiffness of the reference problem and \hat{K} is the spatially varying deviation from the reference. Defining the corresponding stress tensors $\bar{\sigma}$ and $\hat{\sigma}$ through (2.2), we arrive at the decomposition

$$\sigma(u) = \bar{\sigma}(u) + \hat{\sigma}(u). \quad (2.32)$$

We use available solution methods for the reference problem to create a series expansion

$$u = u_0 + u_2 + u_3 + \cdots, \quad (2.33)$$

where the terms of the series follow from the iterative solution of the reference problem,

$$\begin{array}{ll} \operatorname{div} \bar{\sigma}(u_0) = -f, & \llbracket u_0 \rrbracket = b \\ \operatorname{div} \bar{\sigma}(u_1) = -\operatorname{div} \hat{\sigma}(u_0), & \llbracket u_1 \rrbracket = 0 \\ \operatorname{div} \bar{\sigma}(u_2) = -\operatorname{div} \hat{\sigma}(u_1), & \llbracket u_2 \rrbracket = 0 \\ \vdots & \vdots \\ \operatorname{div} \bar{\sigma}(u_n) = -\operatorname{div} \hat{\sigma}(u_{n-1}) - f, & \llbracket u_n \rrbracket = b. \end{array} \quad (2.34)$$

[21] showed that the series converge if the spatial variability of \hat{K} is small. It is easy to see that the converged solution u meets all requirements of (2.1), and hence forms a proper solution of the problem. In practice the series will be truncated at a finite number to yield an approximate solution.

Note that all systems other than the first (recall that f is zero in practice) are inhomogeneous. [21] proposed to solve these systems by convolution of a Green's function, and derived several expressions to illustrate the technique and verify the results. However, [16] later showed many of these expressions to contain serious errors. Though their corrections significantly changed many of the originally presented results, the main geodetic conclusion regarding the bias introduced by unaccounted for heterogeneity was upheld.

2.4.2 Topography

Secondly, we consider the method by [70] for incorporating moderate topography Γ_1 in a homogeneous, isotropic medium. Let the topography be described by the function $h(x_1, x_2)$. Like before, we aim to use the homogeneous halfspace as a problem for which we have solution methods available. To evaluate the traction at the topographic surface we consequently require an extrapolation operator. To this end we define

$$T_0 f = (x_1, x_2) \rightarrow f(x_1, x_2, 0), \quad (2.35a)$$

$$T_h f = T_0 f + h T_0 \left(\frac{\partial f}{\partial x_3} \right). \quad (2.35b)$$

Here T_0 restricts any halfspace solution to the surface of the halfspace; T_h extrapolates this to the topographic height by first order Taylor expansion.

Assuming that the vertical topographic variations are small in comparison to the horizontal length scales we can approximate the surface normal by

$$\tilde{n} = \left(-\frac{\partial h}{\partial x_1}, -\frac{\partial h}{\partial x_2}, 1 \right). \quad (2.36)$$

Given a halfspace stress solution, we can now take (2.35) to extrapolate the stress upward, and use the normal vector of (2.36) to compute a first order approximation of the traction at the topographic surface. The result can be rewritten using the equilibrium condition of (2.1a), to

$$T_h \sigma(u) \tilde{n} = T_0 \sigma(u) e_3 - h T_0 f - \tau(u), \quad (2.37)$$

where u is a halfspace solution, e_i is the i -th unit vector, and where we have defined

$$\tau(u) = \sum_{i=1,2} \frac{\partial h T_0 \sigma(u) e_i}{\partial x_i} \quad (2.38)$$

We observe that the traction at the topographic surface is approximated to first order by the traction at the halfspace surface, reduced by the body forces integrated to the topographic surface, and reduced by horizontal derivatives of stress and height. By (2.1d) this surface traction should be zero.

We proceed as before by forming the series of (2.33), of which the terms are formed by iteratively solving the homogeneous halfspace problem with surface tractions,

$$\begin{array}{lll} \operatorname{div} \sigma(u_0) = -f, & \sigma_n(u_0) = h T_0 f, & \llbracket u_0 \rrbracket = b \\ \operatorname{div} \sigma(u_1) = 0, & \sigma_n(u_1) = \tau(u_0), & \llbracket u_1 \rrbracket = 0 \\ \operatorname{div} \sigma(u_2) = 0, & \sigma_n(u_2) = \tau(u_1), & \llbracket u_2 \rrbracket = 0 \\ \vdots & \vdots & \vdots \\ \hline \operatorname{div} \sigma(u) = -f, & \sigma_n(u) = \tau(u) + h T_0 f & \llbracket u \rrbracket = b, \end{array} \quad (2.39)$$

where $\sigma_n(u)$ denotes the traction at the halfspace surface Γ_1 . Typically, in absence of body forces f , the first iteration involves application of any of the techniques covered in this paper, while the following involve techniques for general traction boundary problems, not covered here.

If the series converge, the approximate traction of (2.37) equals zero, thereby satisfying the traction-free boundary condition at Γ_1 . Because of the truncations involved in the extrapolation of the traction vector, however, no accuracy is gained beyond the first iteration. For higher accuracy a higher quality extrapolation is required. Truncating at first order, the extrapolated displacements at the topographic surface become

$$u_h = T_h u_0 + T_0 u_1, \quad (2.40)$$

or equivalently,

$$u_h(x_1, x_2) = u_0(x_1, x_2, 0) + u_1(x_1, x_2, 0) + h(x_1, x_2) \frac{\partial u_0}{\partial x_3}(x_1, x_2, 0). \quad (2.41)$$

2.5 Finite Elements

Analytical methods are fast and accurate, but are often limited to problems of moderate complexity. In this section we explore the Finite Element Method (FEM), which has the ability to solve the problem defined in Section 2.1 for any set of conditions including topography, anisotropy, and heterogeneity. We begin with a derivation of the weak formulation, which is an essential element of Finite Element computations, followed by an outline of the method itself. Fig. 2.9 shows a flow chart of the resulting procedure. For a more extensive background to FEM we refer to reference works such as [39].

2.5.1 Weak formulation

The Finite Element Method is a generic approximation method for differential equations. It depends on a weak reformulation of the problem, which we shall derive first. For the problem considered it is obtained by multiplying (2.1a) by a test function v and integrating over the domain,

$$\int_{\Omega \setminus \mathcal{F}} v \cdot [\operatorname{div} \sigma(u) - f] \, d\Omega = 0, \quad (2.42)$$

combined with the condition that (2.42) must hold for any sufficiently smooth function v . Per usual, this is proceeded by partial integration and substitution of boundary conditions, by which we arrive at the weak formulation. An uncommon element is the internal boundary \mathcal{F} which gives rise to an extra boundary integral, as formalized previously in (2.8). Substituting this result in (2.42) yields

$$a(u, v) = \int_{\mathcal{F}} [\llbracket v \rrbracket \cdot \{\sigma_\nu(u)\} + \{v\} \cdot \llbracket \sigma_\nu(u) \rrbracket] \, dS + \int_{\Gamma} v \cdot \sigma_n(u) \, dS + \int_{\Omega \setminus \mathcal{F}} v \cdot f \, dV, \quad (2.43)$$

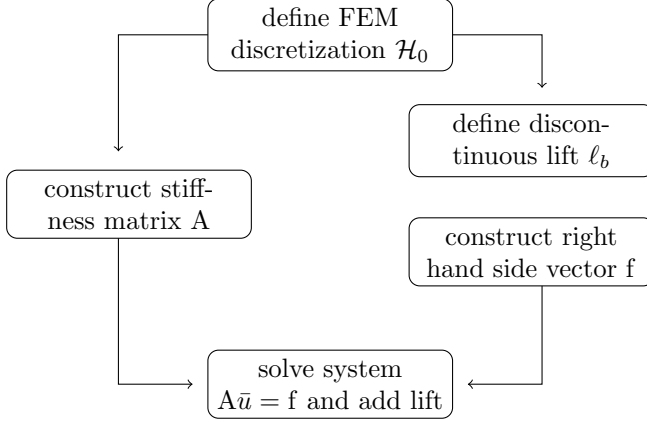


Figure 2.9 – Flow chart representation of the finite element method.

where a is the bilinear form defined in (2.6). We recognize terms from boundary conditions (2.1c) and (2.1d). Substituting these, we arrive at

$$a(u, v) = \int_{\mathcal{F}} \llbracket v \rrbracket \cdot \{\sigma_\nu(u)\} dS + \int_{\Gamma_2} v \cdot \sigma_n(u) dS + \int_{\Omega \setminus \mathcal{F}} v \cdot f dV. \quad (2.44)$$

The traction terms σ_ν and σ_n are unbounded and need to be eliminated from the formulation. This is achieved by restricting the set of admissible test functions to

$$\mathcal{H}_0 := \{v \in \mathcal{H}^1(\Omega \setminus \mathcal{F}) : v|_{\Gamma_2} = 0, \llbracket v \rrbracket = 0\}. \quad (2.45)$$

In the first condition we recognize the restriction that is commonly imposed at Dirichlet-type boundaries. The second condition is the less common, but analogous restriction for imposed jumps, where it restricts the test function to be continuous over \mathcal{F} . For test functions $v \in \mathcal{H}_0$, (2.44) transforms to

$$a(u, v) = \mathcal{L}(v), \quad (2.46)$$

where we introduced the linear functional

$$\mathcal{L}(v) := \int_{\Omega} v \cdot f dV. \quad (2.47)$$

Note that the domain of integration is extended to include \mathcal{F} , which is permitted because both v and f are bounded.

Still lacking in the formulation at this point are boundary conditions (2.1b) and (2.1e), which are direct restrictions on the admissible solutions u . In order to enforce the jump condition we introduce a *lift* function ℓ_b , that features the desired dislocation $\llbracket \ell_b \rrbracket = b$ at \mathcal{F} , and is smooth on the remainder $\Omega \setminus \mathcal{F}$. With ℓ_b fixed, we have defined an (affine) space of admissible solutions of the form

$$u = \ell_b + u_0, \quad u_0 \in \mathcal{H}_0. \quad (2.48)$$

Substituting solutions of this form in (2.46), the weak formulation of (2.1) becomes: to find $u_0 \in \mathcal{H}_0$, such that

$$\forall v \in \mathcal{H}_0 : a(v, u_0) = \mathcal{L}(v) - a(v, \ell_b). \quad (2.49)$$

This is the problem that the Finite Element Method solves in discrete sense.

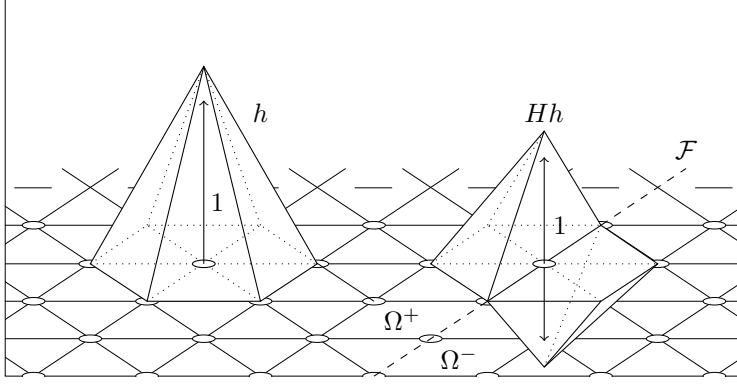


Figure 2.10 – Representation of a scalar-valued finite element shape function on a 2D triangular mesh. Left: a typical first order shape function, which is nonzero at a limited set of elements and continuous everywhere. Right: the same shape function multiplied by the Heaviside function (2.56), which introduces a discontinuity along \mathcal{F} .

2.5.2 Finite Element Method

The first step in any Finite Element computation, as indicated in Fig. 2.9, is to define a suitable discretization for function space \mathcal{H}_0 . Practically this is done by dividing the domain Ω in tetrahedral or hexahedral shaped *elements*, which are equipped with piecewise polynomial, vector valued *shape functions* $h_1, h_2, \dots, h_N : \Omega \rightarrow \mathbb{R}^3$. For first order polynomials, a particular shape function on a two-dimensional mesh may look like the leftmost, hat-shaped function of Fig. 2.10. The shape functions span the following approximation space:

$$\mathcal{H}_0 \supset \mathcal{H}'_0 := \left\{ \sum_{n=1}^N h_n v_n : v \in \mathbb{R}^N \right\}. \quad (2.50)$$

For the inclusion to hold, shape functions that have support on the Dirichlet boundary Γ_2 are eliminated, such that $h_i|_{\Gamma_2} = 0$; the jump condition $\llbracket h_i \rrbracket$ is automatically satisfied by continuity over Ω . We now define the discretized *trial function* u_0 and *test function* v by assigning weights from vectors u and v ,

$$u_0(x) = \sum_{n=1}^N h_n(x) u_n, \quad v(x) = \sum_{n=1}^N h_n(x) v_n. \quad (2.51)$$

A discrete system is now directly obtained via substitution in (2.49), the discrete analogue of which becomes

$$\forall v \in \mathbb{R}^N : \sum_{i=1}^N \sum_{j=1}^N v_i A_{ij} u_j = \sum_{i=1}^N v_i f_i, \quad (2.52)$$

where

$$\begin{aligned} A_{ij} &:= a(h_i, h_j), \\ f_i &:= \mathcal{L}(h_i) - a(h_i, \ell_b). \end{aligned} \quad (2.53)$$

To verify that this equality holds for all vectors \mathbf{v} it is sufficient to verify that it holds for all unit vectors. This turns (2.52) into a linear system of the form

$$\mathbf{A}\mathbf{u} = \mathbf{f}. \quad (2.54)$$

Note finally that as a consequence of the reciprocity principle (2.10), the *stiffness matrix* \mathbf{A} is symmetric.

The last open issue at this point is the lift ℓ_b . Mathematically, any choice is valid that satisfies the jump condition $[[\ell_b]] = b$. Practically, the natural and by far most common choice is to reuse the shape functions for this purpose. Because these are inherently continuous, we introduce a discontinuity at \mathcal{F} by multiplication with a Heaviside function:

$$\ell_b(x) := \sum_{n=1}^N H(x) h_n(x) b_n. \quad (2.55)$$

Here \mathbf{b} is a weights vector and H the Heaviside function. The Heaviside relies on a partitioning of the domain in subdomains Ω^+ and Ω^- with a shared boundary that contains \mathcal{F} . At this extended dislocation it introduces a unit step:

$$H(x) := \begin{cases} +\frac{1}{2} & x \in \Omega^+ \\ -\frac{1}{2} & x \in \Omega^-. \end{cases} \quad (2.56)$$

The jump introduced by the lift function then becomes

$$[[\ell_b]] = \sum_{n=1}^N (h_n|_{\mathcal{F}}) b_n, \quad (2.57)$$

where $h_n|_{\mathcal{F}}$ denotes the restriction of h_n to the dislocation. The weights \mathbf{b} are chosen such that this jump closely follows b , which is typically a good approximation for smooth slip distributions and which converges under mesh refinement. Shape functions h_n that have no support on \mathcal{F} are assigned zero weights, localizing the lift to the dislocation. Adding the resulting lift to the solution of (2.54), the final displacement field becomes

$$u(x) = \sum_{n=1}^N h_n(x) [\mathbf{u}_n + H(x) \mathbf{b}_n]. \quad (2.58)$$

So far we have not made any assumption about the location of the dislocation in connection with the finite element mesh. When these are aligned such that the dislocation \mathcal{F} coincides with finite element edges we arrive at the “split nodes” method originally introduced by [51] — although it must be noted that this is the modern interpretation of the original idea, including, most notably, the use of the term ‘lift’. The primary benefit of the alignment is that the lift remains piecewise polynomial, which is a requirement for efficient numerical integration. This closely resembles standard Dirichlet boundary conditions, which involve a similar finite element based lift and for which theory is well established. It also greatly helps integration of the method in existing finite element codes. Practically the discretization corresponds to a discontinuously transformed domain such as shown in Fig. 2.11. The figure also visually motivates the term ‘split nodes’.

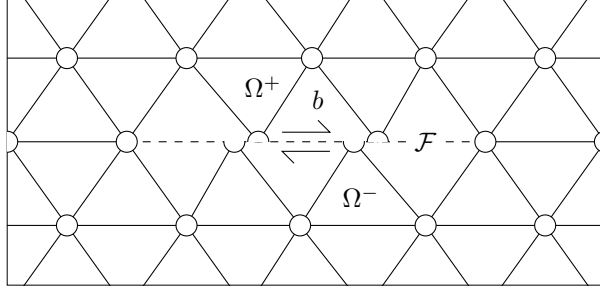


Figure 2.11 – Schematic representation of a dislocated 2D mesh, illustrating the “split node” method described by [51] where the dislocation conforms to element boundaries. Split nodes are displaced by an a priori offset. This affects the solution in neighboring elements (marked + and -), but has no effect on any of the remaining nodes.

Elegance and efficiency are what make the split nodes technique popular. An important downside of the method, however, is the requirement that element edges and dislocation should coincide, which closely links the finite element mesh to the dislocation geometry. This makes it computationally very expensive to consider many different dislocation geometries. For these situations it is possibly more beneficial to choose a mesh-independent lift function. It is worth mentioning in that context that the lift defined in (2.55) bears great similarities to expressions found in Partition of Unity and XFEM methods. Relevant work in this direction is conducted for example by [34].

Lastly, although the finite element method is very powerful, it has one fundamental drawback in its restriction to finite domains. Physically, domains such as the earth are, obviously, finite. Practically, however, the computational domain will always be truncated around a region of interest, which introduces a truncation error via the non-physical boundary. The common solution is to make the domain large enough for boundary effects to be negligible, or to use techniques such as infinite element boundaries, such as described by [8]. In comparison, analytical methods that apply to halfspace domains have the advantage of not introducing any boundary effects.

2.6 Conclusions

This paper is a general review of the mathematics of dislocation theory. We started by formally defining the dislocation problem, and presented a concise and uniform overview of available solution methods that exist for this problem. We showed that a predominant number of analytical methods apply to the subclass of halfspace problems, which are limited to a semi-infinite medium and a flat surface. The finite element method is in many ways the most powerful method; its only real restriction is the limitation to finite domains, which is easily remedied by choosing the domain suitably large. Other methods may have their limitations similarly relaxed, for example by using Taylor expansions around analytical solutions. A truly general solution method for the full class of Volterra dislocation problems as defined in (2.1), however, does not exist.

In terms of accuracy, all presented methods can be made arbitrarily accurate, with the exception of the topographic expansion of Section 2.4. The analytical solutions such as Okada's and Comninou's are exact for the specific problems they describe. In practice, however, a linear superposition is usually required to approximate a continuous slip distribution or a curved geometry. A finer distribution leads to improved accuracy. In Taylor series expansions, a higher order expansion leads to improved accuracy. In finite element methods, a finer mesh leads to improved accuracy. Of course, a higher accuracy comes at a greater computational effort.

In terms of work the methods differ considerably. Analytical methods such as Okada's equations can be evaluated relatively cheaply, but the work scales linearly with the number of point evaluations required. Propagator matrix methods are more expensive because of the act of propagation, which scales the work with the number of vertical layers. Series expansion methods add work to the method that underlies the expansion. Finite element methods are at the expensive end of the spectrum, but differ in that the work is virtually independent of the number of point evaluations. This cascade represents the main trade-off in dislocation modelling practice. A low computational cost comes at the expense of restrictions in domain and material properties.

In practice the dislocation problem is often part of an inverse algorithm, which searches a parameter space to fit the model to observations. If this parameter space includes the geometry and positioning of the fault plane this inverse problem is non-linear, and it will typically require many iterations of the forward problem to converge to a solution. In these cases, the trade-off between computational cost and modelling accuracy often isolates the cheapest model as the only viable choice, even though a more powerful method would arguably lead to more accurate results. Looking forward, this suggests there is room for development of a method that manages to strike a better balance between modelling power and computational efficiency.

3 | The Weakly-enforced Slip Method

Tectonic faults are commonly modelled as Volterra or Somigliana dislocations in an elastic medium. Various solution methods exist for this problem. However, the methods used in practice are often limiting, motivated by reasons of computational efficiency rather than geophysical accuracy. A typical geophysical application involves inverse problems for which many different fault configurations need to be examined, each adding to the computational load. In practice, this precludes conventional finite-element methods, which suffer a large computational overhead on account of geometric changes.

This paper presents a new non-conforming finite-element method based on weak imposition of the displacement discontinuity. The weak imposition of the discontinuity enables the application of approximation spaces that are independent of the dislocation geometry, thus enabling optimal reuse of computational components. Such reuse of computational components renders finite-element modeling a viable option for inverse problems in geophysical applications. A detailed analysis of the approximation properties of the new formulation is provided. The analysis is supported by numerical experiments in 2D and 3D.

This chapter is a verbatim copy of “Discontinuities without discontinuity: The Weakly-enforced Slip Method” by G.J. van Zwieten, E.H. van Brummelen, K.G. van der Zee, M.A. Gutiérrez and R.F. Hanssen; Computer Methods in Applied Mechanics and Engineering, 2014 [77].

The world is perpetually reminded of the fact that seismic hazard is still beyond reach of prediction — as it was most recently by the disaster that struck Japan. The difficulty is not just to predict the exact moment of failure, which, as argued by some [33], might never reach a level of practicality. It is also the nature of the risk, and the extent to which stress is accumulating, that turns out to be surprisingly difficult to constrain. The 2011 Tohoku-Oki earthquake demonstrated a great lack of understanding of ongoing tectonics [43]. Arguably, a better understanding could have reduced the secondary effects if such information had led to more apt measures and regulations.

The main reason for this poor state of information can be traced to the absence of direct measurements. The primary quantities of interest, being the magnitude and orientation of the stress tensor in the earth's crust, can be obtained only through tedious, expensive, point-wise measurements. A viable broad scale method to directly measure the global stress field does not exist. For this reason information is obtained mostly from secondary observables, earthquakes themselves being an important source. Earthquakes represent significant, near instantaneous changes in the global stress field. By accurately determining the location of the segment of the fault that collapsed, one can progressively update the stress field and evolve it in time. This way the tectonic evolution is monitored, and hazardous areas can be identified as regions where stress accumulates. For successful tracking of stress development, however, it is essential to understand the tectonic mechanism behind every earthquake. This includes the location and geometry of the section of the fault that collapsed, and the direction and magnitude of fault slip. It is increasingly popular to base such analyses on local co-seismic surface displacements. This type of information has become available since the nineties with the advent of space borne interferometric SAR measurements of the earth's surface, and with the widespread availability of GPS measurements [66]. Analysis of this data has in recent years seen rapid adoption and is now routinely performed for all major earthquakes.

A mechanical model is required to connect observations to physics. Most commonly (if not exclusively) used is an elastic dislocation model, based on the assumption that on short time scales, nonlinear (plastic) effects are negligible. The model embeds a displacement discontinuity of given location and magnitude in an elastic medium, causing the entire medium to deform under the locked-in stress. Many different solution methods have been developed for this particular problem, based on analytical solutions or numerical approximations; see for instance [78] for an overview of the most prominent methods. However, methods founded on analytical solutions generally dictate severe model simplifications, such as elastic homogeneity or generic geometries, which restricts their validity. The computational complexity of methods based on numerical approximation, on the other hand, is typically prohibitive in practical applications. Because in practice the surface displacements are given, and the dislocation parameters are the unknowns, the computational setting is always that of an inverse problem. A typical inversion requires several thousands of evaluations of the forward model, and therefore computational efficiency is a key requirement. Moreover, the forward problems in the inversion process are essentially identical, except for the fault geometry. Reuse of computational components, such as approximate factors of the system matrix, is imperative for efficiency of the inversion. Current numerical methods for seismic problems do not offer such reuse options.

Finite-element methods provide a class of numerical techniques that are particularly versatile in terms of modeling capabilities in geophysics. Finite-element methods allow for elastic heterogeneity, anisotropy, and topography; all things that can not well be accounted for with currently used analytical and semi-analytical methods. In geophysical practice, finite-element methods are however often rejected for reasons of computational cost. The high computational cost can be retraced to the condition that the geometry of the fault coincides with element edges, which is a requirement engendered by the strong enforcement of the dislocation; see [51]. Consequently, the mesh geometry depends on the fault, which in turn implies that mesh-dependent com-

ponents such as the stiffness matrix and approximate factorizations of that matrix cannot be reused for different fault geometries and must be recomputed whenever the geometry of the fault changes. The recomputation of these components in each step of a nonlinear inversion process leads to a prohibitive overall computational complexity.

To overcome the complications of standard finite-element techniques in nonlinear inversion processes in tectonophysics, this paper introduces the *Weakly-enforced Slip Method* (WSM), a new numerical method in which displacement discontinuities are weakly imposed. The WSM formulation is similar to Nitsche’s variational principle for enforcing Dirichlet boundary conditions [54]. The weak imposition of the discontinuity in WSM decouples the finite element mesh from the geometry of the fault, which renders the stiffness matrix and derived objects such as approximate factors independent of the fault and enables reuse of these objects. Therefore, even though the computational work required for a single realisation of the fault geometry is comparable to that of standard FEM, reuse of components makes WSM significantly more efficient when many different fault geometries are considered. This makes finite-element computations based on WSM a viable option for nonlinear inverse problems.

A characteristic feature of WSM is that it employs standard continuous finite-element approximation spaces, as opposed to the conventional FEM split-node approach [51] which introduces actual discontinuities in the approximation space. Instead, WSM approximations feature a ‘smeared out’ jump with sharply localized gradients. We will establish that the error in the WSM approximation converges only as $O(h^{1/2})$ in the \mathbf{L}^2 -norm as the mesh width h tends to zero and that the error diverges as $O(h^{-1/2})$ in the energy norm. In addition, however, we will show that the WSM approximation displays optimal local convergence in the energy norm, i.e., optimal convergence rates are obtained on any subdomain excluding a neighborhood of the dislocation. The numerical experiments convey that WSM also displays optimal local convergence in the \mathbf{L}^2 -norm.

The remainder of this manuscript is organized as follows. Section 3.1 presents strong and weak formulations of Volterra’s dislocation problem, and derives the corresponding lift-based finite-element formulation. Section 3.2 introduces the Weakly-enforced Slip Method based on two formal derivations, viz., by collapsing the support of the lift onto the fault and by application of Nitsche’s variational principle. In Section 3.3, we examine the approximation properties of the WSM formulation. Section 3.4 verifies and illustrates the approximation properties on the basis of numerical experiments for several 2D and 3D test cases. In addition, to illustrate the generality of WSM, in the numerical experiments we consider several test cases that violate the conditions underlying the error estimates in Section 3.3, such as discontinuous slip distributions and rupturing dislocations. Section 3.5 presents concluding remarks.

3.1 Problem formulation

In this section we define Volterra’s dislocation problem. We will postulate the strong formulation in 3.1.1, followed by a derivation of the weak formulation in 3.1.2. The latter will serve as a basis for the derivation of the Finite Element approximations, for which we lay foundations in Section 3.1.3.

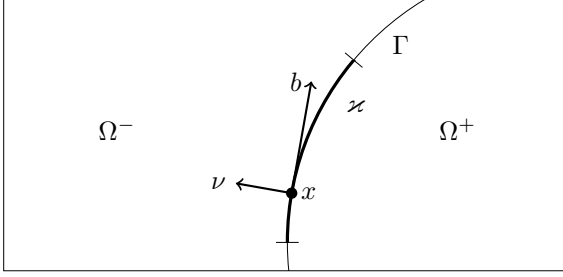


Figure 3.1 – Definition of the domain Ω , normal vector ν , slip vector b , fault Γ and dislocation κ .

3.1.1 The strong form

We start by defining the geometric setup. We consider an open bounded domain $\Omega \subset \mathbb{R}^N$ ($N = 2, 3$) with Lipschitz boundary $\partial\Omega$. An $(N-1)$ -dimensional Lipschitz manifold Γ , referred as the fault, divides the domain in two disjoint open subdomains Ω^+ and Ω^- , such that $\Omega = \text{int}(\overline{\Omega^+} \cup \overline{\Omega^-})$. We equip Γ with a unit normal vector $\nu : \Gamma \rightarrow \mathbb{R}^N$ directed into the subdomain Ω^- . The fault supports a slip distribution $b : \Gamma \rightarrow \mathbb{R}^N$, corresponding to a dislocation. The fault is referred to as a non-rupturing fault if the dislocation $\kappa = \text{supp}(b)$ is compact in Γ , and a rupturing fault otherwise. Let us note that in tectonics, rupturing faults correspond to intersections of the slip plane with the surface of the earth. Figure 3.1 illustrates this setup for $N = 2$. It is to be noted that b need not be tangential to the fault. If b has a non-vanishing normal component then the fault is opening, such as may be caused by an intruding material.

The displacement field generated by the dislocation is represented by $u : \Omega \setminus \Gamma \rightarrow \mathbb{R}^N$. For convenience, we restrict our considerations to linear elasticity. We denote by the map $u \mapsto \epsilon(u)$ the strain tensor corresponding to the displacement field u , according to

$$\epsilon(u) := \frac{1}{2} [\nabla u + (\nabla u)^T], \quad (3.1)$$

under the assumption that u is differentiable on $\Omega \setminus \Gamma$. The constitutive behavior corresponds to Hooke's law:

$$\sigma(u) := A : \epsilon(u), \quad (3.2)$$

i.e., $\sigma_{ij}(u) = A_{ijkl}\epsilon_{kl}(u)$, where we adhere to the convention on summation on repeated indices. The tensors σ and A are referred to as the stress tensor and the elasticity tensor, respectively. The elasticity tensor is subject to the usual symmetries $A_{ijkl} = A_{ijlk} = A_{klij}$. Moreover, we assume that it is bounded and satisfies a strong positivity condition, i.e., there exist positive constants $\bar{c}_A > 0$ and $\underline{c}_A > 0$ such that:

$$\underline{c}_A e_{ij}e_{ij} \leq A_{ijkl}e_{ij}e_{kl} \leq \bar{c}_A e_{ij}e_{ij} \quad (3.3)$$

for all tensors e . The elasticity tensor is in principle allowed to vary over the domain Ω , subject to the above conditions. Auxiliary smoothness conditions on A will be introduced later.

To facilitate the formulation, we denote by $\sigma_n(u) := \sigma(u) \cdot n$ the traction on a boundary corresponding to the displacement field u . Moreover, we introduce the jump

operator $\llbracket \cdot \rrbracket$ and average operator $\{\cdot\}$ according to:

$$\llbracket v \rrbracket : \Gamma \ni x \mapsto v^+(x) - v^-(x), \quad (3.4a)$$

$$\{v\} : \Gamma \ni x \mapsto \frac{1}{2}[v^+(x) + v^-(x)], \quad (3.4b)$$

where v^+ and v^- represent the traces of v from within Ω^+ and Ω^- , respectively. Given a partition of the boundary $\partial\Omega$ into $\mathcal{D} \neq \emptyset$ and \mathcal{N} such that $\mathcal{D} \cap \mathcal{N} = \emptyset$, we define the Volterra dislocation problem as follows:

Strong formulation: given a body force $f : \Omega \rightarrow \mathbb{R}^N$, displacement $g : \mathcal{D} \rightarrow \mathbb{R}^N$, traction $h : \mathcal{N} \rightarrow \mathbb{R}^N$, and slip $b : \Gamma \rightarrow \mathbb{R}^N$, find displacement field $u : \Omega \setminus \Gamma \rightarrow \mathbb{R}^N$ such that

$$-\operatorname{div} \sigma(u) = f \text{ in } \Omega \setminus \Gamma \quad (3.5a)$$

$$\llbracket u \rrbracket = b \text{ on } \Gamma \quad (3.5b)$$

$$\llbracket \sigma_\nu(u) \rrbracket = 0 \text{ on } \Gamma \quad (3.5c)$$

$$u = g \text{ on } \mathcal{D} \quad (3.5d)$$

$$\sigma_n(u) = h \text{ on } \mathcal{N} \quad (3.5e)$$

Equation (3.5a) is the usual equilibrium condition, which applies everywhere in Ω except on the manifold Γ . Equations (3.5b) and (3.5c) respectively express that the displacements at the boundaries of Ω^+ and Ω^- differ by the slip vector b , and that the tractions at the boundaries of Ω^+ and Ω^- are in static equilibrium, i.e., equal and opposite. It is to be remarked that this condition corresponds to a standard linear approximation in the small-slip limit, as the traction equilibrium at the fault occurs in fact in the deformed configuration. The boundary conditions (3.5d) and (3.5e) correspond to Dirichlet and Neumann boundary conditions, representing a prescribed displacement and a prescribed traction, respectively.

To facilitate the presentation, we note that the solution to the Volterra dislocation problem (3.5) can be separated into a discontinuous part $u_0 : \Omega \setminus \Gamma \rightarrow \mathbb{R}^N$ with homogeneous data and a continuous part $u_1 : \Omega \rightarrow \mathbb{R}^N$ with inhomogeneous data:

$$\begin{aligned} -\operatorname{div} \sigma(u_0) &= 0 & \text{in } \Omega \setminus \Gamma & & -\operatorname{div} \sigma(u_1) &= f & \text{in } \Omega \\ u_0 &= 0 & \text{on } \mathcal{D} & & u_1 &= g & \text{on } \mathcal{D} \\ \sigma_n(u_0) &= 0 & \text{on } \mathcal{N} & & \sigma_n(u_1) &= h & \text{on } \mathcal{N} \\ \llbracket u_0 \rrbracket &= b & \text{on } \Gamma & & & & \\ \llbracket \sigma_\nu(u_0) \rrbracket &= 0 & \text{on } \Gamma & & & & \end{aligned} \quad (3.6)$$

The sum $u_0 + u_1$ satisfies (3.5). Therefore, the inhomogeneous data f, g, h in (3.5) can be treated separately in a standard continuous elasticity problem, and without loss of generality we can restrict our consideration to homogeneous data. We retain $f := 0$ to identify the right member of (3.5a).

For rupturing faults, some compatibility conditions arise with respect to the boundary conditions. In particular, an intersection of the dislocation with the boundary of the domain is only admissible at the Neumann boundary \mathcal{N} . Otherwise, an inadmissible incompatibility between the jump condition $\llbracket u_0 \rrbracket = b$ and the homogeneous Dirichlet condition $u_0 = 0$ ensues.

3.1.2 The weak form

To derive the weak formulation of (3.5), we note that for any piecewise smooth function v from $\Omega \setminus \Gamma$ into \mathbb{R}^N , we have the identities:

$$\begin{aligned} \int_{\Omega \setminus \Gamma} v \cdot \operatorname{div} \sigma(u) &= \oint_{\partial\Omega^-} v \cdot \sigma_n(u) - \int_{\Omega^-} \sigma(u) : \nabla v + \oint_{\partial\Omega^+} v \cdot \sigma_n(u) - \int_{\Omega^+} \sigma(u) : \nabla v \\ &= \oint_{\partial\Omega} v \cdot \sigma_n(u) - \int_{\Omega \setminus \Gamma} \sigma(u) : \nabla v + \int_{\Gamma} \llbracket v \rrbracket \cdot \{\sigma_\nu(u)\} + \{v\} \cdot \llbracket \sigma_\nu(u) \rrbracket \end{aligned} \quad (3.7)$$

The first identity results from integration by parts. The second identity follows from a rearrangement of the boundary terms and

$$\begin{aligned} \int_{\Gamma} v^+ \cdot \sigma_n^+(u) + v^- \cdot \sigma_n^-(u) &= \int_{\Gamma} v^+ \cdot \sigma_\nu^+(u) - v^- \cdot \sigma_\nu^-(u) \\ &= \int_{\Gamma} (v^+ - v^-) \cdot \frac{1}{2} (\sigma_\nu^+(u) + \sigma_\nu^-(u)) + \int_{\Gamma} \frac{1}{2} (v^+ + v^-) \cdot (\sigma_\nu^+(u) - \sigma_\nu^-(u)) \end{aligned} \quad (3.8)$$

In the weak formulation, the admissible displacement fields will be insufficiently regular to ensure the existence of the tractions $\sigma_n(\cdot)$. Hence, the terms involving these tractions in (3.7) must be eliminated by means of the boundary conditions and auxiliary conditions on v . The traction term on the Neumann boundary can be eliminated by means of (3.5e). To remove the traction term on $\partial\Omega$, we stipulate that v vanishes on \mathcal{D} . The traction average in the final term of (3.7) is eliminated by requiring that v be continuous. The traction jump in the final term is deleted by means of the traction-continuity condition (3.5c).

Summarizing, we find that a solution u of (3.5) satisfies

$$a_{\Gamma}(u, v) = l(v) \quad (3.9)$$

for all sufficiently smooth functions $v : \Omega \rightarrow \mathbb{R}^N$ that vanish on \mathcal{D} , where

$$a_{\Gamma}(u, v) = \int_{\Omega \setminus \Gamma} \sigma(u) : \nabla v \quad (3.10a)$$

$$l(v) = \int_{\Omega} v \cdot f \quad (3.10b)$$

Note that v is assumed to be smooth on Ω and, in particular, that it is continuous across the fault Γ .

To furnish a functional setting for the weak formulation of (3.5) based on (3.9), we denote by $\mathbf{H}^k(\Omega)$ the usual Sobolev space of square-integrable functions from Ω into \mathbb{R}^N with square-integrable distributional derivatives of order $\leq k$, equipped with the inner product

$$(u, v)_{k, \Omega} = \sum_{|\alpha| \leq k} \int_{\Omega} D^{\alpha} u \cdot D^{\alpha} v$$

and the corresponding norm $\|\cdot\|_{k, \Omega}$ and semi-norm $|\cdot|_{k, \Omega}$. For the square-integrable functions and the corresponding norm and inner-product, we introduce the condensed notation $\mathbf{L}^2(\Omega) := \mathbf{H}^0(\Omega)$, $\|\cdot\|_{\Omega} := \|\cdot\|_{0, \Omega}$ and $(\cdot, \cdot)_{\Omega} := (\cdot, \cdot)_{0, \Omega}$. We denote by

$\mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ the subspace of $\mathbf{H}^1(\Omega)$ of functions that vanish on $\mathcal{D} \subseteq \partial\Omega$. To accommodate the discontinuity corresponding to the dislocation, we introduce the lift operator $\ell(\cdot)$, which assigns to any suitable slip $b : \Gamma \rightarrow \mathbb{R}^N$ a function ℓ_b in $\mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$ such that $[\![\ell_b]\!] = b$. A precise specification of conditions on the slip distribution is given in Section 3.3. The weak formulation of (3.5) based on (3.9) writes

Weak formulation: given the lift $\ell_b \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$, find $u \in \ell_b + \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ such that

$$a_\Gamma(u, v) = l(v) \quad \forall v \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega). \quad (3.11)$$

The bilinear form $a_\Gamma : \mathbf{H}^1(\Omega \setminus \Gamma) \times \mathbf{H}^1(\Omega \setminus \Gamma) \rightarrow \mathbb{R}$ and linear form $l : \mathbf{H}^1(\Omega \setminus \Gamma) \rightarrow \mathbb{R}$ in (3.11), are the extensions (by continuity) of the corresponding forms in (3.10). The treatment of the dislocation by means of a lift operator in (3.11) is analogous to the treatment of inhomogeneous Dirichlet boundary conditions in weak formulations; see, e.g., [26, p.113]. Let us note that in the weak formulation (3.11), we have identified $\{u \in \mathbf{H}^1(\Omega \setminus \Gamma) : [\![u]\!] = 0\}$ with $\mathbf{H}^1(\Omega)$. This identification is unambiguous, on account of a one-to-one correspondence between the functions in these spaces.

To analyze the weak formulation (3.11), and to prepare the presentation of the Weakly-enforced-Slip method in Section 3.2, we note that (3.11) is to be interpreted in the following manner: find $u := \bar{u} + \ell_b$ with $\bar{u} \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ such that

$$a(\bar{u}, v) = l(v) - a_\Gamma(\ell_b, v) \quad \forall v \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega). \quad (3.12)$$

where the bilinear form $a : \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}$,

$$a(u, v) = \int_{\Omega} \sigma(u) : \nabla v, \quad (3.13)$$

corresponds to the restriction of $a_\Gamma(\cdot, \cdot)$ to $\mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega)$. Indeed, for all pairs $(u, v) \in \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega)$, the function $\sigma(u) : \nabla v$ is Lebesgue integrable on Ω , and because the manifold Γ corresponds to a set of N -Lebesgue measure zero, the integrals of $\sigma(u) : \nabla v$ on $\Omega \setminus \Gamma$ and on Ω coincide. It is important to note that the restriction of the bilinear form $a_\Gamma(\cdot, \cdot)$ to $\mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega)$ in (3.13) is independent of the fault Γ . The function \bar{u} in (3.11) is referred to as the continuous complement of the solution u with respect to the jump lift ℓ_b and, indeed, it resides in $\mathbf{H}_{0,\mathcal{D}}^1(\Omega)$.

For the assumed linear-elastic behavior according to (3.1) and (3.2), it follows straightforwardly that

$$|a_\Gamma(u, v)| \leq \bar{c}_A |u|_{1,\Omega \setminus \Gamma} |v|_{1,\Omega \setminus \Gamma} \quad (3.14a)$$

$$|a_\Gamma(u, u)| \geq \underline{c}_A |u|_{1,\Omega \setminus \Gamma}^2 \quad (3.14b)$$

for all $u, v \in \mathbf{H}^1(\Omega \setminus \Gamma)$, where \bar{c}_A and \underline{c}_A denote the continuity and strong-positivity constants of the elasticity tensor, respectively, and $|\cdot|_{1,\Omega \setminus \Gamma}$ represents the usual \mathbf{H}^1 -seminorm. By virtue of Poincaré's inequality (see, e.g., [12, Theorem 5.3.5]) there exists a bounded positive constant C_P such that

$$\|u\|_{1,\Omega \setminus \Gamma} \leq C_P |u|_{1,\Omega \setminus \Gamma} \quad \forall u \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma) \quad (3.15)$$

Hence, the \mathbf{H}^1 -norm and \mathbf{H}^1 -semi norm are equivalent on $\mathbf{H}_{0,\mathcal{D}}^1(\Omega)$. Equations (3.14) and (3.15) imply that the bilinear forms $a_\Gamma(\cdot, \cdot)$ and, accordingly, $a(\cdot, \cdot)$, are continuous and coercive. Moreover, it is easily verified that the linear form $l(\cdot) - a_\Gamma(\ell_b, \cdot) : \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}$ in the right member of (3.12) is continuous. Problem (3.12) therefore complies with the conditions of the Lax-Milgram lemma (see, for instance, [12, Theorem 2.7.7]) and, hence, it is well posed.

It is interesting to note that (3.12) allows for a physical interpretation that is very close to Volterra's classical construction for dislocations, popularly known as the 'Volterra knife': to make a cut in the material, displace the two sides and hold them while welding the seam, and finally release the sides so the material assumes its state of self-stressed equilibrium. The initial cut and displacement is represented by the lift ℓ , which is not in equilibrium and is hence maintained by an external load. The addition of \bar{u} represents the transition to a state of equilibrium, by removing the external load but leaving the displacement intact.

3.1.3 Finite element approximation

Galerkin finite-element approximation methods for Volterra's dislocation problem (3.11) are generally based on a restriction of the weak formulation to a suitable finite dimensional subspace. The general structure of finite-element methods can be found in many textbooks, for instance, [40, 74, 26, 19]. We present here the main concepts and definitions for the ensuing exposition.

The approximation spaces in finite-element methods are generally subordinate to a mesh \mathcal{T}_h , viz., a cover of the domain by non-overlapping element domains $\kappa \subset \Omega$. The subscript $h > 0$ indicates the dependence of the mesh on a resolution parameter, for instance, the diameter of the largest element in the mesh. In general, we impose some auxiliary conditions on the mesh, such as shape-regularity of the elements and conditions on the connectivity between elements; see, for instance, [26, 57] for further details. A finite-element approximation space $\mathbf{V}_h^p \subset \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ subordinate to \mathcal{T}_h can then be defined, for instance, as the subspace of vector-valued continuous element-wise polynomials of degree $\leq p$ which vanish on \mathcal{D} :

$$\mathbf{V}_h^p = \{v_h \in C^0(\bar{\Omega}, \mathbb{R}^N) : (v_h)_i|_\kappa \in \mathbb{P}^p \text{ for all } \kappa \in \mathcal{T}_h, i = 1, 2, \dots, N, v_h|_{\mathcal{D}} = 0\} \quad (3.16)$$

with \mathbb{P}^p the N -variate polynomials of degree p . Below, our interest is generally restricted to the h -dependence of the approximation space and, accordingly, we will suppress p . The finite-element approximation of (3.11) based on an approximation space \mathbf{V}_h writes: find $u_h := \bar{u}_h + \ell_b$ with $\bar{u}_h \in \mathbf{V}_h$ subject to

$$a(\bar{u}_h, v_h) = l(v_h) - a_\Gamma(\ell_b, v_h) \quad \forall v_h \in \mathbf{V}_h. \quad (3.17)$$

We refer the right-most term in the right member of (3.17) as the lift term.

Approximation properties of the Finite Element Method are generally investigated on the basis of a sequence of meshes $\mathcal{T}_{\mathcal{H}} := (\mathcal{T}_h)_{h \in \mathcal{H}}$, parametrized by a decreasing sequence of mesh parameters $\mathcal{H} = \{h_1, h_2, \dots\}$ with 0 as only accumulation point. A sequence of meshes is called quasi uniform if there exist positive constants \underline{C} and \bar{C} , independent of h , such that $\underline{C}h \leq \text{diam}(\kappa) \leq \bar{C}h$ for all $\kappa \in \mathcal{T}_h$ and all $h \in \mathcal{H}$. Standard interpolation theory in Sobolev spaces (see, for instance, [26, 12]) conveys

that a sequence of approximation spaces $\mathbf{V}_{\mathcal{H}}$ of the form (3.16) based on quasi-uniform meshes possesses the following approximation property: there exists a positive constant \mathcal{C}^\ddagger independent of h such that for all $h \in \mathcal{H}$, all $k \geq 0$ and both $m \in \{0, 1\}$, it holds that

$$\inf_{v_h \in \mathbf{V}_h} \|v - v_h\|_{m, \Omega} \leq \mathcal{C} h^{l+1-m} |v|_{l+1, \Omega} \quad \forall v \in \mathbf{H}^{k+2}(\Omega) \cap \mathbf{H}_{0, \mathcal{D}}^1(\Omega) \quad (3.18)$$

with $l = \min\{p, k+1\}$. The estimate (3.18) imparts that for all sufficiently smooth $v \in \mathbf{H}_{0, \mathcal{D}}^1(\Omega)$, the $\|\cdot\|_{m, \Omega}$ -norm of the best approximation in \mathbf{V}_h in that norm decays as $O(h^{p+1-m})$ as $h \rightarrow 0$.

The lift ℓ_b in (3.17) is in principle arbitrary. However, the use of an arbitrary lift carries severe algorithmic complexity, as one has to explicitly construct the lift and evaluate integrals involving products of (gradients of) the lift and finite-element shape functions. Moreover, the evaluation of these integrals by a suitable numerical integration scheme generally leads to a high computational complexity, because the fault is allowed to intersect elements, and there are no efficient quadrature schemes to integrate the discontinuous function that arises. Therefore, in practice, it is convenient to integrate the lift in the finite-element setting. Provided that the fault coincides with element edges, a lift ℓ_{b_h} is then constructed in the broken approximation space:

$$\hat{\mathbf{V}}_h = \{u \in C^0(\bar{\Omega} \setminus \Gamma) : u|_{\Omega^\pm} \in (\mathbf{V}_h)|_{\Omega^\pm}\} \quad (3.19)$$

It is to be noted that the slip does not generally reside in $[\hat{\mathbf{V}}_h]$ and, accordingly, b is to be replaced by a suitable interpolant b_h . Moreover, if the fault does not coincide with element edges, then it is to be replaced by an approximation subject to this condition. The aforementioned approach corresponds to the split-node method by Melosh [51], where the adjective ‘split’ refers to the discontinuities between the elements in Ω^+ and Ω^- contiguous to Γ . The split-node approach is analogous to the standard treatment of Dirichlet boundary conditions; see, for instance, [26, §3.2.2]. The split-node approach bypasses the aforementioned complications of an arbitrary-lift approach and the evaluation of the lift term comes essentially free of charge as part of the regular stiffness matrix. A fundamental disadvantage of the split-node approach, however, is that it requires that the fault coincides with element edges, which connects the fault geometry to the geometry and, generally, the topology of the mesh. As a result, computational primitives such as the stiffness matrix and preconditioners for the stiffness matrix, which are contingent on the mesh, cannot be reused for analyses of alternative fault geometries. This is a prohibitive restriction if many dislocation geometries have to be considered, for instance, in inverse problems.

3.2 The Weakly-enforced Slip Method

In section 3.1.3 we substantiated that the treatment of the lift term in the split-node approach, which is the natural counterpart of the standard treatment of Dirichlet boundary conditions in finite-element approximations, is unsuitable if many fault

[‡]We use \mathcal{C} to denote a generic positive constant, of which the value and connotation may change from one instance to the next, even within a single chain of expressions.

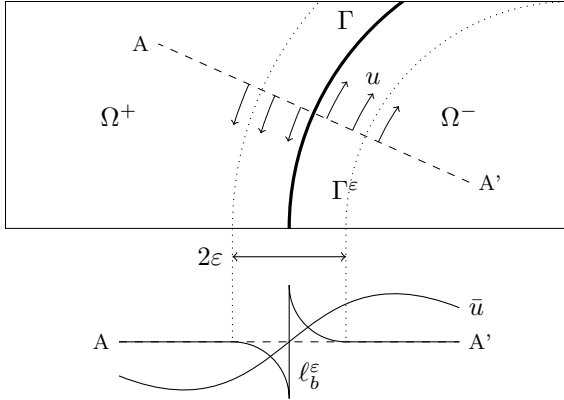


Figure 3.2 – Schematic representation of an ε -local lift $\text{supp}(\ell_b^\varepsilon) \in \Gamma^\varepsilon$, which is added to a continuous \bar{u} to form the discontinuous displacement field u .

geometries have to be analysed, on account of the inherent dependence of the finite-element mesh on the fault geometry.

In this section we propose a new and fundamentally different treatment of the lift term that retains mesh independence: the *Weakly-enforced Slip Method* (WSM). Below, we present two different formal derivations of the WSM formulation. The derivation in Section 3.2.1 relies on a limit procedure. In Section 3.2.2, we derive the WSM formulation on the basis of Nitsche’s variational principle for enforcing Dirichlet-type boundary conditions [54].

3.2.1 Collapsing the lift

Our aim is to derive a tractable finite-element approximation of Volterra’s dislocation problem (3.11), in which the finite-element space and the bilinear form and, accordingly, the stiffness matrix are independent of the fault geometry.

In principle, the lift-based Galerkin formulation (3.17) already exhibits the appropriate form. However, as elaborated in Section 3.1.3, the corresponding finite-element formulation is intractable for general lift operators. One can infer, however, that the complications engendered by a general lift operator can be avoided by collapsing the support of the lift on the fault. The integration of a discontinuous function in Ω then reduces to the integration of a smooth function on the dislocation. Numerical evaluation of integrals on the dislocation is feasible given a parametrization of the fault. Moreover, the intricate explicit construction of a lift is obviated, and only the slip distribution itself is required, which is presented as part of the problem specification. A further advantage of collapsing the lift is that of localization: instead of having to evaluate the lift term for all shape functions of which the support intersects with the support of the lift, only the shape functions of which the support intersects with the dislocation have to be considered.

To derive the lift term corresponding to a collapsed lift, we consider a symmetric lift ℓ_b^ε as illustrated in Fig. 3.2, with compact support in an ε -neighborhood of the fault:

$$\Gamma^\varepsilon := \{x \in \Omega : \text{dist}(x, \Gamma) < \varepsilon\}. \quad (3.20)$$

By virtue of the compact support of ℓ_b^ε in Γ^ε , the lift term of (3.17) evaluates to

$$a_\Gamma(\ell_b^\varepsilon, v) = \int_{\Gamma^\varepsilon \setminus \Gamma} \sigma(v) : \nabla \ell_b^\varepsilon = \int_\Gamma b \cdot \{\sigma_\nu(v)\} - \int_{\Gamma^\varepsilon \setminus \Gamma} \ell_b^\varepsilon \cdot \operatorname{div} \sigma(v). \quad (3.21)$$

The first identity follows from the symmetry of the bilinear form in (3.10a). The second identity results from integration-by-parts and $[\![\ell_b^\varepsilon]\!] = b$ and $\{\ell_b^\varepsilon\} = 0$. Let us note that the second identity is formal in the sense that it requires more regularity of v than is actually provided by $\mathbf{H}^1(\Omega)$. We shall momentarily ignore this aspect, but it manifests itself in the analysis of the approximation properties of the WSM formulation in Section 3.3. Without loss of generality, we can assume ℓ_b^ε to be bounded independent of ε . The second term in the final expression in (3.21) then vanishes if Γ^ε collapses on Γ . Therefore, formally passing to the limit in (3.21), we obtain

$$a_\Gamma(\ell_b^\varepsilon, v) \xrightarrow{\varepsilon \rightarrow +0} \int_\Gamma b \cdot \{\sigma_\nu(v)\} \quad (3.22)$$

According to (3.22), the lift term reduces to an integral on Γ in the limit of collapsing the support of the lift onto the fault. The WSM formulation corresponds to replacing the lift term $a_\Gamma(\ell_b, \cdot)$ in the right member of (3.17) by the limit functional according to (3.22):

Weakly-enforced Slip Method: given a slip distribution $b : \Gamma \rightarrow \mathbb{R}^N$, find $u_h \in \mathbf{V}_h$ such that

$$a(u_h, v_h) = l(v_h) - \int_\Gamma b \cdot \{\sigma_\nu(v_h)\} \quad \forall v_h \in \mathbf{V}_h. \quad (3.23)$$

The nomenclature *Weakly-enforced Slip Method* serves to indicate that in (3.23) the slip discontinuity is weakly enforced in the right-hand side only, and does not appear in the approximation space.

It is to be noted that although the WSM formulation (3.23) is derived from the lift-based formulation (3.17) by collapsing the lift, in contrast to (3.17) we do not add a lift to the continuous complement u_h . Because $\mathbf{V}_h \subset C^0(\Omega, \mathbb{R}^N)$, WSM thus yields a continuous approximation to the discontinuous solution of the Volterra dislocation problem (3.11). This implies that the approximation near the dislocation will inevitably be inaccurate. We will however show in Section 3.3 that away from the dislocation, the error in the WSM approximation converges optimally under mesh refinement.

3.2.2 Alternative derivation via Nitsche's variational principle

To further elucidate the WSM formulation, we present in this section an alternative derivation of (3.23) based on *Nitsche's Variationsprinzip* [54]. Nitsche presented in [54] a variational principle for weakly imposing Dirichlet boundary conditions in finite-element approximations of elliptic problems, i.e., without incorporating such

essential boundary conditions in the approximation space. Nitsche's variational principle can be extended to the Volterra dislocation problem (3.11) to weakly impose the slip discontinuity. To specify this extension, we consider a suitable broken space $\hat{\mathbf{V}}(h)$ which encapsulates the broken approximation space $\check{\mathbf{V}}_h$ and contains the solution u to the Volterra dislocation problem (3.11). We define the quadratic functional $J : \hat{\mathbf{V}}(h) \rightarrow \mathbb{R}$:

$$J(w) = \frac{1}{2}a_\Gamma(w, w) - \int_\Gamma \llbracket w \rrbracket \cdot \{\sigma_\nu(w)\} + \frac{1}{2}\psi \int_\Gamma \llbracket w \rrbracket^2, \quad (3.24)$$

for some suitable constant $\psi > 0$, generally dependent on h . Let $\check{\mathbf{V}}_h$ denote either the broken approximation space $\check{\mathbf{V}}_h$ or the continuous approximation space \mathbf{V}_h and consider the approximation $\check{u}_h \approx u$ according to:

$$\check{u}_h := \arg \inf_{v_h \in \check{\mathbf{V}}_h} J(u - v_h) \quad (3.25)$$

Equation (3.25) implies that \check{u}_h satisfies the Kuhn-Tucker optimality condition $J'(u - \check{u}_h)(v_h) = 0$ for all $v_h \in \check{\mathbf{V}}_h$, where $v \mapsto J'(w)(v)$ denotes the Fréchet derivative of J at w . For J according to (3.24), the optimality condition implies that $\check{u}_h \in \check{\mathbf{V}}_h$ satisfies:

$$\begin{aligned} a_\Gamma(\check{u}_h, v_h) - \int_\Gamma \llbracket \check{u}_h \rrbracket \cdot \{\sigma_\nu(v_h)\} - \int_\Gamma \llbracket v_h \rrbracket \cdot \{\sigma_\nu(\check{u}_h)\} + \psi \int_\Gamma \llbracket \check{u}_h \rrbracket \cdot \llbracket v_h \rrbracket \\ = a_\Gamma(u, v_h) - \int_\Gamma \llbracket u \rrbracket \cdot \{\sigma_\nu(v_h)\} - \int_\Gamma \llbracket v_h \rrbracket \cdot \{\sigma_\nu(u)\} + \psi \int_\Gamma \llbracket u \rrbracket \cdot \llbracket v_h \rrbracket \\ = l(v_h) - \int_\Gamma b \cdot \{\sigma_\nu(v_h)\} + \psi \int_\Gamma b \cdot \llbracket v_h \rrbracket \quad \forall v_h \in \check{\mathbf{V}}_h. \end{aligned} \quad (3.26)$$

The final identity follows by invoking integration-by-parts on $a_\Gamma(u, v_h)$, a rearrangement of terms and the strong formulation of Volterra's dislocation problem in (3.5).

If the broken approximation space $\hat{\mathbf{V}}_h$ is inserted for $\check{\mathbf{V}}_h$, the optimality condition (3.26) can be reinterpreted as a symmetric-interior-penalty (SIP) discontinuous-Galerkin-type formulation; see, for instance, [57, Sec. 4.2]. In contrast to standard discontinuous Galerkin formulations, the slip terms in the right-hand side, i.e., the terms containing b in the ultimate expression in (3.26), enforce the jump discontinuity at the fault. Convergence results for this formulation can be established in a similar manner as in [54]. For suitable stabilization parameters ψ , the functional J in (3.24) is equivalent to $\|\cdot\|_{1,\Omega \setminus \Gamma}$ and (3.25) implies quasi-optimal convergence of \check{u}_h .

If the continuous approximation space \mathbf{V}_h is inserted for $\check{\mathbf{V}}_h$, the terms containing $\llbracket \check{u}_h \rrbracket$ and $\llbracket v_h \rrbracket$ vanish, and we obtain the WSM formulation (3.23). Hence, WSM can indeed be interpreted as an extension of Nitsche's variational principle to the Volterra dislocation problem with continuous approximation spaces. Furthermore, in view of $\hat{\mathbf{V}}_h \supset \mathbf{V}_h$, WSM can also be regarded as a SIP discontinuous Galerkin formulation, based on a continuous subspace. One can infer that the WSM approximation retains the quasi-optimal approximation property in $\|\cdot\|_{1,\Omega \setminus \Gamma}$. However, since the continuous approximation spaces applied in WSM are not dense in $H_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$, the immediate significance of this quasi-optimality for the approximation properties of WSM is limited.

3.3 Approximation properties of WSM

An analysis of the approximation properties of the Weakly-enforced Slip Method is non-trivial, owing to the fact that in WSM one considers approximations in $\mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ -conforming subspaces, while the solution itself resides in $\mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$, and the embedding of $\mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ in $\mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$ is non-dense. Essentially, we attempt to approximate a discontinuous function by a continuous one and, in doing so, we incur an error that does not vanish under mesh refinement. Standard techniques to assess global approximation properties, on all of $\Omega \setminus \Gamma$, viz., Céa's lemma or the Strang lemmas [26, Lems. 2.25-27], therefore provide only partial information; see Section 3.3.2.

To provide a foundation for analyzing the approximation provided by WSM, we first recall some aspects of traces and tractions in Section 3.3.1. Section 3.3.2 investigates the global approximation properties of WSM, i.e., on the entire domain. Section 3.3.3 establishes the local approximation properties of WSM, i.e., on the domain excluding a neighborhood of the fault.

3.3.1 Traces and tractions

To enable an analysis of the approximation behavior of WSM, some elementary aspects of trace theory are required. For a comprehensive overview, we refer to [61]. To make the theory applicable to the Volterra dislocation problem, we must impose auxiliary smoothness conditions on the elasticity tensor. In particular, we assume:

$$A_{ijkl} \in C^1(\overline{\Omega}) \quad (3.27)$$

It is to be noted that (3.27) implies that the elasticity tensor is C^1 continuous on the domain, including the boundary, and across the fault, including the dislocation. The C^1 continuity on the subdomains $\overline{\Omega}^+$ and $\overline{\Omega}^-$ ensures that tractions are well defined. The C^1 continuity across the fault is required to establish global convergence of the WSM approximation in the \mathbf{L}^2 -norm and local convergence in the \mathbf{H}^1 -norm; see Sections 3.3.2 and 3.3.3 below. Let us note that in tectonophysics the elasticity tensor is generally assumed to be uniformly constant in the domain.

Let $\omega \subset \mathbb{R}^N$ denote an arbitrary connected domain with Lipschitz boundary. In particular, recalling the partition of Ω into the complementary subsets Ω^\pm , we envisage $\omega \in \{\Omega^+, \Omega^-\}$. We denote by $\widehat{\gamma}$ the restriction of a function in $\mathbf{C}^1(\overline{\omega})$ to the boundary $\partial\omega$. By virtue of the density of $\mathbf{C}^1(\omega)$ in $\mathbf{H}^1(\omega)$, the operator can be extended to a linear continuous *trace operator*, denoted by γ , from $\mathbf{H}^1(\omega)$ into $\mathbf{L}^2(\partial\omega)$. The image of γ is denoted by $\mathbf{H}^{1/2}(\partial\omega)$. Considering a subset $\varkappa \subset \partial\omega$, we denote by $\gamma_\varkappa := (\gamma(\cdot))|_\varkappa$ the composition of the trace operator and the restriction to \varkappa . The image of γ_\varkappa restricted to the class of functions $\mathbf{H}_{0,\partial\omega \setminus \varkappa}^1(\omega)$ that vanish on $\partial\omega \setminus \varkappa$ is indicated by $\mathbf{H}_0^{1/2}(\varkappa)$:

$$\mathbf{H}_0^{1/2}(\varkappa) = \{\gamma_\varkappa u : u \in \mathbf{H}_{0,\partial\omega \setminus \varkappa}^1(\omega)\}. \quad (3.28)$$

The space $\mathbf{H}_0^{1/2}(\varkappa)$ can be endowed with the norm:

$$\|\lambda\|_{1/2,\varkappa} := \inf \{ \|u\|_{1,\omega} : u \in \mathbf{H}_{0,\partial\omega \setminus \varkappa}^1(\omega), \gamma_\varkappa u = \lambda \}. \quad (3.29)$$

with the obvious extension to $\mathbf{H}^{1/2}(\partial\omega)$. There exist continuous right inverses

$$\gamma^{-1} : \mathbf{H}^{1/2}(\partial\omega) \rightarrow \mathbf{H}^1(\omega), \quad \gamma_{\varkappa}^{-1} : \mathbf{H}_0^{1/2}(\varkappa) \rightarrow \mathbf{H}_{0,\partial\omega \setminus \varkappa}^1(\omega), \quad (3.30)$$

of γ and γ_{\varkappa} . Such a right inverse is called a *lifting* (or *lift*) of the trace. It is to be noted that lift operators are generally non-unique.

We denote by $\widehat{\sigma}_n : u \mapsto n \cdot \gamma(\sigma(u))$ the traction of a function in $\mathbf{C}^2(\overline{\omega})$ on $\partial\omega$, where n denotes the exterior unit normal vector on $\partial\omega$. We define

$$\mathbf{H}_{\text{div}\sigma}^1(\omega) := \{u \in \mathbf{H}^1(\omega) : \text{div } \sigma(u) \in \mathbf{L}^2(\omega)\}. \quad (3.31)$$

Applying index notation for transparency, the chain rule yields

$$\partial_j \sigma_{ij}(u) = (\partial_j A_{ijkl}) \epsilon_{kl}(u) + A_{ijkl} (\partial_j \epsilon_{kl}(u)) \quad (3.32)$$

Therefore, the condition $\text{div } \sigma(u) \in \mathbf{L}^2(\omega)$ provides a meaningful condition on u if $\partial_j A_{ijkl} \in L^\infty(\omega)$. For $\omega \in \{\Omega^+, \Omega^-\}$, this auxiliary condition on the elasticity tensor is satisfied under the standing assumption (3.27). The vector space $\mathbf{H}_{\text{div}\sigma}^1(\omega)$ is a Hilbert space under the inner-product associated with the norm $(\|\cdot\|_{1,\omega}^2 + \|\text{div } \sigma(\cdot)\|_\omega^2)^{1/2}$. The traction $\widehat{\sigma}_n$ can be extended to a bounded linear operator, denoted by σ_n , from $\mathbf{H}_{\text{div}\sigma}^1(\omega)$ into $\mathbf{H}^{-1/2}(\partial\omega) := [\mathbf{H}^{1/2}(\partial\omega)]'$, the dual space of $\mathbf{H}^{1/2}(\partial\omega)$. For each $u \in \mathbf{H}_{\text{div}\sigma}^1(\omega)$, the functional $\sigma_n(u)$ acts on functions in $\mathbf{H}^{1/2}(\partial\omega)$ by means of the following duality pairing:

$$\langle \sigma_n(u), \lambda \rangle = \int_\omega \text{div } \sigma(u) \cdot \gamma^{-1}(\lambda) + \int_\omega \sigma(u) : \nabla \gamma^{-1}(\lambda) \quad (3.33)$$

One may note that for functions in $\mathbf{C}^2(\overline{\omega})$, Equation (3.33) corresponds to a standard integration-by-parts identity. Continuity of the operator σ_n thus defined follows from the sequence of bounds:

$$\begin{aligned} |\langle \sigma_n(u), \lambda \rangle| &\leq \|\text{div } \sigma(u)\|_\omega \|\gamma^{-1}(\lambda)\|_\omega + \|\sigma(u)\|_\omega \|\gamma^{-1}(\lambda)\|_{1,\omega} \\ &\leq \left(\|\text{div } \sigma(u)\|_\omega^2 + \|\sigma(u)\|_\omega^2 \right)^{1/2} \left(\|\gamma^{-1}(\lambda)\|_\omega^2 + \|\gamma^{-1}(\lambda)\|_{1,\omega}^2 \right)^{1/2} \\ &\leq (1 + \bar{c}_A^2)^{1/2} (\|u\|_{1,\omega}^2 + \|\text{div } \sigma(u)\|_\omega^2)^{1/2} \|\gamma^{-1}(\lambda)\|_{1,\omega} \end{aligned} \quad (3.34)$$

and the continuity of the lifting of the trace from $\mathbf{H}^{1/2}(\partial\omega)$ into $\mathbf{H}^1(\omega)$. The dual space $[\mathbf{H}^{1/2}(\partial\omega)]'$ is a Banach space under the norm

$$\|v\|_{-1/2,\partial\omega} = \sup_{\lambda \in \mathbf{H}^{1/2}(\partial\omega)} \frac{\langle v, \lambda \rangle}{\|\lambda\|_{1/2,\partial\omega}}.$$

The restriction of the traction $(\widehat{\sigma}_n(\cdot))|_{\varkappa}$ to a subset $\varkappa \subset \partial\omega$ of the boundary can be extended to a bounded linear operator $\sigma_{n,\varkappa}$ from $\mathbf{H}_{\text{div}\sigma}^1(\omega)$ into $\mathbf{H}^{-1/2}(\varkappa) := [\mathbf{H}_0^{1/2}(\varkappa)]'$. The functional $\sigma_{n,\varkappa}(u)$ acts on functions in $\mathbf{H}_0^{1/2}(\varkappa)$ via the duality pairing:

$$\langle \sigma_{n,\varkappa}(u), \lambda \rangle = \int_\omega \text{div } \sigma(u) \cdot \gamma_{\varkappa}^{-1}(\lambda) + \int_\omega \sigma(u) : \nabla \gamma_{\varkappa}^{-1}(\lambda) \quad (3.35)$$

Continuity of the operator $\sigma_{n,\varkappa}$ thus defined follows in a similar manner as in (3.34).

For non-rupturing faults, it holds that $b \in \mathbf{H}_0^{1/2}(\Gamma)$ and the above definitions apply without revisions. The slip discontinuity (3.5b) and the traction discontinuity (3.5c) are then to be understood in the sense of traces and tractions outlined above. However, for rupturing faults, i.e., if the dislocation intersects with the boundary of the domain, then $b \notin \mathbf{H}_0^{1/2}(\Gamma)$, and further consideration is required. We can accommodate b in the space:

$$\widetilde{\mathbf{H}}^{1/2}(\Gamma) := \{\gamma_\Gamma u : u \in \mathbf{H}^1(\omega)\} \quad (3.36)$$

with $\omega \in \{\Omega^+, \Omega^-\}$. The space $\widetilde{\mathbf{H}}^{1/2}(\Gamma)$ is a Banach space under the norm

$$\|\widetilde{\lambda}\|_{1/2,\Gamma} = \inf \{\|u\|_{1,\omega} : u \in \mathbf{H}^1(\omega), \gamma_\Gamma u = \lambda\} \quad (3.37)$$

The principal complication pertaining to rupturing faults, is that the corresponding slip vectors cannot be lifted into $\mathbf{H}_{0,\partial\omega \setminus \Gamma}^1(\omega)$, as traces of functions in $\mathbf{H}^1(\omega)$ do not admit the discontinuity that would otherwise arise at the intersection of $\bar{\Gamma}$ and $\partial\omega \setminus \bar{\Gamma}$. Hence, we cannot use (3.35) to define an extension of the restriction of the traction, $(\widehat{\sigma}_n(\cdot))|_\Gamma$, to a bounded linear operator from $\mathbf{H}_{\text{div}\sigma}^1(\omega)$ into $[\widetilde{\mathbf{H}}^{1/2}(\Gamma)]'$. However, there exists a continuous right inverse $\widetilde{\gamma}_\Gamma^{-1} : \widetilde{\mathbf{H}}^{1/2}(\Gamma) \rightarrow \mathbf{H}^1(\omega)$ of the operator γ_Γ , for instance,

$$\widetilde{\gamma}_\Gamma^{-1}(\lambda) = \arg \inf \left\{ \|u\|_{1,\omega} : u \in \mathbf{H}^1(\omega), \gamma_\Gamma u = \lambda \right\} \quad (3.38)$$

Let us note that the image of the lift operator $\widetilde{\gamma}_\Gamma^{-1}$ corresponds to a harmonic function subject to inhomogeneous Dirichlet conditions on Γ with data λ and a homogeneous Neumann condition on $\partial\omega \setminus \Gamma$. The lift operator $\widetilde{\gamma}_\Gamma^{-1}$ can be modified to include homogeneous Dirichlet conditions on $\mathcal{D} \subset \partial\omega \setminus \Gamma$ in the codomain, if necessary. The lift operator $\widetilde{\gamma}_\Gamma^{-1}$ enables us to extend $(\widehat{\sigma}_n)|_\Gamma$ to a continuous linear operator:

$$\widetilde{\sigma}_{n,\Gamma} : \left\{ u \in \mathbf{H}_{\text{div}\sigma}^1(\omega) : \|\sigma_{n,\partial\omega \setminus \Gamma}(u)\|_{-1/2,\partial\omega \setminus \Gamma} = 0 \right\} \rightarrow [\widetilde{\mathbf{H}}^{1/2}(\Gamma)]' \quad (3.39)$$

The functional $\widetilde{\sigma}_{n,\Gamma}(u)$ acts on functions in $\widetilde{\mathbf{H}}^{1/2}(\Gamma)$ via the duality pairing:

$$\langle \widetilde{\sigma}_{n,\Gamma}(u), \lambda \rangle = \int_\omega \text{div } \sigma(u) \cdot \widetilde{\gamma}_\Gamma^{-1}(\lambda) + \int_\omega \sigma(u) : \nabla \widetilde{\gamma}_\Gamma^{-1}(\lambda) \quad (3.40)$$

Essentially, in (3.39) and (3.40), we have defined the extension $\widetilde{\sigma}_{n,\Gamma}$ of the restriction of the traction to the dislocation, $(\widehat{\sigma}_n(\cdot))|_\Gamma$, by restricting the domain of the extended operator to functions for which the traction vanishes on $\partial\omega \setminus \Gamma$. This restriction in the definition is consistent with the standing assumption that an intersection of the dislocation with the boundary of the domain can only occur at Neumann boundaries.

In the analysis below, we restrict ourselves to non-rupturing faults. The analysis in Section 3.3.2 however extends to rupturing faults by replacing the spaces and trace and traction operators for non-rupturing faults with those for rupturing faults.

3.3.2 Global approximation properties of WSM

To assess the global approximation properties of WSM, we first construct an upper bound on the functional $(b, \{\sigma_\nu(\cdot)\})_\Gamma : \mathbf{V}_h \rightarrow \mathbb{R}$ in the right member of the WSM

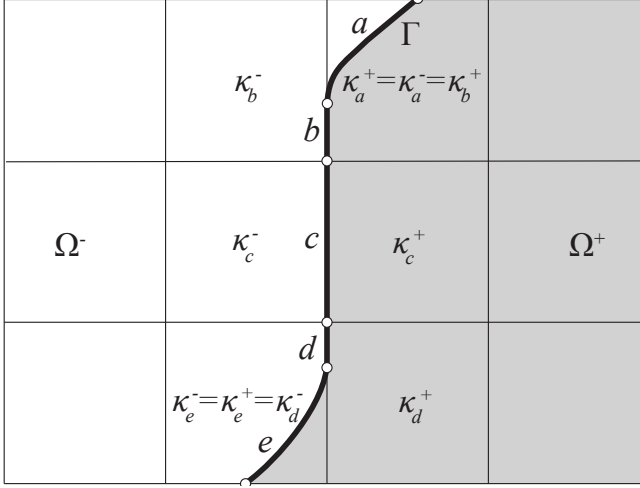


Figure 3.3 – Illustration of the covering of the fault Γ by segments $\mathcal{S}_h = \{a, b, c, d, e\}$, and the corresponding elements $\{\kappa_{(\cdot)}^+, \kappa_{(\cdot)}^-\}$. Because b, c, d coincide with element boundaries, $\kappa_{(\cdot)}^+ \neq \kappa_{(\cdot)}^-$ for these segments. The segments a, e are interior to elements and, accordingly $\kappa_{(\cdot)}^+ = \kappa_{(\cdot)}^-$ for these segments.

formulation. It is to be noted that, in general, $\mathbf{V}_h \not\subset \mathbf{H}_{\text{div } \sigma}^1(\Omega)$. Hence, the functional $(b, \{\sigma_\nu(\cdot)\})_\Gamma : \mathbf{V}_h \rightarrow \mathbb{R}$ does not admit an interpretation as a duality pairing according to (3.35). Because \mathbf{V}_h is piecewise polynomial, however, an upper bound can be constructed on the basis of inverse and trace inequalities. We refer to [57] for a comprehensive treatment of this subject. Inverse and trace inequalities can generally be derived under suitable (sufficient) regularity conditions on the finite-element mesh; see [57, Chap. 1]. A detailed treatment of the conditions underlying inverse and trace inequalities is beyond the scope of this work. Instead, we shall directly assume that a suitable discrete trace inequality holds. To formulate the assumption, for a given sequence of partitions \mathcal{T}_h and for each $h \in \mathcal{H}$, we denote by \mathcal{S}_h a dense cover of the fault by means of open intersections of the fault with element interiors or with element faces, i.e.,

$$\begin{aligned} \mathcal{S}_h = & \{s \subset \Gamma : s = \Gamma \cap \kappa \neq \emptyset \text{ for some } \kappa \in \mathcal{T}_h\} \cup \\ & \{s \subset \Gamma : s = \text{int}(\Gamma \cap \partial\kappa_0 \cap \partial\kappa_1) \neq \emptyset \text{ for some } \kappa_0, \kappa_1 \in \mathcal{T}_h, \kappa_0 \neq \kappa_1\} \end{aligned} \quad (3.41)$$

See Figure 3.3 for an illustration. The separate treatment of element boundaries serves to ensure that subsets of Γ that coincide with element faces are separately included in \mathcal{S}_h . To each segment $s \in \mathcal{S}_h$, we associate a pair of contiguous elements $\{\kappa_s^+, \kappa_s^-\}$ such that $s \subset \kappa^\pm \cup \partial\kappa^\pm$ and $\kappa^\pm \cap \Omega^\pm \neq \emptyset$. If $s \subset \partial\kappa$ (resp. $s \subset \kappa$) for some κ then κ^+ and κ^- will be distinct (resp. identical). We assume that the following discrete trace inequality holds for all $h \in \mathcal{H}$, all $s \in \mathcal{S}_h$ and all element-wise polynomial functions v_h of degree at most p :

$$(\text{diam}(\kappa))^{1/2} \|v_h\|_s \leq C_\Gamma \|v_h\|_\kappa, \quad (3.42)$$

for both $\kappa \in \{\kappa_s^+, \kappa_s^-\}$, for some $C_\Gamma > 0$ independent of h ; cf. [57, Lemma 1.46]. The constant C_Γ is allowed to increase with the polynomial order p .

Lemma 1 (Continuity of the WSM linear form). *Consider a manifold $\Gamma \subset \Omega \subset \mathbb{R}^N$, a slip distribution $b \in \mathbf{H}_0^{1/2}(\Gamma)$ and a sequence of partitions \mathcal{T}_h such that for all $h \in \mathcal{H}$, the discrete trace inequality (3.42) holds for all $s \in \mathcal{S}_h$ and all element-wise polynomial*

functions on \mathcal{T}_h and

$$\|b\|_{\mathcal{T}_h, \Gamma} := \left(\sum_{s \in \mathcal{S}_h} h_s^{-1} \|b\|_s^2 \right)^{1/2} < \infty \quad (3.43)$$

with h_s the harmonic average of the diameters of the elements adjacent to s ,

$$\frac{1}{h_s} = \frac{1}{\text{diam}(\kappa_s^+)} + \frac{1}{\text{diam}(\kappa_s^-)} \quad (3.44)$$

Then for all $h \in \mathcal{H}$, the linear form $(b, \{\sigma_\nu(\cdot)\})_\Gamma : \mathbf{V}_h \rightarrow \mathbb{R}$ is continuous and

$$|(b, \{\sigma_\nu(v)\})_\Gamma| \leq 2^{-1/2} \bar{c}_A C_\Gamma M^{1/2} \|b\|_{\mathcal{T}_h, \Gamma} \|v\|_{1, \Omega} \quad (3.45)$$

with \bar{c}_A the continuity constant of the elasticity tensor in (3.3), C_Γ the constant in the discrete trace inequality (3.42) and M the maximum multiplicity of the multiset $\{\kappa \in \{\kappa_s^+, \kappa_s^-\} : s \in \mathcal{S}_h\}$.

Remark. The maximum multiplicity M in Lemma 1 indicates the maximum number of occurrences of any one element in connection to any segment $s \in \mathcal{S}_h$ as a member of the set $\{\kappa_s^+, \kappa_s^-\}$. For instance, in Figure 3.3, the element $\kappa_a^+ = \kappa_a^- = \kappa_b^+$ has multiplicity 3. One can infer that M is bounded by the maximum number of faces of any element in the mesh, increased by 2 for interior segments.

Proof. We first separate the integral on Γ into a sum of contributions from the segments and apply (3.3) and the Cauchy-Schwarz inequality to obtain

$$\begin{aligned} |(b, \{\sigma_\nu(v_h)\})_\Gamma| &= \left| 2^{-1} \sum_{s \in \mathcal{S}_h} (b, \nu \cdot \gamma_\Gamma^+(\sigma(v_h)) + \nu \cdot \gamma_\Gamma^-(\sigma(v_h)))_s \right| \\ &\leq 2^{-1} \bar{c}_A \sum_{s \in \mathcal{S}_h} \|b\|_s \left(\|\gamma_\Gamma^+(\nabla v_h)\|_s + \|\gamma_\Gamma^-(\nabla v_h)\|_s \right) \end{aligned} \quad (3.46)$$

where $\gamma_\Gamma^\pm(\cdot)$ denotes the trace of (\cdot) from within Ω^\pm . Noting that $v_h \in \mathbf{V}_h$ is element-wise polynomial, we deduce from the discrete trace inequality (3.42), the arithmetic-geometric mean inequality and (3.44):

$$\begin{aligned} |(b, \{\sigma_\nu(v_h)\})_\Gamma| &\leq 2^{-1} \bar{c}_A C_\Gamma \sum_{s \in \mathcal{S}_h} h_s^{-1/2} \|b\|_s (\|\nabla v_h\|_{\kappa_s^+} + \|\nabla v_h\|_{\kappa_s^-}) \\ &\leq 2^{-1/2} \bar{c}_A C_\Gamma \left(\sum_{s \in \mathcal{S}_h} h_s^{-1} \|b\|_s^2 \right)^{1/2} \left(\sum_{s \in \mathcal{S}_h} \|\nabla v_h\|_{\kappa_s^+}^2 + \|\nabla v_h\|_{\kappa_s^-}^2 \right)^{1/2} \\ &\leq 2^{-1/2} \bar{c}_A C_\Gamma M^{1/2} \|b\|_{\mathcal{T}_h, \Gamma} \|v_h\|_{1, \Omega} \end{aligned} \quad (3.47)$$

□

To determine the global approximation properties of the WSM formulation in the \mathbf{H}^1 -norm, we note that the WSM formulation is *inconsistent*: The solution of the weak formulation of the Volterra dislocation problem (3.11) violates the WSM weak form (3.23) by $(b, \{\sigma_\nu(v)\})_\Gamma$, for all $v \in \mathbf{V}_h$. The second Strang lemma [26, Lemma 2.25] then provides the following characterization of the global approximation properties of WSM:

Theorem 2 (Global approximation properties of WSM in the \mathbf{H}^1 -norm). *Assume that the conditions of Lemma 1 hold. Let $u \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$ denote the solution to the Volterra dislocation problem (3.11) and let $u_h \in \mathbf{V}_h$ denote its WSM approximation according to (3.23). It holds that*

$$\|u - u_h\|_{1,\Omega \setminus \Gamma} \leq (1 + C_P \underline{c}_A^{-1} \bar{c}_A) \inf_{v_h \in \mathbf{V}_h} \|u - v_h\|_{1,\Omega \setminus \Gamma} + C_P \underline{c}_A^{-1} \sup_{v_h \in \mathbf{V}_h \setminus \{0\}} \frac{|(b, \{\sigma_\nu(v_h)\})_\Gamma|}{\|v_h\|_{1,\Omega}} \quad (3.48)$$

Proof. We first recall that the bilinear form $a_\Gamma(\cdot, \cdot)$ is bounded on $\mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma) \times \mathbf{V}_h$ and coercive on $\mathbf{V}_h \times \mathbf{V}_h$; see (3.14) and (3.15). For arbitrary $v_h \in \mathbf{V}_h$, we have the chain of inequalities:

$$\begin{aligned} \|u_h - v_h\|_{1,\Omega \setminus \Gamma}^2 &\leq C_P \underline{c}_A^{-1} |a_\Gamma(u_h - v_h, u_h - v_h)| \\ &= C_P \underline{c}_A^{-1} |a_\Gamma(u_h - u, u_h - v_h) + a_\Gamma(u - v_h, u_h - v_h)| \\ &\leq C_P \underline{c}_A^{-1} (|(b, \{\sigma_\nu(u_h - v_h)\})_\Gamma| + \bar{c}_A |u - v_h|_{1,\Omega \setminus \Gamma} |u_h - v_h|_{1,\Omega \setminus \Gamma}) \end{aligned} \quad (3.49)$$

which leads to

$$\|u_h - v_h\|_{1,\Omega \setminus \Gamma} \leq C_P \underline{c}_A^{-1} \sup_{w_h \in \mathbf{V}_h \setminus \{0\}} \frac{|(b, \{\sigma_\nu(w_h)\})_\Gamma|}{\|w_h\|_{1,\Omega}} + C_P \underline{c}_A^{-1} \bar{c}_A \|u - v_h\|_{1,\Omega \setminus \Gamma} \quad (3.50)$$

The bound (3.48) then follows from the triangle inequality. \square

For quasi-uniform meshes, Lemma 1 and Theorem 2 lead to a simple asymptotic characterization of the global approximation properties of WSM in the \mathbf{H}^1 -norm. This characterization is detailed in the following corollary:

Corollary 3. *Assume that the conditions of Lemma 1 hold and that the sequence of partitions \mathcal{T}_h is quasi-uniform with respect to the mesh parameter, i.e., for all $h \in \mathcal{H}$ there exist constants $\bar{C} > 0$ and $\underline{C} > 0$ independent of h such that $\underline{C}h \leq \text{diam}(\kappa) \leq \bar{C}h$ for all $\kappa \in \mathcal{T}_h$. It then holds that*

$$\|u - u_h\|_{1,\Omega \setminus \Gamma} \leq \mathcal{C} h^{-1/2} \quad (3.51)$$

as $h \rightarrow 0$.

Proof. Subject to the quasi-uniformity condition on the sequence of partitions, we have

$$\|b\|_{\mathcal{T}_h, \Gamma} = \left(\sum_{s \in \mathcal{S}_h} h_s^{-1} \|b\|_s^2 \right)^{1/2} \leq 2\underline{C}^{-1/2} h^{-1/2} \left(\sum_{s \in \mathcal{S}_h} \|b\|_s^2 \right)^{1/2} = 2\underline{C}^{-1/2} h^{-1/2} \|b\|_{\Gamma} \quad (3.52)$$

It then follows from (3.47) and (3.48) that

$$\begin{aligned} \|u - u_h\|_{1,\Omega \setminus \Gamma} &\leq (1 + C_P \underline{c}_A^{-1} \bar{c}_A) \inf_{v_h \in \mathbf{V}_h} \|u - v_h\|_{1,\Omega \setminus \Gamma} + 2^{-1/2} C_P \underline{c}_A^{-1} \bar{c}_A C_\Gamma M^{1/2} \|b\|_{\mathcal{T}_h, \Gamma} \\ &\leq \mathcal{C} + \mathcal{C} h^{-1/2} \end{aligned} \quad (3.53)$$

and the assertion follows in the limit $h \rightarrow 0$. \square

The potential divergence of the bound in Corollary 3 does of course not immediately imply that the error in the WSM approximation itself increases as $h \rightarrow 0$. However, Theorem 4 below asserts that $\|u - u_h\|_\Omega$ vanishes as $h \rightarrow 0$. From this result, the continuity of u_h and the discontinuity of u , one can infer that, indeed, $\|u - u_h\|_{1,\Omega \setminus \Gamma}$ must diverge as $h \rightarrow 0$.

Theorem 2 conveys that for each $h > 0$, it holds that $u - u_h \in \mathbf{L}^2(\Omega)$. Based on the Aubin-Nitsche Lemma, we can therefore construct an estimate of the error in the WSM approximation in the \mathbf{L}^2 -norm. The estimate is formulated in Theorem 4 below. Some auxiliary conditions on the domain Ω are required, as specified in the premises of the theorem.

Theorem 4 (Global approximation properties of WSM in the \mathbf{L}^2 -norm). *Assume that the conditions of Lemma 1 hold. In addition, assume that Ω is convex or of class C^2 . Let u denote the solution to (3.11) and let $u_h \in \mathbf{V}_h$ denote the WSM approximation according to (3.23). Let \mathbf{T}_h denote the intersection of $\gamma_\Gamma(\mathbf{V}_h)$ with $\mathbf{H}_0^{1/2}(\Gamma)$. It holds that*

$$\begin{aligned} |(u - u_h, \varphi)| &\leq \mathcal{C} \left(\|u - u_h\|_{1,\Omega \setminus \Gamma} + \|b\|_{\mathcal{T}_h, \Gamma} \right) \inf_{v_h \in \mathbf{V}_h} |z_\varphi - v_h|_{1,\Omega} \\ &+ \mathcal{C} \inf_{\lambda_h \in \mathbf{T}_h} \left(\|\varphi\|_\Omega \|b - \lambda_h\|_\Gamma + (\|\lambda_h\|_{1/2,\Gamma} + \|\lambda_h\|_{\mathcal{T}_h, \Gamma}) \inf_{w_h \in \mathbf{V}_h} |z_\varphi - w_h|_{1,\Omega \setminus \Gamma} \right) \end{aligned} \quad (3.54)$$

for arbitrary $\varphi \in \mathbf{L}^2(\Omega)$ and a corresponding $z_\varphi \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ such that $\|z_\varphi\|_{2,\Omega} \leq \mathcal{C} \|\varphi\|_\Omega$, with $\|\cdot\|_{\mathcal{T}_h, \Gamma}$ defined by (3.43).

Proof. For arbitrary $\varphi \in \mathbf{L}^2(\Omega)$, consider the dual problem (in the sense of distributions):

$$-\operatorname{div} \sigma(z_\varphi) = \varphi \quad \text{in } \Omega \quad (3.55a)$$

$$z_\varphi = 0 \quad \text{on } \mathcal{D} \quad (3.55b)$$

$$\sigma_n(z_\varphi) = 0 \quad \text{on } \mathcal{N} \quad (3.55c)$$

or, equivalently, in weak form:

$$z \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega) : \quad a_\Gamma(v, z_\varphi) = (v, \varphi)_\Omega \quad \forall v \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega) \quad (3.56)$$

By virtue of the smoothness conditions on the elasticity tensor (3.27), the conditions on the domain, and the smoothness of the (homogeneous) boundary data in (3.55b) and (3.55c), the dual problem (3.55) possesses an elliptic-regularity property (see, e.g., [27, 11]). The regularity property implies that the dual solution resides in $\mathbf{H}^2(\Omega) \cap \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ and satisfies the estimate $\|z_\varphi\|_{2,\Omega} \leq \mathcal{C} \|\varphi\|_\Omega$.

Denoting by $\ell_b \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$ a suitable lift of b such that $[\![\ell_b]\!] = b$, it holds that $u - \ell_b \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$. Moreover, the WSM approximation u_h resides in $\mathbf{H}_{0,\mathcal{D}}^1(\Omega)$. The dual problem (3.56) therefore gives:

$$\begin{aligned} (u - u_h, \varphi)_\Omega &= (u - \ell_b, \varphi)_\Omega + (\ell_b, \varphi)_\Omega - (u_h, \varphi)_\Omega \\ &= a_\Gamma(u - \ell_b, z_\varphi) + (\ell_b, \varphi)_\Omega - a_\Gamma(u_h, z_\varphi) \\ &= a_\Gamma(u, z_\varphi) - a_\Gamma(u_h, z_\varphi) - (a_\Gamma(\ell_b, z_\varphi) - (\ell_b, \varphi)_\Omega) \end{aligned} \quad (3.57)$$

By means of (3.33), the term in parenthesis in the ultimate expression can be identified as the weak formulation of $(b, \{\sigma_\nu(z_\varphi)\})_\Gamma$. However, for $z_\varphi \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$, the trace theorem asserts that $\sigma_\nu(z_\varphi) \in \mathbf{L}^2(\Gamma)$ and $(b, \{\sigma_\nu(z_\varphi)\})_\Gamma$ coincides with its extension to a duality pairing. From the weak form of the Volterra dislocation problem (3.11) it moreover follows that $a_\Gamma(u, z_\varphi) = l(z_\varphi)$. The Galerkin-orthogonality property of the WSM approximation (3.23) then yields the following identities:

$$\begin{aligned} (u - u_h, \varphi)_\Omega &= l(z_\varphi - v_h) - a_\Gamma(u_h, z_\varphi - v_h) + (b, \{\sigma_\nu(v_h)\})_\Gamma - (b, \{\sigma_\nu(z_\varphi)\})_\Gamma \\ &= a_\Gamma(u - u_h, z_\varphi - v_h) - (b, \{\sigma_\nu(z_\varphi)\} - \{\sigma_\nu(v_h)\})_\Gamma \end{aligned} \quad (3.58)$$

which are valid for all $v_h \in \mathbf{V}_h$.

For the first term in the final expression in (3.58) we can construct an appropriate bound without further digression. However, the second term, pertaining to the difference between the average traction of $z \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_{0,\mathcal{D}}^1(\Omega)$ and the average of the direct evaluation of the traction of the finite-element function $v_h \in \mathbf{V}_h$, is more difficult to estimate. The essential complication in constructing an estimate, is that the direct evaluation of the traction of the Galerkin finite-element approximation of (3.56) in \mathbf{V}_h does not coincide with the weak formulation, because $\mathbf{V}_h \notin \mathbf{H}_{\text{div } \sigma}^1(\Omega)$; see also [14]. Approximation results are available for the weak formulation of the traction (see, for instance, [13]) but, to our knowledge, not for the direct formulation.

To estimate the second term in (3.58), we will use an auxiliary Nitsche-type approximation [54] to the dual problem. We consider the broken approximation spaces $\hat{\mathbf{V}}_h$ according to (3.19). Moreover, we define $\hat{\mathbf{V}}(h) := (\mathbf{H}^2(\Omega \setminus \Gamma) \cap \mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)) + \hat{\mathbf{V}}_h$. We equip $\hat{\mathbf{V}}(h)$ with the mesh-dependent inner product

$$(u, v)_{\hat{\mathbf{V}}(h)} = a_\Gamma(u, v) + \zeta \sum_{s \in \mathcal{S}_h} h_s^{-1} (\llbracket u \rrbracket, \llbracket v \rrbracket)_s \quad (3.59)$$

and the corresponding norm $\|\cdot\|_{\hat{\mathbf{V}}(h)}$. We note that the embedding of $\mathbf{H}^2(\Omega \setminus \Gamma)$ into $\hat{\mathbf{V}}(h)$ is continuous, i.e., $\|v\|_{\hat{\mathbf{V}}(h)} \leq \mathcal{C} \|v\|_{2,\Omega \setminus \Gamma}$ for all $v \in \mathbf{H}^2(\Omega \setminus \Gamma)$. We define the bilinear operator $\hat{a}_\Gamma : \mathbf{V}(h) \times \mathbf{V}_h \rightarrow \mathbb{R}$ according to:

$$\hat{a}_\Gamma(u, v) = a_\Gamma(u, v) - (\llbracket v \rrbracket, \{\sigma_\nu(u)\})_\Gamma + \theta \sum_{s \in \mathcal{S}_h} (\llbracket u \rrbracket, \llbracket v \rrbracket)_s \quad (3.60)$$

with θ a suitable constant. For a suitable choice of the constants ζ and θ , the bilinear form in the left member of (3.61) is continuous and coercive on $\hat{\mathbf{V}}_h \times \hat{\mathbf{V}}_h$ and $\hat{a}_\Gamma(u, \cdot) : \hat{\mathbf{V}}_h \rightarrow \mathbb{R}$ represents a continuous linear functional for all $u \in \mathbf{V}(h)$. Let $z_{\varphi_h} \in \hat{\mathbf{V}}_h$ denote the solution to the following Nitsche-type projection problem:

$$z_{\varphi_h} \in \hat{\mathbf{V}}_h : \quad \hat{a}_\Gamma(z_{\varphi_h}, w_h) = \hat{a}_\Gamma(z_\varphi, w_h) \quad \forall w_h \in \hat{\mathbf{V}}_h \quad (3.61)$$

By virtue of the continuity and coercivity of \hat{a}_Γ on $\hat{\mathbf{V}}_h \times \hat{\mathbf{V}}_h$ and the continuity of $\hat{a}_\Gamma(z_\varphi, \cdot)$ on $\hat{\mathbf{V}}_h$, the projection problem (3.61) defines a unique element $z_{\varphi_h} \in \hat{\mathbf{V}}_h$ which satisfies

$$\|z_\varphi - z_{\varphi_h}\|_{\hat{\mathbf{V}}(h)} \leq \mathcal{C} \inf_{w_h \in \hat{\mathbf{V}}_h} \|z_\varphi - w_h\|_{\hat{\mathbf{V}}(h)} \leq \mathcal{C} \inf_{w_h \in \mathbf{V}_h} |z_\varphi - w_h|_{1,\Omega} \quad (3.62)$$

The first inequality in (3.62) is a straightforward consequence of Céa's lemma. The second inequality follows from $\mathbf{V}_h \subset \hat{\mathbf{V}}_h$, the continuity of functions in \mathbf{V}_h and (3.14a).

By adding a suitable partition of zero to the second term in (3.58) and applying the triangle inequality, we obtain:

$$|(b, \{\sigma_\nu(z_\varphi)\} - \{\sigma_\nu(v_h)\})_\Gamma| \leq |(b, \{\sigma_\nu(z_\varphi - z_{\varphi_h})\})_\Gamma| + |(b, \{\sigma_\nu(v_h - z_{\varphi_h})\})_\Gamma| \quad (3.63)$$

For the first term in the right-member of (3.63), we derive from (3.61) and (3.62):

$$\begin{aligned} & |(b, \{\sigma_\nu(z_\varphi - z_{\varphi_h})\})_\Gamma| = \\ & |(b - \llbracket w_h \rrbracket, \{\sigma_\nu(z_\varphi - z_{\varphi_h})\})_\Gamma + a_\Gamma(w_h, z_\varphi - z_{\varphi_h}) - \theta \sum_{s \in S_h} h_s^{-1} (\llbracket w_h \rrbracket, \llbracket z_{\varphi_h} \rrbracket)_s| \\ & \leq \|b - \llbracket w_h \rrbracket\|_\Gamma (\|\{\sigma_\nu(z_\varphi)\}\|_\Gamma + \|\{\sigma_\nu(z_{\varphi_h})\}\|_\Gamma) \\ & \quad + \bar{c}_A |w_h|_{1, \Omega \setminus \Gamma} |z_\varphi - z_{\varphi_h}|_{1, \Omega \setminus \Gamma} + \theta \|\llbracket w_h \rrbracket\|_{\mathcal{T}_h, \Gamma} \|\llbracket z_{\varphi_h} \rrbracket\|_{\mathcal{T}_h, \Gamma} \end{aligned} \quad (3.64)$$

for all $w_h \in \hat{\mathbf{V}}_h$. The trace theorem implies $\|\{\sigma_\nu(z_\varphi)\}\|_\Gamma \leq \mathcal{C} \|z_\varphi\|_{2, \Omega} \leq \mathcal{C} \|\varphi\|_\Omega$. The continuity of \hat{a}_Γ and a_Γ implies that $\|\{\sigma_\nu(z_{\varphi_h})\}\|_\Gamma \leq \mathcal{C} \|z_{\varphi_h}\|_{\hat{\mathbf{V}}(h)}$. Moreover, by virtue of (3.62) and the continuity of the embedding of $\mathbf{H}^2(\Omega)$ into $\hat{\mathbf{V}}(h)$ it holds that $\|z_{\varphi_h}\|_{\hat{\mathbf{V}}(h)} \leq \mathcal{C} \|z_\varphi\|_{\hat{\mathbf{V}}(h)} \leq \mathcal{C} \|z_\varphi\|_{2, \Omega} \leq \mathcal{C} \|\varphi\|_\Omega$. Hence, we have $\|\{\sigma_\nu(z_\varphi)\}\|_\Gamma + \|\{\sigma_\nu(z_{\varphi_h})\}\|_\Gamma \leq \mathcal{C} \|\varphi\|_\Omega$. Noting that $\mathbf{T}_h = \gamma_\Gamma(\mathbf{V}_h) \cap \mathbf{H}_0^{1/2}(\Gamma)$ coincides with $\{\llbracket w_h \rrbracket : w_h \in \hat{\mathbf{V}}_h\}$, for all $\lambda_h \in \mathbf{T}_h$ there exists a $w_h \in \hat{\mathbf{V}}_h$ such that $\llbracket w_h \rrbracket = \lambda_h$ and $|w_h|_{1, \Omega \setminus \Gamma} \leq \mathcal{C} \|\lambda\|_{1/2, \Gamma}$. From (3.62)–(3.64) we deduce:

$$\begin{aligned} |(b, \{\sigma_\nu(z_\varphi - z_{\varphi_h})\})_\Gamma| & \leq \mathcal{C} (\|\varphi\|_\Omega \|b - \lambda_h\|_\Gamma \\ & \quad + (\|\lambda_h\|_{1/2, \Gamma} + \|\lambda_h\|_{\mathcal{T}_h, \Gamma}) \inf_{w_h \in \hat{\mathbf{V}}_h} |z_\varphi - w_h|_{1, \Omega \setminus \Gamma}) \end{aligned} \quad (3.65)$$

for all $\lambda_h \in \mathbf{T}_h$, with $\|\cdot\|_{\mathcal{T}_h, \Gamma}$ according to (3.43). For the second term in the right-member of (3.63) we deduce from (3.47), the triangle inequality and (3.62):

$$\begin{aligned} |(b, \{\sigma_\nu(v_h - z_{\varphi_h})\})_\Gamma| & \leq 2^{-1/2} \bar{c}_A C_\Gamma M^{1/2} \|b\|_{\mathcal{T}_h, \Gamma} |v_h - z_{\varphi_h}|_{1, \Omega \setminus \Gamma} \\ & \leq \mathcal{C} \|b\|_{\mathcal{T}_h, \Gamma} \left(|z_\varphi - v_h|_{1, \Omega} + \inf_{w_h \in \hat{\mathbf{V}}_h} |z_\varphi - w_h|_{1, \Omega} \right) \end{aligned} \quad (3.66)$$

with \mathcal{C} independent of h . Collecting the results in (3.58)–(3.66) and taking the infimum with respect to v_h and λ_h , one obtains the estimate in (3.54). \square

Under slightly stronger conditions on the regularity of the slip distribution b , we can derive from Theorem 4 a straightforward characterization of the asymptotic approximation properties of WSM in the \mathbf{L}^2 -norm for quasi-uniform meshes in the limit as $h \rightarrow 0$:

Corollary 5. *Assume that the conditions of Corollary 3 and Theorem 4 hold and, moreover, $b \in \mathbf{H}^1(\Gamma) \cap \mathbf{H}_0^{1/2}(\Gamma)$. Let u denote the solution to (3.11) and let $u_h \in \mathbf{V}_h$ denote the WSM approximation according to (3.23). It holds that*

$$\|u - u_h\|_\Omega \leq \mathcal{C} h^{1/2} \quad (3.67)$$

as $h \rightarrow 0$.

Proof. Based on the estimate $\|z_\varphi\|_{2,\Omega} \leq \mathcal{C}\|\varphi\|_\Omega$, it follows from standard interpolation theory in Sobolev spaces (see, for instance, [26, 12]) that $\inf_{v_h \in \mathbf{V}_h} \|z - v_h\|_{1,\Omega} \leq \mathcal{C}h\|\varphi\|_\Omega$. Similarly, for b in $\mathbf{H}^1(\Gamma) \cap \mathbf{H}_0^{1/2}(\Gamma)$, we have $\inf_{\lambda_h \in \mathbf{T}_h} \|b - \lambda_h\|_\Gamma \leq \mathcal{C}h\|b\|_{1,\Gamma}$. Setting $\lambda_h^* = \arg \inf_{\lambda_h \in \mathbf{T}_h} \|b - \lambda_h\|_\Gamma$, it holds that $\|\lambda_h^*\|_{1/2,\Gamma} \leq \mathcal{C}$ and $\|\lambda_h^*\|_{\mathcal{T}_h,\Gamma} \leq \mathcal{C}h^{-1/2}$. From (3.52) we obtain $\|b\|_{\mathcal{T}_h,\Gamma} \leq \mathcal{C}h^{-1/2}$ and, in turn, it follows from Theorem 2 that $\|u - u_h\|_{1,\Omega} \leq \mathcal{C}h^{-1/2}$ as $h \rightarrow 0$. Inserting the above estimates into (3.54), we derive:

$$\begin{aligned} \|u - u_h\|_\Omega &= \sup_{\varphi \in \mathbf{L}^2(\Omega) \setminus \{0\}} \frac{|(u - u_h, \varphi)_\Omega|}{\|\varphi\|_\Omega} \\ &\leq \sup_{\varphi \in \mathbf{L}^2(\Omega) \setminus \{0\}} \frac{1}{\|\varphi\|_\Omega} \left(\mathcal{C}h^{-1/2}h\|\varphi\|_\Omega + \mathcal{C}h\|\varphi\|_\Omega + (1 + h^{-1/2})h\|\varphi\|_\Omega \right) \end{aligned} \quad (3.68)$$

as $h \rightarrow 0$. The estimate in (3.67) then follows straightforwardly by combining terms. \square

Remark. The reinforced regularity condition $b \in \mathbf{H}^1(\Gamma) \cap \mathbf{H}_0^{1/2}(\Gamma)$ on the slip distribution ensures that $\inf_{\lambda_h \in \mathbf{T}_h} \|b - \lambda_h\|_\Gamma \leq \mathcal{C}h$ as $h \rightarrow 0$. Noting that the weaker estimate $\inf_{\lambda_h \in \mathbf{T}_h} \|b - \lambda_h\|_\Gamma \leq \mathcal{C}h^{1/2}$ suffices to obtain the result in Corollary 5, one is lead to question whether the reinforced regularity condition on b is actually necessary. If the condition is dismissed, however, an interpolation estimate in the fractional Sobolev space $\mathbf{H}_0^{1/2}(\Gamma)$ is required. Interpolation estimates in fractional Sobolev spaces are technical (see, for instance, [23]) and the particular result required here is to our knowledge not available.

It is noteworthy that the asymptotic convergence behavior according to Corollary 5 is consistent with the notion that the continuous WSM approximation incurs an $O(1)$ error, pointwise, in the $O(h)$ neighborhood of the discontinuity composed of the intersected elements as $h \rightarrow 0$. In particular, denoting by Γ_h the union of the elements for which the intersection of the fault with the closure of the element is non-empty, it holds that

$$\|1\|_{\Gamma_h} = (\text{meas}_N(\Gamma_h))^{1/2} = O(h^{1/2}) \quad (3.69)$$

as $h \rightarrow 0$, with $\text{meas}_N(\Gamma_h)$ the N -Lebesgue measure of Γ_h .

3.3.3 Local approximation properties

Next, we consider the local approximation properties of WSM, i.e., on Ω excluding a small neighborhood of the fault. To this end, we will relate the WSM approximation to the standard Galerkin approximation associated with a locally supported lift, outside the support of the lift.

For arbitrary $\varepsilon > 0$, let $\mathcal{X}^\varepsilon := \{x \in \Omega : \text{dist}(x, \mathcal{K}) < \varepsilon\}$ denote the open ε -neighborhood of the dislocation fault. We consider a lift $\ell_b \in \mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$ of b with compact support in \mathcal{X}^ε . Let $\bar{u}_h \in \mathbf{V}_h$ denote the Galerkin approximation of the continuous complement of the solution with respect to the jump lift ℓ_b according to (3.17). As a straightforward consequence of Céa's lemma, it follows that $\bar{u}_h + \ell_b$ is endowed with the quasi-optimal approximation property:

$$\|u - (\bar{u}_h + \ell_b)\|_{1,\Omega \setminus \Gamma} \leq \mathcal{C} \inf_{v_h \in \mathbf{V}_h} \|u - (v_h + \ell_b)\|_{1,\Omega \setminus \Gamma} \quad (3.70)$$

A meaningful characterization of the local approximation properties of WSM is therefore provided by an estimate for the deviation $\|\bar{u}_h + \ell_b - u_h\|_{1,\Omega \setminus \mathcal{K}^\varepsilon}$ between the WSM approximation u_h and the local-lift-based approximation $\bar{u}_h + \ell_b$. Note that $\mathcal{K}^\varepsilon \supset \text{supp}(\ell_b)$ implies that ℓ_b vanishes on $\Omega \setminus \mathcal{K}^\varepsilon$ and, hence, the estimate pertains to the deviation between $\bar{u}_h \in \mathbf{V}_h$ and $u_h \in \mathbf{V}_h$.

The characterization of the local approximation properties of WSM in the \mathbf{H}^1 -norm in Theorem 7 below is based on the interpolation of a particular extension of functions in \mathbf{V}_h onto \mathcal{K}^ε . Specifically, we consider an extension corresponding to the operator $E : \mathbf{V}_h \rightarrow \mathbf{H}^1(\Omega)$:

$$E(v_h) = v_h \quad \text{in } \Omega \setminus \mathcal{K}^\varepsilon \quad (3.71a)$$

$$-\text{div } \sigma(E(v_h)) = 0 \quad \text{in } \mathcal{K}^\varepsilon \quad (3.71b)$$

in the sense of distributions, and its optimal approximation in the norm defined by $a_\Gamma(\cdot, \cdot)$:

$$E_h(v_h) = \arg \min_{w_h \in v_h + \mathbf{V}_{h,\mathcal{K}^\varepsilon}} a_\Gamma(E(v_h) - w_h, E(v_h) - w_h) \quad (3.72)$$

where $\mathbf{V}_{h,\mathcal{K}^\varepsilon} := \{v_h \in \mathbf{V}_h : \text{supp}(v_h) \subseteq \mathcal{K}^\varepsilon\} \neq \emptyset$ denotes the class of approximation functions of which the support is confined to \mathcal{K}^ε . Note that $\mathbf{V}_{h,\mathcal{K}^\varepsilon} \neq \emptyset$ implies $\varepsilon \geq \mathcal{C}h$.

Equations (3.71a) and (3.72) imply that $E(v_h)$ and $E_h(v_h)$ coincide with v_h outside \mathcal{K}^ε . Inside \mathcal{K}^ε , the extension $E(v_h)$ is defined by the homogeneous elasticity problem (3.71b). From $E(v_h) \in \mathbf{H}^1(\Omega)$ it follows that $E(v_h)$ is continuous at $\partial\mathcal{K}^\varepsilon$ in the trace sense, which implies that (3.71b) is complemented by the Dirichlet boundary condition $E(v_h) = v_h$ on $\partial\mathcal{K}^\varepsilon$. By virtue of the smoothness condition on the elasticity tensor in (3.27), the extension operator according to (3.71) exhibits an interior regularity property on \mathcal{K}^ε . In particular, for all $v_h \in \mathbf{V}_h$ it holds that $E(v_h) \in \mathbf{H}_{\text{loc}}^2(\mathcal{K}^\varepsilon)$ (i.e., $\phi E(v_h) \in \mathbf{H}^2(\mathcal{K}^\varepsilon)$ for any $\phi \in C^\infty(\Omega)$ with compact support in \mathcal{K}^ε) and the following estimate holds for each open subset $K \Subset \mathcal{K}^\varepsilon$:

$$\|E(v_h)\|_{2,K} \leq \mathcal{C} \|E(v_h)\|_{\mathcal{K}^\varepsilon} \quad \forall v_h \in \mathbf{V}_h; \quad (3.73)$$

see, for instance, [27, §6.3.1] or [46, §2.3] for further details. It is important to note that estimate (3.73) in conjunction with the trace theorem implies that

$$\|\{\sigma_\nu(E(v_h))\}\|_{\mathcal{K}} \leq \mathcal{C} \|E(v_h)\|_{1,\mathcal{K}^\varepsilon} \quad \forall v_h \in \mathbf{V}_h, \quad (3.74)$$

for some $\mathcal{C} > 0$ independent of v_h , provided that there exists an open subset $K \Subset \mathcal{K}^\varepsilon$ such that $\mathcal{K} \subset K$. This provision implies that the dislocation must be properly contained in Ω . The proof of Theorem 7 involves an estimate of the difference between the average traction of $E(v_h)$ and its approximation $E_h(v_h)$ according to (3.72). In general, we can estimate $(b, \{\sigma_\nu(E(v_h) - E_h(v_h))\})_\Gamma$ in the same manner as in the proof of Theorem 4. The bound (3.74) in combination with the optimal-approximation property of $E_h(v_h)$ in (3.72) however suggests that in this case a sharper estimate can be established. The derivation of such a refined estimate is intricate, however, as $(b, \{\sigma_\nu(\cdot)\})_\Gamma$ is unbounded on $\mathbf{H}^1(\Omega)$ and the estimate involves the difference between $E(v_h)$ and $E_h(v_h)$, which reside in different subspaces of $\mathbf{H}^1(\Omega)$. We therefore formulate the refined estimate in the form of a conjecture:

Conjecture 6. Assume that there exist an $\varepsilon^* > 0$ and an open subset $K \Subset \mathcal{K}^\varepsilon$ such that $\varkappa \subset K$. Consider a sequence of approximation spaces $\mathbf{V}_\mathcal{H}$, the extension operator according to (3.71) and its approximation according to (3.72). For all $\varepsilon \geq \varepsilon^*$ and all $v_h \in \mathbf{V}_h$, it holds that

$$\left\| \{\sigma_\nu(E(v_h) - E_h(v_h))\} \right\|_\varkappa \leq \mathcal{C} \|E(v_h) - E_h(v_h)\|_{1, \mathcal{K}^\varepsilon} \quad (3.75)$$

for some constant $\mathcal{C} > 0$ independent of h .

Theorem 7 below presents a general characterization of the local approximation properties of WSM, independent of Conjecture 6, and a refinement, which is contingent on the conjecture.

Theorem 7 (Local approximation properties of WSM in the \mathbf{H}^1 -norm). Assume that there exist an $\varepsilon^* > 0$ and an open subset $K \Subset \mathcal{K}^\varepsilon$ such that $\varkappa \subset K$. For arbitrary $\varepsilon \geq \varepsilon^*$, let $\ell_b \in \mathbf{H}_{0,D}^1(\Omega \setminus \Gamma)$ denote a lift of b such that $\text{supp}(\ell_b) \subset \mathcal{K}^\varepsilon$. Given a sequence of approximation spaces $\mathbf{V}_\mathcal{H}$, for each $h \in \mathcal{H}$ let \bar{u}_h denote the approximation of the continuous complement corresponding to ℓ_b in (3.17) and let $u_h \in \mathbf{V}_h$ denote the WSM approximation according to (3.23). Consider the extension operator in (3.71) and its approximation in \mathbf{V}_h according to (3.72). It holds that

$$\begin{aligned} \|\bar{u}_h + \ell_b - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon}^2 &\leq \mathcal{C} \|\ell_b\|_{1, \mathcal{K}^\varepsilon \setminus \Gamma} \|E(\bar{u}_h - u_h) - E_h(\bar{u}_h - u_h)\|_{1, \mathcal{K}^\varepsilon} \\ &\quad + \mathcal{C} \left| (b, \{\sigma_\nu(E(\bar{u}_h - u_h) - E_h(\bar{u}_h - u_h))\})_\Gamma \right| \end{aligned} \quad (3.76)$$

for some constant $\mathcal{C} > 0$ independent of h and ε . If, in addition, Conjecture 6 holds, then

$$\|\bar{u}_h + \ell_b - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon} \leq \mathcal{C} (\|\ell_b\|_{1, \mathcal{K}^\varepsilon \setminus \Gamma} + \|b\|_\Gamma) \sup_{v_h \in \mathbf{V}_h} \inf_{w_h \in v_h + \mathbf{V}_{h, \mathcal{K}^\varepsilon}} \frac{\|E(v_h) - w_h\|_{1, \mathcal{K}^\varepsilon}}{\|v_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon}} \quad (3.77)$$

for some constant $\mathcal{C} > 0$ independent of h and ε .

Proof. We use the condensed notation $E_h := E_h(\bar{u}_h - u_h)$. Noting that (3.72) implies that E_h coincides with $\bar{u}_h - u_h$ on $\Omega \setminus \mathcal{K}^\varepsilon$ and that $\text{supp}(\ell_b) \subset \mathcal{K}^\varepsilon$, it holds that

$$\|\bar{u}_h + \ell_b - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon} = \|\bar{u}_h - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon} \leq \|E_h\|_{1, \Omega} \quad (3.78)$$

Using the Poincaré inequality (3.15) and the strong positivity of $a_\Gamma(\cdot, \cdot)$ according to (3.14b), we obtain the following bound:

$$\|E_h\|_{1, \Omega}^2 \leq (1 + C_p) \underline{c}_A^{-1} |a_\Gamma(E_h, E_h)| \quad (3.79)$$

By introducing a suitable partition of zero, we derive from the WSM formulation (3.23) and the lift-based Galerkin approximation (3.17):

$$\begin{aligned} a_\Gamma(E_h, E_h) &= a_\Gamma(\bar{u}_h - u_h, E_h) + a_\Gamma(E_h - (\bar{u}_h - u_h), E_h) \\ &= (b, \{\sigma_\nu(E_h)\})_\Gamma - a_\Gamma(\ell_b, E_h) + a_\Gamma(E_h - (\bar{u}_h - u_h), E_h) \end{aligned} \quad (3.80)$$

By virtue of (3.71b), $a_\Gamma(v_h, E(\bar{u}_h - u_h))$ vanishes for all $v_h \in \mathbf{V}_{h, \mathcal{K}^\varepsilon}$. The optimality condition associated with (3.72) therefore reduces to:

$$a_\Gamma(v_h, E_h) = 0 \quad \forall v_h \in \mathbf{V}_{h, \mathcal{K}^\varepsilon}. \quad (3.81)$$

Noting that $E_h - (\bar{u}_h - u_h) \in \mathbf{V}_{h, \mathcal{K}^\varepsilon}$, Equation (3.81) implies that the right-most term in the ultimate expression in (3.80) vanishes. By virtue of the interior regularity of $E(\bar{u}_h - u_h)$, the following identity holds:

$$a_\Gamma(\ell_b, E(\bar{u}_h - u_h)) = (b, \{\sigma_\nu(E(\bar{u}_h - u_h))\})_\Gamma \quad (3.82)$$

Let us note that (3.82) corresponds to the (admissible) identification of the duality pairing between $b \in \mathbf{H}_0^{1/2}(\mathcal{K})$ and $\{\sigma_\nu(E(\bar{u}_h - u_h))\} \in \mathbf{H}^{-1/2}(\mathcal{K})$ to an \mathbf{L}^2 inner product. From (3.80)–(3.82) and the triangle inequality we then deduce that:

$$|a_\Gamma(E_h, E_h)| \leq |a_\Gamma(\ell_b, E(\bar{u}_h - u_h) - E_h)| + |(b, \{\sigma_\nu(E(\bar{u}_h - u_h) - E_h)\})_\Gamma| \quad (3.83)$$

We recall that $E(\bar{u}_h - u_h) - E_h$ vanishes on $\Omega \setminus \mathcal{K}^\varepsilon$. Estimate (3.76) then follows straightforwardly from (3.78), (3.79), (3.83) and the continuity of $a_\Gamma(\cdot, \cdot)$ according to (3.14a).

To prove the auxiliary assertion (3.77), we note that $\|E(v_h) - E_h(v_h)\|_{1, \Omega}$ in fact depends only on the trace of v_h on $\partial\mathcal{K}^\varepsilon$ and, by the trace theorem, it holds that $\|E(v_h) - E_h(v_h)\|_{1, \Omega} \leq \mathcal{C}\|v_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon}$. Therefore, if Conjecture 6 holds, we obtain from (3.76):

$$\begin{aligned} \|\bar{u}_h - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon}^2 &\leq \mathcal{C}(\|\ell_b\|_{1, \mathcal{K}^\varepsilon \setminus \Gamma} + \|b\|_\Gamma) \|E(\bar{u}_h - u_h) - E_h(\bar{u}_h - u_h)\|_{1, \mathcal{K}^\varepsilon} \\ &\leq \mathcal{C}(\|\ell_b\|_{1, \mathcal{K}^\varepsilon \setminus \Gamma} + \|b\|_\Gamma) \sup_{v_h \in \mathbf{V}_h} \frac{\|E(v_h) - E_h(v_h)\|_{\mathcal{K}^\varepsilon}}{\|v_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon}} \|\bar{u}_h - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon} \end{aligned} \quad (3.84)$$

Estimate (3.77) follows directly from the identity in (3.78), the ultimate bound in (3.84) and the definition of E_h in (3.72). \square

Theorem 7 essentially implies that if the lift-based approximation (3.17) displays optimal global convergence, then the WSM approximation (3.23) displays optimal local convergence.

Corollary 8. *Assume that Conjecture 6 holds and that there exist an $\varepsilon^* > 0$ and an open subset $K \Subset \mathcal{K}^{\varepsilon^*}$ such that $\mathcal{K} \subset K$. Assume that $A_{ijkl} \in C^{k+1}(\bar{\Omega})$ for some integer $k \geq 0$ and that Ω is convex or of class C^{k+2} . Assume that b admits a sufficiently smooth local lifting, in particular, that there exists an ℓ_b such that $\{\sigma_\nu(\ell_b)\} = 0$, $\operatorname{div} \sigma(\ell_b) \in \mathbf{H}^k(\Omega)$ and $\operatorname{supp}(\ell_b) \subset \mathcal{K}^\varepsilon$. Let $\mathbf{V}_\mathcal{H}$ denote a sequence of $\mathbf{H}^1(\Omega)$ -conforming piecewise polynomial approximation spaces of degree $p \geq 1$ with approximation property (3.18). Let u denote the solution to the Volterra dislocation problem (3.11) and let $u_\mathcal{H}$ denote the sequence of WSM approximations (3.23) corresponding to the approximation spaces $\mathbf{V}_\mathcal{H}$. It holds that*

$$\|u - u_h\|_{1, \Omega \setminus \mathcal{K}^\varepsilon} \leq \mathcal{C}h^l \quad (3.85)$$

as $h \rightarrow 0$ with $\mathcal{C} > 0$ independent of h and $l = \min\{p, k+1\}$.

Proof. Note that the following integration-by-parts identity holds for all $v \in \mathbf{H}^1(\Omega)$:

$$\begin{aligned} a_\Gamma(\ell_b, v) &= \int_{\partial(\Omega \setminus \Gamma)} v \cdot \sigma_n(\ell_b) - \int_{\Omega \setminus \Gamma} v \cdot \operatorname{div} \sigma(\ell_b) \\ &= \int_\Gamma v \cdot \{\sigma_\nu(\ell_b)\} - \int_{\Omega \setminus \Gamma} v \cdot \operatorname{div} \sigma(\ell_b) \end{aligned} \quad (3.86)$$

The second identity follows by rearranging terms and applying (3.8). The first term in the ultimate expression in (3.86) vanishes by virtue of the conditions on ℓ_b . Moreover, because $l(v) = (f, v) = 0$, the right member of (3.12) corresponds to

$$l(v) - a_\Gamma(\ell_b, v) = - \int_{\Omega \setminus \Gamma} v \cdot \operatorname{div} \sigma(\ell_b) \quad (3.87)$$

By virtue of $\operatorname{div} \sigma(\ell_b) \in \mathbf{H}^k(\Omega)$ and the conditions on the elasticity tensor and the domain, it then holds that the continuous complement \bar{u} in (3.12) resides in $\mathbf{H}^{k+2}(\Omega)$; see, for instance, [27, Theorem 6.3.5]. The quasi-optimality of the Galerkin approximation (3.17) in combination with the optimal approximation properties of \mathbf{V}_h according to (3.18) implies that $\|\bar{u} - \bar{u}_h\|_{1,\Omega} \leq \mathcal{C}h^l$. Moreover, the conditions on the elasticity tensor imply that $E(v_h) \in \mathbf{H}_{\operatorname{loc}}^{k+2}(\mathcal{X}^\varepsilon)$ and, in turn, (3.18) yields:

$$\sup_{v_h \in \mathbf{V}_h} \inf_{w_h \in v_h + \mathbf{V}_{h,\mathcal{X}^\varepsilon}} \frac{\|E(v_h) - w_h\|_{1,\mathcal{X}^\varepsilon}}{\|v_h\|_{1,\Omega \setminus \mathcal{X}^\varepsilon}} \leq \mathcal{C}h^l \quad (3.88)$$

The estimate (3.85) then follows from

$$\|u - u_h\|_{1,\Omega \setminus \mathcal{X}^\varepsilon} \leq \|\bar{u} - \bar{u}_h\|_{1,\Omega} + \|\bar{u}_h + \ell_b - u_h\|_{1,\Omega \setminus \mathcal{X}^\varepsilon} \quad (3.89)$$

and Theorem 7. □

3.4 Numerical results

We will assess the approximation properties of the Weakly-enforced Slip Method on the basis of three different test cases. All three test cases are designed to have analytical results available. This allows us to compare the exact solution of boundary value problem (3.5) with the WSM solution (3.23) and study the behavior of errors and convergence under refinement of the finite finite-element mesh.

The considered test cases are:

- I. a two-dimensional infinite domain loaded in plane strain, with a straight, finite dislocation and smooth slip distribution,
- II. a three-dimensional semi-infinite domain with traction-free surface, a planar, non-rupturing dislocation and constant slip, and
- III. a three-dimensional semi-infinite domain with traction-free surface, a planar, surface rupturing dislocation and constant slip.

All three test cases are in principle set on infinite domains. To make the problems amenable to treatment by the finite-element method, we truncate the domains, and restrict the analyses to suitably large, but finite, neighbourhoods of the dislocation. In order to focus on the treatment of the dislocations, boundary truncation errors are controlled by constraining the finite-element approximation to the analytical solution at the artificial lateral boundaries.

Let us note that of the three test cases, only the first one is by design in full accordance with the theory developed in Section 3.3. Test case II and III feature a

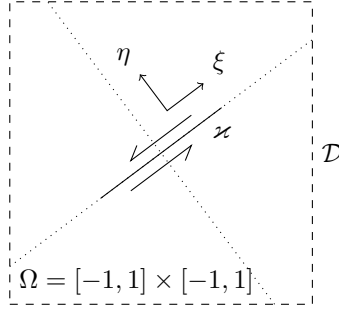


Figure 3.4 – Schematic overview of test case I. The dashed line marks the computational domain, which is a truncation of the actual (infinite) domain. A unit-length dislocation \varkappa is placed at an $\arctan(3/4)$ angle with the horizontal axis at the center of the computational domain.

piecewise constant slip to match available analytical solutions, violating the condition of Lemma 1 which states that the slip can be lifted into $\mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$, i.e., $b \in \mathbf{H}_0^{1/2}(\varkappa)$. Test case III moreover considers a rupturing fault and, accordingly, it has nonzero slip at the intersection with the domain boundary. The results below convey that the main results of the theory nonetheless uphold, suggesting that the conditions under which they apply can be relaxed.

3.4.1 Test case I: 2D plane strain

The first test case is a line dislocation in plane strain in an infinite two-dimensional isotropic domain. The dislocation is straight and of unit length, with a smoothly varying slip that is tangent to the dislocation line, making it a pure shear dislocation. Figure 3.4 shows a schematic of the computational setup.

Introducing an arclength coordinate ξ along the fault, we consider a smooth, piecewise quadratic slip distribution according to:

$$b : \mathbb{R} \rightarrow b_0 \begin{cases} \frac{3}{2} + 6\xi + 6\xi^2 & -\frac{1}{2} < \xi < -\frac{1}{6} \\ 1 - 12\xi^2 & -\frac{1}{6} < \xi < \frac{1}{6} \\ \frac{3}{2} - 6\xi + 6\xi^2 & \frac{1}{6} < \xi < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \quad (3.90)$$

The dislocation, which corresponds to the support of the slip, is located in the interval $\xi \in [-\frac{1}{2}, \frac{1}{2}]$. The slip is symmetric, with zero displacement at the tips and smoothly opening and closing. We set the scaling factor $b_0 = 0.1$. The value of b_0 is however non-essential on account of the linearity of the problem. An analytical solution to this problem can be constructed for an infinite, homogeneous, isotropic domain, by adapting the well known solution for edge dislocations [38]. Introducing a perpendicular coordinate η , the resulting displacement along the (ξ, η) coordinates is expressed in terms of Lamé parameters λ and μ as

$$u(\xi, \eta) = b_0 [6U(\xi - \frac{1}{2}, \eta) - 18U(\xi - \frac{1}{6}, \eta) + 18U(\xi + \frac{1}{6}, \eta) - 6U(\xi + \frac{1}{2}, \eta)] \quad (3.91)$$

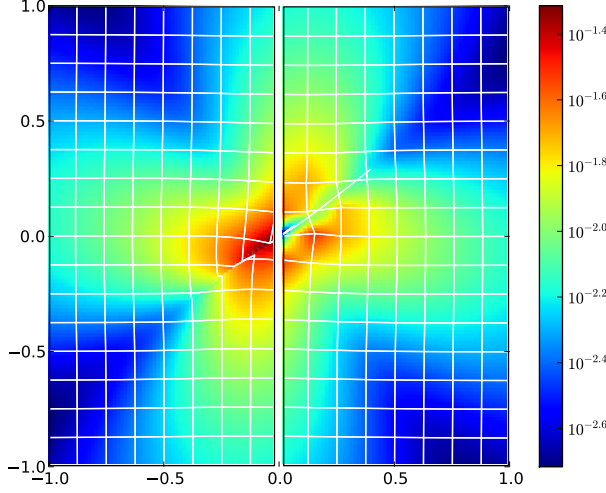


Figure 3.5 – Side by side view of the discontinuous exact solution of test case I (left) with the continuous WSM approximation on a 16×16 mesh and linear shape functions (right). Color represents displacement magnitude on a log scale. Mesh lines on the left are introduced to indicate displacements and to facilitate comparison.

where

$$U(\xi, \eta) = \left(-\frac{\frac{2\lambda+3\mu}{\lambda+2\mu}\xi\eta}{\frac{\mu}{\lambda+2\mu}\xi^2 - \frac{2\lambda+\mu}{\lambda+2\mu}\eta^2} \right) \log \sqrt{\xi^2 + \eta^2} + \left(\frac{\frac{3\lambda+4\mu}{\lambda+2\mu}\eta^2 - \xi^2}{2\frac{\lambda}{\lambda+2\mu}\xi\mu} \right) \arctan \frac{\xi}{\mu}.$$

A finite sized problem is obtained by truncating the domain to $\Omega = (-1, 1)^2$ and introducing a Dirichlet condition at the boundary $\partial\Omega = \mathcal{D}$ with data corresponding to the exact solution. The fault is located in the center of the domain at an angle of $\arctan(3/4)$. All numerical results are generated for Lamé parameters $\lambda = 1, \mu = 1$.

Figure 3.5 shows the exact solution of Eqn. (3.91) (left) in the deformed configuration, side by side with the WSM approximation for linear shape functions on a 16×16 mesh (right). The colors indicate displacement magnitude on a log scale. The grid-line pattern represents the distortion of a mesh that is uniform in the undeformed configuration. The lines coincide with element edges for the finite element computation on the right, and are matched on the left to facilitate visual comparison. Because the test case is symmetric, we expect the images to be approximately rotationally symmetric. We observe that the rotational symmetry breaks at the dislocation, where the exact solution is discontinuous, while the WSM approximation is continuous throughout the domain.

The global symmetry of the displacement pattern of Figure 3.5 indicates that the error induced by the WSM approximation diminish with distance from the dislocation. To quantify this observation, Figure 3.6 shows the errors for a 16×16 mesh (left), and for a sequence of meshes along the line B–B' perpendicular to the fault (right). The latter shows that for every mesh the error decays exponentially with increasing distance from the fault, with the exception of a narrow zone in the vicinity of the dislocation, where the error displays a transition to the constant error $\frac{1}{2}b_0$ at the dislocation. One may observe that the error at the dislocation is independent of element size. Away from the dislocation, errors are seen to decrease with element size, approximately reducing by $10^{0.6} \approx 4$ whenever the mesh width is halved. This error reduction provides a first indication that local L^2 convergence is optimal at $O(h^{p+1})$ as $h \rightarrow 0$.

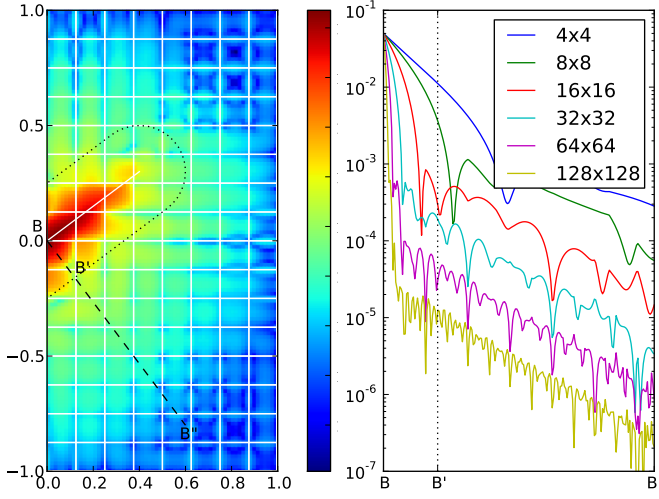


Figure 3.6 – Absolute value of the displacement error for the 16×16 mesh result of Figure 3.5 (left), side by side with error evaluations along the line B-B'' for a sequence of meshes of increasing density. The errors shown on the left thus correspond with the 3rd (red) curve on the right.

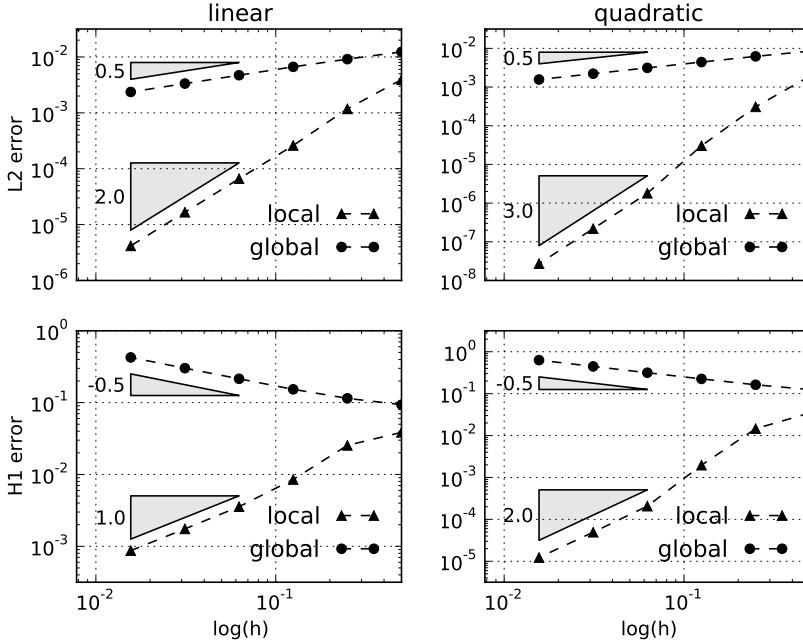


Figure 3.7 – Mesh convergence of WSM applied to test case I, showing the L^2 -norm (top) and H^1 -norm (bottom) of the error for linear (left) and quadratic (right) shape functions. The markers corresponds to mesh sequences of $\{4 \times 4, 8 \times 8, \dots, 128 \times 128\}$ elements are considered. The global error is computed by integration over the entire computational domain Ω , the local error by integration over the >0.1 distance exterior around ζ , bounded by the dotted line in Figure 3.6.

Convergence of the WSM approximation under mesh refinement is further examined in Figure 3.7, which displays the \mathbf{L}^2 -norm and \mathbf{H}^1 -norm of the error for both linear ($p = 1$) and quadratic ($p = 2$) shape functions. The curve marked ‘global’ shows the error integrated over the entire computational domain. The curve marked ‘local’ shows the error integrated over the domain excluding an 0.1-neighborhood of the dislocation, corresponding to the dotted area in Figure 3.6. We observe that for both linear and quadratic approximations, the global \mathbf{H}^1 -norm of the error diverges as $O(h^{-1/2})$ as $h \rightarrow 0$, while the global \mathbf{L}^2 -norm of the error converges as $O(h^{-1/2})$, independent of the order of approximation. This asymptotic behavior is in accordance with the estimates in Corollaries 3 and 5. Figure 3.7 moreover corroborates that the local \mathbf{H}^1 -norm of the error converges as $O(h^p)$ as $h \rightarrow 0$, in agreement with the estimate in Corollary 8. The local \mathbf{L}^2 -norm of the error, for which no theory was developed, also displays an optimal convergence rate of $O(h^{p+1})$ as $h \rightarrow 0$.

3.4.2 Test case II: 3D traction-free halfspace

The second test case is a planar dislocation buried in a semi-infinite, three dimensional, homogeneous, isotropic domain with a flat, traction-free surface. The dislocation is a rectangular plane, the sides of which are displaced over a constant distance in both strike and in dip direction, such that in geodetic terms the setting is that of an oblique left-lateral thrust fault. Figure 3.8 shows a schematic of the computational setup. Analytical solutions to this problem have been derived by Okada [55, 56]. We use homogeneous Lamé constants $\lambda = 1, \mu = 1$ for all subsequent computations.

Computations are performed on a truncated domain $\Omega = [-1, 1] \times [-1, 1] \times [-1, 0]$, with Dirichlet conditions enforcing the Okada solution at the five truncation planes. Figure 3.9 shows displacements along the intersection planes A-A’ and B-B’ indicated in Figure 3.8. The topmost figure shows the exact solution according to Okada’s equations. The middle and bottom figures display the WSM approximation of the displacement on a $16 \times 16 \times 8$ mesh, for linear and quadratic shape functions, respectively. The displacement fields in the WSM approximations are seen to be continuous everywhere, though they become highly irregular at the dislocation. At further distances from the dislocation, however, the approximations exhibit very good agreement with the exact solution, despite the coarseness of the considered mesh.

Figure 3.10 examines convergence of the global and local norms of the error under mesh refinement. The results confirm the $O(h^{-1/2})$ divergence and the $O(h^{1/2})$ convergence of the global \mathbf{H}^1 -norm and global \mathbf{L}^2 -norm, respectively, in agreement with the estimates in Corollaries 3 and 5. Furthermore, we observe $O(h^p)$ convergence of the local \mathbf{H}^1 -norm in accordance with the estimate in Corollary 8 and $O(h^{p+1})$ convergence for the local \mathbf{L}^2 -norm. Figure 3.10 moreover displays the \mathbf{L}^2 -norm of the displacement error at the surface, which exhibits an optimal convergence rate of $O(h^{p+1})$. It is to be noted that the agreement of the observed convergence rates for the global \mathbf{H}^1 -norm and \mathbf{L}^2 -norm and the local \mathbf{H}^1 -norm with the estimates in Section 3.3 is non-obvious, as a piecewise constant slip cannot be lifted into $\mathbf{H}_{0,\mathcal{D}}^1(\Omega \setminus \Gamma)$, and the test case under consideration therefore fails to meet the conditions imposed in Section 3.3.

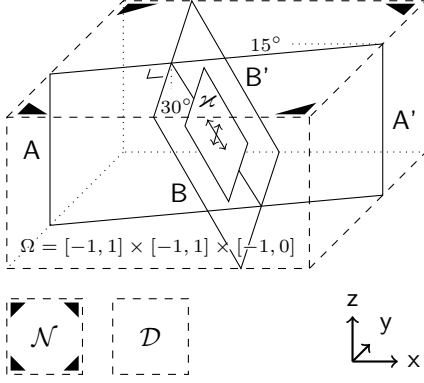


Figure 3.8 – Schematic overview of test case II. A $1 \times 3^{-1/2}$ sized dislocation plane is positioned at 15° strike and 30° dip, spanning a 0.25–0.75 depth range. The surface is traction free. Dashed faces mark the computational domain, which is a truncation of the actual (semi-infinite) domain. The fault is represented by the B–B’ plane, where at \varkappa the medium is dislocated by a constant 0.2 displacement in strike direction and a constant 0.1 displacement in dip direction. The perpendicular intersection plane A–A’ serves visualization purposes only; see Figures 3.9 and 3.12.

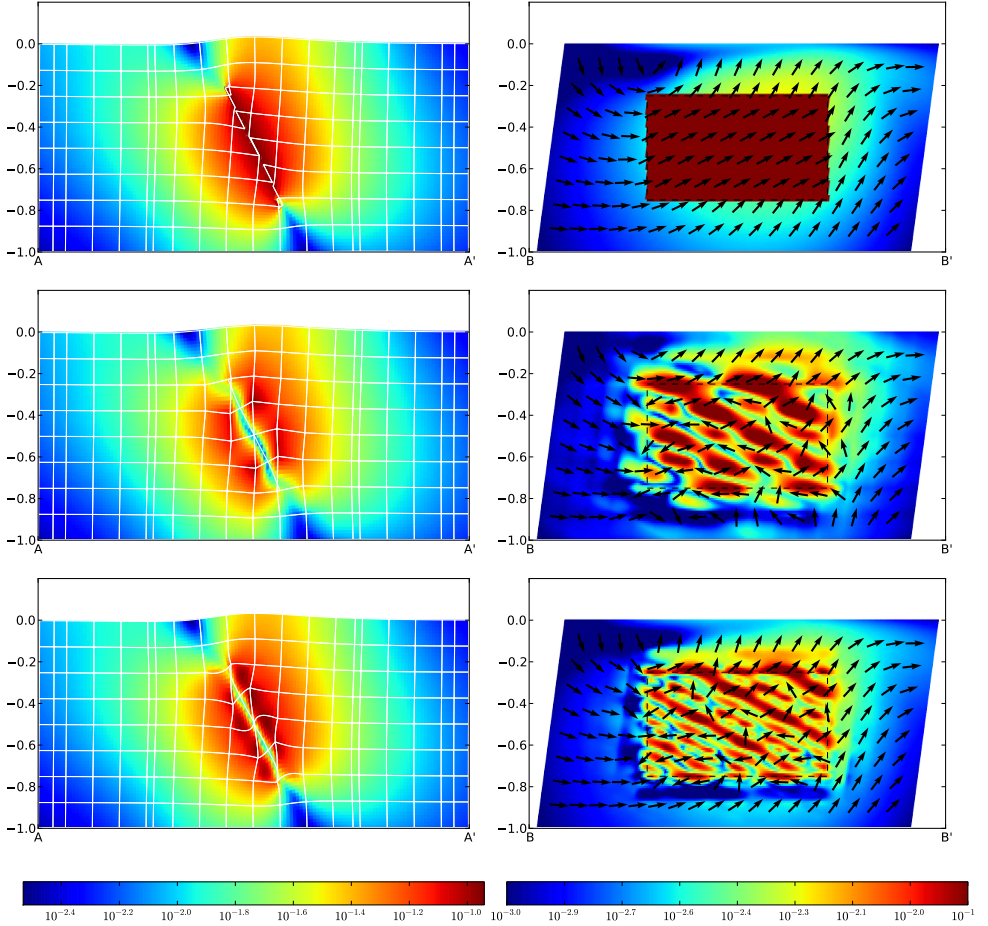


Figure 3.9 – Cross-section views of the displacement field of test case II, corresponding to the A–A’ and B–B’ intersection planes of Figure 3.8: exact displacement as derived by Okada [56] (top) and WSM approximation on a $16 \times 16 \times 8$ mesh with linear (middle) and quadratic shape functions (bottom). Colors indicate the displacement magnitude on a logarithmic scale analogous to Figure 3.5. Arrows indicate direction only.

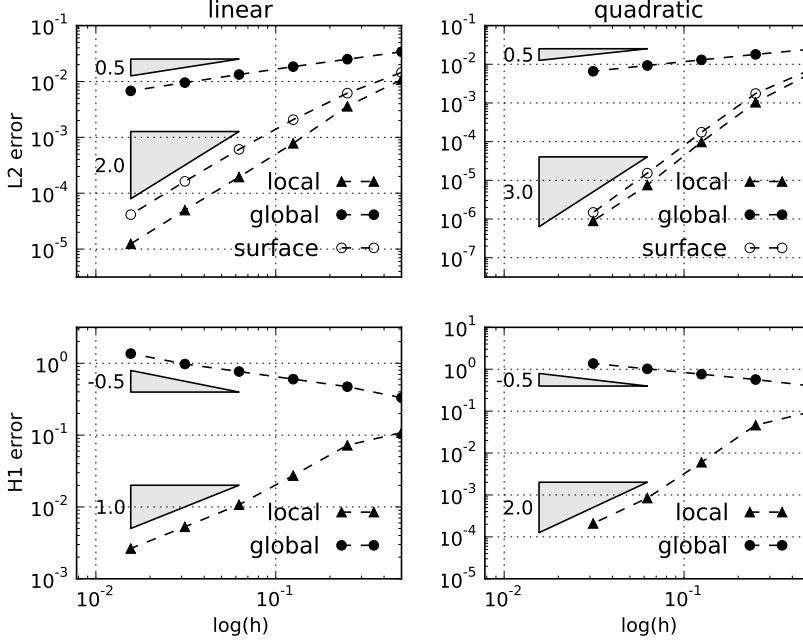


Figure 3.10 – Mesh convergence of WSM for test case II. The figures are analogous to Figure 3.7, with the addition of the L^2 -norm of the error in the displacement at the traction-free surface.

3.4.3 Test case III: 3D traction-free rupturing halfspace

The third test case is in everything equal to the second, except that the dislocation is now extended in vertical direction to span a depth range of 0–0.75, and the dip slip direction is reversed to form the geodetic equivalent of a rupturing reverse fault; see Figure 3.11. Figure 3.12 shows the cross section displacements, and Figure 3.13 the norms of the displacement-error under mesh refinement. We observe that the global (resp. local) H^1 -norm and L^2 -norm of the error are again proportional to $h^{-1/2}$ and $h^{1/2}$ (resp. h^p and h^{p+1}), respectively. In addition to the L^2 -norm of the error in the displacement field at the traction-free boundary, Figure 3.13 also presents the local L^2 -norm of the error at the traction-free boundary, i.e., the error on the surface excluding an 0.1-neighborhood of the dislocation. The global L^2 -norm of the surface error converges with an asymptotic rate of $O(h^{1/2})$. This suboptimal convergence behavior is caused by the fact that in this case the discontinuity reaches the surface. The local L^2 -norm of the error at the surface again display optimal convergence at a rate of $O(h^{p+1})$. It is to be noted that the convergence results for the global H^1 -norm and L^2 -norm and the local H^1 -norm agree with the estimates in Section 3.3, despite the fact that the analysis in Section 3.3 is restricted to non-rupturing faults.

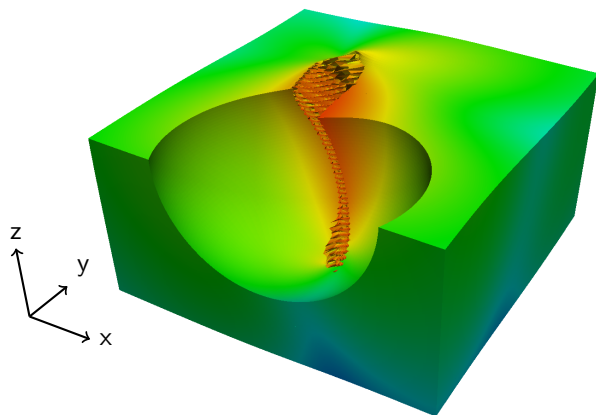


Figure 3.11 – Perspective projection of a WSM computation of test case III, using linear shape functions on a $128 \times 128 \times 64$ mesh. Deformations are amplified by a factor 2 for visualization purposes; colors represent displacement magnitude on a log scale. The orientation aligns roughly with that of Figure 3.8. A spherical cut-out exposes part of the interior of the domain, revealing a zone of large displacements local to the dislocation.

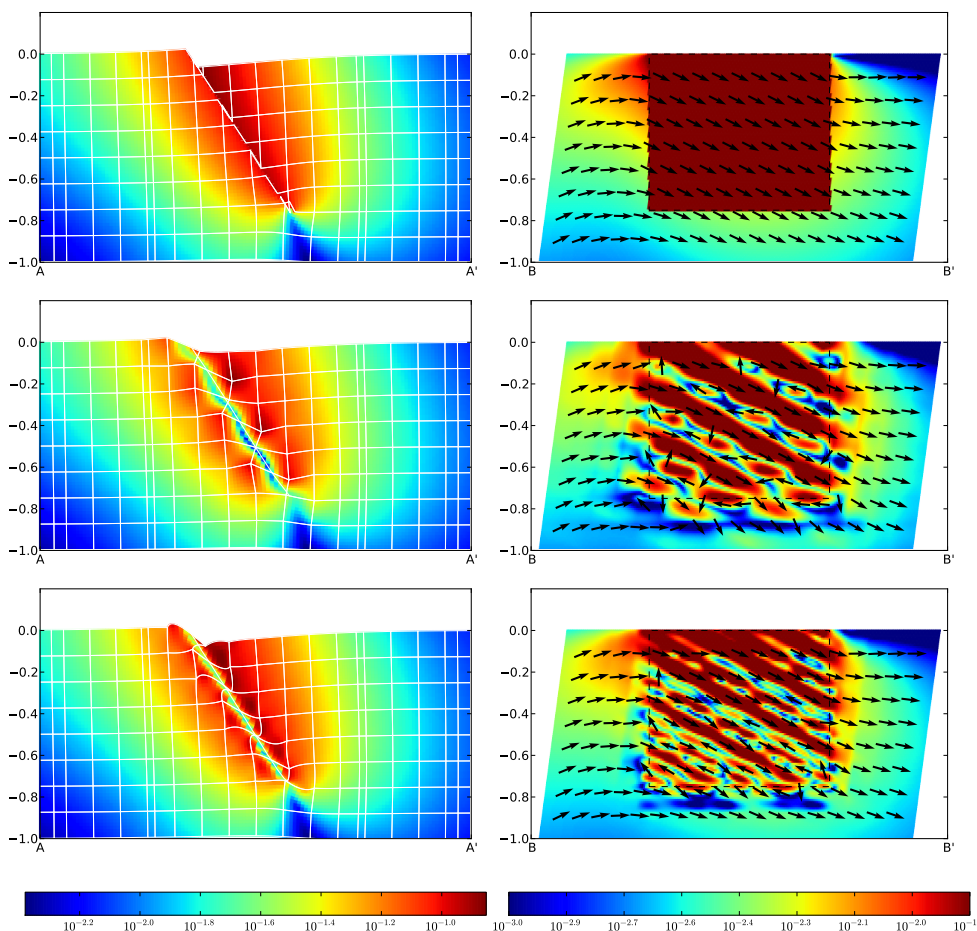


Figure 3.12 – Cross-section views of the displacement field of test case III. The figures are analogous to Figure 3.9.

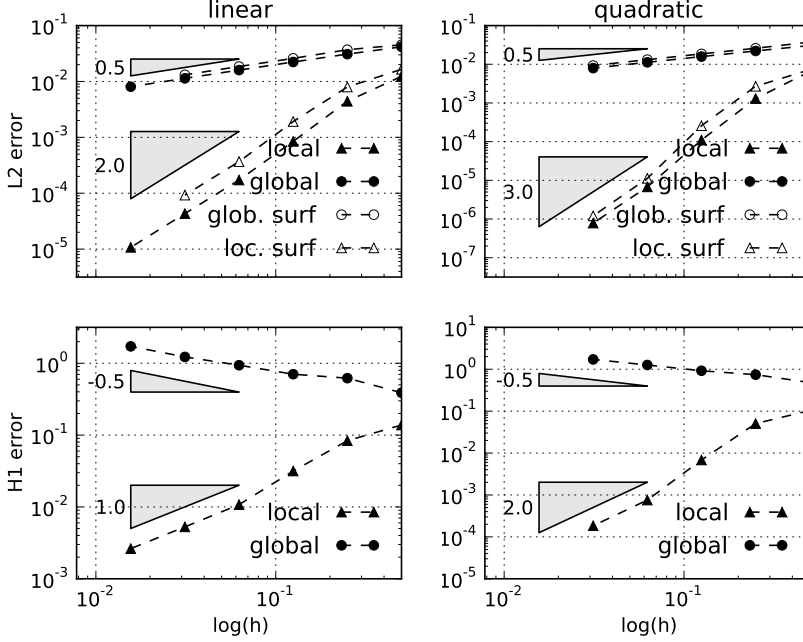


Figure 3.13 – Mesh convergence of WSM applied to test case III. The figures are analogous to Figure 3.10, except that the L^2 -norm of the error on the surface is now indicated by ‘global surface’ and additionally a ‘local surface’ L^2 -norm of the error has been plotted, excluding an 0.1-neighborhood of the dislocation.

3.5 Conclusions

To solve Volterra’s dislocation problem by standard finite-elements techniques, the dislocation is required to coincide with element edges. This requirement links the finite-element mesh with the fault geometry, which prohibits the reuse of computational components in situations where multiple geometries have to be considered. In particular, it renders the finite-element method infeasible in nonlinear inversion problems.

To overcome the problems of standard finite-element techniques in nonlinear inversion processes, in this paper we introduced the *Weakly-enforced Slip Method* (WSM), a new finite-element approximation for Volterra’s dislocation problem in which the slip discontinuity is weakly imposed in the right-hand-side load functional. Accordingly, the bilinear form in the formulation and, hence, the stiffness matrix are independent of the fault geometry. The method is summarized by the following weak formulation:

$$u \in \mathbf{V}_h : \quad a(u, v) = - \int_{\Gamma} b \cdot \{\sigma_{\nu}(v)\} \quad \forall v \in \mathbf{V}_h.$$

The stiffness matrix depends on properties of the continuous domain only, namely, variations in the elasticity and topology and geometry of the domain, and remains independent of fault geometry. Fault dependence manifests in the right-hand-side load functional only. The load functional is formed by integrating over the fault, which is

allowed to cut through elements. The integration along the fault is a non-standard operation in finite-element methods, but we expect that it can be incorporated in most existing finite-element toolkits with minor effort. We further note that no approximations are required regarding fault geometry and slip distribution, unlike lift-based methods, which require a parametrization for both.

We established that as a consequence of the continuous approximation in WSM, the dislocation is not resolved. Accordingly, the WSM approximation displays suboptimal convergence in the \mathbf{L}^2 -norm under mesh refinement. In particular, the \mathbf{L}^2 -norm of the error decays only as $O(h^{1/2})$ as the mesh width h tends to 0, independent of the order of approximation. Furthermore, the \mathbf{H}^1 -norm of the error generally diverges as $O(h^{-1/2})$ as $h \rightarrow 0$, independent of the order of approximation. In addition, we however proved that WSM has outstanding local approximation properties, and that the method generally displays optimal convergence in the \mathbf{H}^1 -norm on the domain excluding an arbitrarily small neighborhood of the dislocation. In particular, for any ε -neighborhood \mathcal{N}^ε of the dislocation, the $\|\cdot\|_{1,\Omega \setminus \mathcal{N}^\varepsilon}$ norm of the error in the WSM approximation generally converges as $O(h^p)$ as $h \rightarrow 0$, with p the polynomial degree of the finite-element space.

Numerical experiments in 2D and 3D were conducted to verify and scrutinize the approximation properties of WSM. The asymptotic error estimates for the global \mathbf{H}^1 -norm, the global \mathbf{L}^2 -norm and the local \mathbf{H}^1 -norm of the error were confirmed in all cases, despite the fact that two of the test cases violate some of the conditions underlying the asymptotic error estimates. In particular, the numerical results indicate that the error estimates extend to rupturing faults, where the dislocation fissures the traction-free surface. The numerical experiments moreover conveyed that the WSM approximation displays optimal local convergence in the \mathbf{L}^2 -norm, i.e., the $\|\cdot\|_{\Omega \setminus \mathcal{N}^\varepsilon}$ -norm of the error decays as $O(h^{p+1})$ as $h \rightarrow 0$. For the non-rupturing-fault test case, the error in the surface displacement converges optimally in the \mathbf{L}^2 -norm at $O(h^{p+1})$. For the rupturing-fault test case, the \mathbf{L}^2 -norm of the error at the surface converges at $O(h^{1/2})$, while the local \mathbf{L}^2 -norm excluding a small neighborhood of the dislocation again converges optimally at $O(h^{p+1})$. Overall, the approximation obtained via WSM is very well behaved, and the method proves remarkably robust.

Given the compelling properties of WSM one might be tempted to seek application in other than our intended field of tectonophysics. Dislocation plasticity comes to mind as one heavily relying on superposition of elastic dislocations. For many problems, however, the location of the fault will be known a-priori, in which case there is no reason to avoid strong imposition. Moreover, often the internal stress is an important quantity to be resolved, which with WSM suffers from inaccuracies in regions close to the fault. In tectonophysics the primary observable is the displacement of the free surface, which WSM is very well capable of resolving.

We believe that WSM can play an important, if specific, role in the application of tectonic fault plane inversions. Being able to precompute the stiffness matrix and a quality preconditioner, for any fault geometry and slip distribution under study, it remains only to integrate over the 2D manifold and solve the system. This makes it feasible to use finite elements in a direct nonlinear inversion. In the hands of geophysicists, this tool will allow all available in-situ knowledge to be made part of the forward model. We hope this will help to improve the accuracy of future co-seismic analyses.

4 | Inverting elastic dislocations using the WSM

Earthquakes cause lasting changes in static equilibrium, resulting in global deformation fields that can be observed. Consequently, deformation measurements such as those provided by satellite based InSAR monitoring can be used to infer an earthquake’s faulting mechanism. This inverse problem requires a numerical forward model that is both accurate and fast, as typical inverse procedures require many evaluations. The Weakly-enforced Slip Method (WSM) was developed to meet these needs, but it was not before applied in an inverse problem setting. Consequently, it was unknown what effect particular properties of the WSM, notably its inherent continuity, have on the inversion process.

Here we show that the WSM is able to accurately recover slip distributions in a Bayesian-inference setting, provided that data points in the vicinity of the fault are removed. In a representative scenario, an element size of 1 km was found to be sufficiently fine to generate a posterior probability distribution that is close to the theoretical optimum. For rupturing faults a masking zone of 10 km sufficed to avoid numerical disturbances that would otherwise be induced by the discretization error. These results demonstrate that the WSM is a viable forward method for earthquake inversion problems. While our synthesized scenario is basic for reasons of validation, our results are expected to generalize to the wider gamut of scenarios that finite element methods are able to capture. This has the potential to bring modeling flexibility to a field that is often forced to impose model restrictions in a concession to computability.

This chapter is a verbatim copy of “Inverting elastic dislocations using the Weakly-enforced Slip Method” by G.J. van Zwieten, E.H. van Brummelen and R.F. Hanssen; International Journal for Numerical and Analytical Methods in Geomechanics, 2022 [76].

With the advent of satellite based interferometry, or InSAR, routine measurements of the earth’s surface deformation have become available, providing a wealth of in-

formation about subsurface processes [9]. One of these processes is tectonic faulting, along with its violent manifestation, earthquakes. While the dynamics of the quake itself cannot be measured from space in the way that seismometers do, what can be measured to great accuracy is the lasting adjustments of static equilibrium to the defect that results from the relative displacement, or *slip*, of two adjacent masses along a fault.

Given two measurements of the earth’s surface covering the area of an earthquake, of which one taken just prior, the other just after, it can be reasonably assumed that the differences can be wholly attributed to the process of tectonic faulting [58]. Separation of *co-seismic* and *post-seismic* signal can be controlled by shortening the satellite revisit time [25]. However, to establish the details of the faulting mechanism that corresponds to the observed co-seismic deformation, such as the location and orientation of the fault, and the amount, depth and direction of the slip that occurred, we require an understanding of the physics bridging the two.

It is generally assumed that on the near-instantaneous time scale of co-seismic deformation, the earth behaves to a good approximation elastically [62]. The default model connecting the fault mechanism and surface observations, therefore, is that of static, elastic dislocations. While a formal definition of this problem will be presented in Section 4.1, sufficing at present is that this type of problem has been studied for well over a century, during which a range of solution methods has been devised [78].

Of the many solution methods available, the most powerful is arguably the Finite Element Method (FEM) [75], which is capable of incorporating all available knowledge of material heterogeneity and surface topography and thus capturing the system to the greatest detail. While FEM is occasionally used to establish slip [1], which is a linear problem, inversion studies that aim to establish the fault plane geometry as well as the slip distribution are commonly forced to the other end of the complexity spectrum for reasons of computational cost. This is why such analyses are often based on the assumption of a homogeneous halfspace [63, 28, 71], for which cheaply evaluable analytical expressions are available. It is with this situation in mind that we proposed the Weakly-enforced Slip Method (WSM) [77] to strike a better balance between modelling complexity and computational cost.

The key strength of the WSM is that it decouples the computational mesh from the geometry of the fault, so that the mesh does not need to be updated for every distinct fault geometry as would be the case in ordinary FEM. Crucially, this means that the stiffness matrix, as well as derived products such as LU factors, can be constructed once and be reused throughout the inversion process. While the remaining vector that needs to be constructed for a given fault geometry is nonstandard, it should be relatively straightforward to integrate this in existing FEM softwares, as well as combine it with domain specific innovations such as Agata et al.’s Fast FEA [1].

The efficiency of WSM does come at a price: the displacement field it produces is continuous, and therefore unable to capture the jump at the fault. It was however shown [77] that the error thus incurred decays exponentially with distance to the fault. It was therefore hypothesized that this drawback is of little consequence in the context of satellite observations, as most surface measurements are sufficiently far removed from the dislocation. Only rupturing or near-rupturing faults will cause numerical errors that significantly interfere with the observables, and even that only locally. As InSAR data tends to be of low quality in these regions with damage leading

to decorrelation, it is hoped that a reduction of numerical accuracy in this region will be of little consequence.

It bears repeating that the above considerations are speculative. While the mathematical soundness of the WSM was proved [77], it is in the present paper that we set out to thoroughly test the utility of the WSM in the problem setting for which it was devised. To this end a number of synthetic but otherwise fully representative case studies are presented and analysed using the WSM, as well as validated against the exact solutions. For the sake of this validation the scenarios will be restricted to homogeneous half spaces so that exact solutions are available in the form of analytical expressions, but it is to be understood that the presented methodology is valid for the wider class of problems including material heterogeneity and topography.

The first two sections will lay the theoretical groundwork for the analysis. Section 4.1 presents the forward problem of determining surface displacements for a given fault location and slip distribution, focusing both on the construction of an exact solution and on approximating it using the WSM. Section 4.2 then considers the inverse problem, of determining fault location and slip distribution for given surface displacements, using the framework of Bayesian inference. Section 4.3 will define the methodology and introduce case studies, and Section 4.4 will present the results of the comparative study.

4.1 Forward Problem: Linear Elastic Dislocation

To formally define the dislocation problem that will stand as the model to the earth's response to tectonic faulting, we will denote by Ω the solid domain, by d its spatial dimension, and by $u : \Omega \rightarrow \mathbb{R}^d$ the deformation field, i.e. the displacement of the solid compared to its reference configuration. We assume that it is possible to create a mapping from the deformation field u to the corresponding state of internal stress σ . In particular, we assume that the stress depends linearly on the deformation gradient, leading to the well known constitutive relation

$$\sigma(u) = C : \nabla u, \quad (4.1)$$

where $C : \Omega \rightarrow \mathbb{R}^{d \times d \times d \times d}$ is the stiffness tensor representing local material properties, subject to symmetry relations $C_{ijkl} = C_{klij} = C_{jikl} = C_{ijlk}$. For the medium to be at rest, Newton's second law states that the divergence of stress must be in balance with the applied loading. Rather than incorporating the loading conditions of the earth's gravitational field, we use the linearity of the stress-strain relation of Equation (4.1) to have u represent only the deviations relative to the existing equilibrium. The earth's gravitational field being constant in time, this means the stress field (4.1) is divergence free.

Through the remainder of this document we will use a consistent notation for the four spaces that form the basis of our mathematical framework. By \mathcal{X} we denote the space of local fault plane coordinates $\xi \in \mathbb{R}^{d-1}$. By \mathcal{M} we denote the space of all possible fault geometries $m : \mathcal{X} \rightarrow \mathbb{R}^d$ that position the manifold in physical space. By \mathcal{B} we denote the space of slip distributions pulled back to \mathcal{X} , the slip vector $b(\xi)$ being the jump in the displacement field when passing from one side of the manifold to the other. By \mathcal{D} we denote the space of surface displacements $d : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$ as measured

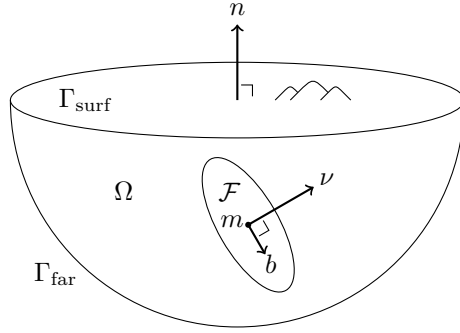


Figure 4.1 – Adapted from Van Zwieten et al. [78], visualization of the forward problem and all entities of Equation (4.2).

in line of sight to the satellite. With this notation in place we can reformulate the forward problem as follows:

(Forward problem) *Determine the surface observations $d \in \mathcal{D}$ corresponding to a given manifold $m \in \mathcal{M}$ and slip distribution $b \in \mathcal{B}$.*

We denote the fault plane $\mathcal{F} = m(\mathcal{X})$ and the domain boundary Γ as shown in Figure 4.1. At any fault point $x \in \mathcal{F}$ we demand that the displacement field u jumps discontinuously by a displacement $b(m^{-1}(x))$. Since this makes the displacement field locally non-differentiable, the stress is locally not defined, and our general equilibrium condition does not apply. Instead, Newton's second law transforms into a jump condition for the traction, stating that the tractions on either side of the fault plane must be in balance. At the earth's surface Γ_{surf} we assume traction-free conditions, and the remaining domain boundary Γ_{far} is assumed to be sufficiently far away for the relative displacements to be zero. Taken together, this results in the following system of equations for the forward problem:

$$\left\{ \begin{array}{ll} \operatorname{div} \sigma(u) = 0 & \text{at } \Omega \setminus \mathcal{F} \text{ (static equilibrium, continuum)} \\ \llbracket u \rrbracket = b \circ m^{-1} & \text{at } \mathcal{F} \text{ (fault slip)} \\ \llbracket \sigma(u) \rrbracket \nu = 0 & \text{at } \mathcal{F} \text{ (static equilibrium across fault)} \\ \sigma(u)n = 0 & \text{at } \Gamma_{\text{surf}} \text{ (traction-free surface)} \\ u = 0 & \text{at } \Gamma_{\text{far}} \text{ (zero displacement at far field)} \end{array} \right. \quad (4.2)$$

Before looking into solution strategies for this system, it is readily apparent that solutions to this problem are linear in b . Hence, for any manifold $m \in \mathcal{M}$, there exists a linear map from the space of slip distributions to the corresponding observations:

$$F_m : \mathcal{B} \rightarrow \mathcal{D}. \quad (4.3)$$

If \mathcal{B} and \mathcal{D} are finite-dimensional and endowed with a basis, we can identify every linear operator with a matrix, $F_m \in \mathbb{R}^{\#\mathcal{B} \times \#\mathcal{D}}$, where $\#(\cdot)$ denotes the cardinality of set (\cdot) .

4.1.1 Analytical solutions

The general expression for solutions to (4.2) was presented in integral form by Volterra [67]:

$$u_n(y) = \sum_{ijkl} \int_{\xi \in \mathcal{X}} b_i(\xi) \nu_j(\xi) C_{ijkl}(m(\xi)) \frac{\partial u_k^n}{\partial x_l}(m(\xi), y) \det \left| \frac{\partial m}{\partial \xi} \right| d\xi, \quad (4.4)$$

where the Green's function $u_k^n(x, y)$ is the k -th component of the displacement vector in x due to a unit point force at y in direction n . Note that (4.4) has an equivalent alternative form owing to the symmetry relation $u_k^n(x, y) = u_n^k(y, x)$, which is a direct result of Betti's reciprocal theorem [47].

To obtain closed form expressions for the Green's function we need to place additional constraints on our system. Firstly, we require that the elastic properties of our medium are homogeneous and isotropic, thus reducing the constitutive model to having two independent parameters. Choosing as our parameters the Young's modulus E and Poisson's ratio ν , the constitutive tensor becomes

$$C_{ijkl} = \frac{E}{2 + 2\nu} \left[\frac{2\nu}{1 - 2\nu} \delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right]. \quad (4.5)$$

Since the displacement resulting from a unit force is inversely proportional to the Young's modulus, we observe that the displacement field resulting from Volterra's equation (4.4) depends only on the Poisson's ratio of the medium.

Secondly, for closed form expressions for the Green's functions $u_n^k(x, p)$ to be available, we require the physical domain Ω to be a half space, that is, to have an infinite, flat surface as its free surface Γ_{surf} and the far field boundary Γ_{far} at infinity. The Green's functions for the 2D half space were derived by Melan [50], those for the 3D half space by Mindlin [52].

Building on Mindlin's results, closed form expressions for Volterra's equation have been presented by Yoffe [73] and Okada [56], subject to further restrictions in terms of fault planes and slip distributions. While it is these results that are typically used in practical applications, for the purposes of our study we shall apply Volterra's equation directly so as not to incur additional and unnecessary restrictions to our case studies that would diminish the value of the present study.

4.1.2 The Weakly-enforced Slip Method

The Weakly-enforced Slip Method is a special case of the Finite Element Method, which is in turn a Galerkin method, employing shape functions to construct a finite system of equations that can be solved numerically. Contrary to classical finite element treatments of Equation (4.2), in which the domain must be discretized such that the mesh conforms to the manifold m , the defining property of the Weakly-enforced Slip Method is that the finite element mesh can be formed independent of m .

Foregoing derivations and proofs, which are presented in detail in Van Zwieten et al. [77], we present the WSM only in terms of its core result. Given a finite element discretization for the computational domain Ω , and generating from it a discrete, vector-valued function space \hat{V} , the WSM solution to Equation (4.2) is the field

$\hat{u} \in \hat{V}$ that satisfies, for all test functions $\hat{v} \in \hat{V}$,

$$\int_{\Omega} \nabla \hat{u} : C : \nabla \hat{v} = \sum_{ijkl} \int_{\xi \in \mathcal{X}} b_i(\xi) \nu_j(\xi) \left\{ C_{ijkl}(m(\xi)) \frac{\partial \hat{v}_k}{\partial x_l}(m(\xi)) \right\} \det \left| \frac{\partial m}{\partial \xi} \right| d\xi. \quad (4.6)$$

Here $\{\cdot\}$ is the mean operator, which takes effect only in case m coincides with an element boundary, making $\sigma(v_h)$ multi-valued; in the general case it reduces to a single evaluation. It is noteworthy that upon substitution of the Green's function $u_k^n(m, y)$ for the test function $\hat{v}_k(m)$, the left-hand-side reduces to $\hat{u}_n(y)$ and we obtain Volterra's Equation (4.4).

The advantage of constructing the discrete solution space \hat{V} independently of m is directly apparent from Equation (4.6): the stiffness matrix, which results from the left hand side of the equation, as well as solution primitives such as LU factors, are independent of m and can thus be created once and reused for many different faulting scenarios. The right hand side vector, resulting from the right hand side of the equation, while dependent of m , is constructed by integrating over the fault plane alone and is therefore considerably cheaper to construct.

The disadvantage, as touched upon before, is that in constructing function space \hat{V} independently of m it cannot possibly allow for discontinuities at any subsequently defined manifold. The main qualitative difference between the exact solution u and the WSM solution \hat{u} , therefore, is that where the former is discontinuous, jumping at the manifold by a distance d , the latter exhibits a smeared out transition extended in the adjacent area. It was shown [77] that the size of the transition zone can be controlled through mesh refinement, and that the error decreases exponentially with distance to the fault. Moreover, the method is shown to have optimal convergence for any subdomain that excludes the manifold.

4.2 Inverse Problem: Bayesian Formulation

The endeavour to infer the fault geometry and slip distribution from observed displacements is known as the inverse problem. With notation as introduced in Section 4.1, this problem can be formulated as follows:

(Inverse problem) *Determine the manifold $m \in \mathcal{M}$ and slip distribution $b \in \mathcal{B}$ corresponding to given observations $d \in \mathcal{D}$.*

While this problem was proved [4] to have a unique solution (subject to regularity assumptions) when observations are exact and cover an entire surface patch, real world observations are inherently discrete and noisy, and no statement of this kind can be made in practice. On the contrary, inverse problems tend to be highly sensitive to overfitting and often require some means of regularization for stability. To this effect we follow the typical strategy of adding *a priori* information, in the form of quantified notions on expected fault geometries and slip distributions, and use it to select the most likely solution out of a multitude of good fits.

Observing that noisy data and prior likelihoods are stochastic concepts, the problem naturally presents itself in terms of random variables, and we select Bayesian inference [65, Ch.1] as a suitable framework for solving stochastic problems of this kind in an approach that is similar to that of Xu et al. [71]. Our stochastic framework

consists of three random variables: manifold $M \in \mathcal{M}$, slip distribution $B \in \mathcal{B}$, and line-of-sight surface measurements $D \in \mathcal{D}$. The probability density of finding a manifold $M = m$ and slip distribution $B = b$ given that we observe surface displacements $D = d$ is given by Bayes' theorem as being proportional to the likelihood of observing $D = d$ given $M = m$ and $B = b$, and the prior probability of $M = m$ and $B = b$ in the absence of observations:

$$\overbrace{f_{MB|D}(m, b, d)}^{\text{posterior}} = \overbrace{f_{D|MB}(d, m, b)}^{\text{likelihood}} \overbrace{f_{MB}(m, b)}^{\text{prior}} / \overbrace{f_D(d)}^{\text{marginal}} . \quad (4.7)$$

The marginal represents the probability of observing $D = d$. Its distribution follows directly from the likelihood and the prior, owing to the fact that the posterior probability density integrates to one. We present this relation for completeness, though we will not need to evaluate it for our purposes:

$$f_D(d) = \int_{m \in \mathcal{M}} \int_{b \in \mathcal{B}} f_{D|MB}(d, m, b) f_{MB}(m, b) \quad (4.8)$$

This means that the only terms that require further elaboration are the prior, the likelihood, and the posterior, which we shall explore in the following sections.

4.2.1 Prior Distribution

The prior $f_{MB}(m, b)$ is the probability density of finding a manifold m and fault slip b absent any observations. It is a quantification of all prior knowledge relating to manifolds, slip distributions, and the local tectonic setting. The prior distribution could be based on a general understanding of physics, or it could introduce knowledge from other sources, such as seismic data. The latter is done for instance by Xu et al. [71], who construct a Gaussian prior around the moment magnitude.

Constructing a truly accurate quantification of prior knowledge is difficult, if not impossible. For this reason the distribution is in practice often parameterized with what is known as hyperparameters: tunable constants that are optimized using additional criteria such as Akaike's Bayesian Information Criterion (ABIC) [72, 31]. Since the present work does not deal with real world data, we are in the special position that prior knowledge can be considered a known quantity. It should be emphasized that this circumstance relates only to the design of the current study, and is in no way implied by the use of WSM as a forward model.

A useful first step in the construction of our prior is to decompose it. Since fault slip is defined on the manifold, a natural, universally valid decomposition is the following:

$$f_{MB}(m, b) = f_M(m) f_{B|M}(b, m), \quad (4.9)$$

where $f_M(m)$ is the prior probability density of the manifold, and $f_{B|M}(b, m)$ the prior probability density of the fault slip conditional to the manifold.

Before constructing a prior for the manifold we must identify \mathcal{M} with a parameter space. For instance, three coordinates, two angles and two lengths define a rectangular plane. Additional parameters can encode curvature, forks, or other irregularities as appropriate. Once defined, the simplest prior is constructed by taking all parameters to be uncorrelated, and every parameter either normally distributed around an

expected value or uniformly distributed within a chosen interval based on available in-situ information. In case actual data is available about a-priori correlations and distributions, this can directly be translated into a high quality prior. If limited information is available, this can be encoded in a weakly informative prior, e.g. a normal distribution with large standard deviation.

Regarding the probability density $f_{B|M}$, the conditionality of the slip b on the manifold m is immediately apparent from the observation that beyond certain dip angles there is a strong predominance of normal faults over reverse faults, which translates in an asymmetric probability on the sign of the slip distribution. Real world inversions would benefit greatly from encoding this and other mechanisms in a high quality prior, though at the expense of making the slip inversion nonlinear. For reasons of simplicity, however, we shall consider the manifold and fault slip to be independent:

$$f_{B|M}(b, m) = f_B(b). \quad (4.10)$$

Adopting this simplification, we further wish the slip vectors to be strongly correlated at points that are close together, and weakly correlated at points that are spaced far apart, based on a general understanding of the physics underlying slip events. These notions are formalized in the positive semi-definite autocovariance function $K : \mathcal{X} \otimes \mathcal{X} \rightarrow \mathbb{R}^{d-1 \times d-1}$:

$$K_{ij}(\xi_1, \xi_2) = \text{cov}(B_i(\xi_1), B_j(\xi_2)), \quad (4.11)$$

which we are free to design in any way that reflects existing prior knowledge. Typically, $K(\xi_1, \xi_2)$ depends only on ξ_1, ξ_2 via their Euclidean distance in relation to a specified correlation length.

For practical reasons we cannot operate on the infinite dimensional space \mathcal{B} , but will instead operate on a finite dimensional subspace and discrete random variables $\hat{B} \in \hat{\mathcal{B}} \subset \mathcal{B}$. To aid its construction we define a vector-valued basis $h = \{h_1, h_2, \dots, h_N\}$ for the discrete space of slip distributions, and thus associate with any random slip $\hat{B} \in \hat{\mathcal{B}}$ a random vector $B^h \in \mathbb{R}^N$ such that $\hat{B}(\xi) = \sum_{n=1 \dots N} B_n^h h_n(\xi)$. We now take B^h to be normally distributed with covariance matrix Σ_B^h , which we aim to construct in such a way that (4.11) still holds to good approximation, i.e., $K(\xi_1, \xi_2) \approx h^T(\xi_1) \Sigma_B^h h(\xi_2)$. To this end we multiply both sides of the equation by $h(\xi_1)h(\xi_2)^T$ and integrate over the domain to form the following projection:

$$\underbrace{\int_{\xi_1 \in \mathcal{X}} \int_{\xi_2 \in \mathcal{X}} h(\xi_1) K(\xi_1, \xi_2) h(\xi_2)^T}_{H_K} = \underbrace{\left(\int_{\xi_1 \in \mathcal{X}} h(\xi_1) h(\xi_1)^T \right)}_{H_\delta} \Sigma_B^h \underbrace{\left(\int_{\xi_2 \in \mathcal{X}} h(\xi_2) h(\xi_2)^T \right)}_{H_\delta} \quad (4.12)$$

The extent to which the projection $\Sigma_B^h = H_\delta^{-1} H_K H_\delta^{-1}$ approximates the autocovariance function depends on the details of the autocovariance function K in relation to the approximation properties of the basis, but can in general be controlled fully by adding basis vectors, i.e. increasing the dimension of $\hat{\mathcal{B}}$.

One remaining issue with the above construction is that the resulting Σ_B^h may not be positive semi-definite, which is a requirement for it to qualify as a covariance matrix. We therefore proceed by diagonalizing the result as $\Sigma_B^h = V \Lambda V^T$, where Λ and V are the real-valued eigenvalues resp. H_δ -orthogonal eigenvectors of the generalized eigenvalue problem $H_K V = H_\delta V \Lambda$. Eliminating the negative eigenvalues

and corresponding eigenvectors we arrive at the covariance matrix that approximates our autocovariance function K . In fact, we could go a step further and eliminate small positive eigenvalues as well, as these modes are seen to not contribute much to the overall expansion (more details on this in Section 4.3.2) — this process has the potential to greatly reduce the dimension of $\hat{\mathcal{B}}$ and hence improve numerical efficiency.

For the purposes of construction we were forced to make the difference explicit between the true space of slip distributions, \mathcal{B} , and the finite dimensional subspace that we will use for the analysis, $\hat{\mathcal{B}}$. We note that, even though we did not need to make this formal, a similar distinction applies to all spaces: our parametric space \mathcal{M} is really a finite dimensional subspace of the much larger space of possible manifolds, and the observation space \mathcal{D} is arguably a discrete subspace of a continuous signal space. Since our analysis is finite dimensional, however, we consider only (sufficiently rich) finite dimensional subspaces. For this reason we shall also drop this distinction for slip distributions, and have \mathcal{B} denote the finite dimensional space going forward.

Finally, note that the covariance matrix is specific to the chosen basis h . We could therefore still strive to create a basis h' in such a way that the corresponding covariance matrix $\Sigma_B^{h'}$ becomes an identity and all slip coefficients become independent random variables. Indeed, the diagonalization provides us with the tools we need in the form of the recombination matrix $V\Lambda^{1/2}$, post removal of unwanted modes. The resulting basis is known as a Karhunen-Loeve expansion [48], and it is what we shall be using in our practical implementation. However, while we shall drop the suffix h from here on, we shall continue to write Σ_B , rather than I , in the interest of preserving structure and keeping our derivations general.

4.2.2 The Likelihood

Given a manifold m and slip distribution b , using a linear map F_m of the type of Equation (4.3) we expect surface observations to equal $F_m b$. Due to model errors and measuring noise, we take the likelihood of observing d to be normally distributed around this expected value with covariance Σ_D :

$$f_{D|MB}(d, m, b) = G_{\Sigma_D}(d - F_m b), \quad (4.13)$$

with the Gaussian probability density function defined as

$$G_{\Sigma}(x) = \frac{\exp(-\frac{1}{2}x^T \Sigma^{-1}x)}{\sqrt{\det |2\pi\Sigma|}}. \quad (4.14)$$

Since the sum of independent, normally distributed random variables is in turn normal, the covariance matrix Σ_D can be seen as the superposition of several noise mechanisms. Spatially uncorrelated noise resulting directly from the properties of the InSAR measurement system contributes to the diagonal, with entries possibly varying to reflect dependence on distance or incidence angle, or local factors such as those caused by damage or other sources of temporal decorrelation. Off-diagonal terms may be added to account for spatially-correlated noise, such as errors caused by atmospheric delay.

Furthermore, it is through the covariance that we may account for the quality of the forward model itself. In the context of the WSM we expect a large error in

locations where the continuous solution space is not able to follow local discontinuities. Making the variance locally large is a convenient way of downweighing the data in this area, while the extreme case of making it locally infinite effectively masks out the area, keeping only the intermediate and far field data for the inversion.

4.2.3 Posterior Distribution

Substituting the prior probability distribution and the likelihood into Bayes' theorem (4.7), we can rework terms to obtain the following result:

$$\begin{aligned}
 f_{MB|D}(m, b, d) &= \underbrace{f_{D|MB}(d, m, b)}_{G_{\Sigma_D}(d - F_m b)} \underbrace{f_B(b)}_{G_{\Sigma_B}(b)} f_M(m) / \underbrace{f_{D|M}(d, m)}_{f_D(d)} \\
 &= \underbrace{G_{\Sigma'_B(m)}(b - b'(m, d))}_{f_{B|MD}(b, m, d)} \underbrace{G_{\Sigma'_D(m)}(d) f_M(m) / f_D(d)}_{f_{M|D}(m, d)} \quad (4.15)
 \end{aligned}$$

with the posterior covariance and expected value of B conditional to M and D ,

$$\Sigma'_B(m)^{-1} = \Sigma_B^{-1} + F_m^T \Sigma_D^{-1} F_m \quad (4.16)$$

$$b'(m, d) = \Sigma'_B(m) F_m^T \Sigma_D^{-1} d. \quad (4.17)$$

and the posterior covariance of D conditional to M ,

$$\Sigma'_D(m)^{-1} = \Sigma_D^{-1} - \Sigma_D^{-1} F_m \Sigma'_B(m) F_m^T \Sigma_D^{-1} \quad (4.18)$$

The identity of Equation (4.15) is verified through direct substitution of the posterior covariances and expected value in the Gaussian probability density function (4.14). Of particular use in this exercise is the Weinstein-Aronszajn identity, $\det |I + AB| = \det |I + BA|$, which, taking $A = F_m$ and $B = -\Sigma'_B(m) F_m^T \Sigma_D^{-1}$, results in the following useful relationship:

$$\det |\Sigma_B| \det |\Sigma_D| = \det |\Sigma'_B(m)| \det |\Sigma'_D(m)| \quad (4.19)$$

While the identity of Equation (4.15) is itself entirely algebraic, the interpretation of the individual terms as conditional probabilities $f_{M|D}$, $f_{B|MD}$ and $f_{D|M}$ is not immediately apparent. The first follows from marginalizing over \mathcal{B} : since the marginal of $G_{\Sigma'_B(m)}$ is 1 by definition, (4.15) directly leads to the identity

$$f_{M|D}(m, d) = G_{\Sigma'_D(m)}(d) f_M(m) / f_D(d). \quad (4.20)$$

Interpretation of the remaining conditional probabilities then follows readily from the conditional probability relation $f_{B|MD}(b, m, d) = f_{MB|D}(m, b, d) / f_{M|D}(m, d)$, and from Bayes' theorem, $f_{D|M}(d, m) = f_{M|D}(m, d) f_D(d) / f_M(m)$.

The result of Equation (4.20) is particularly useful as it allows us to evaluate the total probability density of manifold m conditional to observations d , leaving the study of the slip b to a separate, later stage. To aid this evaluation we use definition

(4.14) of the multi-variate Gaussian, together with identities (4.17), (4.18) and (4.19), to expand the marginal distribution (4.20) into

$$f_{M|D}(m, d) = \underbrace{\left(\frac{\exp(-\frac{1}{2}d^T \Sigma_D^{-1}d)}{f_D(d) \sqrt{\det[2\pi \Sigma_D] \det[\Sigma_B]}} \right)}_{\text{independent of } m} \frac{f_M(m) \exp(\frac{1}{2}d^T \Sigma_D^{-1} F_m b'(m, d))}{\sqrt{\det[\Sigma'_B(m)]^{-1}}}. \quad (4.21)$$

While the expression contains the marginal f_D , which though a known quantity is impractical to evaluate, this inconvenience is circumvented by using sampling techniques that are scaling invariant, meaning that $f_D(d)$ and other m -independent terms can be ignored. An example of such a technique is Markov Chain Monte Carlo (MCMC), which allows one to obtain low order moments of the conditional probability density of m using a feasibly low number of evaluations.

Using any technique to single out a particular manifold, the probability density of slip b conditional to measurements d and manifold m is normally distributed around expected value $b'(m, d)$ with posterior covariance $\Sigma'_B(m)$. Interestingly, the latter is independent of the measurements, meaning we can evaluate a-priori how a certain combination of measurement noise properties, satellite viewing geometry and manifold position results in a variance of the estimated slip along the length of the manifold.

In the expected value of Equation (4.17) we recognize the solution to a weighted least squares problem with Tikhonov regularization:

$$b'(m, d) = \arg \min_{b \in \mathcal{B}} \left(\|F_m b - d\|_{\Sigma_D^{-1}}^2 + \|b\|_{\Sigma_B^{-1}}^2 \right). \quad (4.22)$$

While this is a standard method for solving ill-posed problems, the stochastic interpretation thus obtained helps us in three ways. Firstly, it provides a confidence measure of the result in the form of a posterior covariance matrix. Secondly, it lends meaning to Σ_B and Σ_D that helps us to construct the required matrices. And lastly, it enables one to design numerical experiments that match the stochastic underpinnings of the method.

4.3 Methodology

To test the WSM-based inversion of tectonic faulting, we synthesize a deformation field for certain fault parameters m and slip distribution b , and then try to estimate the fault parameters \hat{m} and slip distribution \hat{b} from noisy line-of-sight data using both the WSM and Volterra's equation. The process can be divided in five steps:

1. Select fault parameters m from M ;
2. Draw a slip distribution b from B (or construct it manually);
3. Draw observation data d from D conditional to M and B using Volterra's equation (4.4) as the forward model according to Sections 4.1.1 and 4.1.2;
4. Evaluate the posterior expected value $\hat{m} = E(M|D = d)$ and covariances using either the WSM or Volterra's equation as the forward model;

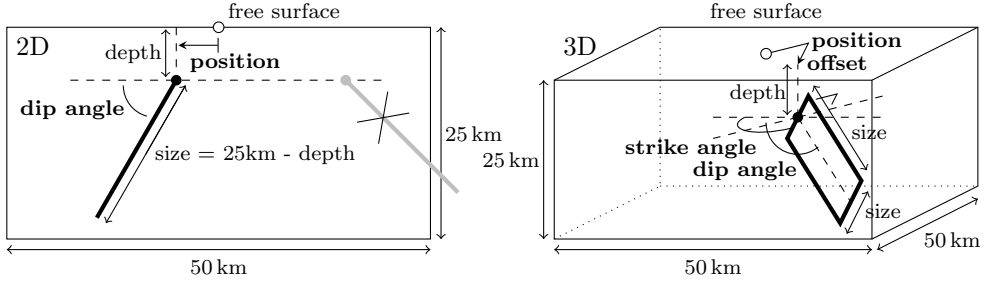


Figure 4.2 – Fault parameters and their representation in 2D and 3D scenarios. The invalid parameter combination shown in gray in the 2D scenario is not a member of \mathcal{M} .

- 4* Alternatively set $\hat{m} = m$ to study a linear-only inversion limited to B .
- 5. Evaluate the posterior expected value $\hat{b} = E(B|M = \hat{m}, D = d)$ and covariances using the same forward model as in 4.

In the following we will elaborate on each of these steps.

4.3.1 Constructing fault parameters

Though any real world situation naturally concerns three-dimensional space, it is advantageous to study a two-dimensional analogue as well as this allows us to study the entire work flow in a setting that is less expensive and easier to visualize. We will therefore construct two different sets of fault parameters, one for one-dimensional faults in two-dimensional space, the other for two-dimensional faults in three-dimensional space.

We limit ourselves to the space of straight faults of fixed dimensions, that are placed anywhere in a box of 50 km width, 50 km breadth (for 3D scenarios) and 25 km depth. In order to distinguish between the class of rupturing and non-rupturing faults (a distinction that can often be made on the basis of field observations) we set the minimum depth to remain fixed. This leaves two parameters in 2D space and four parameters in 3D space to parameterize the entire space, as summarized in Figure 4.2.

We note that while the fixed dimensions of the fault plane form an upper bound for the dimensions of the fracture zone, the support of the slip distribution can still localize within these confines. Introducing additional parameters for length and width would therefore not add actual degrees of freedom but rather ambiguities between the two spaces \mathcal{M} and \mathcal{B} , manifested in additional expenses for the nonlinear inversion due to the increased dimension of \mathcal{M} . While omitting dimensions from the parameter space means that we require expert judgement to define what size is sufficiently large for a given situation, we can verify the validity of this assumption a posteriori, for example by testing if the inverted slip is sensitive to fault plane enlargement.

Similar considerations apply to fault location, where in-plane variations can to some degree be captured by \mathcal{B} . This is what allows us to fix the depth, while the actual onset of slip might be deeper still. Relatedly, in the 3D scenario we encode the location as a ‘position’ that is normal to strike, and an ‘offset’ along strike. While we could conceivably eliminate the latter and rely entirely on \mathcal{B} to capture the in-plane

component, we choose to keep the offset in \mathcal{M} , as the fault plane would otherwise have to be undesirably large in order to still cover the search box in all orientations. However, we anticipate that its posterior variance will be significantly larger than that of the position due to the remaining ambiguities between \mathcal{M} and \mathcal{B} .

The fault size is the largest size that fits the box given the vertical offset. Rupturing faults are thus of 25 km length, while faults that close 5 km below the surface are of 20 km length. For the 3D scenario both dimensions of the fault are always kept equal. The mapping to physical space $m(\xi)$ is an affine transformation, supporting the assumption that the slip distribution $b(\xi)$ can be considered independently. As it is for our purposes important that the entire fault fits inside the search box, we define \mathcal{M} to contain no positions that fall outside of it, nor any angle that causes the fault to intersect its boundary. Figure 4.2 shows an example of this in the 2D scenario. On the resulting oblique domain we take the prior distribution of M to be uniform.

4.3.2 Constructing the slip distribution

To construct a space of slip distributions we take the local fault plane coordinates to be the unit line or unit square, $\mathcal{X} = [0, 1]^{d-1}$, mapping through m onto a fault plane of dimensions L . We take slip components in orthogonal directions to be independent, which is a non-restrictive assumption in practice. This reduces the distribution of the vector field to a series of identically distributed scalar fields, for which we define an exponential autocorrelation function with correlation length ℓ and maximum slip amplitude β :

$$K(\xi_1, \xi_2) = \beta^2 w(\xi_1) w(\xi_2) \exp\left(-\frac{1}{2} \frac{|\xi_1 - \xi_2|^2}{(\ell/L)^2}\right). \quad (4.23)$$

Here w is a window function. The window sets the variance to zero on all boundaries for non-rupturing faults, or for all but the surface edge for rupturing faults. In 2D space (with a 1D fault) these window functions are $w_{\text{closed}}(\xi) = 4\xi(1 - \xi)$ and $w_{\text{open}}(\xi) = \xi(2 - \xi)$, respectively. In 3D space (with a 2D fault) they are the tensor product of $w_{\text{closed}} \otimes w_{\text{closed}}$ and $w_{\text{closed}} \otimes w_{\text{open}}$.

A Karhunen-Loève expansion is constructed for this autocovariance function via the projection of Equation (4.12) using a truncated trigonometric series for the basis h — though we remark that a mesh-based construction can be used instead in case more flexibility is required. Similar to the window functions, we distinguish the non-rupturing and the rupturing situations. For non-rupturing faults in one dimension, we use the orthonormal sine series $h_n(\xi) = \sqrt{2} \sin(\xi n \pi)$. For rupturing faults we use a modified cosine series $h_n(\xi) = \sum_{i=0 \dots n} \alpha_{ni} \cos(\xi i \pi)$, with coefficients α_{ni} chosen such that $h_n(1) = 0$ and the basis functions are orthonormal. For two-dimensional faults we use the outer products to form a scalar basis on the unit square, preserving orthonormality.

While orthonormality is not a requirement, it is a convenient property as we no longer need to form H_δ (now an identity) and the generalized eigenvalue problem reduces to a conventional eigenvalue problem $H_K V = V \Lambda$ — the size of which depends on the truncation point of the trigonometric series. Selecting the n largest eigenvalues, the Karhunen-Loève expansion is formed by recombining the modes by $V \Lambda^{1/2}$ causing the corresponding covariance matrix to reduce to an identity; drawing a sample from the distribution then amounts to independently drawing n coefficients from a standard normal distribution. Since orthogonal slip components are taken to be independent

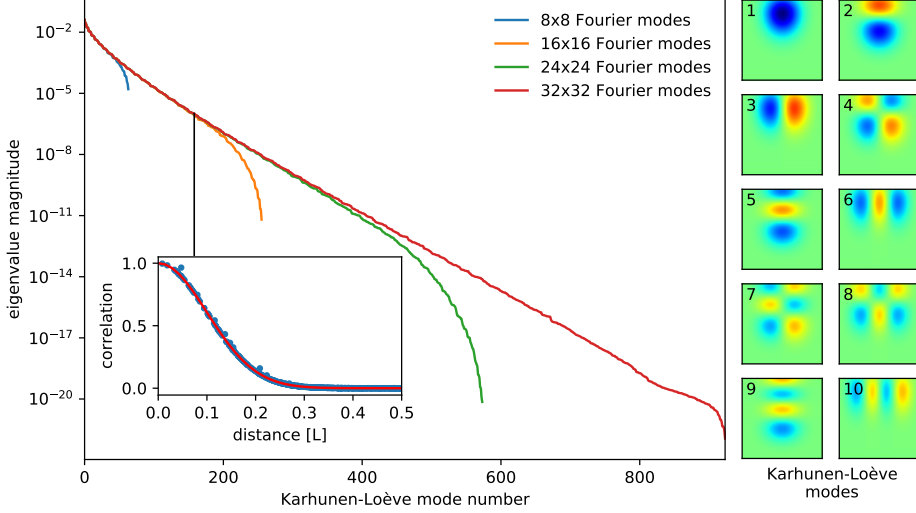


Figure 4.3 – The eigenvalue spectrum Λ of the H_K matrix for a rupturing fault in three-dimensional space and a correlation length $\ell = 0.1L$, computed at different truncation points of the trigonometric series. The inset shows the autocorrelation of the 153-mode Karhunen-Loève expansion based on 1024 randomly selected point pairs (blue) along with the target autocorrelation function (red). Shown to the right are the first 10 Karhunen-Loève modes $V\Lambda^{1/2}$.

in our choice of autocovariance function, it suffices to form a scalar basis for each of the components of the vector. Figure 4.3 shows the spectrum Λ for first nine scalar bases functions for illustration.

For our experiments we set the correlation length to $\ell = 2.5$ km and the slip amplitude to $\beta = 1$ m. These values are representative of actual geophysical conditions but otherwise arbitrary. To determine the number of Karhunen-Loève modes to retain, we set a maximum error of $\|H_K - V\Lambda V^T\| < 10 \text{ mm}^2$ measured in the Frobenius norm. By this process we establish that $n = 12$ Karhunen-Loève modes are sufficient for the 2D scenario and $n = 153$ modes for the 3D scenario. Repeating this process for different truncation sizes of the original trigonometric series we find that this result is stable beyond 16 modes per dimension, and decide based on this that 32 basis functions per dimension is a sufficiently rich starting point for the expansion.

In a practical application one may wonder if the constructed prior distribution is an accurate representation of reality, and indeed if the true slip distribution is even an element of \mathcal{B} at all. To make sure that a slip distribution drawn from the prior does not represent an artificial best case scenario with little real world value, we additionally construct a slip distribution that has local support at a select area of the fault. In addition to testing the robustness of the method to incorrect assumptions, the local support also allows us to test whether fault dimensions can indeed be captured via the slip, rather than via additional fault parameters.

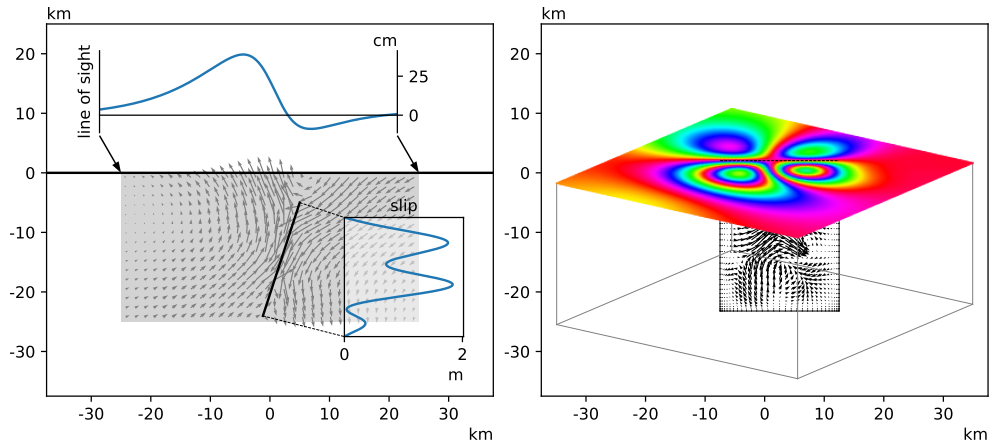


Figure 4.4 – Visualisation of the search box containing a non-rupturing fault at position 5 km ($0.2L$) and dip angle 72° (0.4π), and in 3D additionally offset 2.5 km ($0.1L$) and strike angle 54° (0.3π) (Section 4.3.1), a slip distribution drawn from the prior distribution (Section 4.3.2) and corresponding synthesized deformation data at a 30° incidence angle (Section 4.3.3). Left: a 2D scenario. Right: a 3D scenario in orthographic projection normal to the fault, showing line of sight deformations modulo 2.8 cm to mimic a typical C-band interferogram.

4.3.3 Synthesizing observation data

Selecting a fault plane m from the space defined in Section 4.3.1, and drawing a random slip b from the distribution defined in Section 4.3.2, we obtain the 2D and 3D configurations shown in Figure 4.4. While it is readily apparent that neither outcome is physically plausible, we note that this merely highlights the limitations of the constructed prior. In the 2D scenario we observe the hanging wall moving upward relative to the footwall, at an angle at which reverse faulting is not likely to occur in nature. This is a direct consequence of the assumption made in Section 3.1 that the random variables M and B are independent, so that the direction of slip cannot be made to depend on the orientation of the fault. The 3D scenario shows a fairly wild slip distribution that is equally unlikely to occur in practice, underscoring the fact that a prior distribution based solely on correlation lengths is a poor quantification of actual expert knowledge. Nevertheless, since the inverse problem measures likelihood only by the prior distribution from which the scenarios are drawn, these considerations of non-physicality are not important for the purposes of this study.

We proceed by synthesizing a displacement field based on the assumption that the free surface is flat and infinite, and the material properties are homogeneous, fixing Poisson’s ratio at $\nu = 0.25$ for all experiments. In this situation we have a fundamental solution available in the form of Melan’s (2D) and Mindlin’s (3D) solution, which means we can synthesize the displacement field by evaluating Volterra’s Equation (4.4). The integral is evaluated numerically by means of Clenshaw-Curtis quadrature up to a truncation error that is well below the selected noise level. The displacement field is evaluated in a uniformly spaced grid at a 50 meter resolution, covering the top of the search box defined in Section 4.3.1. These displacements are

then projected onto a vector at a 30 degrees incidence angle, simulating line-of-sight observations as obtained from a typical satellite mission. Finally, Gaussian noise is added to the line-of-sight data that is spatially uncorrelated (thus ignoring atmospheric delays) and normally distributed with a standard deviation of 1 mm. Note that, while the displacement in Figure 4.4 is displayed without noise, the assumed noise is so small relative to the magnitude of the signal that adding it would not make a visible difference.

A difference between the raw synthesised data and deformation data received from a satellite mission is that the SAR sensor provides deformation data modulo the sensor's semi-wave length (for details see e.g. Hanssen [36]), measuring $d + k\lambda/2$ for an unknown integer value of k . While this is typically solved through phase unwrapping (under the assumption that $|\Delta d| < \lambda/4$ for adjacent points) it leaves the data inherently lacking an absolute reference. As we aim for this synthesized study to be representative of real world situations, we need to make sure that our measurement data is similarly relative. To this end we select one measurement point as a reference and subtract its deformation from that of the other measurement points. The inversion is then performed based on the differenced data, using a forward model that reflects the identical differencing procedure.

To formalize this procedure we introduce the unit vector $r \in \mathbb{R}^n$ that selects the reference point, and $e \in \mathbb{R}^n$ the vector of ones. With that we can express the differencing operator as

$$D = I - er^T. \quad (4.24)$$

The differencing is followed by a restriction operation that removes the reference point from the data, for which we introduce the operator $R \in \mathbb{R}^{n \times n-1}$. Note that r and R are related via $r^T R = 0$ and $RR^T + rr^T = I$.

Regardless of the covariance of the measurement noise, all differenced data is correlated due to the shared reference point: $E((d_i - d_{\text{ref}})(d_j - d_{\text{ref}})) \neq 0$ in general due to the nonzero variance of d_{ref} . Using the differencing and restriction operators we can express the covariance matrix of the differenced data in terms of that of the original measurements, which we shall denote henceforth as $\check{\Sigma}_D$:

$$\Sigma_D = R^T D \check{\Sigma}_D D^T R. \quad (4.25)$$

The additional covariances render the matrix fully dense. Fortunately, we note that we can express the inverse covariance of the differenced data in terms of that of the original measurements. Therefore, if the original covariance matrix can be inverted efficiently (for instance if it is diagonal) then this property carries over to the new covariance matrix:

$$\Sigma_D^{-1} = R^T X \check{\Sigma}_D^{-1} R, \quad (4.26)$$

where X is a rank-1 update defined as

$$X = I - \frac{\check{\Sigma}_D^{-1} e e^T}{e^T \check{\Sigma}_D^{-1} e}. \quad (4.27)$$

While our choice of spatially uncorrelated noise implies that the covariance matrix is diagonal, (4.26) holds for any $\check{\Sigma}_D$ and is easily verified using the identities $D^T X = X$,

$DX^T = D$ and $X\check{\Sigma}_D^{-1} = \check{\Sigma}_D^{-1}X^T$, together with $DR = R$, $RR^TD = D$ and $R^TR = I$. Using the same identities it further follows that

$$D^TR\Sigma_D^{-1}R^TD = X\check{\Sigma}_D^{-1}. \quad (4.28)$$

This last result (4.28) is noteworthy for two reasons. Firstly, in Equations (4.16) and (4.17) the inverse covariance occurs only surrounded by either the forward model or the data, both of which need to be differenced as $F_m = R^TD\tilde{F}_m$ and $d = R^TD\tilde{d}$. The result shows that it is not necessary to perform these operations explicitly, and that a rank-1 update of the original covariance matrix inverse is all it takes to switch from an inversion of absolute measurements to that of relative measurements. Secondly, the absence of R and r in (4.27) proves that the inversion is entirely insensitive to the chosen reference point — indeed, we do not need to make any choice at all.

4.3.4 Sampling the posterior distribution

We are interested in evaluating the expected value and auto-covariance of $M|D$, i.e. the posterior distribution of fault parameters given the measurements at the surface. The probability density function $f_{M|D}(m, d)$ was presented in (4.20). However, evaluation of this expression for a particular value of m is problematic because of the marginal $f_D(d)$ contained within, which, while defined in Equation (4.8), does not have a closed form expression. As such we cannot feasibly evaluate the Lebesgue integral to compute the desired quantities.

We find a solution in the class of Markov Chain Monte Carlo (MCMC) methods, which provides an algorithm for drawing samples from $M|D$ while relying only on the ratio of the probability density function at two points m_1 and m_2 , thereby cancelling the marginal and other factors that are independent of m . As the sample sequence $\{m_1, m_2, \dots\}$ thus produced has an empirical probability measure that coincides with the posterior distribution, the expected value and higher moments can be evaluated using Monte Carlo integration

$$E(g(M|D)) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1 \dots N} g(m_i), \quad (4.29)$$

for any g , which we can truncate at any N depending on the desired level of accuracy.

Building on the result of (4.21), we observe that we can additionally isolate the prior distribution f_M as being independent of m , as we choose it to be uniformly distributed in our setup. This leaves the following linearity relation to feature in our MCMC method:

$$f_{M|D}(m, d) \sim \frac{\exp\left(\frac{1}{2}d^T\Sigma_D^{-1}F_mb'(m, d)\right)}{\sqrt{\det|\Sigma'_B(m)|^{-1}}} \quad (4.30)$$

Every evaluation of this function involves a linear inversion of the slip for given m , as is seen directly from the presence of the posterior expected value $b'(m, d)$ and covariance matrix $\Sigma_B^{-1}(m)$. The construction of these and of the forward model F_m will be discussed in Section 4.3.5, which details the linear slip inversion process. Of note presently is that when the required inversion of $\Sigma_B^{-1}(m)$ is performed via a Cholesky decomposition, then the trace of the Cholesky matrix conveniently equals the square-root determinant in Equation (4.30).

The particular MCMC method selected for our purpose is the Metropolis-Hastings [37] algorithm, which performs a random walk through the sample space \mathcal{M} using a proposal distribution to generate candidates, combined with an acceptance/rejection step based on the ratio of probability densities. The algorithm as it is employed here consists of the following steps:

1. initialize $m_0 \in \mathcal{M}$
2. for $i = 1, 2, \dots, N$:
 - (a) draw a random update vector $\delta_m \in \mathcal{M}$ from the proposal distribution
 - (b) draw a uniform random number $u \in [0, 1]$
 - (c) set $m_i = m_{i-1} + \begin{cases} \delta_m & \text{if } f_d(m_{i-1} + \delta_m) \geq u f_d(m_{i-1}) \text{ (accept update)} \\ 0 & \text{otherwise (reject update)} \end{cases}$

Note that condition of step 2c implies that the update is always accepted if it leads to a state of higher probability. Note further that if an update is rejected the current state is repeated in the sequence, thereby adding weight to the empirical distribution and subsequent Monte Carlo integration (4.29).

What remains is to define the starting vector and the proposal distribution. While the proven convergence of the Markov chain suggests that both can be chosen arbitrarily if we take N sufficiently large, this approach is not feasible in practice as the number of iterations required to escape from a local maximum can be prohibitively large. Instead, we require the starting point to be reasonably close to where f_d takes its maximum, and the proposal distribution to be locally similar to f_d in order to have a reasonable acceptance rate.

Aiming for the global maximum, we select the starting vector m_0 using a grid search to find local maxima followed by the Nelder-Mead uphill simplex method. The proposal distribution is taken to be Gaussian with covariance Σ_P , which we would like to resemble the distribution of f_d local to m_0 . Aiming to use a projection in logarithmic space, we wish to form a symmetric matrix A such that, for all \hat{m} in the vicinity of m_0 ,

$$-\frac{1}{2}(\hat{m} - m_0)^T A (\hat{m} - m_0) \approx \log f_d(\hat{m}) - \log f_d(m_0). \quad (4.31)$$

As this relation is linear in the matrix coefficients A_{ij} we can optimize it using the weighted linear least squares method, in which we reuse the sequence $\{\hat{m}_i\}$ of Nelder-Mead iterates as data points, and $f_d(\hat{m}_i)$ as weights in order to downweigh the tails of the distribution. We note that, while it is convenient to reuse available data, we are at liberty to augment the series with extra evaluations in the vicinity of the optimum to increase the quality of the projection, even though we have found no need to do so for the cases considered.

Finally, we use the optimal scaling result of Roberts et al. [59] to form the covariance matrix of our proposal distribution,

$$\Sigma_P = \frac{(2.38)^2}{\#\mathcal{M}} A^{-1}. \quad (4.32)$$

4.3.5 Evaluating the posterior expected slip

The expected value $E(B|MD)$ of the slip distribution given a fault m and surface measurements d is provided in closed form by Equation (4.17). The posterior covariance matrix is defined in Equation (4.16), in which Σ_B is the identity matrix owing to the properties of the Karhunen-Loève expansion. Crucially, both involve the formation of the forward model F_m , which maps the coefficient vector that encodes the slip distribution onto the corresponding vector of surface deformation gradients. It follows that the rows of the matrix F_m are formed by the surface deformations corresponding to the slip distribution that is represented by the individual basis vectors $h_n(\xi)$ that we constructed in Section 4.3.2.

If the selected forward model is Volterra’s equation then F_m is formed by repeated evaluation of Equation (4.4). If the selected model is the WSM then constructing F_m involves constructing a finite element matrix and solving it for a block of right-hand-side vectors. Having discussed the evaluation of Volterra’s equation in Section 4.3.3, we will use the remainder of this section to elaborate on details of the latter.

A first step in any finite element computation is the formation of the computational mesh on which the discrete basis is formed, in our case to describe the displacement field. Recall from Section 4.3.1 that all fault planes \mathcal{M} will be confined in a rectangular box of given size. We now create a regularly spaced grid of elements spanning this search box, allowing us to cheaply trace any physical coordinate inside the box to the containing element and its element-local coordinate which will greatly aid the efficiency of the fault plane integration. We note, however, that efficient lookup procedures exist for other mesh types as well, for example using quad trees [45] or alternating digital trees [10].

Since our computational domain is a halfspace we have no boundary conditions to place on the walls of the search box, except for the free surface which is traction free. Instead we take the infinite element approach of extending our mesh with several rows of extra elements and using a geometric map to continuously stretch the elements outside the search box towards infinity. Specifically, if $2L$ is the width of the box, $2n_{\text{box}}$ the number of elements spanning the box and $2n_{\text{inf}}$ the number of elements spanning infinity, we apply the following piecewise hyperbolic map to every spatial dimension:

$$x_i(e) = \frac{L}{n_{\text{box}}} \begin{cases} e - (n_{\text{box}} + e)^2 / (n_{\text{inf}} + e) & -n_{\text{inf}} < e < -n_{\text{box}} \\ e & -n_{\text{box}} \leq e \leq +n_{\text{box}} \\ e + (n_{\text{box}} - e)^2 / (n_{\text{inf}} - e) & +n_{\text{box}} < e < +n_{\text{inf}} \end{cases} \quad (4.33)$$

Note that this includes the depth direction, in which case we take $-n_{\text{inf}} < e \leq 0$. Unless stated otherwise we will select a infinity-to-box ratio of $n_{\text{inf}}/n_{\text{box}} = \frac{3}{2}$, which means that in 3D the treatment of the far field increases the number of elements by a factor $(\frac{3}{2})^3 \approx 3.38$ relative to the number of elements in the search box.

In creating the discrete function space V_n we make use of the fact that our mesh is structured by creating a C^1 quadratic spline basis, also known as isogeometric analysis [39], which we showed [77] to have better accuracy to degrees of freedom, and we remove the outermost basis functions to impose the far field constraint. With that we are in a position to evaluate the left hand side of Equation (4.6) to form the stiffness matrix. Since this matrix will be reused many times we also invest the

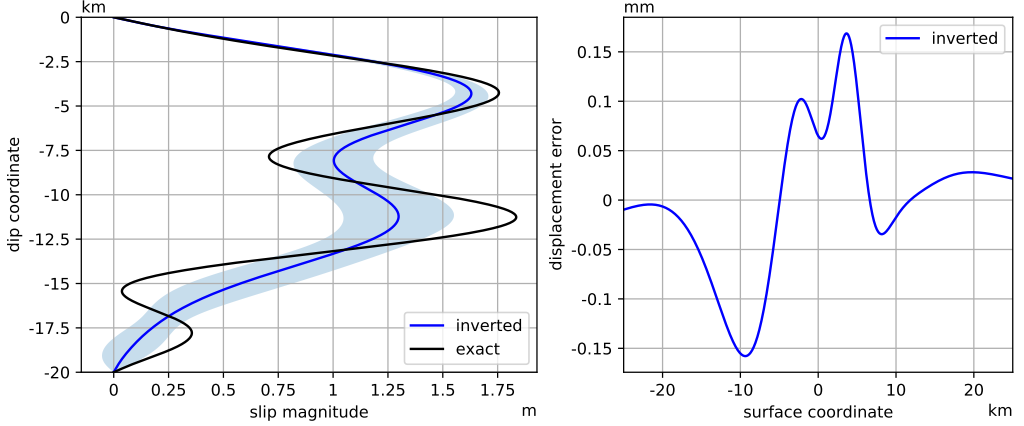


Figure 4.5 – Linear inversion of the slip distribution in the 2D non-rupturing scenario of Figure 4.4 using Volterra’s equation. Left: The exact slip distribution b , the inverted slip distribution $b'(m, d)$, and the one standard deviation, 68% confidence interval $\pm \sqrt{h^T \Sigma'_B(m) h}$. Right: The difference between the exact surface deformation and the deformation that corresponds to the inverted slip distribution $b'(m, d)$ using Volterra’s equation as the forward model.

time to construct a high quality preconditioner, opting, in fact, to form a complete Cholesky decomposition.

The right-hand side of Equation (4.6) involves an integral over \mathcal{X} for which we use the same Clenshaw-Curtis quadrature scheme that we used for the synthetization of observation data in Section 4.3.3, while making use of the rectilinearity of the mesh in the search box to locate the corresponding element coordinates necessary to evaluate the basis functions. Having formed the right-hand side vector we can solve the system and form the discrete solution $u_h \in V_n$, which can then be evaluated in any point of our choosing.

4.4 Results

We will present several results of the process outlined in Section 4.3, with the aim of illustrating the many variables and their effect on the overall computation. While we are mainly interested in applications in three-dimensional space, we find that most computational aspects appear identically in the two-dimensional analogue. Appreciating the advantages for visualisation we therefore present most of our observations in this setting, adding 3D results mainly to confirm these findings. For structure we will use the scenarios of Figure 4.4 as a baseline test case, with minor modifications where required.

4.4.1 Linear inversion: slip distribution

We will study first the linear inversion process, in which we keep the fault parameters m equal to the exact values and invert the slip distribution b only — referring to the

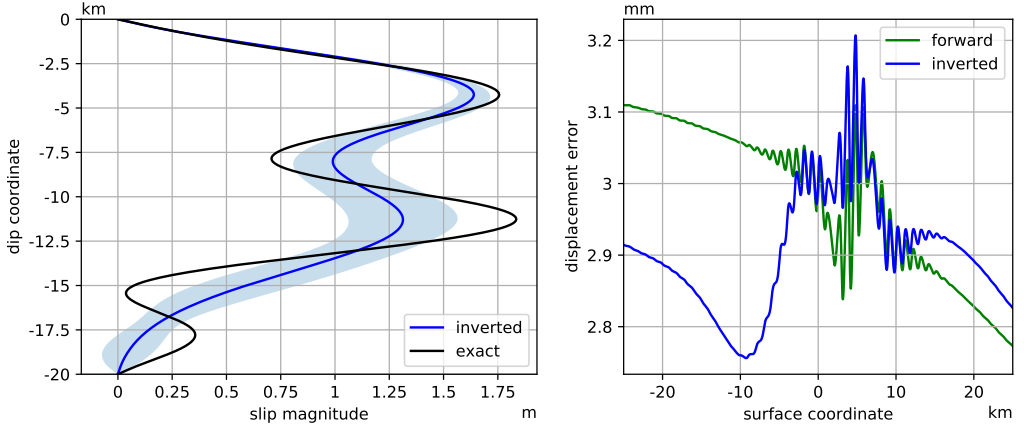


Figure 4.6 – Linear inversion of the slip distribution in the 2D non-rupturing scenario of Figure 4.4 using the WSM forward model on a 76×38 element mesh, of which 50×25 elements form the search box. The graph layout is identical to that of Figure 4.5, with the addition of the ‘forward’ deformation error that corresponds to the exact slip distribution, rather than the inverted slip distribution.

methodology of Section 4.3 we set $\tilde{m} = m$ in step 4. A natural starting point is to establish the best case solution by inverting using Volterra’s equation, the forward model that was used to synthesize the data. The results of this are shown in Figure 4.5. While the inverted slip distribution does not match the exact slip due to the smoothing effect of the prior distribution, we observe in the left panel a reasonable fit that is in keeping with the posterior variance. In the right panel we observe that the deformation error stays well below the 1 mm measurement noise standard deviation.

Repeating this process with the WSM on a 76×38 element mesh we obtain the result of Figure 4.6, showing that the expected slip and standard deviation are almost identical to those obtained using Volterra’s equation. Paradoxically, the corresponding deformation error in the right panel (‘inverted’) is relatively large. It is noteworthy that the deformation error is characterized by a significant offset (approximately 3 mm) with a relatively small variation (approximately 0.2 mm). A similar offset can be seen in the (‘forward’) error that the model produces with the exact slip as input, thus representing the discretization error for this particular computational setting. From this we can conclude that the offset does not result from the inversion process, but is in fact a side effect of the discrete model. Fortunately, by virtue of the differencing approach layed out in Section 4.3.3, the inversion is insensitive to offsets of this kind. We observe that the errors are small relative to the offset, with a peak to peak error range that is well below 1 mm, which explains the perceived paradox.

In addition to the 3 mm offset, the forward error curve of Figure 4.6 shows a distinct spatial trend, dropping by 0.35 mm over the length of the domain. Both aspects of the discretization error are studied in Figure 4.7, which shows two variations of the mesh resolution. On the left we see the effect of increasing the resolution in the far field while keeping that in the search box fixed. Comparing to Figure 4.6, we observe that both the offset and the trend are greatly reduced, indicating that these phenomena are caused largely by the treatment of the far field. At the same time

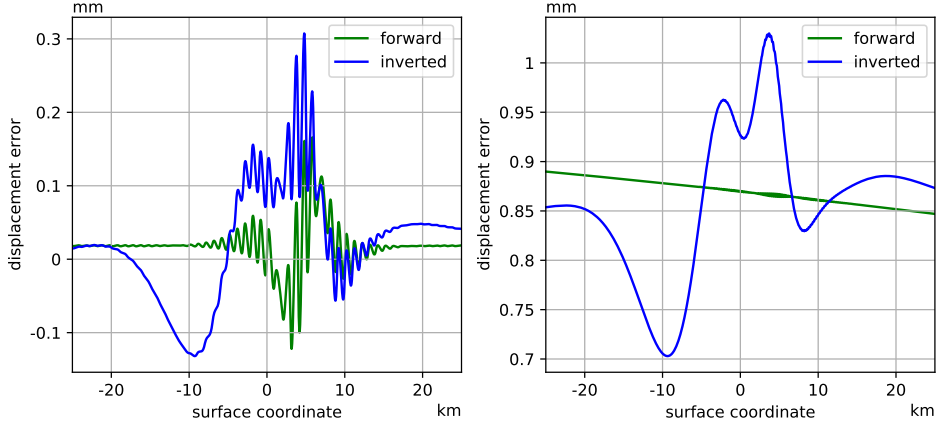


Figure 4.7 – Linear inversion of the slip distribution in the identical setting to Figure 4.6 except for different finite element meshes. Referencing Equation (4.33), the baseline mesh was constructed for $n_{\text{box}} = 25$ and $n_{\text{inf}} = 38$, resulting in a 76×38 element mesh of which 50×25 elements form the search box. Left: $n_{\text{box}} = 25$ and $n_{\text{inf}} = 200$, keeping the search box resolution fixed while adding elements to the far field. Right: $n_{\text{box}} = 100$ and $n_{\text{inf}} = 150$, keeping the infinity-to-box ratio fixed while increasing resolution by a factor 4.

the errors did not change significantly relative to the offset, confirming that this far field-induced error should not strongly affect the inversion. On the right we see how an eight-fold uniform mesh refinement results in a stark reduction of the discretization error, and in an inverted deformation error that closely resembles that of the baseline result of Figure 4.5 modulo the remaining offset and trend.

Taken together, the results of Figure 4.7 uphold our result [77] that the discretization error can be made arbitrarily small via mesh refinement. However, as we have already seen, it is by no means necessary to drive the error orders below the noise level of the deformation measurements. To see what happens in the opposite direction, Figure 4.8 compares the Volterra and WSM-based inversion at a noise level that is 100 times smaller while maintaining the mesh. While Volterra’s equation correctly tightens the error margin around the exact slip, the WSM-based inversion deviates significantly from the exact slip due to the dominant numerical error. At this noise level it takes an eightfold uniform mesh refinement for the WSM-based inversion to return to being indistinguishable from the Volterra based inversion, beyond which the inversion is essentially mesh-independent. Based on these results, we consider that a mesh at which the discretization error does not exceed half the standard deviation of the measurement noise appears to strike a good balance between accuracy and numerical efficiency.

There is one situation where we cannot control the discretization error through mesh refinement, which is in the case of a rupturing fault. As the approximation is inherently continuous, the error at the point of intersection equals half the slip magnitude regardless of element size. Since this violates the established rule that the discretization error may not exceed the measurement noise, care must be taken to avoid the detrimental effects we observed in the right panel of Figure 4.8. Arguably the simplest way to achieve this is to discard measurements close to the rupture and

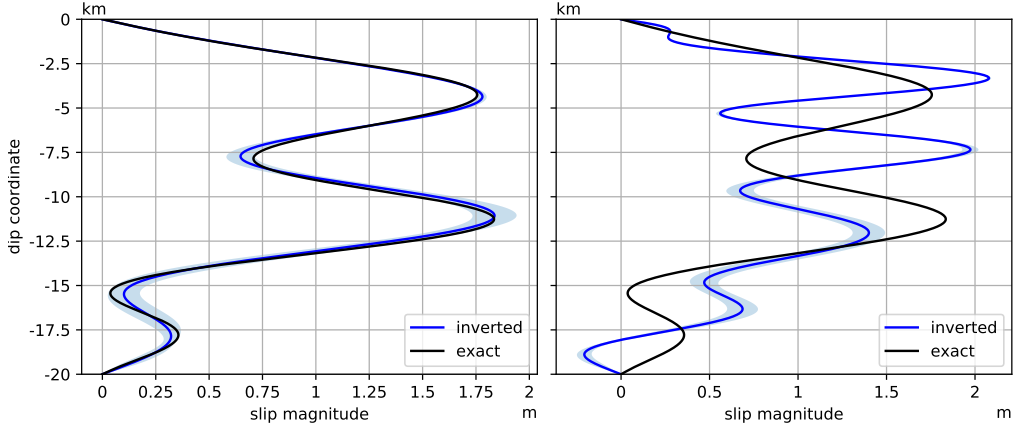


Figure 4.8 – Linear inversion of the slip distribution in the identical setting to Figure 4.6 except for a $100\times$ reduction of the measurement noise to 0.01 mm. Left: Volterra-based inversion. Right: WSM-based inversion on a 76×38 element mesh, of which 50×25 elements form the search box. This is an illustration of the adverse effects on the inversion when the noise level undercuts the discretization error.

use only the remaining intermediate to far field data. Incidentally, the masking out of data in the rupture zone is not uncommon in the context of SAR interferometry, as local destruction tends to lead to decorrelation of the radar signal. We hypothesize, therefore, that no valuable data need be discarded in practice.

An example of a rupturing fault can be seen in Figure 4.9. The right panel shows the absolute displacement (rather than the displacement error) in which we observe continuous oscillations at the 5 km position where the exact displacement exhibits a discontinuity. The oscillations decay rapidly, reaching sub-millimeter scale amplitudes at a 5 km distance from the surface rupture. The left panel shows the inversion result based on the deformation data outside of this ± 5 km interval that is marked gray in the right panel. The result accurately recovers the exact slip distribution, and is virtually indistinguishable from the Volterra-based inversion (not displayed) subject to the same data mask, confirming that data masking is a suitable strategy to deal with the continuous representation of discontinuities in the WSM. We note that since WSM displays optimal convergence away from the dislocation [77, Cor.7] the required masking zone can be made arbitrarily narrow by sufficiently refining the mesh.

So far we have drawn a slip distribution from the prior distribution, the same that is subsequently used in the inversion procedure to reconstruct the slip from measurements. Since it is difficult in practice to accurately capture prior knowledge in terms of a distribution, it is relevant to study the robustness of the procedure to slip distributions not being elements of our discrete space \mathcal{B} . Examples of this can be seen in Figure 4.10, which shows two Gaussian slip distributions, one rupturing, the other non-rupturing. Though neither is a member of \mathcal{B} , both distributions are recovered with reasonable accuracy, confirming that the methodology has at least some lenience to inadequacies in the choice of the prior. This also confirms our premise that the size of the fault plane need not be an independent parameter if we have reasonable upper bounds, as the areas of zero slip are captured accurately.

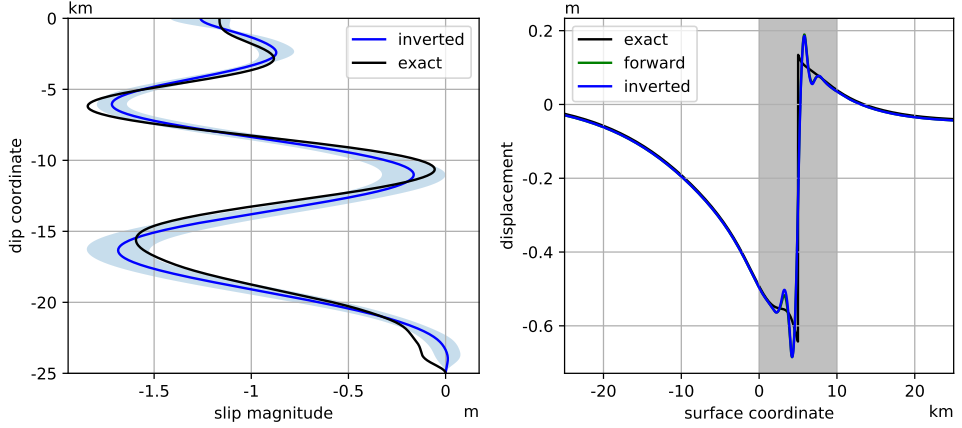


Figure 4.9 – Linear inversion of the slip distribution in a 2D rupturing scenario. The graph layout is similar to that of Figure 4.5, with the difference that the right panel shows the absolute displacements rather than the displacement errors with the black ‘exact’ curve representing the synthesized displacement field. The gray band in the the right panel corresponds to the area that was masked out in order for the locally meter-scale errors not to affect the inversion.

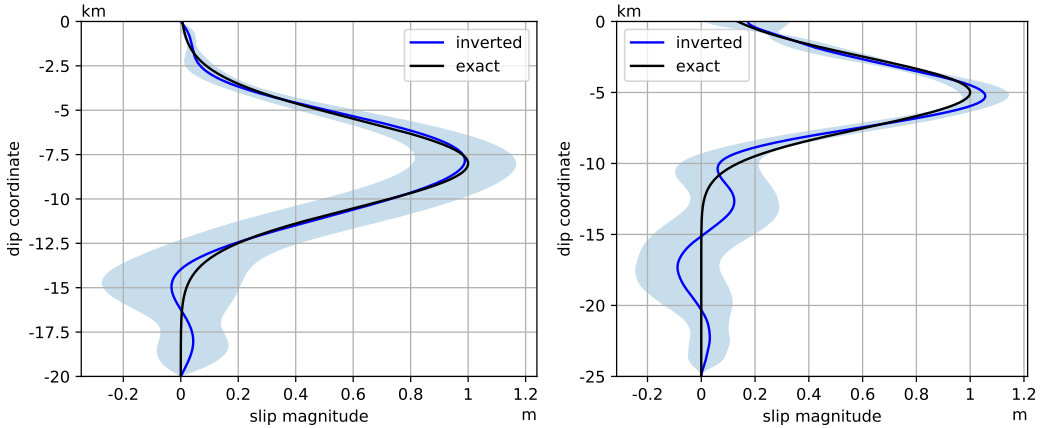


Figure 4.10 – Linear inversion of a manually constructed Gaussian slip distribution in a 2D scenario, solved on a 76×38 element mesh. Left: a non-rupturing scenario with the slip centered at 40% dip. Right: a rupturing scenario with the slip centered at 20% dip and a masked area of 5 km.

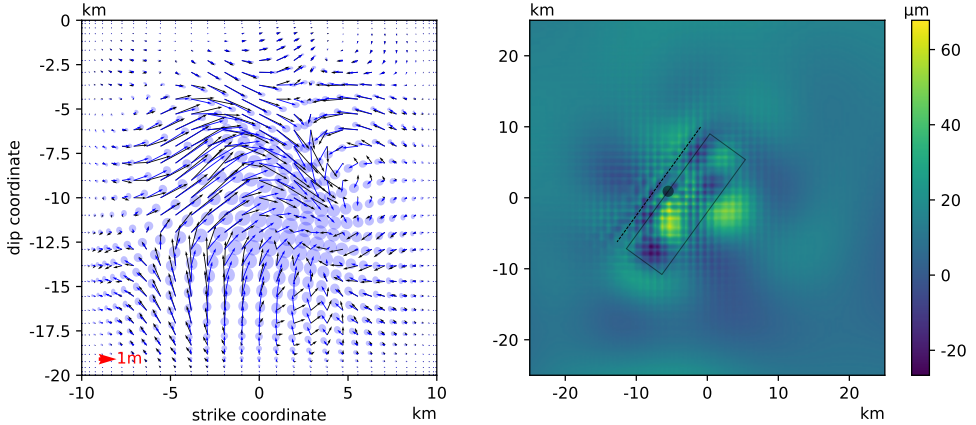


Figure 4.11 – Linear inversion of the slip distribution in a 3D non-rupturing scenario using the WSM forward model on a $76 \times 76 \times 38$ element mesh, of which $50 \times 50 \times 25$ elements form the search box. Left: The exact slip distribution b as black quivers, the inverted slip distribution $b'(m, d)$ as blue quivers, and the local one standard deviation or 47% confidence region resulting from the 2×2 posterior covariance matrix $h^T \Sigma'_B(m) h$ as blue ellipses, centered at the expected value and at matching scale. Right: the vector norm of the deformation gradient error. The solid rectangle shows the outline of the fault plane.

Finally we turn to the 3D scenario (right panel) of Figure 4.4. Employing the identical methodology, Figure 4.11 shows the slip distribution and the corresponding error in deformation gradient, in which we recognize similar patterns to those we observed in the direct 2D equivalent of Figure 4.6. Taking into account the local standard deviation, the inverted and exact slip distributions are in good agreement. We also confirmed that the result is indistinguishable from that obtained via Volterra’s equation. The error is largest in the vicinity of the fault, but stays well clear of the 1 mm noise level. Interestingly, the error offset that we observed in the 2D results is much less pronounced in the 3D situation.

4.4.2 Nonlinear inversion: fault parameters

We proceed by studying the nonlinear inversion of the fault parameters. Referring again to the methodology of Section 4.3, in step 4 we now evaluate the posterior expected value for the fault parameters m using the Metropolis-Hastings MCMC process. This process samples the posterior distribution through repeated evaluation of $f_d(m)$ as defined in Equation (4.30), being proportional to the posterior probability density function $f_{M|D}$. As this entails evaluation of the expected value $b'(m, d)$, all aspects of the linear inversion process as explored in the previous section remain in effect.

Returning to the non-rupturing 2D baseline scenario (left panel) of Figure 4.4, Figure 4.12 shows the results of a MCMC process comparing Volterra’s equation to the WSM. We use the same relatively coarse mesh that we used for Figure 4.6 to see if there are adverse effects in pushing against the boundary of the discretization error. Both distributions are seen to capture the fault parameters correctly, pinpointing the

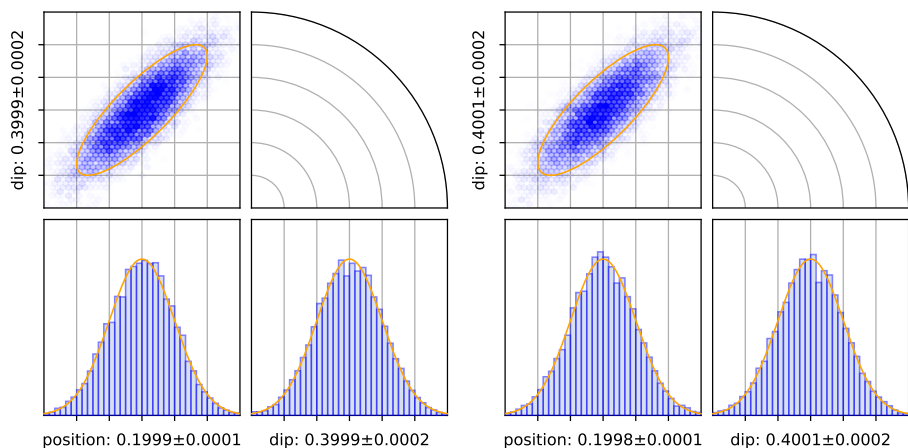


Figure 4.12 – Binned results of a 50,000-sample MCMC process for $M|D$ using Volterra’s equation in the 2D non-rupturing scenario of Figure 4.4, which has fault parameters position= $0.2L$ and dip angle= 0.4π . Left: reference results from Volterra’s equation. Right: results from the WSM on a 48×24 element mesh, of which 32×16 elements form the search box. The bottom row shows the marginalized distributions for fault position (left) and the dip angle (right), with axis labels showing mean value \pm standard deviation or, equivalently, the 68% confidence interval. The orange overlay shows the corresponding normal distribution. Grid lines are spaced at one standard deviation. The top row shows the cross correlation of x coordinate and dip angle, with the orange overlay showing the bivariate normal distribution at two standard deviations or, equivalently, the 91% confidence region.

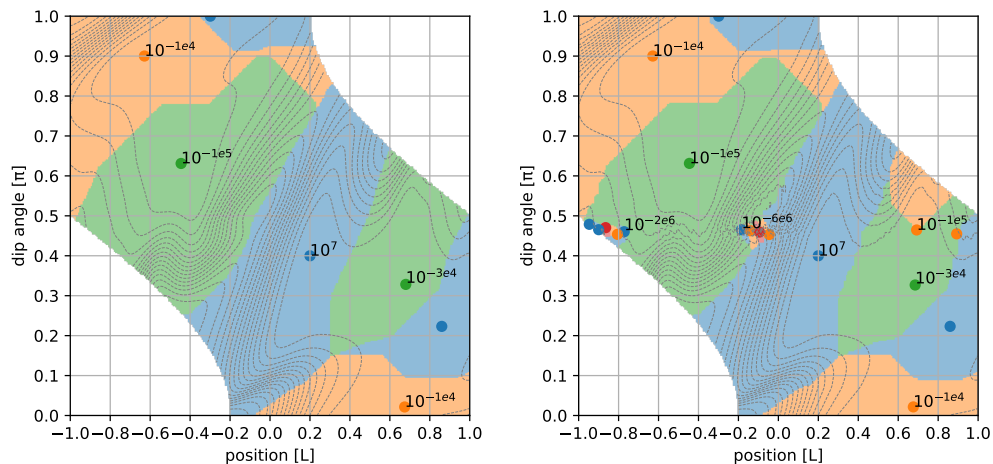


Figure 4.13 – Contour lines and local maxima of $f_{M|D}$ for the same 2D non-rupturing scenario as that of Figure 4.12. The coloured areas represent (inverse) watersheds of the probability landscape, that is, the collection of points from which an uphill gradient method converges to the associated local maximum. Left: reference result from Volterra’s equation. Right: results from the WSM on a 48×24 element mesh. Powers smaller than -100 are shortened using scientific notation, denoting e.g. 10^{-31826} as 10^{-3e4} .

exact values to a high degree of accuracy both in terms of a nearly exact expected value and of the equally narrow confidence interval. Furthermore, both distributions are in excellent agreement to each other, demonstrating the robustness of the method with regard to discretization errors.

We started the MCMC random walk from the global maximum, obtained via the Nelder-Mead uphill simplex method, that in turn was started from where we know the exact solution to be. Knowledge of an exact solution is a luxury that is not available in any practical application, which means a global search algorithm will typically be required to prepare for the final gradient ascent. A relevant question, therefore, is whether the WSM has a disturbing effect in this regard. To explore this, Figure 4.13 compares the posterior probability density function obtained through direct evaluation of Volterra’s equation against that obtained via the WSM, identifying local maxima as well as the associated watersheds. While the WSM introduces some spurious local maxima, the global maximum as well as its watershed appear identical, suggesting that there is no difference with regard to global optimization strategies.

For the rupturing scenario we mask out an area of 7.5 km centered at the point of rupture. While it may seem contradictory to mask out the rupture zone while simultaneously inverting for the rupture coordinate, this process should be understood in the context of having prior information in the form of in situ observations. Even though the precise trajectory of the rupture may not be known to sufficient accuracy, it may well be sufficient to define a masking zone. In fact, though not employed here, we are at liberty to modify the prior distribution to have it reflect this knowledge as well. Selecting the uniformly refined grid of Figure 4.9, the 7.5 km zone is 50% wider than the observed minimum in order to not artificially limit mobility of the rupture coordinate, but rather give it some freedom to find its optimum within the confines of the broader mask.

Figure 4.14 shows the side by side results of Volterra’s equation and the WSM. Note that the mask was applied to Volterra’s equation as well for sake of comparison, even though the method does not require it. The distributions in the rupturing scenario are less precise as a result of data masking, but are otherwise in excellent agreement.

Figure 4.15 again explores the posterior probability density function, where this time we see a very large qualitative difference between Volterra’s equation and the WSM. While both show a clear delineation at $(0.2 \pm 0.3)L$, corresponding to the applied data mask, the WSM produces many more local maxima, clustering in particular at the crossover points and at shallow dip angles. The global maximum still has a fairly large associated watershed, however, suggesting that the multitude of local maxima is not necessarily problematic in a global optimization context. Note also that the global optimization algorithm needs only consider positions inside the masked region — the non-shaded region in the figure — as ruptures outside of the mask are in violation of its premise.

We conclude again with the 3D scenario (right panel) of Figure 4.4. Using the same mesh as was used for the slip inversion of Figure 4.11, Figure 4.16 shows the posterior distribution of the 3D fault parameters. One can observe that the expected values accurately match the parameters that were used to generate the synthetic data. The position along strike has a markedly larger variance than that perpendicular to it, which matches the expectations laid out in Section 4.3.1 relating to ambiguities

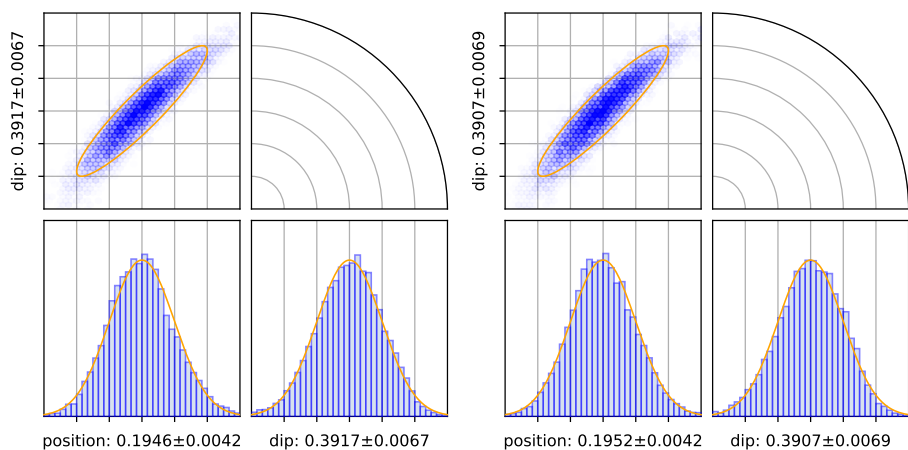


Figure 4.14 – Binned results of a 50,000-sample MCMC process for $M|D$ using the WSM forward model in a 2D rupturing scenario with the same parameters as in Figure 4.12: position= $0.2L$ and dip angle= 0.4π . A radius $0.3L$ data mask is applied centered at $x=0.2L$. Left: reference results from Volterra’s equation. Right: results from the WSM on a 96×48 element mesh, of which 64×32 elements form the search box.

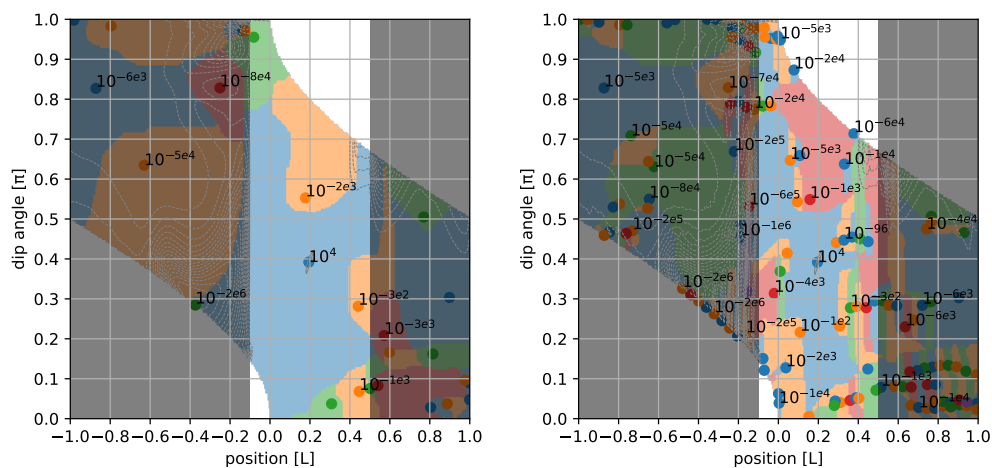


Figure 4.15 – Contour lines, local maxima and watersheds of $f_{M|D}$ for the same 2D non-rupturing scenario as that of Figure 4.14, with an overlay delineating the masked region. Left: reference result from Volterra’s equation. Right: results from the WSM on a 96×48 element mesh. Powers smaller than -100 are shortened using scientific notation, denoting e.g. 10^{-31826} as 10^{-3e4} .

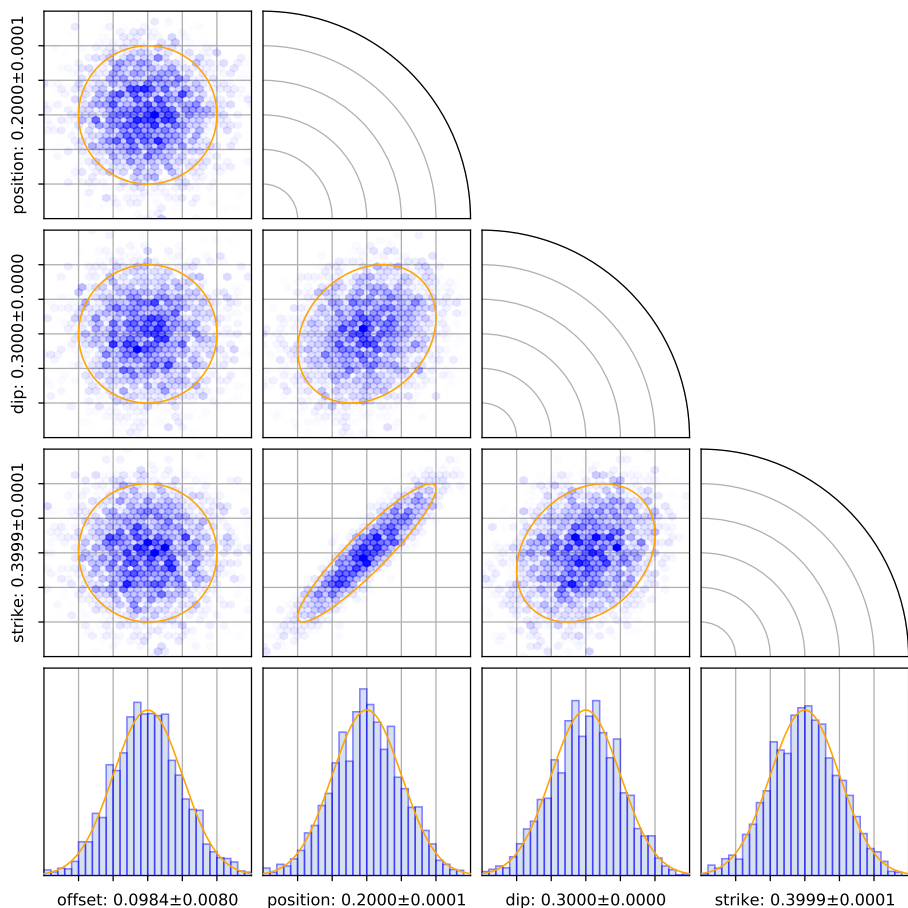


Figure 4.16 – Binned results of a 10.000-sample MCMC process for $M|D$ corresponding to the 3D non-rupturing scenario of Figure 4.4, which has fault parameters $\text{offset}=0.1L$, $\text{position}=0.2L$, $\text{strike angle}=0.3\pi$ and $\text{dip angle}=0.4\pi$. The bottom row shows from left to right the marginalized distributions for fault offset, position, strike and dip, with axis labels showing mean value \pm standard deviation or, equivalently, the 68% confidence interval. The orange overlay shows the corresponding normal distribution. Grid lines are spaced at one standard deviation. The remaining rows show from top to bottom the fault parameters y , strike and dip, thus covering all cross correlations. The orange overlay shows the bivariate normal distribution at two standard deviations or, equivalently, the 91% confidence region.

with the slip distribution. Though the high computational costs involved in evaluating the Volterra solution in a 3D space prohibit running the MCMC process using Volterra’s equation to verify the correctness of this result, we consider the foregoing to be sufficient support to present this as a demonstration of the WSM driving a realistic, nonlinear inversion of a 3D fault plane.

4.5 Conclusions

In this paper we performed for the first time a full inversion of fault plane parameters and fault slip distribution using the Weakly-enforced Slip Method (WSM) that was developed explicitly for this purpose. By restricting the domain to a homogeneous halfspace we were able to synthesize the deformation data and to compare the WSM-based inversion against a reference result obtained from the exact solution. This allowed us to study in detail the effect that the discretization error has on different aspects of the inversion process. To provide our study with a mathematical framework we placed the inverse problem in a Bayesian setting, in which a prior probability distribution is refined through observations into a posterior probability that quantifies the likelihood of various faulting mechanisms.

For linear inversions, the WSM was found to be competitive with Volterra’s equation in terms of accuracy, showing excellent agreement already at coarse meshes (5 elements per 10 km for the situation considered) in case of non-rupturing faults. As a practical rule of thumb for the minimum required mesh density, we demonstrated empirically that the discretization errors must not exceed the standard deviation of the measurement noise in order to avoid large numerical errors. Conversely, increasing the mesh density beyond this point contributes little to the accuracy of the inversion.

A WSM-based inversion of rupturing faults requires additional measures to account for the local smearing out of the discontinuity, but it is argued that similar measures are often required in practice regardless. When a simple data mask is applied to disregard data points in the vicinity of the rupture, the WSM and exact method again show excellent agreement, albeit at a finer mesh that is required to localize the discretization error to the rupture zone. Local to the rupture the error is observed to decay exponentially, at a rate that is inversely proportional to the element size of the computational mesh. It is expected that this relation holds as well in the case of local, rather than uniform, refinements.

While the WSM was originally analyzed on a finite domain with exact boundary conditions [78], real world applications cannot rely on the availability of such data. Instead of introducing artificial boundaries, we opted for a finite-to-infinite mapping for the treatment of the far field in order not to introduce assumptions that might limit the validity of our results. Though this treatment introduces a fairly significant error, we observe that the relative displacement error of two nearby points remains dictated by the local element size. This circumstance fits remarkably well with the fact that satellite-based InSAR observations are inherently relative, which means that treatment of the data must be insensitive to global offsets. While successful in this regard, meshing the far field is arguably an expensive solution, increasing the number of degrees of freedom by a factor 3.38 in the test cases considered. Further study in this direction is therefore warranted.

For non-linear inversions, any global search algorithm followed by a gradient-based optimization is shown to perform equally well for the WSM as it does for the exact forward method, as demonstrated by a full comparison of the posterior probability density. Furthermore, we demonstrated that the iterates of the Nelder-Mead uphill simplex method can be reused in a linear least squares projection to provide a high quality Gaussian proposal distribution for a subsequent exploration of the posterior probability using the Metropolis-Hastings Markov Chain Monte Carlo method.

In an observation that is unrelated to the WSM we have remarked that certain parameters that are customarily added to the space of fault parameters, such as the fault plane dimensions, can instead be captured at lesser cost by the slip distribution. In situations where an ambiguous relationship remains between a fault parameter and the slip distribution, this translates to a large variance of the posterior distribution, as demonstrated by the along-strike offset of the fault plane in the 3D scenario.

In conclusion, we believe that the present work convincingly demonstrates the utility of the WSM in real world applications, combining the power and flexibility of finite element analysis with a highly efficient reuse of computational effort. It also provides a practical framework by which such studies can be performed. While the current experiments have been restricted to homogeneous halfspaces for reasons of verification, none of these restrictions were required by the methodology as it is presented here; nor do we have reason to believe that our findings are limited to these conditions.

5 | Conclusions and Outlook

The objective of this thesis has been to develop a fast but versatile numerical method for the simulation of dislocations in general elastic domains, to be used in the specific inverse setting of the inference of earthquake mechanisms from InSAR data.

This objective was met with the introduction of the Weakly-enforced Slip Method (WSM) in Chapter 3 and its application in a Bayesian inversion setting in Chapter 4. The following section will substantiate this by revisiting the research questions. This is followed by recommendations for future research.

5.1 Research Questions

The research was guided by four research questions, which will be answered here with reference to the relevant sections.

Solution space

Compared to the split-node method, what alternative decisions could be considered in devising a method for the simulation of dislocations while staying within the broader framework of the Finite Element Method?

A key design decision of the split-node method, discussed in Section 2.5.2, is to align the dislocation to element boundaries. The reasons for doing so are compelling. Firstly, the per-element assembly routine is identical to that of a non-dislocated medium, which means that the implementation is near standard with adjustments only in the global assembly. Secondly, as a dislocation implies not only a discontinuity but also discontinuous normal gradients, the C^1 -discontinuities (kinks) of a typical FEM basis at the interface between elements allow for excellent approximation of the displacement field local to the dislocation.

The general finite element formulation was reached in Section 3.1.3, which noted that the lift is “in principle arbitrary”. In other words, the link with element boundaries is not mandated by any underlying theory, and alternative methodologies are possible even though the section concludes that “the use of an arbitrary lift carries severe algorithmic complexity”. For example, the ε -local lift of Figure 3.2 can be constructed independently of the finite element mesh, but the required distance metric is nontrivial for curved faults and the resulting assembly non-standard.

Algorithmic complexity might be a price worth paying, however, when it allows many different fault geometries to be considered on a fixed finite element mesh. Since

the dislocation manifests only in the right-hand side of (3.23), the most expensive steps such as matrix assembly and factorization, as well as mesh generation itself, can be reused between simulations. On balance, therefore, the resulting method will be significantly cheaper in a multi-query setting, even if integration of the lift is more expensive than that of the split-node method.

Speed

Is it possible to devise a computational method that is fast enough to be feasibly used in iterations mandated by the inversion procedure? How does computational effort compare to existing techniques?

As typical inversion procedures require many evaluations of the forward model, a mesh-independent lift will result in an overall significantly cheaper procedure as it allows substantial parts of the computation to be reused, such as the stiffness matrix and its factorization. Remaining steps per evaluation are assembly of the right-hand-side vector and solving of the linear system. To minimize costs of assembly, Section 3.2.1 presents the Weakly-enforced Slip Method (WSM) as the limit of a sequence of lifts with increasingly narrow support, which reduces assembly to an integral over the dislocation manifold.

Assembly of the WSM is still non-standard, involving the evaluation of basis functions in arbitrary domain points, rather than in the usual Gaussian quadrature points. But there exist efficient strategies for this, with Section 4.3.5 providing some pointers. Furthermore, compared to situations dealing with any non-degenerate lift, the required number of quadrature points is greatly reduced owing the fact that the manifold is a lower dimensional structure. Finally, evaluation of the integrand is cheap, involving only gradients of the basis functions, the local stiffness tensor, the normal vector, and the Burgers vector.

To compare the computational effort of the WSM to existing techniques, one should not compare with the split-node method but rather with analytical expressions, such as the widely used Okada's solution detailed in Section 2.2.3. The difference in nature between both methods make a direct comparison difficult. Whereas the cost of analytical expressions scales linearly with the number of surface points evaluations, the WSM is virtually insensitive to the number of surface points, but scales instead with mesh density. As such, there is a cross-over point where the WSM becomes the cheaper method. Complicating matters further, analytical solutions are usually superposed to simulate slip distributions, scaling the costs of per-point evaluations and sliding the cross-over point in favour of the WSM.

Regardless of the above, based on experience gathered in generating the results of Section 4.4 the WSM is estimated to be the slower method in most realistic scenarios. This means that in situations for which, for instance, a homogeneous halfspace is a sufficiently good model, analytical solutions are likely the best choice. Crucially, though, the study of Chapter 4 has proven the WSM to be a feasible method for the inversion of realistic three-dimensional displacement data, meaning that model simplifications are no longer a prerequisite for computability.

Flexibility

Is it possible to make significant speed improvements while keeping the power and flexibility of the Finite Element Method intact? If not, what are the limitations of the new method?

The WSM achieves its speed improvements by severing the connection between the manifold and the finite element mesh, without imposing restrictions of any type in its stead. Any mesh that is suited for the FEM, therefore, can be used for the WSM, granting full flexibility for describing surface topology and other relevant structures. Material heterogeneity, including anisotropy, is supported within the confines of linear elasticity theory. Finally, the dislocation manifold can be arbitrarily shaped, may intersect the surface, and supports any slip distribution including out of plane components.

The approximation proofs of Section 3.3 are subject to moderately restrictive conditions: the elasticity tensor is assumed to be C^1 continuous on the entire domain, including across the dislocation; and the domain is assumed to be Lipschitz, placing regularity constraints on the shape of the dislocation. The regularity constraint is unlikely to be limiting in practice. The continuity condition is not debilitating either, though it may be inconvenient to adhere to. There are reasons to expect that this condition may be relaxed to include sharp and even discontinuous feature transitions, but further study is required to confirm this.

The main limitation of the WSM ties in with its main strength, which is that its computational efficiency derives from the fact that the stiffness matrix can be reused. Geologically, tectonic faults may represent discontinuities not only in the displacement field, but also in the material properties due to the nature of the tectonic process. However, as this would reintroduce a link between the fault and the stiffness matrix, most of the WSM's advantages would vanish if this effect were to be modelled accurately. Instead, feasibility demands that the transition be averaged to the general area where the fault is thought to be.

Errors

If speed gains come at the expense of accuracy, essentially replacing model errors by discretization errors, how will these errors affect the accuracy of a typical inversion?

Corollary 8 of Section 3.3.3 states that the WSM converges optimally on any subdomain that excludes the dislocation manifold. This directly implies that, for any domain point that does not lie directly on the manifold, the displacement error can be made arbitrarily small through mesh refinement, to the point where it will no longer significantly affect the inversion. As this would require an infeasibly fine mesh, however, this is not a useful argument in practice. A characterization of the WSM's discretization error and its influence on the inversion process is therefore warranted.

The absence of a discontinuity introducing lift, combined with the inherent continuity of the discrete solution space, implies that the WSM solution is continuous. Consequently, the error at the dislocation equals half the Burgers vector, regardless of mesh size. At the same time, Corollary 8 states that the error at any distance converges optimally, suggesting the formation of an error cone that narrows with element size. This is experimentally confirmed by Figure 3.6, which shows progressively narrowing error cones, pointwise exponential convergence, as well as spatially

exponential error decay.

The error decay suggests that sufficiently buried faults will not be affected by the peculiar nature of the WSM's discretization error, as the surface measurements are naturally removed from the manifold. This is confirmed numerically in Section 4.4, which found that near-optimal results could be obtained with element sizes as coarse as 1 km. In a Bayesian interpretation, it was found empirically that the largest deformation error should be less than the sensor noise for its influence on the inversion to become insignificant.

Rupturing faults, on the other hand, will require special treatment as the error cone will affect a region around the surface rupture. The proposed solution, to disregard data in this region and perform the inversion based on the remaining intermediate and far field data, was numerically shown to be effective. While masking data a priori may appear conflicting with the goals of establishing a fault geometry, the nature of a rupturing fault is that it is a visible feature, and as such locally well constrained. The approach is further thought reasonable in the context that InSAR data is often less reliable in regions of considerable damage, leading some of it to be discarded, and that the width of the masking zone can be controlled via mesh refinement.

5.2 Recommendations

This thesis recommends four topics for further study.

1. The convergence proofs of Section 3.3.3 presently assume C^1 continuity of the elasticity tensor. Though this assumption can always be satisfied by means of a smooth approximation, doing so would add unwelcome complexity to the method. Particularly relevant are discontinuous material properties resulting from horizontal stratification, which the general Finite Element Method is well equipped to handle though layer-conforming meshing. It is considered worthwhile to investigate if the continuity assumption is a requirement for the WSM, or whether it can be relaxed to accommodate this fairly common situation.
2. Section 4.4 showed that the inversion of rupturing faults requires a masking zone local to the rupture. The width of this zone can be controlled through mesh refinement. However, this thesis considered only uniform refinements, meaning that a narrowing of the masking zone implies a uniform increase in degrees of freedom, regardless of whether the additional detail is desired. It is expected that the same narrowing can be achieved through local refinement of the rupture zone, which would not substantially increase the problem size, but this needs to be confirmed in further study.
3. The WSM has several capabilities that were not explored in this thesis. Most prominently, it allows for material heterogeneity and anisotropy, neither of which were used in the case studies considered. The fault plane can be arbitrarily shaped, allowing for additional fault plane parameters such as curvature to be made part of the inversion. And the WSM allows for opening modes, making it relevant for other geophysical processes such as dyke intrusions. Further research is warranted to explore these topics in practice.

4. Finally, the inversion methodology of Section 4.3 took elastic properties to be a known quantity, which in most locations is not a realistic assumption. The Stochastic Finite Element Method is developed to solve problems that involve uncertainty, and it is reasonable to assume that its methods apply to the WSM as well. It appears to be worthwhile, therefore, to explore an extension to the Bayesian approach set forth in Section 4.2 which accounts for uncertainties in the material properties, so as to obtain more accurate posterior distributions.

Bibliography

- [1] R. Agata, T. Ichimura, K. Hirahara, M. Hyodo, and M. Hori T. Hori. Several hundred finite element analyses of an inversion of earthquake fault slip distribution using a high-fidelity model of the crustal structure. *Procedia Computer Science*, 29:877–887, 2014. doi:10.1016/j.procs.2014.05.079.
- [2] K. Aki and P.G. Richards. *Quantitative seismology*. University Science Books, 2002.
- [3] R.M. Allen, P. Gasparini, O. Kamigaichi, and M. Bose. The status of earthquake early warning around the world: An introductory overview. *Seismological Research Letters*, 80(5):682–693, 2009. doi:10.1785/gssrl.80.5.682.
- [4] A. Aspri, E. Beretta, and A.L. Mazzucato. Dislocations in a layered elastic medium with applications to fault detection. *Journal of the European Mathematical Society*, in press.
- [5] G.M. Atkinson. An overview of developments in seismic hazard analysis. *13th World Conference on Earthquake Engineering*, 2004.
- [6] E. Attema, Y.L. Desnos, and G. Duchossois. Synthetic aperture radar in Europe: ERS, Envisat, and beyond. *Johns Hopkins APL Technical Digest*, 21(1), 2000.
- [7] D.J. Barber, H.R. Wenk, G. Hirth, and D.L. Kohlstedt. Dislocations in minerals. In J.P. Hirth and L. Kubin, editors, *Dislocations in Solids*, volume 16, chapter 95, pages 171–232. Elsevier, 2010. doi:10.1016/S1572-4859(09)01604-0.
- [8] P. Bettess. Infinite elements. *International Journal for Numerical Methods in Engineering*, 11:53–64, 1977. doi:10.1002/nme.1620110107.
- [9] J. Biggs and T.J. Wright. How satellite InSAR has grown from opportunistic science to routine monitoring over the last decade. *Nature Communications*, 11(3863), 2020. doi:10.1038/s41467-020-17587-6.
- [10] J. Bonet and J. Peraire. An alternating digital tree (ADT) algorithm for 3D geometric searching and intersection problems. *Numerical Methods in Engineering*, 31(1), 1991. doi:10.1002/nme.1620310102.
- [11] S.C. Brenner and C. Carstensen. Finite element methods. In E. Stein, R. de Borst, and T.J.R. Hughes, editors, *Encyclopedia of Computational Mechanics*, volume 1: Fundamentals, chapter 4, pages 73–118. John Wiley & Sons, Ltd., 2004.

- [12] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, Berlin, 2nd edition, 2002.
- [13] F. Brezzi, T.J.R. Hughes, and E. Süli. Variational approximation of flux in conforming finite-element methods for elliptic partial differential equations: a model problem. *Rend. Mat. Acc. Lincei*, 9(12):159–166, 2001.
- [14] E.H. van Brummelen, K.G. van der Zee, V.V. Garg, and S. Prudhomme. Flux evaluation in primal and dual boundary-coupled problems. *Journal of Applied Mechanics*, 79(1), 2011. doi:10.1115/1.4005187.
- [15] D. Capecchi. *History of Virtual Work Laws*. Springer Milano, 1st edition, 2012. doi:10.1007/978-88-470-2056-6.
- [16] P. Cervelli, S. Kenner, and P. Segall. Correction to “dislocations in inhomogeneous media via a moduli perturbation approach: General formulation and two-dimensional solutions” by Yijun Du, Paul Segall, and Huajian Gao. *Journal of Geophysical Research*, 104(B10):23,271–23,277, 1999. doi:10.1029/1999JB900229.
- [17] M.A. Chinnery. The deformation of the ground around surface faults. *Bulletin of the Seismological Society of America*, 51(3), 1961.
- [18] M.A. Chinnery and D.B. Jovanovich. Effect of earth layering on earthquake displacement fields. *Bulletin of the Seismological Society of America*, 62(6), 1972.
- [19] P.G. Ciarlet. Basic error estimates for elliptic problems. In *Finite Element Methods (Part 1)*, volume 2 of *Handbook of Numerical Analysis*, chapter 2, pages 17–351. Elsevier, 1991. doi:10.1016/S1570-8659(05)80039-0.
- [20] M. Comninou. The angular dislocation in a half space. *Journal of Elasticity*, 5:203–216, 1975. doi:10.1007/BF00126985.
- [21] Y. Du, P. Segall, and H. Gao. Dislocations in inhomogeneous media via a moduli perturbation approach: General formulation and two-dimensional solutions. *Journal of Geophysical Research*, 99(B7):13767–13779, 1994. doi:10.1029/94JB00339.
- [22] A.M. Duncan, D.K. Chester, and J.E. Guest. Mount Etna volcano: Environmental impact and problems of volcanic prediction. *The Geographical Journal*, 147(2):164–178, 1981. doi:10.2307/634532.
- [23] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Mathematics of Computation*, 34(150):441–463, 1980. doi:10.1090/S0025-5718-1980-0559195-7.
- [24] A.M. Dziewonski, T.A. Chou, and J.H. Woodhouse. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal for Geophysical Research: Solid Earth*, 86(B4):2825–2852, 1981. doi:10.1029/JB086iB04p02825.

- [25] J. Elliott, R. Walters, and T. Wright. The role of space-based observation in understanding and responding to active tectonics and earthquakes. *Nature Communications*, 7(13844), 2016. doi:10.1038/ncomms13844.
- [26] A. Ern and J.L. Guermond. *Theory and Practice of Finite Elements*. Number 159 in Applied Mathematical Sciences. Springer-Verlag, New York, 2004.
- [27] L.C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 2009.
- [28] Y. Fialko, D. Sandwell, M. Simons, and et al. Three-dimensional deformation caused by the Bam, Iran, earthquake and the origin of shallow slip deficit. *Nature*, 435:295–299, 2005. doi:10.1038/nature03425.
- [29] G. Fichera. The Italian contribution to the mathematical theory of elasticity. *Meccanica*, 19(4), 1984. doi:10.1007/BF01556321.
- [30] United Nations Office for Disaster Risk Reduction. *Living with Risk: A global review of disaster reduction initiatives*. United Nations, New York, NY, 2004.
- [31] Y. Fukahata and T.J. Wright. A non-linear geodetic data inversion using ABIC for slip distribution on a fault with an unknown dip angle. *Geophysical Journal International*, 173(2):353–364, 2008. doi:10.1111/j.1365-246X.2007.03713.x.
- [32] R.J. Geller. Earthquake prediction: a critical review. *Geophysical Journal International*, 131(3):425–450, 1997. doi:10.1111/j.1365-246X.1997.tb06588.x.
- [33] R.J. Geller, D.D. Jackson, Y.Y. Kagan, and F. Mulargia. Earthquakes cannot be predicted. *Science*, 275(5306), 1997. doi:10.1126/science.275.5306.1616.
- [34] R. Gracie, G. Ventura, and T. Belytschko. A new fast finite element method for dislocations based on interior discontinuities. *International Journal for Numerical Methods in Engineering*, 69(2):423–441, 2007. doi:10.1002/nme.1896.
- [35] R.F. Hanssen. Atmospheric heterogeneities in ERS tandem SAR interferometry. *DEOS Report*, 98(1), 1998.
- [36] R.F. Hanssen. *Radar Interferometry: Data Interpretation and Error Analysis*. Remote Sensing and Digital Image Processing. Springer, 2001. doi:10.1007/0-306-47633-9.
- [37] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi:10.1093/biomet/57.1.97.
- [38] J.P. Hirth and J. Lothe. *Theory of dislocations*. Krieger Pub. Co., 4th edition, 1982.
- [39] T. Hughes and J.A. Evans. Isogeometric analysis. Technical Report 10-18, ICES, 2010.
- [40] T.J.R. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover, 2000.

- [41] J. Jackson. Fatal attraction: living with earthquakes, the growth of villages into megacities, and earthquake vulnerability in the modern world. *Philosophical Transactions of the Royal Society*, 364(1845):1911–1925, 2006. doi:10.1098/rsta.2006.1805.
- [42] H. Kanamori and E.E. Brodsky. The physics of earthquakes. *Physics Today*, 54(6), 2001. doi:10.1063/1.1387590.
- [43] R.A. Kerr. New work reinforces megaquake’s harsh lessons in geoscience. *Science*, 332(6032), 2011. doi:10.1126/science.332.6032.911.
- [44] G.C.P. King, R.S. Stein, and J. Lin. Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America*, 84(3):935–953, 1994. doi:10.1785/BSSA0840030935.
- [45] R. Krause and E. Rank. A fast algorithm for point-location in a finite element mesh. *Computing*, 57:49–62, 1996. doi:10.1007/BF02238357.
- [46] J.L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications I*. Springer-Verlag, Berlin, 1972.
- [47] A.E.H. Love. *A treatise on the mathematical theory of elasticity*. Dover publications, 4th edition, 1927.
- [48] M. Loève. *Probability Theory I*, chapter Elementary probability theory, pages 1–52. Springer, 1977. doi:10.1007/978-1-4684-9464-8.
- [49] D. Massonnet, M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl, and T. Rabaute. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, 364:138–142, 1993. doi:10.1038/364138a0.
- [50] E. Melan. Der Spannungszustand der durch eine Einzelkraft im Innern beanspruchten Halbscheibe. *Zeitschrift für Angewandte Mathematik und Mechanik*, 12(6):343–346, 1932. doi:10.1002/zamm.19320120603.
- [51] H.J. Melosh and A. Raefsky. A simple and efficient method for introducing faults into finite element computations. *Bulletin of the Seismological Society of America*, 71(5), 1981.
- [52] R.D. Mindlin. Force at a point in the interior of a semi-infinite solid. *Journal of Applied Physics*, 7(5):195–202, 1936. doi:10.1063/1.1745385.
- [53] R. Nason and J. Weertman. A dislocation theory analysis of fault creep events. *Journal of Geophysical Research*, 78(32):7745–7751, 1973. doi:10.1029/JB078i032p07745.
- [54] J.A. Nitsche. Über ein Variationsprinzip zur Lösung Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 36:9–15, 1971.
- [55] Y. Okada. Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, 75(4), 1985.

- [56] Y. Okada. Internal deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, 82(2):1018–1040, 1992.
- [57] D.A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer Verlag, 2012.
- [58] W. Prescott. Seeing earthquakes from afar. *Nature*, 364:100–101, 1993. doi:10.1038/364100d0.
- [59] G.O. Roberts, A. Gelman, and W.R. Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1):110–120, 1997. doi:10.1214/aoap/1034625254.
- [60] K. Rybicki. The elastic residual field of a very long strike-slip fault in the presence of a discontinuity. *Bulletin of the Seismological Society of America*, 61(1), 1971.
- [61] F.J. Sayas. Weak normal derivatives, normal and tangential traces and tangential differential operators on Lipschitz boundaries. Unpublished notes, available from this author on request, 2009.
- [62] P. Segall. *Earthquake and volcano deformation*. University Presses of California, Columbia and Princeton, 2010.
- [63] M. Simons, Y. Fialko, and L. Rivera. Coseismic deformation from the 1999 mw 7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations. *Bulletin of the Seismological Society of America*, 92(4):1390–1402, 2002. doi:10.1785/0120000933.
- [64] J.A. Steketee. On Volterra’s dislocations in a semi-infinite elastic medium. *Canadian Journal of Physics*, 36, 1958. doi:10.1139/p58-024.
- [65] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.
- [66] A.A. Tronin. Remote sensing and earthquakes: A review. *Physics and Chemistry of the Earth*, 31(4-9):138–142, 2006. doi:10.1016/j.pce.2006.02.024.
- [67] V. Volterra. Sur l’équilibre des corps élastiques multiplement connexes. *Annales scientifiques de l’École Normale Supérieure*, 24:401–517, 1907.
- [68] S.N. Ward. Quasi-static propagator matrices: creep on strike-slip faults. *Tectonophysics*, 120(1-2):83–106, 1985. doi:10.1016/0040-1951(85)90088-5.
- [69] R. Weldon, K. Scharer, T. Fumal, and G. Biasi. Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work. *GSA Today*, 14(9), 2004. doi:10.1130/1052-5173(2004)014<4:WATECW>2.0.CO;2.
- [70] C.A. Williams and G. Wadge. An accurate and efficient method for including the effects of topography in three-dimensional elastic models of ground deformation with applications to radar interferometry. *Journal of Geophysical Research*, 105(B4):8103–8120, 2000. doi:10.1029/1999JB900307.

- [71] W. Xu, R. Dutta, and S. Jónsson. Identifying active faults by improving earthquake locations with InSAR data and Bayesian estimation: The 2004 tabuk (saudi arabia) earthquake sequence. *Bulletin of the Seismological Society of America*, 105(2A):765–775, 2015. doi:10.1785/0120140289.
- [72] T. Yabuki and M. Matsu’ura. Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. *Geophysics Journal International*, 109(2):363–375, 1992. doi:10.1111/j.1365-246X.1992.tb00102.x.
- [73] E.H. Yoffe. The angular dislocation. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 5(50):161–175, 1960. doi:10.1080/14786436008243299.
- [74] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method*, volume 2: Solid Mechanics. Butterworth Heinemann, 5th edition, 2000.
- [75] O.C. Zienkiewicz, R.L. Taylor, and J.Z. Zhu. *The Finite Element Method: Its Basis and Fundamentals*. Butterworth-Heinemann, 2013.
- [76] G.J. van Zwieten, E.H. van Brummelen, and R.F. Hanssen. Inverting elastic dislocations using the weakly-enforced slip method. *International Journal for Numerical and Analytical Methods in Geomechanics*, 2022. doi:10.1002/nag.3365.
- [77] G.J. van Zwieten, E.H. van Brummelen, K.G. van der Zee, M.A. Gutiérrez, and R.F. Hanssen. Discontinuities without discontinuity: The weakly-enforced slip method. *Computer Methods in Applied Mechanics and Engineering*, 271:144–166, 2014. doi:10.1016/j.cma.2013.12.004.
- [78] G.J. van Zwieten, R.F. Hanssen, and M.A. Gutiérrez. Overview of a range of solution methods for elastic dislocation problems in geophysics. *Journal for Geophysical Research: Solid Earth*, 118(4):1721–1732, 2013. doi:10.1029/2012JB009278.

Summary

A vital component in the management of seismic hazard is the study of past seismic events. Classically, this has been the domain of seismology, which studies the dynamic manifestations of the event to infer properties such as epicenter and moment magnitude. More recently it has become possible to perform similar analyses on the basis of the static consequences of a seismic event, as satellite borne Synthetic Aperture Radar (SAR) data allows us to compare the local surface geometries before and after a seismic event. The locality of the deformation data promises reconstructions with greater detail and subject to fewer model uncertainties.

With current technology, it is not possible to use SAR to their full potential. The non-linearity of the static dislocation problem that links faulting mechanisms to observed deformations causes any inverse method to require many evaluations of the forward model. This poses limits on the permissible cost of solving the dislocation problem, restricting most approaches to simplified model assumptions such as material homogeneity and absence of topography. In situations where more accurate information is available, this presents a clear opportunity for improvement by accelerating the computational methods instead.

This thesis presents the Weakly-enforced Slip Method (WSM), a modification of the Finite Element Method (FEM), as a fast approach for solving static dislocation problems. While the computational cost of the WSM is similar to that of the FEM for single dislocations, the WSM is significantly faster when many different dislocation geometries are considered, owing to the reuse of computationally expensive components such as matrix factors. This property makes the method ideally suited for inverse settings, opening the way to incorporating all available in situ data in a forward model that is simultaneously flexible and cheaply evaluable. Moreover, we prove that the WSM retains the essential convergence properties of the FEM.

A limitation of the WSM is that it produces continuous displacement fields, which implies a large error local to the dislocation. We show that this error decreases rapidly with distance, and that in a typical scenario the majority of deformation data has a discretization error that is smaller than observational noise, particularly when a fault is buried. In the case of shallow or rupturing faults, neighbouring data needs to be discarded from the analysis to avoid disruption. With this measure in place, we show via Bayesian inference of synthesized datasets that the discretization errors of the WSM do not significantly affect the inverse problem.

Samenvatting

Een belangrijk onderdeel van het in kaart brengen van seismisch risico is de bestudering van de lokale historie van seismische gebeurtenissen. Dit is van oudsher het domein van seismologie: de discipline die de dynamiek van een beving analyseert om hieruit karakteristieken te destilleren als epicentrum en magnitude. Met het beschikbaar komen van satelliet-gebaseerde Synthetic Aperture Radar (SAR) is het mogelijk geworden om soortgelijke studies uit te voeren op basis van de statische gevolgen van een seismische gebeurtenis, door de blijvende vervorming van het aardoppervlak te nemen als primaire waarneming. Het lokale karakter van deze data belooft reconstructies met groter detail op basis van minder ingrijpende modelaannames.

Met de huidige technologie is het niet mogelijk om SAR ten volle te benutten. De niet-lineariteit van het statische dislocatieprobleem, nodig om aarbevingsmechanismen te koppelen aan waargenomen vervormingen, maakt dat elke inverse-methode zeer veel evaluaties van het voorwaartse model vereist. Dit stelt praktische grenzen aan de kosten van het dislocatieprobleem, dat om deze reden doorgaans sterk wordt versimpeld, leidend tot veelal onrealistische aannames als materiële homogeniteit en afwezigheid van topografie. In situaties waarin nauwkeurigere informatie beschikbaar is ligt hier evident potentie voor verbetering, door in plaats van het model te versimpelen de rekenmethode te versnellen.

Dit proefschrift presenteert de Weakly-enforced Slip Method (WSM), een modificatie van de eindige elementen-methode (Finite Element Method, FEM) voor het efficient oplossen van statische dislocatieproblemen. Hoewel de rekenkosten van de WSM vergelijkbaar zijn met die van de FEM bij de analyse van individuele dislocaties, is de WSM aanzienlijk sneller wanneer een veelvoud van dislocatiegeometrieën wordt beschouwd door mogelijk hergebruik van computationeel kostbare componenten als matrixfactoren. Deze eigenschap maakt de WSM bij uitstek geschikt voor inverse-problemen, die hiermee meer in-situ data in het model kunnen betrekken. Bovendien bewijzen we dat essentiële convergentie-eigenschappen van de FEM behouden blijven.

Een belangrijke beperking van de WSM is dat het verplaatsingsvelden genereert die continu zijn over het domein, wat een grote fout impliceert ter plaatse van de dislocatie. We laten zien dat deze fout snel afneemt met de afstand tot de dislocatie, en op de meeste plaatsen typisch kleiner is dan de waarnemingsruis, zeker wanneer het breukvlak niet tot de oppervlakte reikt. Wanneer dit wel het geval is kan meetdata uit de directe omgeving uit de analyse worden verwijderd om verstoringen te voorkomen. Op basis van een vergelijkend onderzoek met gesynthetiseerde gegevens laten we zien dat deze ingreep afdoende is om te voorkomen dat de discretisatiefouten van de WSM het inverseprobleem significant beïnvloeden.

Curriculum Vitae

Gerrit Johannes (Gertjan) van Zwieten
born July 22, 1982 in Moordrecht, The Netherlands.

Education and experience

- 2013–present Founder, Evalf (formerly HvZ Engineering).
- 2010–2013 Researcher, faculty of Mechanical Engineering, section Multiscale Engineering Fluid Dynamics (MEFD), Eindhoven University of Technology.
- 2006–2010 Ph.D. candidate, faculty of Aerospace Engineering, sections Engineering Mechanics (EM) and Mathematical Geodesy and Positioning (MGP), Delft University of Technology.
- 2004–2006 M.Sc. in Technical Mathematics, Delft University of Technology.
- 2000–2004 B.Sc. in Aerospace Engineering, Delft University of Technology.
- 1994–2000 Coornhert Gymnasium Gouda.

Relevant publications

- 2022 Inverting elastic dislocations using the Weakly-enforced Slip Method, G.J. van Zwieten, E.H. van Brummelen and R.F. Hanssen, International Journal for Numerical and Analytical Methods in Geomechanics [76].
- 2014 Discontinuities without discontinuity: The Weakly-enforced Slip Method, G.J. van Zwieten, E.H. van Brummelen, K.G. van der Zee, M.A. Gutiérrez and R.F. Hanssen, Computer Methods in Applied Mechanics and Engineering [77].
- 2013 Overview of a range of solution methods for elastic dislocation problems in geophysics, G.J. van Zwieten, R.F. Hanssen and M.A. Gutiérrez, Journal for Geophysical Research: Solid Earth [78].